

Universidade de São Paulo  
INSTITUTO DE ASTRONOMIA, GEOFÍSICA E CIÊNCIAS ATMOSFÉRICAS  
DEPARTAMENTO DE CIÊNCIAS ATMOSFÉRICAS

RUBINEI DORNELES MACHADO

Estabelecimento, modelagem e caracterização de regiões críticas de velocidade do vento e temperatura do ar para o despacho de carga sobre o Subsistema Nordeste do Sistema Interligado Nacional

SÃO PAULO

2020

RUBINEI DORNELES MACHADO

Caracterização e modelagem de zonas de estagnação de vento do nordeste brasileiro em  
escala temporal e horizontal

(Versão Corrigida)

Tese de Doutorado apresentada ao Departamento de Ciências  
Atmosféricas para obtenção do título de Doutor em Ciências

Área de Concentração: Meteorologia

Orientador: Prof. Dr. Ricardo Hallak

SÃO PAULO

2020

À Suelen dos Santos Saraiva



## AGRADECIMENTOS

Ao prof. Dr. Ricardo Hallak que com confiança, compreensão e amizade aceitou o desafio dessa orientação mesmo com todas as particularidades que o desenvolvimento deste trabalho impôs.

Ao prof. Sergio Colle pela oportunidade de desenvolvimento profissional e pessoal e, por ter me inserido na comunidade científica internacional por meio da BSRN. Obrigado por ser o meu melhor exemplo de cientista mesmo em um país com tantas adversidades.

À minha mãe, Maria Cecília Dorneles Machado e, ao meu pai, Rubinei Silva Machado (*in memoriam*) por todo o incentivo e dificuldades enfrentadas para possibilitar a minha formação.

À minha esposa, Suelen Saraiva, pelo suporte emocional e financeiro para o desenvolvimento desta tese. Sua compreensão, incentivo, companheirismo e amor foram fundamentais.

À CAPES e ao CNPq pelo auxílio financeiro durante a realização do doutorado.

À Coordenação Geral de Meteorologia Aplicada, Desenvolvimento e Pesquisa do Instituto Nacional de Meteorologia, Sr. Expedito Rebello e Dra. Andrea Malheiros Ramos, pelo fornecimento dos dados da rede de estações meteorológicas automáticas.

À Seção de Informática e secretárias do DCA pelo suporte durante o desenvolvimento desta tese. E, em especial, ao Sebastião pelas excelentes discussões sobre o cenário político nacional.

Ao LEPTEN/UFSC por todo o suporte, infraestrutura e compreensão durante os longos períodos de ausência. Aos amigos e colegas engenheiros Júlio Scussel, Allan Starke, Leonardo Lacerda e Luigi Passos, técnicos Maycon Macedo e Mateus Stahelin e à Rosângela Avi e Gabriel Bravo por todo apoio durante esse período.

Ao amigo e professor Dr. José Cardemil pelo convite e interação com a Universidade do Chile.

Aos professores Dr. Michel Nobre Muza, Dr. Mário Quadro e Dr. Luiz Nakamura pelas contribuições e discussões científicas durante o desenvolvimento deste trabalho.

À toda família Os de Paula – Levi, Fabíola, Vinicius, Fabiana, Eduardo, Edson, Andinho Matos, Guilherme, Carol e Pâmela por me incluírem nos seus dias a dia e me ensinarem a amar esta cidade. Todos os debates que permearam os diversos tipos de pré-conceitos existentes na sociedade e a nossa negritude me tornaram uma pessoa melhor. Vocês foram fundamentais e serão inesquecíveis.



## RESUMO

Machado, Rubinei D. **Estabelecimento, modelagem e caracterização de regiões críticas de velocidade do vento e temperatura do ar para o despacho de carga sobre o Subsistema Nordeste do Sistema Interligado Nacional**. 2020. 154p. Tese (Doutorado em Ciências) – Instituto de Astronomia, Geofísica e Ciências Atmosféricas, Universidade de São Paulo, São Paulo, 2020. Versão Corrigida.

A temperatura do ar e a velocidade do vento interferem diretamente no despacho de carga em Linhas Aéreas de Transmissão (LT). Conhecer o comportamento destas variáveis possibilita ao setor elétrico avaliar o melhor aproveitamento das linhas existentes bem como direcionar novos projetos de construção. Assim, a presente tese investiga o regime de vento e temperatura do ar próximo à superfície com o objetivo de identificar, caracterizar e simular as zonas de calmaria de vento e de máximos de temperatura. Esta caracterização foi realizada por meio da análise de cluster para a qual foram utilizados dados horários medidos entre 2005 e 2018 em 125 estações meteorológicas do INMET instaladas sobre o subsistema Nordeste do SIN. Ao final desta análise agruparam-se 35 estações em 7 clusters. Para preencher os dados faltantes das séries temporais destas estações avaliou-se a imputação de dados em falhas geradas artificialmente para o ano de 2017 utilizando 4 diferentes métodos. Os resultados apontaram que o sucesso no preenchimento de séries temporais incompletas utilizando a interpolação linear pode subsidiar estudos de ampacidade a partir de valores mais fidedignos das condições críticas ao longo de uma LT. Adicionalmente, os métodos ERA5 e WRF utilizados na avaliação do preenchimento também podem ser utilizados com eficiência para estimar a temperatura do ar quando não há dados medidos por estações meteorológicas próximas as LT. A caracterização da frequência de ocorrência das máximas temperaturas do ar demonstrou que o período de maior restrição para o despacho de carga trata-se da primavera e durante a tarde. Para a velocidade do vento apenas o ciclo diurno apresenta variações significativas entre os períodos analisados. E a maior restrição considerando o vento ocorre para todos os períodos do ciclo diurno com exceção da tarde. Desta forma, a capacidade operativa de LT pode ser flexibilizada na escala sazonal sobre o subsistema Nordeste do SIN. Adicionalmente, o estabelecimento das regiões críticas de mínimos de vento e máximas temperaturas e a caracterização destas pode contribuir para o desenvolvimento de projetos de novas LT que considerem condições meteorológicas mais representativas da região de interesse.

Palavras-chave: Variabilidade sazonal. Regimes de vento. Análise de agrupamento. Imputação de dado faltante. Linhas Aéreas de Transmissão. Ampacidade.





## Abstract

Machado, Rubinei D. **Establishment, modeling and characterization of critical regions of wind speed and air temperature for power dispatch over the Northeast Subsystem of the National Interconnected System.** 2020. 154p. Thesis (Ph.D.) – Department of Atmospheric Sciences, University of São Paulo, São Paulo. 2020. Revision version.

The air temperature and wind speed are limiting factors for the power dispatch in power transmission lines (TL). And determining the behavior of these variables allows the Electric Sector to evaluate the efficient use of the TL in operation and support the development new TL projects. Thus, the present work investigated the low wind speed at 10 meters and high air temperature at 2 meters in order to identify, characterize and simulate the regimes of this variables. This study used cluster analysis to find similar behavior of the meteorological conditions (low wind speed and high temperature) between 125 weather stations from the INMET distributed about Northeast SIN subsystem. The time series have hourly temporal resolution and about 13 years (2005-2018) of data. The seven clusters were obtained from 35 meteorological stations. In order to fill in the missing data from the time series of these stations, the imputation of data in artificially generated gaps for the year 2017 was evaluated using 4 different methods. The results show that the success in filling artificial gaps using linear interpolation may contribute to support studies of ampacity from more reliable values of critical conditions close to TL. Additionally, the ERA5 and WRF methods used in the gap filling evaluation can also be used efficiently to estimate the air temperature when there is no data measured by meteorological stations close to the LT. The characterization of the frequency of occurrence of maximum air temperature demonstrated the period of greatest restriction for the power dispatch was found to spring and during the afternoon. For the wind speed only diurnal cycle show significant variations between the periods analyzed. The threshold of 1 m/s for wind speed during the diurnal cycle was only exceeded in the afternoon. In other words, the other periods of the day are critical for TL due to the low wind speeds. Thus, the TL operating capacity can be optimized on a seasonal scale over the Northeastern SIN subsystem. In addition, the establishment of critical regions of minimum wind and maximum temperatures and their characterization may contribute to the development of new TL projects that consider more representative meteorological conditions in the region of interest.

Keywords: Seasonal variability. Wind regime. Cluster analysis. Imputation of missing values. Power transmission lines. Ampacity.



## SUMÁRIO

<b>1 INTRODUÇÃO</b> .....	18
1.1 Objetivos da tese .....	20
1.1.1 Objetivos Específicos .....	21
<b>2 REVISÃO DE LITERATURA</b> .....	22
2.1 Sistema elétrico brasileiro – cenário atual.....	22
2.1.1 Sistema Interligado Nacional .....	23
2.2 Sistemas atmosféricos que atuam sobre o subsistema nordeste do SIN.....	33
<b>3 DADOS E METODOLOGIA</b> .....	42
3.1 Área de estudo .....	42
3.2 Dados.....	43
3.2.1 Dados meteorológicos medidos .....	43
3.2.2 Série temporal sintética de dados meteorológicos .....	46
3.2.3 Banco de dados relacional .....	48
3.3 Controle de qualidade dos dados .....	50
3.4 Análise de cluster.....	50
3.5 Preenchimento de falhas .....	57
3.5.1 Imputação de dados .....	59
3.5.2 Avaliação de desempenho .....	62
<b>4 RESULTADOS</b> .....	66
4.1 Análise de preenchimento de falhas artificiais.....	66
4.1.1 Cluster 1 .....	66
4.1.2 Cluster 2 .....	73
4.1.3 Cluster 3.....	78
4.1.4 Cluster 4.....	85
4.1.5 Cluster 5.....	93
4.1.6 Cluster 6.....	99
4.1.7 Cluster 7.....	107
4.2 Caracterização dos mínimos de vento e máximos de temperatura do ar .....	113
4.3 Síntese dos resultados .....	125

<b>5 CONCLUSÕES</b> .....	133
<b>6 SUGESTÕES PARA TRABALHOS FUTUROS</b> .....	137
<b>7 REFERÊNCIAS</b> .....	139

## LISTA DE TABELAS

Tabela 3.1 – Estações meteorológicas do INMET com mediana da mínima velocidade diária do vento $\leq 1$ m/s e máxima temperatura diária do ar $\geq 31,5$ °C, localizadas no subsistema nordeste do SIN em 2018.....	45
Tabela 3.2 – Correlação cofenética dos pares de distância métrica e métodos de ligação utilizados para a construção de diferentes árvores hierárquicas utilizando o conjunto de variável V e T.....	55
Tabela 3.3 – Agrupamentos formados por meio da análise de cluster considerando a similaridade da mediana das máximas temperaturas diárias do ar e da mediana das mínimas velocidades do vento.....	56
Tabela 3.4 – Exemplo hipotético das informações das falhas artificiais geradas para 10% dos dados de uma variável meteorológica qualquer pertencente a uma estação meteorológica. ...	57
Tabela 3.5 - Tabela de contingência hipotética para determinação da performance dos estimadores.....	63
Tabela 4.1 – Estações meteorológicas agrupadas no cluster 1 e a mediana das máximas temperaturas diárias do ar ( $T_d$ ) e das mínimas velocidades diárias do vento ( $V_d$ ).....	67
Tabela 4.2 – Métricas de verificação do desempenho do preenchimento de falhas geradas artificialmente para o cluster-1. ....	70
Tabela 4.3 – Estações meteorológicas agrupadas no cluster-2 e a mediana das máximas temperaturas diárias do ar ( $T_{med}$ ) e das mínimas velocidades diárias do vento ( $V_{med}$ ). ....	73
Tabela 4.4 – Métricas de verificação do desempenho do preenchimento de falhas geradas artificialmente para o cluster-2. ....	76
Tabela 4.5 – Estações meteorológicas agrupadas no cluster-3 e a mediana das máximas temperaturas diárias do ar ( $T_{med}$ ) e das mínimas velocidades diárias do vento ( $V_{med}$ ). ....	79
Tabela 4.6 – Métricas de verificação do desempenho do preenchimento de falhas geradas artificialmente para o cluster-3. ....	83
Tabela 4.7 - Estações meteorológicas agrupadas no cluster-4 e a mediana das máximas temperaturas diárias do ar ( $T_{med}$ ) e das mínimas velocidades diárias do vento ( $V_{med}$ ). ....	86
Tabela 4.8 - Métricas de verificação do desempenho do preenchimento de falhas geradas artificialmente para o cluster-4. ....	90

Tabela 4.9 - Estação meteorológica do cluster-5 e a mediana das máximas temperaturas diárias do ar ( $T_{med}$ ) e das mínimas velocidades diárias do vento ( $V_{med}$ ). .....	93
Tabela 4.10 – Métricas de verificação do desempenho do preenchimento de falhas geradas artificialmente para o cluster-5. ....	97
Tabela 4.11 – Estações meteorológicas agrupadas no cluster-6 e a mediana das máximas temperaturas diárias do ar ( $T_{med}$ ) e das mínimas velocidades diárias do vento ( $V_{med}$ ). ....	100
Tabela 4.12 - Métricas de verificação do desempenho do preenchimento de falhas geradas artificialmente para o cluster-6. ....	104
Tabela 4.13 – Estações meteorológicas agrupadas no cluster-7 e a mediana das máximas temperaturas diárias do ar ( $T_{med}$ ) e das mínimas velocidades diárias do vento ( $V_{med}$ ). ....	107
Tabela 4.14 – Métricas de verificação do desempenho do preenchimento de falhas geradas artificialmente para o cluster-7. ....	111
Tabela 4.15 - Métodos de preenchimento das séries temporais brutas da temperatura do ar e velocidade do vento de acordo com o índice de probabilidade de detecção. ....	128

## LISTA DE FIGURAS

Figura 2.1 – Evolução da matriz energética do subsistema Nordeste.....	24
Figura 2.2 – Sistema Interligado Nacional atual e projeções.....	26
Figura 2.3 – Estimativa de investimentos em construção de linhas de transmissão e capacidade de transformação em cada ciclo quinquenal do PAR de 2009 a 2023. ....	26
Figura 2.4 – Distâncias esquemáticas de cabo-solo $D_{NP}$ , $D_{LP}$ e $D_{mP}$ e as respectivas temperaturas $T_{NP}$ , $T_{LP}$ e $T_{mP}$ de um cabo condutor. ....	28
Figura 2.5 – Influência das condições atmosféricas na temperatura do cabo condutor e na corrente elétrica despachada em uma linha de transmissão.....	29
Figura 2.6 – Balanço de energia em um cabo condutor.....	30
Figura 2.7 – Variação da ampacidade em função da velocidade do vento e da temperatura do ar.....	31
Figura 2.8 – Localização das circulações locais sobre o subsistema Nordeste do SIN. ....	36
Figura 2.9 – Desenvolvimento esquemático da circulação de brisa marítima. Letra Q indica superfície mais aquecida do que a superfície F (fria). ....	37
Figura 2.10 – Fluxo esquemático da brisa vale-montanha durante (a) o período da manhã e (b) da tarde sobre a cidade de Petrolina-PE.....	38
Figura 2.11 – Influência do planalto da Borborema no transporte de umidade associado a brisa vale-montanha. ....	39
Figura 3.1- Linhas de transmissão planejadas (linhas contínuas magenta), existentes (linhas contínuas verde) e estações meteorológicas (pontos vermelhos) situadas no Subsistema Nordeste do SIN em 2019.....	42
Figura 3.2 - Mediana das mínimas velocidades diárias do vento indicadas pelo tamanho dos círculos e das máximas temperaturas diárias do ar representadas pelas cores.....	44
Figura 3.3 – Localização das estações meteorológicas com mediana da velocidade mínima do vento igual ou menor a 1 m/s e temperatura máxima do ar maior ou igual a 31,5 °C.....	45
Figura 3.4 – Domínio da simulação com a grade de 5 km.....	48
Figura 3.5 – Tabela <i>station</i> contendo o metadados de doze estações meteorológicas do INMET.....	49
Figura 3.6 – Dendrograma formado utilizando a distância métrica (a) <i>cityblock</i> e método de ligação <i>ward</i> e (b) a distância euclidiana e o método de pesos.....	54

Figura 3.7 – Dendrograma construído utilizando a distância euclidiana como medida de similaridade e método de ligação <i>média</i> aplicado ao conjunto de variáveis das 35 estações meteorológicas presentes na tabela 3.1.....	55
Figura 3.8 – Distribuição espacial dos agrupamentos formados a partir do método hierárquico considerando 35 estações meteorológicas da rede INMET.....	56
Figura 3.9 – Domínio da grade e interpolação espacial utilizando o método ordinário kriging para estimar o dado de (a) velocidade do vento e (b) temperatura do ar para a estação meteorológica Barra (triângulo invertido) a partir de 28 estações meteorológicas mais próximas (sinal de positivo). .....	61
Figura 4.1 – Localização das estações meteorológicas que formaram o cluster-1.....	67
Figura 4.2 – Séries temporais brutas da (a) temperatura do ar e (b) velocidade do vento das estações do INMET pertencentes ao cluster-1. ....	68
Figura 4.3 – Média dos valores removidos artificialmente (linha tracejada) e estimados por 4 diferentes métodos para (a) temperatura do ar e (b) velocidade do vento para o cluster-1.....	69
Figura 4.4 – Diagrama de dispersão e reta de regressão da Probabilidade de Detecção média dos valores médios removidos artificialmente das séries temporais do cluster-1 para (a) temperatura do ar e (b) velocidade do vento considerando diferentes tamanhos de falhas.....	71
Figura 4.5 – Localização das estações meteorológicas que formaram o cluster-2.....	73
Figura 4.6 – Séries temporais brutas da (a) temperatura do ar e (b) velocidade do vento das estações do INMET pertencentes ao cluster-2. ....	74
Figura 4.7 – Média dos valores removidos artificialmente (linha tracejada) e estimados por 4 diferentes métodos para (a) temperatura do ar e (b) velocidade do vento para o cluster-2.....	76
Figura 4.8 – Diagrama de dispersão e reta de regressão da Probabilidade de Detecção média dos valores médios removidos artificialmente das séries temporais do cluster-2 para (a) temperatura do ar e (b) velocidade do vento considerando diferentes tamanhos de falhas.....	78
Figura 4.9 – Localização das estações meteorológicas que formaram o cluster-3.....	79
Figura 4.10 – Séries temporais brutas da (a) temperatura do ar e (b) velocidade do vento das estações do INMET pertencentes ao cluster-3. ....	80
Figura 4.11 – Média dos valores removidos artificialmente (linha tracejada) e estimados por 4 diferentes métodos para (a) temperatura do ar e (b) velocidade do vento para o cluster-3.....	81
Figura 4.12 – Diagrama de dispersão e reta de regressão da Probabilidade de Detecção média dos valores médios removidos artificialmente das séries temporais do cluster-3 para (a) temperatura do ar e (b) velocidade do vento considerando diferentes tamanhos de falhas.....	84
Figura 4.13 - Localização das estações meteorológicas que formaram o cluster-4. ....	86



Figura 4.14 - Séries temporais brutas da (a) temperatura do ar e (b) velocidade do vento das estações do INMET pertencentes ao cluster-4.....	88
Figura 4.15 - Média dos valores removidos artificialmente (linha tracejada) e estimados por 4 diferentes métodos para (a) temperatura do ar e (b) velocidade do vento para o cluster-4. ....	90
Figura 4.16 - Diagrama de dispersão e reta de regressão da Probabilidade de Detecção média dos valores médios removidos artificialmente das séries temporais do cluster-4 para (a) temperatura do ar e (b) velocidade do vento considerando diferentes tamanhos de falhas.....	92
Figura 4.17 - Localização da estação meteorológica que representou o cluster-5.....	93
Figura 4.18 - Série temporal bruta da (a) temperatura do ar e (b) velocidade do vento das estações do INMET pertencente ao cluster-5.....	94
Figura 4.19 – Média dos valores removidos artificialmente (linha tracejada) e estimados por 4 diferentes métodos para (a) temperatura do ar e (b) velocidade do vento para o cluster-5. ....	96
Continuação – Figura 4.20 – Média dos valores removidos artificialmente (linha tracejada) e estimados por 4 diferentes métodos para (a) temperatura do ar e (b) velocidade do vento para o cluster-5.....	97
Figura 4.21 - Diagrama de dispersão e reta de regressão da Probabilidade de Detecção média dos valores médios removidos artificialmente da série temporal do cluster-5 para (a) temperatura do ar e (b) velocidade do vento considerando diferentes tamanhos de falhas.....	98
Continuação – Figura 4.22 - Diagrama de dispersão e reta de regressão da Probabilidade de Detecção média dos valores médios removidos artificialmente da série temporal do cluster-5 para (a) temperatura do ar e (b) velocidade do vento considerando diferentes tamanhos de falhas. ....	99
Figura 4.23 – Localização das estações meteorológicas que formaram o cluster-6. ....	100
Figura 4.24 - Séries temporais brutas da (a) temperatura do ar e (b) velocidade do vento das estações do INMET pertencente ao cluster-6.....	101
Continuação – Figura 4.25 - Séries temporais brutas da (a) temperatura do ar e (b) velocidade do vento das estações do INMET pertencente ao cluster-6.....	102
Figura 4.26 - Média dos valores removidos artificialmente (linha tracejada) e estimados por 4 diferentes métodos para (a) temperatura do ar e (b) velocidade do vento para o cluster-6. ...	104
Figura 4.27 - Diagrama de dispersão e reta de regressão da Probabilidade de Detecção média dos valores médios removidos artificialmente das séries temporais do cluster-6 para (a) temperatura do ar e (b) velocidade do vento considerando diferentes tamanhos de falhas... ..	106
Figura 4.28 - Localização das estações meteorológicas que formaram o cluster-7.....	107

Figura 4.29 - Séries temporais brutas da (a) temperatura do ar e (b) velocidade do vento das estação do INMET pertencente ao cluster-7.....	108
Figura 4.30 - Média dos valores removidos artificialmente (linha tracejada) e estimados por 4 diferentes métodos para (a) temperatura do ar e (b) velocidade do vento para o cluster-7....	110
Figura 4.31 – Diagrama de dispersão e reta de regressão da Probabilidade de Detecção média dos valores médios removidos artificialmente das séries temporais do cluster-7 para (a) temperatura do ar e (b) velocidade do vento considerando diferentes tamanhos de falhas....	112
Figura 4.32 – Distribuição de frequência das temperaturas máximas diárias na escala de tempo sazonal e diurna para o (a) cluster-1, (b) cluster-2, (c) cluster-3, (d) cluster-4, (e) cluster-5, (f) cluster-6 e (g) cluster-7.....	116
Figura 4.33 – Distribuição de frequência das mínimas velocidades diárias do vento na escala de tempo sazonal e diurna para o (a) cluster-1, (b) cluster-2, (c) cluster-3, (d) cluster-4, (e) cluster-5, (f) cluster-6 e (g) cluster-7.....	122
Figura 4.34 – Diagrama de Taylor para (a) temperatura do ar e (b) velocidade do vento considerando os quatro métodos de estimava para cada cluster e usando os valores removidos artificialmente (medido) como referência. ....	127

## 1 INTRODUÇÃO

As características do vento em superfície tais como direção e velocidade média, rajadas, persistência e variabilidade contribuem para estabelecer as potenciais atividades econômicas, sociais e de desenvolvimento de uma região. O conhecimento do regime do vento de uma localidade pode determinar a viabilidade da instalação de usinas eólicas para a geração de energia, de linhas de transmissão para o aumento do despacho de carga, de indústrias no tocante a dispersão de poluentes, de plataformas *offshore* para exploração de petróleo, do material a ser utilizado em edificações e a ocupação territorial da população.

No tocante ao sistema elétrico brasileiro que se divide em geração, transmissão, distribuição e comercialização de energia, o vento também tem um papel fundamental. Isso porque a ampliação de parques eólicos é uma alternativa à forte dependência das hidroelétricas. Por exemplo, em 2018 as usinas hidrelétricas possuíam a capacidade de gerarem 67% da demanda energética do país (ONS, 2019a). Embora essa seja uma fonte limpa e renovável, a variabilidade da precipitação impacta diretamente na variação do preço da energia. Ou seja, em períodos com déficit hídrico há um aumento da participação das usinas térmicas para suprir a demanda energética. E a geração térmica por sua vez além de ser uma fonte poluente, também opera com um custo maior e, portanto, eleva o preço da energia para o mercado consumidor. Dessa forma, aumentar a diversificação da matriz energética a partir da inserção da energia eólica no sistema contribui para uma operação mais robusta e menos dependente.

Adicionalmente, o Brasil conta com uma extensa malha de linhas de transmissão, em sua maioria aéreas, que estão interconectadas e permitem que a energia gerada em uma região possa ser consumida em outra independente da distância. Essa infraestrutura de transmissão em conjunto com a de geração forma o Sistema Interligado Nacional (SIN). É a arquitetura do SIN que garante maior robustez à operação do sistema. Estas interconexões evitam a falta de fornecimento de energia para uma determinada região seja por questões de tempo e clima ou por falhas humanas ou de equipamentos (ONS, 2019a). No entanto, o aumento da capacidade instalada de geração, independentemente do tipo da fonte (hidro, solar, eólica ou outra), também se traduz na maior demanda por despacho da energia produzida.

Por outro lado, a quantidade de energia (também chamada de corrente ou carga) que pode ser transmitida em uma linha é limitada pela temperatura do cabo condutor e pela altura mínima entre o cabo condutor e o solo (FRONTIN et al., 2010). E esses valores são definidos na fase do projeto da construção da linha e, além de fatores físicos, também consideram as condições meteorológicas que influenciam na troca de calor entre o cabo e o meio ambiente. É

basicamente por esse motivo que linhas de transmissão, teoricamente, não podem despachar corrente adicional à estabelecida anteriormente, pois aumenta o risco de elevar a temperatura do cabo condutor definida no projeto a ponto de violar a distância mínima de segurança entre o cabo e o solo aumenta. E como consequências dessa violação se pode mencionar as perdas econômicas geradas pelo desligamento da linha e morte de animais em propriedades rurais ou até mesmo humanas caso ocorra em área urbana ou de manejo rural.

O plano de ação e reforços nas instalações do SIN elaborado pelo ONS (Operador Nacional do Sistema Elétrico) e divulgado em 2019 estima que de 2017 a 2023 a capacidade máxima de geração do país terá um aumento de aproximadamente 25% (ONS, 2019a). E para atender à essa demanda crescente o ONS prevê a construção de cerca de 6.800 km de linhas a um custo de R\$ 13,6 bilhões de reais (ONS, 2019a). No entanto, o relatório de acompanhamento diferenciado da expansão da transmissão do SIN divulgado em março de 2019 pela ANEEL (Agência Nacional de Energia Elétrica) indica que entre 2014 e 2018 o tempo médio para obtenção de licenciamento ambiental para construção de novas linhas foi de 370 dias, enquanto de execução de obras físicas somou 536 dias (ANEEL, 2019). Neste mesmo relatório, a ANEEL aponta que, em março de 2019, 35% dos 141 empreendimentos relativos à construção de linhas estavam com previsão de atraso para conclusão das obras. Esse fato, pode postergar o início da operação de usinas de geração de energia recém-inauguradas, pois limita a capacidade do SIN de despachar a carga adicional.

Portanto, diante das dificuldades e custos apresentados para a implantação de novas linhas de transmissão cabe o seguinte questionamento: o aumento da capacidade de transmissão das linhas existentes poderia ser uma alternativa para 1) diminuir ou postergar os investimentos, os quais tem custo médio de R\$ 2 milhões de reais por quilômetro de linha construída (ONS, 2019a) e 2) diminuir o impacto causado pelo atraso da data de início de operação da linha?

Alguns estudos como DACONTI e LAWRY (2003) têm apontado que é possível aumentar a capacidade de despacho de uma linha existente mantendo-se a confiabilidade e qualidade da prestação do serviço de transmissão. A respeito do quanto pode ser esse aumento NTULI *et al.* (2016) afirmaram que a capacidade de carga que pode ser transmitida em uma linha, excetuando-se as propriedades físico-mecânicas dos cabos e estruturas de transmissão, tem como fator determinante a distância de segurança entre o cabo condutor e o solo. E as condições meteorológicas ao longo de uma linha desempenham um papel termorregulador importante para manter essa distância. Em outras palavras, o estabelecimento da carga possível de ser transmitida deve considerar as condições atmosféricas adversas para o resfriamento de um cabo condutor que ocorrem próximas as linhas de transmissão. Tais condições podem

dificultar o resfriamento do cabo condutor por convecção se houver baixa velocidade do vento e aumentar a temperatura do cabo caso a temperatura do ar e radiação solar estejam elevadas. Nesse sentido, a norma brasileira que estabelece os procedimentos para projeto de linhas aéreas de transmissão de energia elétrica (ABNT, 1985) recomenda que, para determinar a temperatura máxima a qual um cabo condutor pode ser submetido sem violar a distância mínima de segurança entre o cabo e o solo, deve-se considerar as condições ambientais (vento, temperatura do ar e radiação solar), características físicas e mecânicas dos condutores e da corrente elétrica. No entanto, há dificuldade para se obter dados meteorológicos próximos a linhas de transmissão, sobretudo no Brasil. E nesse caso, ainda segundo a ABNT (1985), na falta de dados meteorológicos medidos em linhas a norma recomenda que se deve considerar a velocidade do vento não superior a 1 m/s, a temperatura do ar sendo a média das máximas diárias registradas na estação mais próxima e radiação solar de 1000 W/m<sup>2</sup>.

A adoção dos valores fixos e de ocorrência simultânea de variáveis meteorológicas para a determinação da corrente máxima de um cabo condutor se traduz em carregamentos de energia conservadores (CIGRE, 2004) a fim de garantir a segurança da operação de tal linha. Porém, esse fato pode levar à subutilização de diversas linhas de transmissão e, conseqüentemente, a investimentos em construção de novas linhas que poderiam ser evitados ou postergados.

Em suma, a motivação para o desenvolvimento desta tese surge da demanda crescente do consumo de energia o que leva à construção de novas usinas de geração e à necessidade de despachar a carga adicional gerada. Ainda é motivação deste trabalho a significativa possibilidade de que as informações meteorológicas adequadas auxiliem o setor elétrico nacional no sentido de pospor investimentos ou contornar o problema gerado pelo atraso da construção de novas linhas utilizando processos de recapitação de linhas existentes para isso, mas sem descurar da segurança e confiabilidade do sistema.

## **1.1 OBJETIVOS DA TESE**

O objetivo geral deste trabalho é mensurar, descrever e modelar o regime de vento e temperatura do ar próximo à superfície sobre a região Nordeste do Brasil, motivado pela demanda que o setor elétrico brasileiro, em especial o Operador Nacional do Sistema Elétrico, tem de conhecer o comportamento sazonal de variáveis meteorológicas que interferem diretamente no despacho de carga em Linhas Aéreas de Transmissão (LT). Esse objetivo possui relevância para o setor elétrico, pois cria a possibilidade de o setor avaliar o melhor

aproveitamento das linhas existentes assim como direcionar novos projetos de linhas com valores típicos de vento e temperatura do ar próximos ao observado com importantes ganhos econômicos para o país a partir da modicidade tarifária.

### **1.1.1 Objetivos Específicos**

- a) Estabelecer a localização das zonas de calmaria de vento e de máximos de temperatura do ar sobre regiões de passagens de linhas aéreas de transmissão de energia elétrica.
- b) Determinar quais são as metodologias estatísticas apropriadas para o levantamento e tratamento de dados meteorológicos de superfície medidos em estações automáticas.
- c) Caracterizar os mínimos de velocidade de vento (zonas de estagnação) do subsistema Nordeste do SIN.
- d) Estabelecer uma metodologia para identificação de ocorrência de zonas de estagnação de vento próximo à superfície sobre áreas de passagens de LT (Linhas Aéreas de Transmissão de Energia) no Nordeste do Brasil considerando a influência dos sistemas atmosféricos atuantes e dos aspectos topográficos e corpos d'água.
- e) Caracterizar a frequência de ocorrência, sazonal e diurna, simultânea dos valores críticos de velocidade do vento e de temperatura do ar definidos pela norma da ABNT – NBR 5422 (ABNT, 1985).
- f) Simular numericamente a região-alvo do presente projeto utilizando o modelo WRF (Weather Research and Forecasting) para avaliar o seu desempenho em reproduzir próximo a superfície os regimes de vento e temperatura do ar medido por estações meteorológicas.

## **2 REVISÃO DE LITERATURA**

Neste capítulo serão apresentados o setor de energia elétrica nacional e os órgãos de regulação e operação que atuam para garantir a eficiência e segurança do sistema elétrico. Esses pontos serão fundamentais para compreender a relação existente entre a demanda e a capacidade de despacho de energia em linhas aéreas de transmissão e a meteorologia. Adicionalmente, o clima da região alvo da presente tese será apresentado, assim como os sistemas meteorológicos atuantes que em conjunto com a topografia definem as sub-regiões climáticas e determinam o limite da corrente elétrica que pode ser conduzida em uma linha.

### **2.1 SISTEMA ELÉTRICO BRASILEIRO – CENÁRIO ATUAL**

O atual modelo do setor elétrico nacional foi estabelecido no ano de 2004 sob três pautas principais a saber: modicidade tarifária, desenvolvimento social e planejamento estratégico (ANEEL, 2008). A modicidade tarifária é considerada para estabelecer a tarifa de energia aplicada aos consumidores finais de energia elétrica. É também por meio da modicidade que se define os valores de contratos com as empresas do setor de energia elétrica de tal modo que elas obtenham ganhos suficientes para garantir suas operações com qualidade e continuidade e ainda com uma receita anual atrativa visto o capital investido (SOARES, 2009).

Desta forma, este modelo busca o equilíbrio entre a cadeia produtiva (remuneração adequada) e o consumidor final (pagamento justo da tarifa). No tocante ao desenvolvimento social o setor elétrico também teve um importante papel a partir de 2004, pois foi através dele que o governo federal universalizou o acesso à energia elétrica no meio rural para famílias de baixa renda e cuja significativa parcela reside em terras indígenas ou de quilombolas, assentamentos e reservas extrativistas. (DE FREITAS; ROMARCO DE OLIVEIRA, 2017; FREITAS; SILVEIRA, 2015).

O atual modelo também criou em 2004 a Empresa de Pesquisa Energética (EPE) com o objetivo de prover o governo federal de estudos relacionados a infraestrutura e planejamento da expansão energética. Adicionalmente, há outras instituições que exercem papel fundamental para o desenvolvimento e operação do modelo atual do setor elétrico. Nesta presente revisão, ressalta-se a Agência Nacional de Energia Elétrica (ANEEL), que foi criada em 1996 após o início da desestatização ocorrida no governo do presidente Fernando Henrique Cardoso, com o objetivo de regular e fiscalizar o setor elétrico. E o Operador Nacional do Sistema Elétrico

(ONS) criado em 1998 com o propósito de operar o despacho de carga da rede básica (geração e transmissão) e propor os reforços e expansão dos sistemas de transmissão (REGIS, 2017).

A desestatização (privatização) iniciada na década de 90 também impulsionou a desverticalização das companhias estatais que até então, de maneira geral, estavam presentes em toda a cadeia produtiva (geração, transmissão e distribuição de energia). A partir da separação das empresas em cada segmento do setor abriu-se a possibilidade da compra e venda de energia produzida pelas geradoras sob o intermédio do Mercado Atacadista de Energia Elétrica. E que em 2004, foi substituído pela Câmara de Comercialização de Energia Elétrica (CCEE) que dentre as suas atribuições também define os preços a serem praticados no mercado livre (ANEEL, 2008).

A cadeia produtiva operacional do setor elétrico é composta por 3 diferentes atividades: geração, transmissão e distribuição. A comercialização de energia é outra atividade importante do setor, porém não é tratada no âmbito operacional.

### **2.1.1 Sistema Interligado Nacional**

Nesta revisão o foco será dado apenas ao conjunto que compõe o Sistema Interligado Nacional (SIN): geração e transmissão de energia elétrica da rede básica.

#### **a) Geração de energia**

A matriz energética do SIN é composta em sua maioria por hidrelétricas classificadas como fonte de geração limpa e renovável e termelétricas que utilizam combustíveis fósseis (por exemplo, gás natural, carvão mineral e óleo diesel) que são não-renováveis e poluentes (ONS, 2019b). Ambas em conjunto representaram, em 2018, aproximadamente 80% (67% e 13%, respectivamente) da capacidade instalada de produção de energia elétrica no SIN (ONS, 2019a). Outra característica importante no tocante a matriz de geração é que as termelétricas atuam com o objetivo de complementar a produção de energia quando a capacidade de geração hídrica é afetada pela estiagem. Essa flexibilidade é devida a arquitetura das interligações do SIN que permite a exploração da máxima geração de energia independentemente da localização geográfica e da matriz.

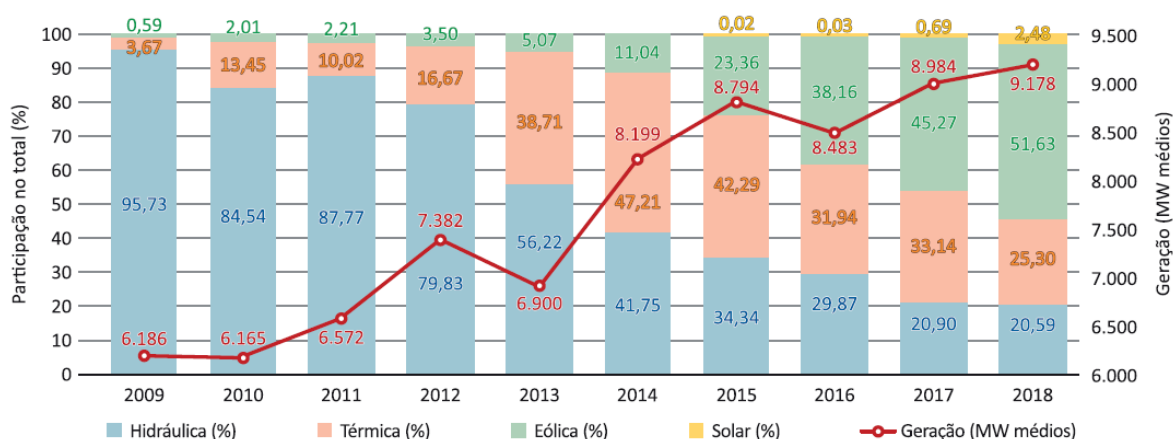
Para diminuir a influência do tempo e do clima na matriz energética e fomentar o mercado interno o governo federal em conjunto com o setor elétrico buscou, especialmente a partir de 2002, com o Programa de Incentivo às Fontes Alternativas de Energia Elétrica (PROINFA) – Lei nº10.438/2002, a diversificação da matriz com o uso de outras fontes



incluindo solar e eólica (ANEEL, 2008). O Plano de Ampliações e Reforços nas Instalações de Transmissão do SIN (PAR) indicou que nos próximos 5 anos (2019 – 2023) haverá um aumento da contribuição da geração eólica e fotovoltaica na capacidade instalada e que atualmente compõem apenas 10% da matriz energética (ONS, 2019a).

Especificamente para o Nordeste do Brasil (excetuando-se o Maranhão por não fazer parte do subsistema Nordeste do SIN) a diversificação da matriz tem experimentado um crescimento expressivo com o uso da energia cinética do vento (figura 2.1). De acordo Bezerra (2019), 86% da capacidade da geração eólica do Brasil está situada no Nordeste sendo que no último ano foi responsável por mais de 50% do total gerado na região. O planejamento do setor para os próximos 5 anos é de um aumento aproximado de 11 GW na capacidade de carga no SIN (ONS, 2019b), dos quais aproximadamente 2 GW estarão no subsistema Nordeste.

Figura 2.1 – Evolução da matriz energética do subsistema Nordeste



Fonte: BEZERRA, (2019).

## b) Transmissão de energia

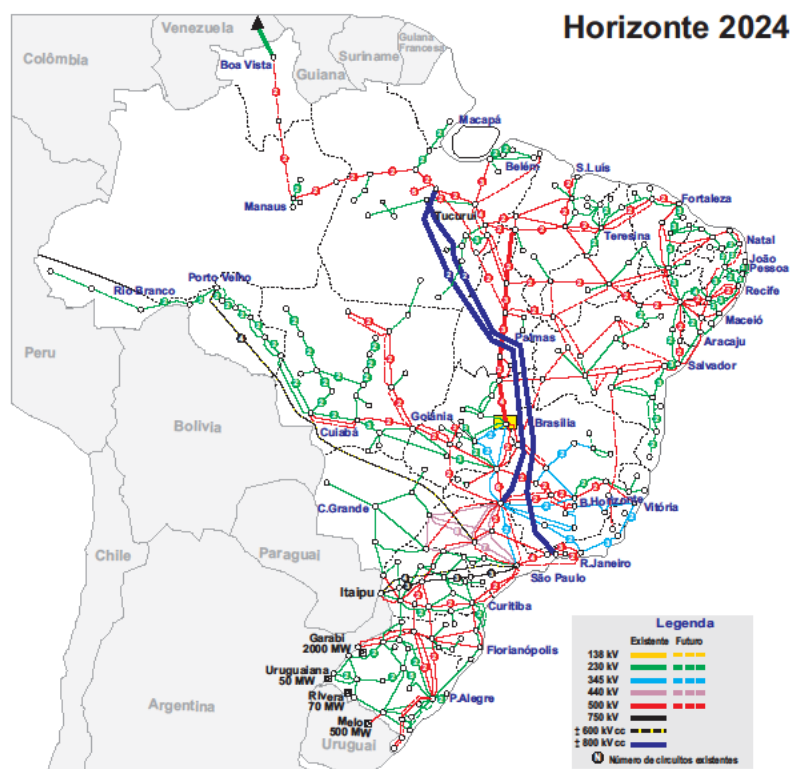
Além das usinas de geração o SIN é composto por aproximadamente 130 mil quilômetros de linhas de transmissão de energia que operam com tensões de 230, 345, 500, 600, 750 e 800 kV (ONS, 2019a) (figura 2.2). Este sistema é subdividido em 4 subsistemas: Sul, Sudeste/Centro-Oeste, Nordeste e Norte. As linhas de transmissão que formam o SIN estão presentes no Distrito Federal e em todos os estados do Brasil, com exceção de Roraima que possui um sistema isolado de geração baseado em termelétricas e é interligada eletricamente a Venezuela. A interconexão dos subsistemas permite que em períodos de menor disponibilidade hídrica a demanda energética possa ser atendida de forma transparente ao mercado consumidor

em qualquer localidade do país. Isso porque a energia pode ser gerada em qualquer localidade do SIN e ser transmitida para qualquer região distante milhares de quilômetros do centro de geração (ANEEL, 2008).

A partir da figura 2.3, que mostra a evolução bienal (2009 – 2020) e quinquenal (2019 – 2023) dos investimentos estimados pelo Operador para ampliações e reforços nas instalações de transmissão, é possível observar a tendência positiva dos investimentos. Nos últimos 10 anos (2009 – 2019) o SIN registrou a expansão de 40 mil km de linhas com o objetivo de atender o aumento da capacidade de geração motivada principalmente pela inserção de fontes renováveis não convencionais no sistema (solar e eólica) (FRONTIN et al., 2010; ONS, 2019a). O custo associado a essa expansão foi de aproximadamente R\$ 83 bilhões (figura 2.3). O PAR estimou que para os próximos 5 anos haverá ainda um investimento na ordem de R\$ 13,6 bilhões de reais para a edificação de novas linhas (ONS, 2019a). Desse montante, R\$ 400 milhões serão destinados a construção de 3 linhas no Nordeste – uma em Pernambuco de 45 km e outras duas na Bahia totalizando 231 km – (ONS, 2019a). De acordo com os números do ONS, pode-se estimar que o investimento no Nordeste terá um custo de R\$ 1,5 milhões/km.

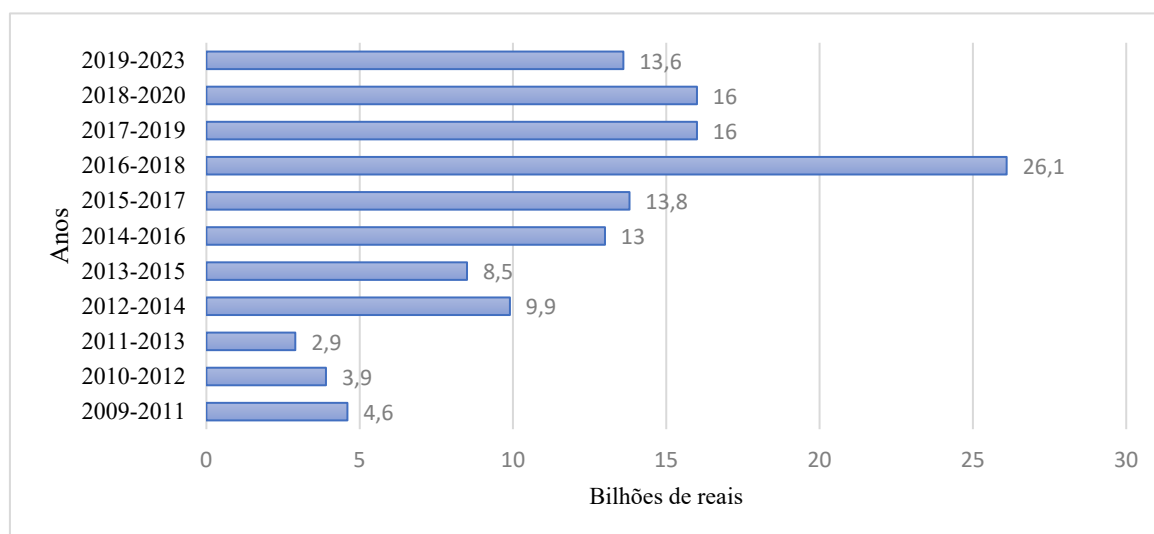
Além do alto custo de construção de novas linhas para atender a demanda crescente de energia, a fiscalização da expansão da transmissão realizada pela ANEEL indicou que em março do presente ano 35% dos 141 empreendimentos em andamento estavam com previsão de atraso (ANEEL, 2019). Embora haja diversas justificativas para os atrasos, o relatório de acompanhamento diferenciado dos empreendimentos de transmissão verificou que para o período de 2014 a 2018 o licenciamento ambiental representou 45% dos motivos para os atrasos (ANEEL, 2019). Ainda de acordo com este relatório, o atraso na obtenção do licenciamento ambiental, que nos últimos 5 anos atingiu a média de 370 dias, pode acarretar o atraso nas demais atividades relacionadas a execução das obras físicas. Assim, entre os anos de 2014 e 2018 o tempo médio para a conclusão e entrega de uma edificação de transmissão de energia foi de 3 anos (ANEEL, 2019).

Figura 2.2 – Sistema Interligado Nacional atual e projeções.



Fonte: (ONS, 2020).

Figura 2.3 – Estimativa de investimentos em construção de linhas de transmissão e capacidade de transformação em cada ciclo quinquenal do PAR de 2009 a 2023.



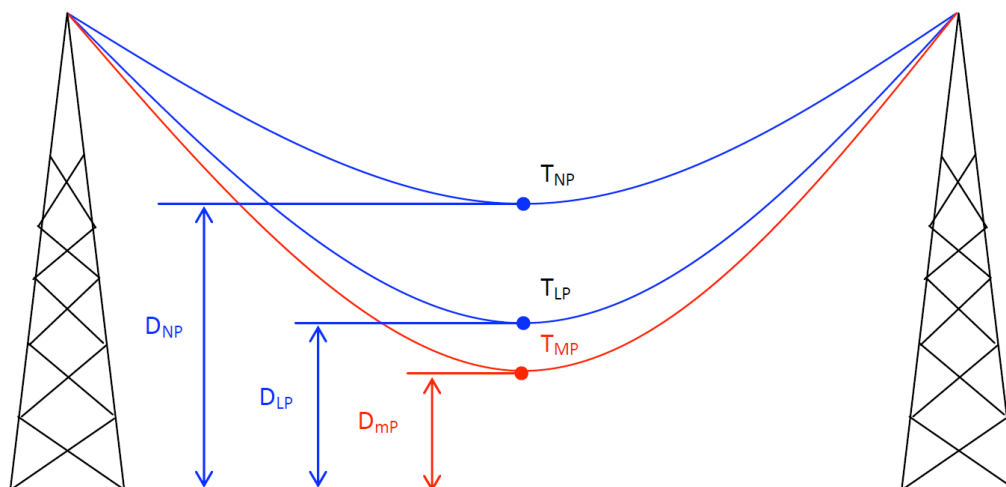
Fonte: ONS (2019a)

Diante do exposto, fica evidente que a construção de novas linhas para atender à crescente demanda energética do país não é um procedimento simples, principalmente devido

ao custo, questões socioambientais e ao tempo de execução física das obras (SILVA, 2009). Diversos trabalhos têm apontado que essas não são as únicas barreiras relacionadas a expansão de linhas e tampouco é um problema restrito ao Brasil. DACONTI e LAWRY (2003) afirmaram que o impacto visual, a suspeita de problemas de saúde relacionada aos efeitos do campo eletromagnético, a impossibilidade do uso da terra em faixas de passagem da linha e a influência no sistema ecológico são fatores que elevam os custos e dificultam o licenciamento ambiental para a construção de novas linhas aéreas de transmissão. CAVASSIN e FERNANDES (2012) indicaram que para diminuir o impacto ecológico tem-se utilizado estruturas cada vez mais altas e muitas vezes desviado o traçado da linha para evitar áreas de conservação ou de preservação ambiental. Essas ações associadas ao pagamento de indenizações para os proprietários das terras onde são constituídas as faixas de passagens das linhas tendem a elevar o custo final das obras.

Entretanto, o aumento da capacidade operativa das linhas de transmissão existentes é uma possível solução para contornar os problemas mencionados (DACONTI; LAWRY, 2003; JARDINETTI, 2014; NTULI et al., 2016). A possibilidade desse aumento se dá porque muitas linhas de transmissão em operação atuam com folga de carregamento elétrico. Isto ocorre devido ao projeto dessas linhas considerarem valores conservativos para determinar a máxima corrente elétrica que garante a segurança da operação (CARVALHO, 2017; ONS, 2017). Essa máxima corrente, também denominada de ampacidade, é aquela à qual um cabo condutor pode ser submetido sem recozer ou exceder a distância cabo-solo (flecha) (YIP et al., 2009). A respeito da flecha é possível observar a partir da figura 2.4 que a elevação da temperatura de um cabo condutor (por exemplo  $T_{MP} > T_{LP} > T_{NP}$ ) resulta no abaixamento da linha ( $D_{mP} < D_{LP} < D_{NP}$ ) devido a dilatação térmica do cabo (LIAO; LIU, 2018). Assim, se o cabo condutor sofrer um aquecimento maior do que a temperatura de projeto haverá a violação da distância de segurança e conseqüentemente falha de isolamento da linha podendo causar mortes e ruptura da transmissão de energia (BEÇA et al., 2018; ENTSOE, 2015; NASCIMENTO, 2009; SILVA, 2009).

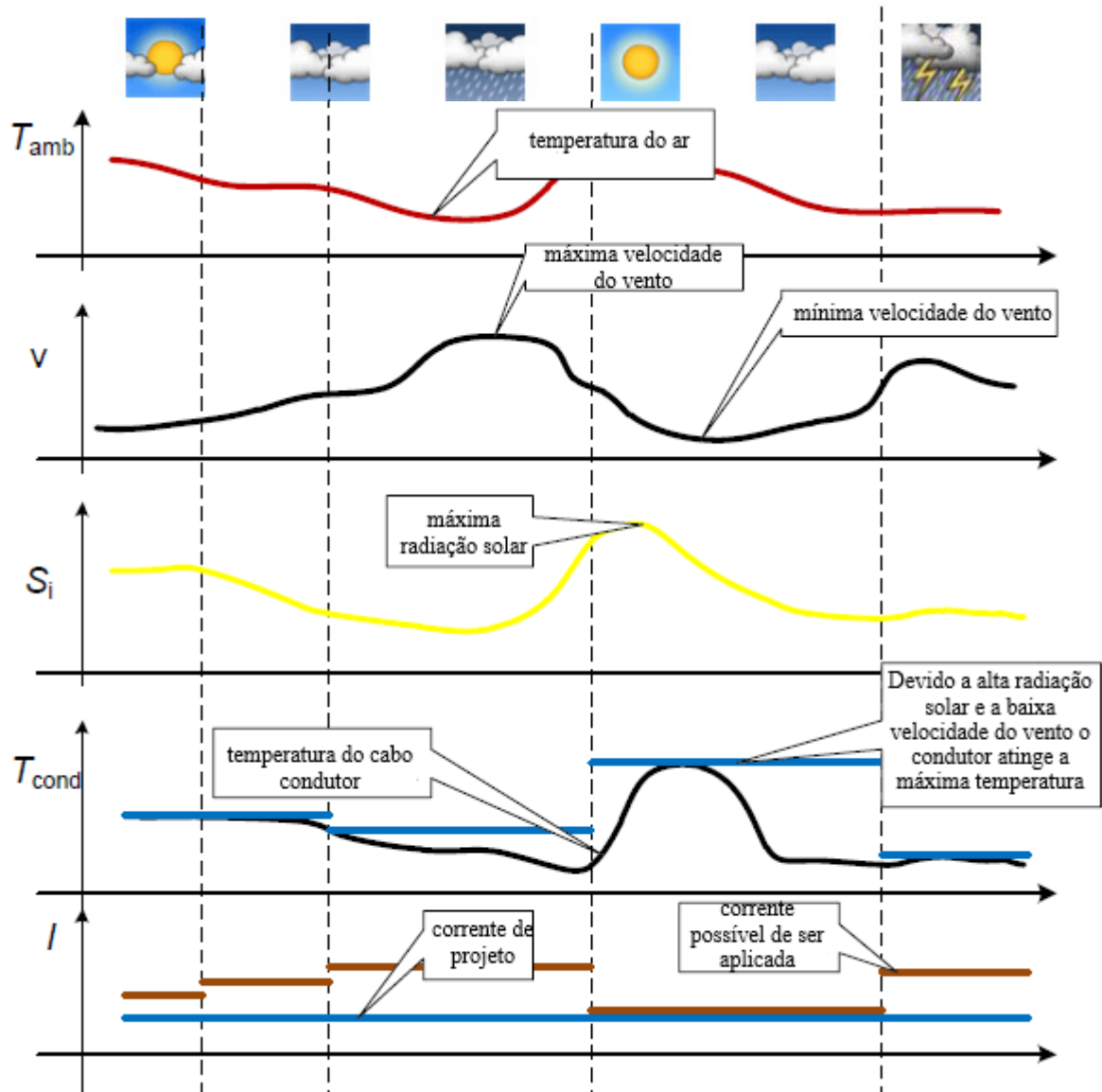
Figura 2.4 – Distâncias esquemáticas de cabo-solo  $D_{NP}$ ,  $D_{LP}$  e  $D_{mP}$  e as respectivas temperaturas  $T_{NP}$ ,  $T_{LP}$  e  $T_{MP}$  de um cabo condutor.



Fonte: Nascimento (2009).

De acordo com a figura 2.5 é possível verificar que a temperatura de um cabo condutor não depende apenas da corrente elétrica passante e das características construtivas do cabo. Notam-se que as diferentes condições atmosféricas hipotéticas apresentadas ainda na mesma figura exercem uma influência expressiva na temperatura do condutor ( $T_{cond}$ ). Por exemplo, para um dia ensolarado onde a temperatura do ar e a radiação são máximas e a velocidade do vento é mínima o cabo condutor atinge a sua máxima temperatura de operação definida no projeto. Enquanto para um dia com ocorrência de tempestade a velocidade do vento em conjunto com baixa radiação solar e temperatura do ar permitem uma operação mais segura da linha inclusive com uma possibilidade de despacho de corrente ( $I$ ) maior do que a de projeto (linha azul).

Figura 2.5 – Influência das condições atmosféricas na temperatura do cabo condutor e na corrente elétrica despachada em uma linha de transmissão.



Fonte: Adaptado de ENTSOE (2015).

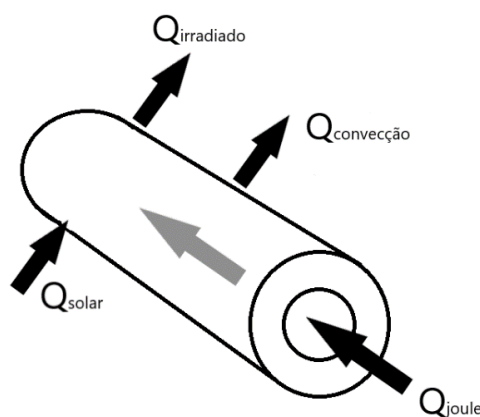
A explicação física da influência das condições atmosféricas na determinação da temperatura do condutor é evidenciada a partir da equação de balanço de energia em um cabo condutor:

$$Q_{joule} + Q_{solar} = Q_{irradiado} + Q_{convecção}$$

Equação 2.1

Essa equação considera quatro termos dos quais dois referem-se ao ganho de calor e os outros dois indicam a perda (figura 2.6). Os que possuem fluxo de calor positivo são:  $Q_{\text{joule}}$  devido ao efeito joule gerado pela corrente elétrica a que o cabo foi submetido e o  $Q_{\text{solar}}$  determinado pela radiação solar. Os fluxos que representam perda de calor são:  $Q_{\text{irradiado}}$  que representa a perda de calor para o ambiente quando a temperatura do cabo condutor é maior que a temperatura ambiente e por fim o  $Q_{\text{convecção}}$  que indica a retirada de calor da superfície do cabo condutor por convecção devido a ação do vento.

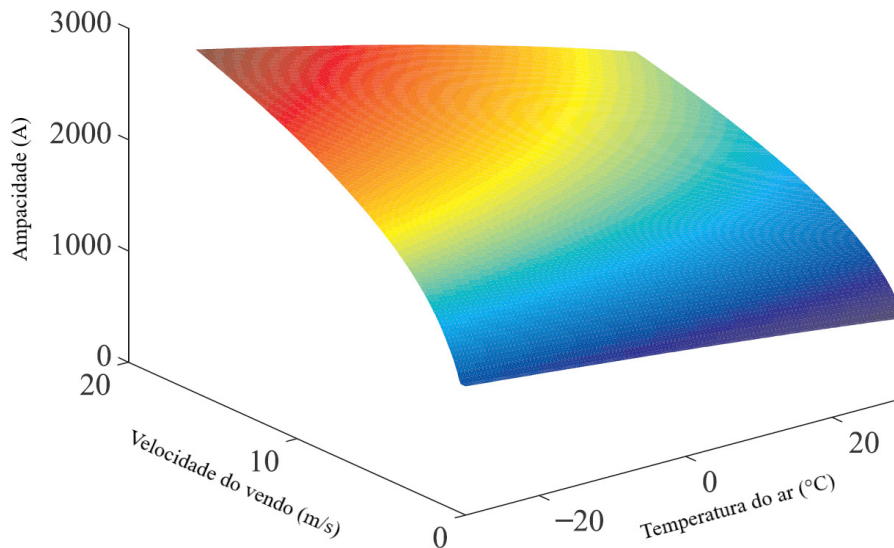
Figura 2.6 – Balanço de energia em um cabo condutor.



Fonte: o autor.

Os autores KARIMI, MUSILEK e KNIGHT (2018) demonstraram para uma linha de transmissão de 240 kV a relação entre a ampacidade e a velocidade do vento e a temperatura do ar (figura 2.7). Esta figura mostra que quanto maior a velocidade do vento maior é a corrente máxima que a linha pode ser submetida sem violar a temperatura de projeto e a distância cabo-solo. Nota-se também que a temperatura tem menor influência do que a velocidade do vento na variação da ampacidade. Por exemplo, para um vento com velocidade constante o aumento da temperatura não diminui de forma acentuada a ampacidade quando comparado com o caso de uma temperatura fixa e decréscimo da velocidade do vento. Portanto, fica evidente que a velocidade do vento é a variável que mais influencia a troca de calor entre o cabo condutor e o ambiente (ENTSOE, 2015; ONS, 2017).

Figura 2.7 – Variação da ampacidade em função da velocidade do vento e da temperatura do ar.



Fonte: KARIMI; MUSILEK; KNIGHT (2018).

Embora tenha sido demonstrado o impacto das condições atmosféricas na ampacidade, ressalta-se que existe a dificuldade de se obter medições meteorológicas próximo as linhas de transmissão. Neste caso, o projeto de linhas de transmissão faz uso da metodologia denominada de Ampacidade Determinística (ONS, 2017; SILVA, 2009). Essa metodologia utiliza parâmetros meteorológicos fixos e representativos das piores condições climáticas da região de operação da linha. Porém, na ausência de dados meteorológicos medidos para determinar a ampacidade de uma linha no Brasil a norma técnica que define as condições básicas para o projeto de linhas aéreas de transmissão recomenda o uso simultâneo dos valores de  $1000 \text{ W/m}^2$  de radiação solar, velocidade do vento igual a  $1 \text{ m/s}$  e a temperatura máxima média registrada na estação meteorológica mais próxima (ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS, 1985). De acordo com o ONS (2017) é devido ao caráter conservador desse método que as linhas do SIN operam com temperatura do condutor na maior parte do tempo distante da temperatura de projeto e que devido a esse fato poderiam ter a capacidade operativa explorada. E outras palavras, a ampacidade calculada pelo método determinístico não reflete a condição de operação real de uma linha de transmissão (CARVALHO, 2017).

Outro método utilizado para determinar a corrente máxima de um cabo condutor é conhecido como Ampacidade Estatística ou Probabilística. Esse método é utilizado quando há



uma base horária de dados climatológicos e de temperatura do condutor (estimada pela equação do balanço térmico) (ANJOS, 2013). A partir desses dados é possível calcular o risco térmico que é definido como a probabilidade da temperatura do condutor exceder a temperatura de projeto e a distância de segurança cabo-solo (SILVA, 2009). De acordo com Furtado (2008) o uso do conceito risco térmico permite a operação da linha de maneira mais realista no tocante as condições atmosféricas e possibilita carregamentos flexíveis sobretudo em caráter de emergência ou manutenções programadas.

De maneira geral os métodos determinísticos e probabilísticos consideram o pior caso em termos de propriedades físico-mecânicas do cabo condutor e condições atmosféricas para definir o limite térmico de operação da linha. Segundo KARIMI, MUSILEK e KNIGHT (2018), uma vez definido o pior cenário de operação, seja para regimes de carregamento de energia de longa duração ou sazonais, esses dois métodos são tratados como Ampacidade Estática. Porém, quando os limites térmicos de operação da linha são definidos em tempo real e a capacidade de despacho de energia varia dinamicamente baseada nas condições atmosféricas atuais tem-se a Ampacidade Dinâmica (LIAO; LIU, 2018). De acordo com (ANJOS, 2013; BALDICK; NEILL, 2009; LARRUSKAIN et al., 2006) esse método faz uso do monitoramento das variáveis meteorológicas e/ou do estado do cabo condutor (temperatura do cabo e distância cabo-solo) para definir a condição atual que uma linha pode ser submetida. Esse monitoramento é realizado de forma indireta a partir da instalação de estações meteorológicas ao longo de uma linha de transmissão ou direta no caso da instalação de um sensor no cabo condutor para medir a corrente elétrica e a temperatura da linha (ALVAREZ et al., 2018; SILVA, 2009). Há outras formas de monitoração direta ou indireta, porém como esse não é o escopo da tese, detalhes podem ser obtidos em GIANNUZZI et al. (2015). Adicionalmente, é possível obter a previsão de ampacidade com horizonte de horas ou de dias a partir de modelos numéricos de previsão do tempo. DUPIN, KARINIOTAKIS e MICHIORRI (2019) indicaram que embora haja poucos trabalhos na literatura tratando desse assunto é possível operar uma linha com ampacidade prevista com um fator de risco aceitável.

Trabalhos como BANERJEE (2015), KARIMI; MUSILEK; KNIGHT (2018) e KHAKI et al. (2010) demonstraram os benefícios quantitativos e qualitativos do uso da Ampacidade Dinâmica em linhas aéreas de transmissão. Estes autores indicaram também a possibilidade de o carregamento dinâmico de linhas reduzir os custos de operação e possibilitar a inserção imediata de energias renováveis no sistema mantendo a segurança desejada. Especificamente, tratando-se do SIN, ANJOS (2013) sugere que é possível reduzir os custos de operação e manter

a segurança do sistema quando as condições atmosféricas são favoráveis ao aumento dinâmico da corrente elétrica passante no cabo condutor.

Por fim, de acordo com NASCIMENTO (2009) reavaliar o método de cálculo da ampacidade utilizado durante o projeto de uma linha de transmissão pode elevar a capacidade de operação mantendo a continuidade e confiabilidade do sistema. Essa reavaliação se faz necessária devido a demanda de despacho, custo de construção de novas linhas e das questões ambientais discutidas anteriormente. Este processo torna-se factível especialmente neste momento onde os avanços tecnológicos, que através da monitoração in loco, permitem a formação de bancos de dados meteorológicos e da condição do cabo condutor.

## **2.2 SISTEMAS ATMOSFÉRICOS QUE ATUAM SOBRE O SUBSISTEMA NORDESTE DO SIN**

Conforme discutido no capítulo anterior a meteorologia tem um papel determinante no equilíbrio térmico do cabo condutor utilizado para transmitir energia elétrica. Assim, a caracterização dos regimes atmosféricos ao longo de uma linha área de transmissão pode permitir o aumento expressivo da capacidade de operação das linhas existentes (CASTILHO, 2010). Adicionalmente, a definição da climatologia de uma determinada região possibilita que engenheiros projetem novas linhas com maior capacidade de despacho de energia ou com menor custo relacionado a especificação técnica dos componentes e estruturas (ANJOS, 2013). E das variáveis meteorológicas consideradas no cálculo da ampacidade, a velocidade do vento e a temperatura do ar são as que mais influenciam o equilíbrio térmico do cabo condutor (NASCIMENTO, 2009).

Para descrever os processos atmosféricos que atuam no subsistema Nordeste do SIN e determinam o regime temporal das variáveis mencionadas será utilizada a classificação proposta por ORLANSKI (1975). Essa classificação propõe a divisão dos fenômenos atmosféricos em três escalas inter-relacionadas que por sua vez são subdivididas em  $\alpha$ ,  $\beta$ ,  $\gamma$ , de acordo com suas dimensões espaço-temporais (EMANUEL, 1986).

A maior escala horizontal e de tempo da classificação de Orlandi refere-se à macroescala. Nela, estão situados os sistemas sinóticos com duração de dias até meses e com dimensões de aproximadamente 10 mil quilômetros (ORLANSKI, 1975). Sobre o subsistema Nordeste do SIN há duas importantes circulações pertencentes à macroescala que determinam as características climáticas denominadas de células de circulação de Hadley e Walker. A primeira é responsável pela formação da Zona de Convergência Intertropical (ZCIT) devido à confluência dos ventos alísios em baixos níveis na região equatorial (FERREIRA; MELLO,

2005). A banda de nebulosidade convectiva formada por esta confluência tem sua variabilidade sazonal, longitudinal e latitudinal devido a anomalia da temperatura da superfície do mar (CARVALHO; OYAMA, 2013). De acordo com CARVALHO e OYAMA (2013) o posicionamento latitudinal da ZCIT varia em torno de 7° atingindo a posição mais austral entre os meses de fevereiro a abril. Assim, a presença do aglomerado convectivo sobre o norte do Nordeste define o período chuvoso da região e conseqüentemente diminui a quantidade de radiação solar à superfície (QUADRO et al., 2012). Adicionalmente, devido a anomalia de radiação há um menor aquecimento diferencial da superfície e por este motivo uma diminuição da velocidade do vento e da temperatura do ar.

De acordo com REBOITA e SANTOS (2015), a célula de Walker é a segunda circulação em larga escala composta de seis células distribuídas zonalmente ao longo do globo e possui variabilidade interanual modulada pelo fenômeno El Niño/Oscilação Sul, que afeta a climatologia do Nordeste do Brasil. Isso porque, quando do enfraquecimento dos ventos alísios, ocorre o aquecimento anômalo da temperatura da superfície do mar no oceano Pacífico Tropical central e leste (episódio de El Niño) (DIAZ; BRADLEL, 2004). E devido a diminuição destes ventos, ocorre o deslocamento do ramo descendente da célula de Walker do oeste da América do Sul, próximo ao equador, para leste sobre a Amazônia, nordeste brasileiro e o oceano Atlântico Tropical (PEREIRA; REBOITA; AMBRIZZI, 2017). Ainda de acordo com esses autores, as correntes descendentes sobre o Nordeste inibem a convecção na região e modulam as condições de secas. Portanto, assim como o posicionamento da ZCIT altera a disponibilidade de radiação solar na superfície, esta altera, por conseguinte o regime de vento e temperatura.

Complementarmente, há outros sistemas de escala sinótica que atuam sobre o Nordeste do Brasil e mais especificamente sobre a área do subsistema Nordeste do SIN. As frentes frias, que são a intersecção ou interface entre massas de ar com características diferentes, também alteram o tempo da região alvo. KOUSKY (1979) investigou a atuação dos sistemas frontais no Nordeste e verificou que as frentes frias avançam sobre a Bahia e atingem o estado do Rio Grande do Norte, e às vezes, até o norte do Ceará durante as estações do outono e inverno. Ainda de acordo com este autor, durante os meses de março a dezembro esses sistemas atingem o sul da Bahia com maior frequência do que em janeiro e fevereiro. Porém, durante esses dois meses os máximos de precipitação são influenciados pelos sistemas frontais (SILVA et al., 2019). Como último resultado, KOUSKY (1979) após utilizar 10 anos de dados medidos por estações meteorológicas de superfície, concluiu também que a variabilidade da precipitação no Nordeste é influenciada pela passagem de frentes frias. Ou seja, esse autor identificou que em

anos com maior passagens de frentes frias o registro de precipitação total também é maior. E assim, tem-se mudança no regime de temperatura do ar e vento à superfície.

Os Vórtices Ciclônicos de Altos Níveis (VCAN) também são sistemas de escala sinótica que modulam o regime de precipitação e consequentemente a distribuição espaço-temporal da temperatura do ar e velocidade do vento no leste do Nordeste (SILVA et al., 2011). Os VCAN que atuam na região são os do tipo Palmer definidos como circulações fechadas na troposfera superior e que não recebem intrusão de ar polar (PALMER, 1951). A presença desses vórtices pode favorecer a formação de ciclones em baixos níveis e produzir convecção na periferia desses sistemas enquanto que os movimentos de subsidência inibem a nebulosidade sobre a área em que o sistema está posicionado (REBOITA et al., 2010). Além disso, durante o deslocamento de um VCAN a precipitação pode ser intensificada pela possível interação com outros sistemas meteorológicos como a ZCIT ou frentes frias (MORAIS, 2016). Este último autor relata ainda que a máxima ocorrência dos VCAN é registrada para o verão do Hemisfério Sul, embora sejam observados durante todo o ano.

Nas áreas tropicais de influência dos ventos alísios há um escoamento de leste para oeste capaz de provocar chuvas desde o estado da Bahia até o Rio Grande do Norte e que em algumas vezes se estende até o centro do Ceará (FERREIRA; MELLO, 2005). Esse escoamento, denominado de Ondas de Leste, é um distúrbio no campo de pressão atmosférica e que se desloca da costa da África até o litoral leste brasileiro (GOMES et al., 2019). Segundo GOMES et al. (2019), esses distúrbios influenciam na climatologia da região alvo da presente tese, pois transportam umidade para áreas geralmente secas, favorecendo o aumento da convergência em baixos níveis e intensificando a precipitação. De acordo com a climatologia entre os meses de março a agosto observa-se a maior influência das Ondas de Leste sobre o Nordeste, embora esses distúrbios atinjam o litoral leste nordestino durante todo o ano (GOMES et al., 2019).

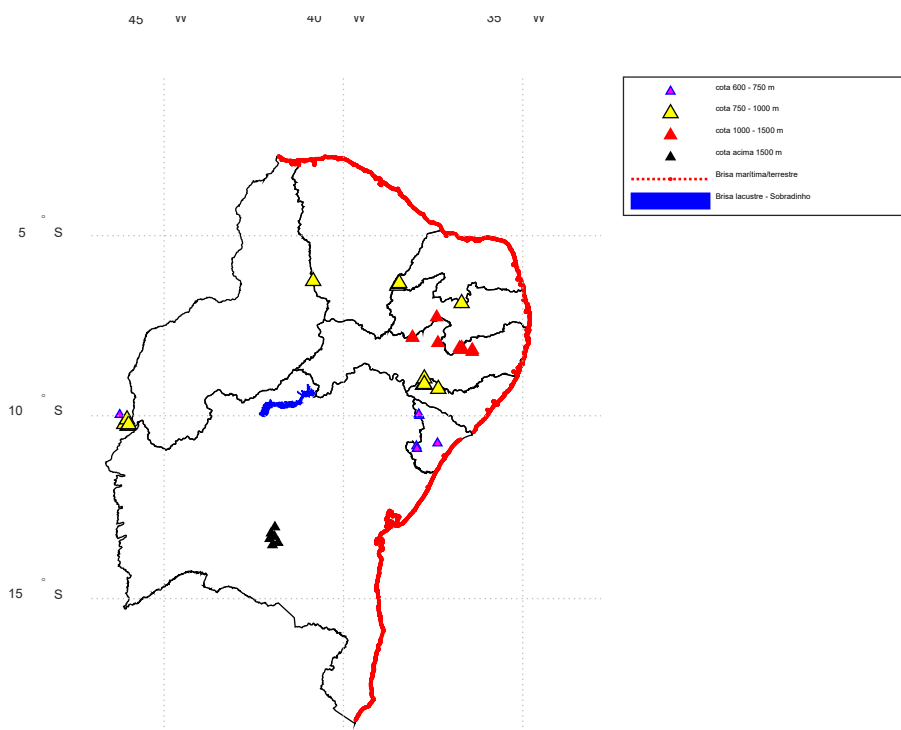
Outro sistema atmosférico que surge a partir da célula de Hadley e transporta calor e umidade do Oceano Atlântico para o leste do Nordeste é o Anticiclone Subtropical do Atlântico Sul (SILVA et al., 2011). Essa circulação que tem comportamento quase estacionário durante o inverno desloca-se para sul e aproxima-se do continente em relação à sua posição climatológica. Nesta situação, a circulação do anticiclone intensifica o transporte de umidade do oceano, favorecendo a precipitação. (REBOITA et al., 2010).

Seguindo a definição de ORLANSKI (1975) o subsistema Nordeste também sofre influência de fenômenos meteorológicos pertencentes a mesoescala. De acordo com este autor, a mesoescala é subdividida em alfa, beta e gama e compreende sistemas atmosféricos com

espaçamento horizontal de 2 mil quilômetros os quais podem se inter-relacionar com a macro e a microescala.

Particularmente sobre o Nordeste as circulações locais com maior importância são a da brisa vale-montanha, a brisa marítima terrestre e lacustre (figura 2.8). Tais circulações se formam devido a heterogeneidade e o aquecimento diferencial da superfície. Desempenham um papel importante no transporte de umidade, regulação da temperatura do ar, intensificação da velocidade do vento e até mesmo na quantidade de precipitação quando associada a outras circulações (SOUZA, 2016).

Figura 2.8 – Localização das circulações locais sobre o subsistema Nordeste do SIN.

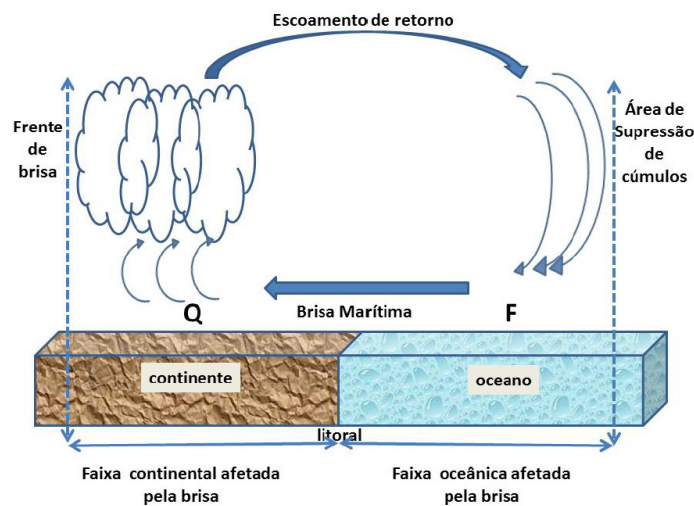


Fonte: o autor.

A circulação do tipo brisa marítima e terrestre é encontrada na costa nordestina e é induzida devido ao aquecimento diferencial entre o continente e oceano (COSTA; LYRA, 2012). Isso ocorre porque a capacidade térmica da água é diferente da superfície continental. Portanto, quando ambas são submetidas a uma mesma radiação solar o continente aquece mais rápido do que o oceano. Por este motivo é que durante o dia o continente sofre um aquecimento maior quando comparado ao oceano. Esse diferencial de temperatura gera um gradiente de pressão horizontal que por sua vez cria um deslocamento de ar (vento) frio do oceano em direção ao continente conhecido com frente de brisa (FREITAS, 2003). Por continuidade, nos níveis mais altos ocorre uma corrente de ar de retorno e assim tem-se a circulação local fechada brisa marítima, conforme observado na figura 2.9 (SOUZA, 2016). Quanto maior for o

aquecimento diferencial entre as superfícies continente – oceano mais intensa será a frente de brisa e, portanto maior a penetração do deslocamento de ar em direção ao interior do continente, podendo atingir centenas de quilômetros (FREITAS et al., 2007). Durante a noite, quando o aquecimento diferencial devido a heterogeneidade das superfícies é interrompido, ocorre a inversão da direção do vento e, então, a frente de brisa desloca-se para o oceano (brisa terrestre) desde que o continente sofre intensa perda radiativa, ao contrário das águas do oceano. A brisa pode ainda ser intensificada por forçantes topográficas ou se houver acoplamento com outras circulações locais, como a ilha de calor (FREITAS et al., 2007).

Figura 2.9 – Desenvolvimento esquemático da circulação de brisa marítima. Letra Q indica superfície mais aquecida do que a superfície F (fria).

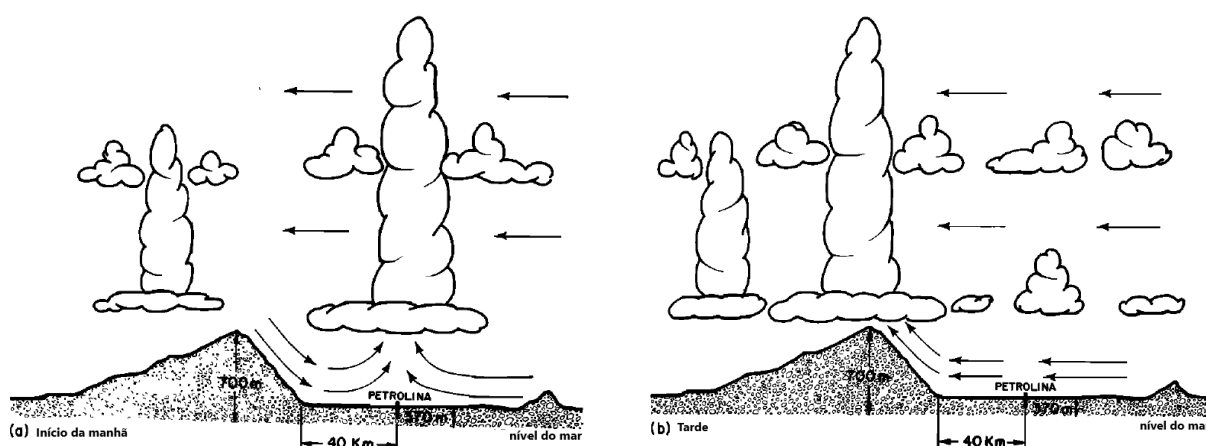


Fonte: SOUZA (2016).

Outras circulações locais podem ser encontradas distantes da costa litorânea e sem relação com a brisa marítima e terrestre. Isso porque no interior do continente quando há descontinuidade marcante na cobertura vegetal, como por exemplo nas regiões de cultivo alagado, pode haver a indução de circulações fechadas durante o ciclo diurno (MELO; CORREIA; SILVA, 2014). De acordo com Alcantara e Souza (2009), em regiões áridas ou semiáridas com a presença de topografia acentuada ou de grandes corpos d'água o aquecimento diferencial entre as superfícies possui dinâmica semelhante à observada na brisa marítima e terrestre (FANG et al., 2018). Portanto, quando uma circulação térmica é induzida por forçante orográfica ou pela presença de um rio ou lago, ambas recebem o nome de brisa vale-montanha e lacustre, respectivamente (BAUMBACH; VOGT, 1999).

RAMOS (1975) verificou influência da topografia em eventos de precipitação que ocorreram na cidade pernambucana de Petrolina localizada no semiárido nordestino. O mesmo autor identificou que durante a noite e o início da manhã o ar mais frio e denso que desce a encosta distante 40 km de Petrolina converge sobre a cidade favorecendo o desenvolvimento de nuvens cumulus profundas e ocasionando alguns episódios de precipitação (figura 2.10a). Segundo RAMOS (1975), durante o período vespertino, devido ao maior aquecimento da superfície, os ventos de leste tornam-se mais intensos e acoplam-se ao fluxo de ar que se inverte montanha acima. De acordo com a figura 2.10b, é possível observar que o ar é forçado a subir a encosta favorecendo a formação de nuvens convectivas com expressivo desenvolvimento vertical a barlavento da montanha e com potencial de precipitação. Por outro lado, o fluxo de ar a sotavento perde umidade pelo processo adiabático ao atravessar a montanha e, sem umidade suficiente não há formação de nuvens com potencial de precipitação (MELO, 2011).

Figura 2.10 – Fluxo esquemático da brisa vale-montanha durante (a) o período da manhã e (b) da tarde sobre a cidade de Petrolina-PE.

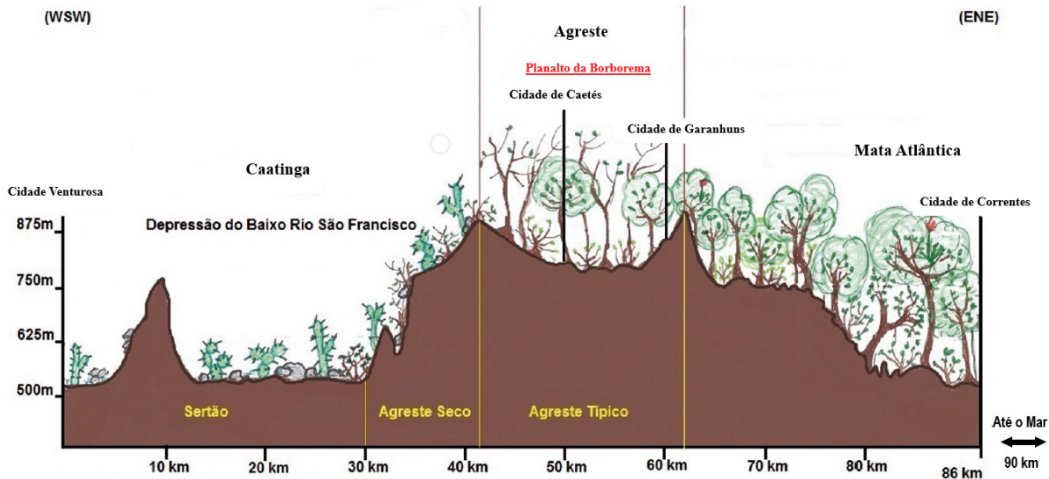


Fonte: adaptado de Ramos (1975).

A vegetação e o clima da região no entorno do Planalto da Borborema, localizado no interior do Nordeste e com altitude máxima de aproximadamente 1200 metros, é um exemplo da dinâmica discutida por RAMOS (1975). A figura 2.11 mostra que a barlavento da Borborema situa-se a Mata Atlântica, caracterizada como uma floresta tropical e com volumes expressivos de precipitação (WANG et al., 2004). A oeste da Zona da Mata, como também é conhecida a Mata Atlântica, encontra-se o Agreste. Essa sub-região mais seca devido à perda de umidade ocasionada pela topografia separa o litoral quente e úmido da sub-região com maior déficit de precipitação do nordeste, denominada de Sertão (SANTANA, 2014). Devido à barreira

orográfica, o Sertão possui uma vegetação única denominada Caatinga, capaz de resistir a ao regime de seca da região (REBOUÇAS, 1997).

Figura 2.11 – Influência do planalto da Borborema no transporte de umidade associado a brisa vale-montanha.



Fonte: Adaptado de Dantas et al. (2015).

A presença de grandes corpos d'água como lagos ou rios também pode induzir a formação da brisa lacustre com a dinâmica de circulação rasa semelhante à marítima e terrestre (COSTA; LYRA, 2012). De acordo com SHEN (1998), esse tipo de circulação pode definir o tipo de cobertura vegetal e as condições climáticas na vizinhança, devido ao transporte de umidade. Os efeitos da brisa lacustre na vizinhança do lago artificial Sobradinho localizado na região semiárida do Nordeste e com extensão de 4500 km<sup>2</sup> e largura máxima entre margens de 50 km, foram estudados por CORREIA, DA SILVA DIAS e DA SILVA ARAGÃO (2006). Estes autores compararam as condições atmosféricas antes da construção do lago e após e observaram a partir de testes de significância que houve mudança no regime da velocidade de vento e da umidade disponível.

Portanto, as circulações de mesoescala termicamente induzidas pelo aquecimento diferencial entre superfícies podem ser um fator importante para estabelecer os regimes de vento, umidade e temperatura do ar, principalmente em regiões com características áridas ou semiáridas (FANG et al., 2018), como é o caso de uma parte do Nordeste do Brasil.

Por fim, os processos atmosféricos pertencentes a escala de tempo inferior a uma hora e de espaçamento horizontal na ordem de metros fazem parte da microescala (ORLANSKI, 1975). Esses processos ocorrem dentro da camada limite atmosférica que é definida como a camada justaposta à superfície terrestre e que se estende até a atmosfera livre (STULL, 1988).



Nesta escala as forçantes associadas à presença da superfície influenciam no escoamento médio do ar próximo ao solo e, por este motivo, a difusão molecular e a turbulência atmosférica são os principais fenômenos de microescala observados no Nordeste do país (FARIA, 2010).

A difusão molecular tem início logo nas primeiras horas do dia com o aquecimento da superfície e, conseqüentemente do ar próximo a ela devido à irradiância solar à superfície. Esse aquecimento desencadeia a convecção que no decorrer do dia, com o aumento da radiação solar, torna-se mais intensa (VIEIRA, 2008). A turbulência, que é o desvio em relação ao valor médio do escoamento do ar, é resultado da convecção e do atrito do ar com a superfície devido à rugosidade do terreno (STULL, 1988). De acordo com FARIA (2010), no Nordeste do Brasil a turbulência é o fenômeno que mais influencia a velocidade do vento na microescala. Isso porque, nessa região a superfície é composta por distintos obstáculos naturais como as chapadas, serras e planaltos, e por regiões com vegetação heterogêneas que alteram o fluxo do escoamento médio do ar e causam a perturbação denominada de turbulência.

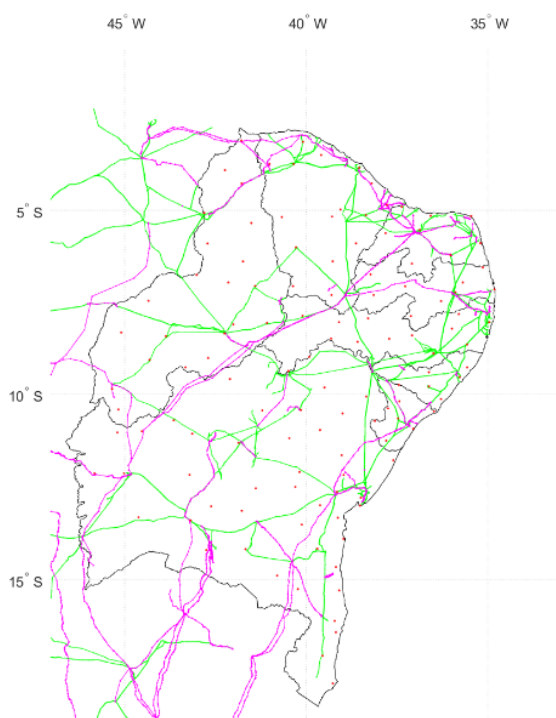


### 3 DADOS E METODOLOGIA

#### 3.1 ÁREA DE ESTUDO

A região alvo da presente tese compreende a área do subsistema Nordeste do Sistema Interligado Nacional que abrange todos os estados do Nordeste com exceção do Maranhão (figura 3.1) (ONS, 2019a). De acordo com ONS (2019b) a geração de energia elétrica a partir do subsistema Nordeste tem sido importante para garantir o suprimento energético para o país, especialmente através da geração eólica. A energia cinética do vento foi responsável por 50% da energia gerada em 2018 (BEZERRA, 2019) e estima-se o crescimento aproximado de 2 GW para os próximos 5 anos (ONS, 2019b). Adicionalmente, outras fontes de energia, como a solar e até mesmo de matriz convencional como hídrica, serão inseridas no subsistema supracitado, o que justifica a relevância e a aplicabilidade imediata deste estudo para a área alvo deste trabalho.

Figura 3.1- Linhas de transmissão planejadas (linhas contínuas magenta), existentes (linhas contínuas verde) e estações meteorológicas (pontos vermelhos) situadas no Subsistema Nordeste do SIN em 2019.



Fonte: o autor.

## 3.2 DADOS

### 3.2.1 Dados meteorológicos medidos

Neste estudo, foram utilizadas inicialmente as medidas de variáveis meteorológicas obtidas em 125 estações meteorológicas automáticas localizadas no subsistema Nordeste do SIN (figura 3.1), pertencentes à rede do Instituto Nacional de Meteorologia (INMET). O INMET iniciou a instalação desta rede no ano de 2005 com o objetivo de monitorar o tempo e o clima no Brasil (MOURA; TADEU; FORTES, 2016). No ano de 2019, a rede de monitoramento totalizou 564 estações em operação sobre o Brasil. Entretanto, ressalta-se que o tamanho da série temporal de cada estação meteorológica pode ser distinto, uma vez que as instalações ocorreram em diferentes datas ou, ainda, que as estações possam ter permanecido fora de operação por um longo período devido a problemas técnicos. As variáveis meteorológicas medidas pela rede são: radiação solar global, velocidade e direção do vento, temperatura e umidade relativa do ar, pressão atmosférica e precipitação. No entanto, apenas a velocidade do vento e a temperatura do ar foram consideradas nesse estudo, pois são as variáveis meteorológicas que mais influenciam no carregamento de linhas de transmissão.

O princípio de medida destas variáveis é o seguinte: a cada 5 segundos o datalogger QML201 (Vaisala) registra os valores instantâneos e a cada minuto realiza a média desses valores amostrados. No final de uma hora é calculada a média dos valores médios de cada minuto e então o valor médio horário é armazenado e distribuído (INMET, 2011).

A velocidade e a direção do vento são medidas pelo sensor WXT521, enquanto a temperatura do ar, pelo QMT103, ambos fabricados pela empresa Vaisala (INMET, 2019).

Neste trabalho, definiu-se como regiões potencialmente críticas para carregamentos de linhas aquelas em que a mediana da velocidade mínima diária do vento medida pelas estações do INMET foi igual ou abaixo a 1 m/s e a mediana da temperatura máxima diária do ar igual ou acima de 31,5 °C. Estes limiares, que indicaram estagnação do vento (calmaria) e máximos de temperatura foram estabelecidos a partir da NBR 5422 da ABNT, que determina que, na ausência de dados medidos, a determinação da temperatura de projeto de uma linha de transmissão deve considerar a velocidade do vento não superior a 1 m/s e a temperatura do ar sendo a média das temperaturas máximas diárias da estação mais próxima em linha reta.

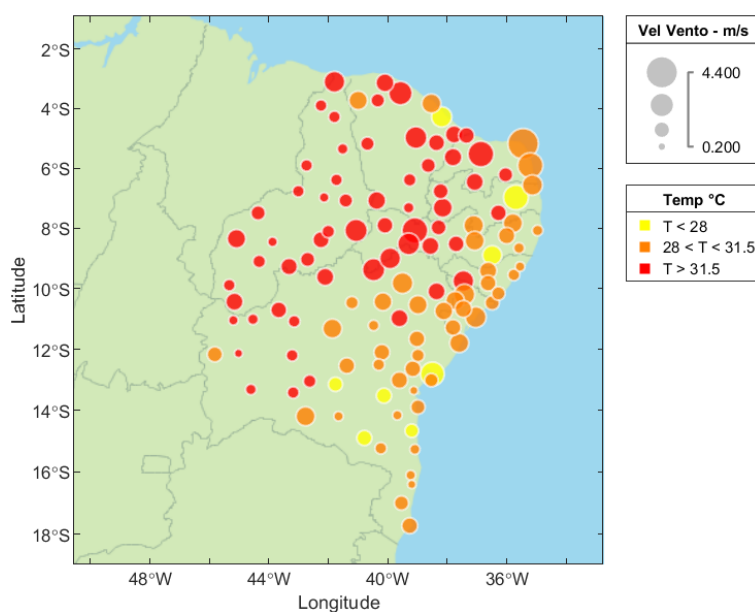
No entanto, para este trabalho este último valor foi obtido a partir do cálculo da mediana das temperaturas máximas diárias registradas por todas as estações automáticas do INMET presentes no subsistema Nordeste do SIN.

A escolha pela mediana em detrimento à média utilizada na NBR 5422 se deu pelo fato de que nenhuma das duas variáveis apresentou distribuição normal. Portanto, a mediana torna-se um valor mais representativo para o objetivo do presente trabalho, pois indica exatamente o percentil 50 do conjunto de dados em questão (WILKS, 2011).

Assim, para identificar em quais regiões há restrição para despacho de carga devido as condições atmosféricas estabelecidas na Norma supracitada calculou-se a mediana das temperaturas máximas diárias e das mínimas velocidades diárias do vento da série temporal de cada estação meteorológica do INMET.

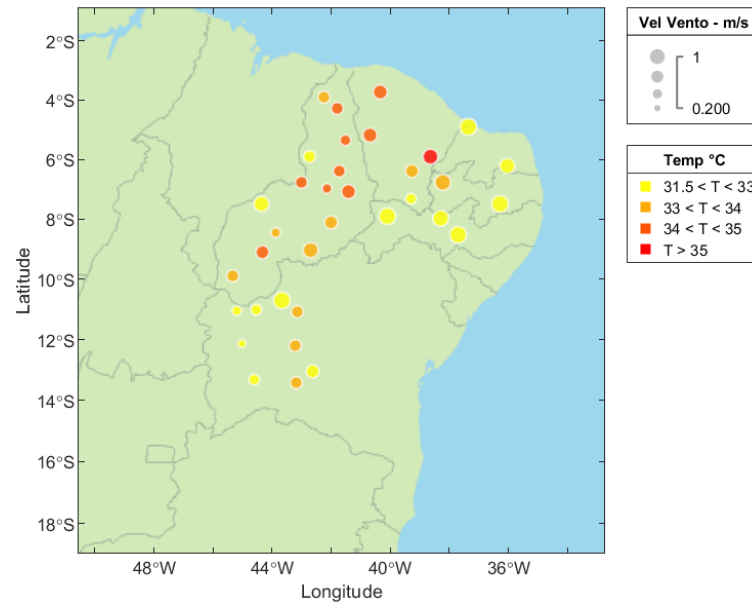
Na figura 3.2 estão plotadas as informações quanto à velocidade do vento, indicadas por meio do tamanho dos círculos e da temperatura do ar, representada por cores. Com o intuito de identificar as regiões com condições climatológicas adversas para o aumento de carga em linhas de transmissão, selecionaram-se as estações meteorológicas cuja mediana das velocidades do vento eram iguais ou inferiores a 1 m/s e as temperaturas máximas do ar iguais ou superiores a 31,5 °C (Figura 3.3). Os círculos presentes na figura 3.3 indicam que a mediana das mínimas velocidades diárias varia entre 0,2 m/s e 1 m/s. As cores quentes indicam o intervalo da mediana das máximas temperaturas diárias.

Figura 3.2 - Mediana das mínimas velocidades diárias do vento indicadas pelo tamanho dos círculos e das máximas temperaturas diárias do ar representadas pelas cores.



Fonte: o autor.

Figura 3.3 – Localização das estações meteorológicas com mediana da velocidade mínima do vento igual ou menor a 1 m/s e temperatura máxima do ar maior ou igual a 31,5 °C.



Fonte: o autor.

Após a aplicação desse critério restaram 35 estações meteorológicas, as quais foram utilizadas para subsidiar as análises da presente tese. Os metadados e as medianas das estações meteorológicas, que identificam as regiões de mínimos de velocidade do vento e máximos de temperatura do ar são apresentados na Tabela 3.1.

Tabela 3.1 – Estações meteorológicas do INMET com mediana da mínima velocidade diária do vento  $\leq 1$  m/s e máxima temperatura diária do ar  $\geq 31,5$  °C, localizadas no subsistema nordeste do SIN em 2018.

Nome	Latitude (°)	Longitude (°)	Altitude (m)	Código INMET	UF	V (m/s)	T (°C)
Barreiras	-12,1	-45,0	470,3	A402	BA	0,2	32
Santa Rita de Cássia	-11,0	-44,5	450,0	A415	BA	0,4	32,8
Correntina	-13,3	-44,6	549,0	A416	BA	0,4	31,6
Bom Jesus da Lapa	-13,4	-43,1	440,0	A418	BA	0,5	33,1
Barra	-11,0	-43,1	403,0	A429	BA	0,5	33,6
Buritirama	-10,7	-43,6	506,5	A432	BA	1,0	32,6
Ibotirama	-12,1	-43,2	430,0	A439	BA	0,5	33,7
Formosa do Rio Preto	-11,0	-45,2	488,0	A452	BA	0,3	32,4
Macaúbas	-13,0	-42,6	574,0	A454	BA	0,6	32,4
Sobral	-3,7	-40,3	109,6	A306	CE	0,7	34,4
Barbalha	-7,3	-39,3	409,0	A315	CE	0,4	31,8
Iguatu	-6,3	-39,2	233,0	A319	CE	0,6	33,0
Crateus	-5,1	-40,6	291,0	A342	CE	0,7	34,3
Jaguaribe	-5,9	-38,6	184,0	A358	CE	0,8	35,1
São Gonçalo	-6,7	-38,2	233,0	A333	PB	0,9	33,4
Cabaceiras	-7,4	-36,2	436,0	A348	PB	1,0	31,8

Continuação – Tabela 3.1 – Estações meteorológicas do INMET com mediana da mínima velocidade diária do vento  $\leq 1$  m/s e máxima temperatura diária do ar  $\geq 31,5$  °C, localizadas no subsistema nordeste do SIN em 2018.

Ibimirim	-8,5	-37,7	448,0	A349	PE	1,0	32,6
Serra Talhada	-7,9	-38,2	461,0	A350	PE	0,9	31,8
Ouricuri	-7,8	-40,1	464,0	A366	PE	1,0	32,0
Florianópolis	-6,7	-43,0	123,2	A311	PI	0,5	34,1
Bom Jesus do Piauí	-9,0	-44,3	415,0	A326	PI	0,6	34
Piripiri	-4,2	-41,7	161,2	A335	PI	0,5	34,2
Alvorada do Gurgueia	-8,4	-43,8	270,0	A336	PI	0,3	33,7
Esperantina	-3,9	-42,2	65,0	A338	PI	0,5	33,9
Picos	-7,0	-41,4	233,0	A343	PI	0,7	34,7
São Raimundo Nonato	-9,0	-42,7	402,0	A345	PI	0,8	33,4
Uruçuí	-7,4	-44,3	393,0	A346	PI	0,8	32,4
Oeiras	-6,9	-42,1	156,0	A354	PI	0,3	34,7
Castelo do Piauí	-5,3	-41,5	286,0	A361	PI	0,4	34,1
São Pedro do Piauí	-5,9	-42,7	287,0	A362	PI	0,5	32,9
Valença do Piauí	-6,3	-41,7	301,0	A363	PI	0,5	34,3
Gilbués	-9,8	-45,3	425,0	A364	PI	0,5	33,0
Canto do Buriti	-8,1	-42,0	308,0	A365	PI	0,6	33,7
Mossoró	-4,9	-37,3	36,0	A318	RN	1,0	32,2
Santa Cruz	-6,2	-36,0	237,0	A367	RN	0,8	32,1

Fonte: o autor.

### 3.2.2 Série temporal sintética de dados meteorológicos

A série sintética dos dados de velocidade do vento e temperatura do ar utilizadas neste trabalho foram produzidas a partir da técnica de *downscaling* dinâmico. Essa técnica foi empregada com o objetivo de obter informações da atmosfera com maior detalhamento quando comparado aos resultados dos modelos de escala global, pois com o aumento da resolução horizontal as forçantes regionais como orografia, linhas de costa, cobertura vegetal e corpos d'água como rios e lagos são melhor representados e, assim, os modelos atmosféricos podem realizar simulações com melhor desempenho (MANZANAS et al., 2018; XU; YANG, 2012).

A versão 4 do modelo regional WRF – ARW (Weather Research and Forecasting - Advanced Research WRF) foi utilizada para o aumento da resolução horizontal a partir dos dados do GFS (Global Forecasting System) desenvolvido pelo NCEP (National Centers for Environmental Prediction). Conforme descrito em SKAMAROCK et al. (2008), o WRF possui coordenada vertical de pressão que acompanha o terreno, esquema de advecção de quinta ordem na horizontal e terceira na vertical, com grade tipo C de Arakawa e integração Runge-Kutta de terceira ordem. Ainda de acordo com o citado, este modelo é compressível e não hidrostático, o que possibilita a integração com alta resolução horizontal. Os usuários deste modelo têm a possibilidade de definir as parametrizações físicas de camada limite planetária, radiação de

onda curta e longa, microfísica e convecção o que permite a integração do WRF de maneira personalizada.

As condições iniciais e de fronteira para a grade do experimento numérico utilizando o WRF são advindas do modelo GFS na versão de  $0,5^\circ \times 0,5^\circ$  de resolução horizontal, frequência temporal de 6 horas (00, 06, 12 e 18 UTC) e com 64 níveis verticais na coordenada sigma-pressão. De maneira geral, a reconstrução do estado da atmosfera no tempo passado produzida pela combinação de dados meteorológicos observados e modelos físicos é denominada de análise ou reanálise (COMPO et al., 2011). O uso desse conjunto de dados garante uma série temporal contínua e sem buracos e com cobertura espacial capaz de fornecer aos modelos regionais o estado da atmosfera em qualquer lugar do globo.

A reanálise ERA5 do ECMWF (European Centre for Medium-range Weather Forecast) foi considerada neste trabalho, pois além de possuir uma resolução horizontal de  $0,25^\circ \times 0,25^\circ$  (~27 km) sobre todo o globo disponibiliza dados na frequência horária (TAREK; BRISSETTE; ARSENAULT, 2019). Esse fato permite avaliar o desempenho do preenchimento de buracos das séries temporais da rede INMET o que até então não era possível com outras análises (por exemplo GFS e Reanalysis II) que disponibilizam dados a cada 6 horas. Ainda a respeito da reanálise do ERA5 o motivo para não a utilizar no processo de downscaling dinâmico se deu pela dificuldade de obtenção dos 12 meses consecutivos de dados necessários para a realização do experimento numérico. Atualmente, só é possível acessar um grande volume de dados do ECMWF utilizando a plataforma Copernicus de distribuição de arquivos binários. E essa plataforma possui uma enorme fila de espera, além de ter apresentado instabilidade e falhas nas inúmeras tentativas de *download* dos dados. Desta forma, decidiu-se pela obtenção apenas dos dados horários das variáveis temperatura do ar a 2 metros e componente zonal e meridional do vento a 10 metros para posterior cálculo da velocidade do vento a 10 metros no ponto de grade mais próximo da localização de cada estação meteorológica do INMET considerada neste estudo.

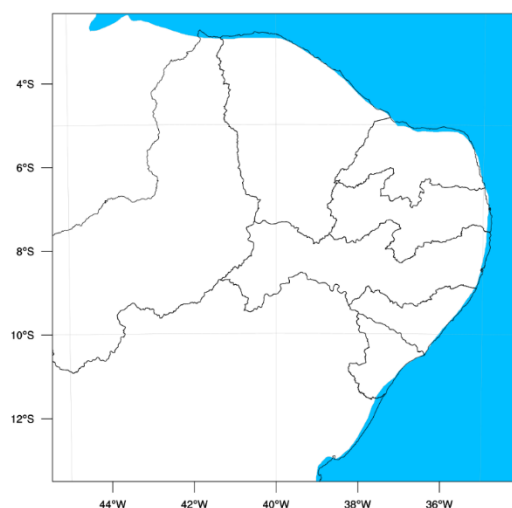
A simulação numérica consistiu em uma integração de 370 dias nos quais os primeiros 5 dias foram descartados devido ao período *spinup*. De acordo com MISRA E KANAMITSU (2004) nos passos iniciais de integração de modelos regionais ocorre o ajuste dos campos atmosféricos e por esse motivo, eliminam-se as primeiras horas ou dias das simulações durante as avaliações do modelo. O domínio da simulação (

Figura 3.4) possui 5 km de espaçamento horizontal e está centrado em  $-8^\circ$  e  $-40^\circ$  de latitude e longitude, respectivamente. A escolha por este domínio está relacionada a demanda de preenchimento das séries temporais das 35 estações meteorológicas que foram selecionadas



durante o processo de análise de agrupamento que será apresentado na seção 3.4 deste trabalho. No tocante às parametrizações físicas utilizaram-se Dudhia para radiação de onda curta e RRTM para de onda longa enquanto para microfísica de nuvens foi considerada a WSM6 e para a camada limite planetária a YSU. Por fim, NOAH-LSM foi utilizada para os processos de superfície e Kain-Fritsch para convecção. A escolha por estas parametrizações se deu depois de uma extensa revisão de literatura. Durante a busca por referências, verificou-se que o trabalho de SOUZA (2016) avaliou o desempenho do modelo WRF relacionado ao conjunto de parametrizações utilizadas para simular brisas no nordeste do Brasil. Portanto, por similaridade (região de integração e variável meteorológica) utilizou-se as parametrizações que apresentaram melhor desempenho nas simulações do trabalho supracitado para realizar a integração do WRF na presente tese.

Figura 3.4 – Domínio da simulação com a grade de 5 km.



Fonte: o autor.

### 3.2.3 Banco de dados relacional

No tocante à meteorologia, o volume de dados diários gerados é evidenciado por dois motivos principais. O primeiro refere-se à expansão das redes de monitoramento e o aumento da frequência temporal das medições que em décadas anteriores se resumia a horários sinóticos, mas que agora acontece até mesmo na escala de minutos. O outro motivo é devido ao avanço tecnológico que possibilita o desenvolvimento de novos modelos e técnicas de simulação e previsão do tempo e clima. Com isso, tem-se o aumento do número de usuários e a demanda por armazenamento e distribuição eficiente dos dados (MA et al., 2010).

A realização de tarefas envolvendo uma enorme coleção de dados tem sido facilitada por operações relacionadas entre si a partir do uso de banco de dados relacionais (PAREDAENS

et al., 1989). No mercado há inúmeras empresas que comercializam bancos de dados como Oracle, Microsoft Access, Microsoft SQL Server, PostgreSQL e MySQL e que possibilitam a manipulação de *big data* de forma relacionada.

Neste trabalho foi utilizado o banco de dados Postgres de código fonte aberto desenvolvido na Universidade da Califórnia, com capacidade ilimitada para armazenamento de dados e que suporta a linguagem estruturada padrão SQL (*Structured Query Language*) (SHARYGIN et al., 2017). De acordo com SHARYGIN et al. (2017), o SQL foi uma das primeiras linguagens estruturadas implementada em banco de dados relacional e que permite inúmeras operações como inserção de dados e consultas complexas com operações matemáticas e lógicas. Adicionalmente, a escolha por esse banco se deu porque diversos softwares estatísticos como o R e o Matlab<sup>®</sup> (utilizado neste trabalho) possuem pacotes que possibilitam acesso direto à base de dados a partir de instruções SQL.

Para o desenvolvimento dessa tese foi criada uma base de dados no Postgres com três tabelas que se relacionam entre si: *station*, *station\_type* e *tb\_registros*. A tabela *station* (figura 3.5) contém o metadados das estações: nome, latitude, longitude, altitude, *timezone*, sigla e região da unidade federativa, código e tipo da estação.

A tabela *station\_type* contém o registro da origem do dado (medido ou simulado), pois para cada ponto de grade mais próximo das estações meteorológicas do INMET foram armazenadas a temperatura do ar e a velocidade do vento extraída da reanálise do ERA5 e do downscaling obtido a partir da integração do WRF, além do dado medido. Por fim, a tabela *tb\_registros* armazena os registros horários medidos pelas estações ou simulados pela reanálise ERA5 ou pelo WRF.

Figura 3.5 – Tabela *station* contendo o metadados de doze estações meteorológicas do INMET.

id	name	latitude	longitude	altitude	code_wmo	uf	region	station_type	tz	
bigint	character varying (30)	character varying (30)	character varying (30)	real	character v	charac	characte	bigint	bigint	
1	2868	ARAPIRACA	-9.80416666	-36.61888888	241	A353	AL	NE	2	-3
2	2869	CORURIBE	-10.14527777	-36.28638888	74	A355	AL	NE	2	-3
3	2871	SAO LUIS DO QUITUNDE	-9.28749999	-35.56583333	19	A356	AL	NE	2	-3
4	2884	IBOTIRAMA	-12.19249999	-43.2125	430	A439	BA	NE	2	-3
5	2885	AMARGOSA	-13.00944444	-39.61694443	407	A434	BA	NE	2	-3
6	2888	BELMONTE	-16.08805555	-39.21555555	88	A447	BA	NE	2	-3
7	2889	BRUMADO	-14.18166666	-41.67222221	470	A433	BA	NE	2	-3
8	2890	BURITIRAMA	-10.72305554	-43.65111111	506.5	A432	BA	NE	2	-3
9	2891	CONDE	-11.8	-37.59972221	31.9	A431	BA	NE	2	-3
10	2892	DELFINO	-10.455	-41.20722222	637	A443	BA	NE	2	-3
11	2893	EUCLIDES DA CUNHA	-10.53722221	-38.99666666	430	A442	BA	NE	2	-3
12	2894	GUANAMBI	-14.20805555	-42.74944444	882	A426	BA	NE	2	-3

Fonte: o autor.

### 3.3 CONTROLE DE QUALIDADE DOS DADOS

O tratamento e as análises estatísticas dos dados foram realizados utilizando o *software* Matlab®. Com o objetivo de garantir a qualidade dos dados medidos pela rede do INMET foi desenvolvido um programa para avaliar a consistência dos dados e remover os valores espúrios das séries temporais das 125 estações meteorológicas utilizadas neste trabalho. Os critérios utilizados foram propostos por FIEBRICH et al. (2010), os quais tratam como inconsistentes os dados de temperatura do ar que não pertencem ao intervalo de -30°C e 50°C ou que em um intervalo de 1 hora não variem no mínimo 0,1°C. Para a velocidade do vento, estes autores definem como dado espúrio aquele fora do intervalo entre 0 e 40 m/s ou aqueles que não variam no intervalo de duas horas.

Adicionalmente, estações meteorológicas com série temporal menor que 12 meses (considerando as falhas) foram removidas deste trabalho por não permitirem a análise sazonal do comportamento do vento e da temperatura do ar. Como consequência destas restrições foram utilizadas 120 estações meteorológicas para as próximas etapas do estudo.

### 3.4 ANÁLISE DE CLUSTER

A análise de cluster é uma técnica exploratória de dados também chamada de análise de agrupamento, que agrupa objetos (variáveis) com base na similaridade ou dissimilaridade entre eles a partir de uma distância métrica (HAIR JR. et al., 2014). Esta técnica tem sido utilizada há muitas décadas para definir ou redefinir zonas climáticas em contraponto às tradicionais classificações, pois permite que regiões climáticas homogêneas sejam encontradas com base em dados meteorológicos observados ou estimados.

Por exemplo, FOVELL; FOVELL (1993) compararam a classificação climática de Köppen produzida para os Estados Unidos a partir dos dados medidos pelo NCDC (*National Climate Data Center*) contra a divisão climática resultante da técnica de agrupamento. Estes autores encontraram diferenças substanciais entre as duas classificações indicando que os clusters possuem mais robustez para representar os diferentes tipos de clima daquele país.

UNAL; KINDAP; KARACA (2003) também empregaram a análise de cluster para redefinir zonas climáticas estabelecidas convencionalmente sobre a Turquia. Para isso, os autores utilizaram dados mensais medidos entre 1951 e 1998 de 113 estações meteorológicas. Os agrupamentos geraram o mesmo número de regiões climáticas quando comparado com as zonas estabelecidas por convenção, entretanto a nova regionalização agrupou estações com

características mais similares no tocante à temperatura do ar e precipitação e com o delineamento mais refinado das fronteiras entre as regiões climáticas distintas.

A respeito da classificação do clima global, NETZEL; STEPINSKI (2016) utilizaram o conjunto de dados de 50 anos da WorldClim (HIJMANS et al., 2005) para criar 32 distintas classificações climáticas baseada em análise de cluster. Os resultados foram confrontados com uma das classificações mais presente na literatura, denominada de Köppen-Geiger (KÖPPEN; GEIGER, 1936). Os autores concluíram que, de todas as classificações produzidas pela técnica estatística multivariada, a melhor foi obtida utilizando temperatura do ar e a amplitude diária (temperatura máxima – temperatura mínima) e a precipitação. Adicionalmente, ETZEL; STEPINSKI (2016) indicaram que a análise de agrupamento apresentou clusters com caráter internamente mais homogêneos e externamente mais distintos quando comparado à de Köppen-Geiger.

Outros autores, como CHAKRABORTY; NAGWANI; DEY (2014) empregaram a análise de cluster para obter a previsão do tempo do dia seguinte considerando os efeitos da poluição do ar na condição atmosférica. Para isso, os autores utilizaram um conjunto de dados contendo 4 variáveis de poluição do ar coletadas entre 2009 e 2010 em Bengala Ocidental, Índia. A partir destes dados os autores criaram agrupamentos para categorizar as condições do tempo de acordo com os poluentes observados. Posteriormente, um processo incremental foi utilizado para associar cada novo dado medido a um cluster existente e assim prever a condição futura do tempo.

Recentemente, GOVENDER; BROOKS; MATTHEWS (2018) também utilizaram técnicas de agrupamento com o objetivo de prever a irradiância solar para o dia seguinte em Durban (África do Sul). A metodologia empregada foi semelhante a proposta por CHAKRABORTY; NAGWANI; DEY (2014), porém os primeiros autores consideraram variáveis solarimétricas como irradiância solar direta e cobertura de nuvens.

Na literatura científica sobre o tema, o uso da análise de cluster é indicado como uma poderosa ferramenta para mineração de dados a qual também é chamada de análise exploratória de dados (ATLURI; KARPATNE; KUMAR, 2018). Na Meteorologia, onde geralmente há um grande volume de dados distribuídos no tempo e no espaço, o uso de agrupamentos permite a minimização deste conjunto de informações preservando importantes elementos e sem que haja deformação na amostra (FAGHMOUS; KUMAR, 2014). Isso permite a extração de padrões de um conjunto de dados auxiliando a análise de condições climáticas e suas interconexões com diversos fenômenos atmosféricos (HADZIMEJLIC; DONKO; HADZIMEJLIC, 2012).

Diante do exposto, a presente tese utilizou-se da análise de cluster para a análise exploratória dos dados das 35 estações meteorológicas presentes na Tabela 3.1 para constituir agrupamentos com condições de velocidade do vento e temperatura do ar homogêneas. Por conseguinte, essa divisão permitiu identificar sub-regiões críticas de carregamento de linhas aéreas de transmissão e então discutir as escalas e os sistemas atmosféricos que influenciam tais regiões.

O primeiro passo foi efetuar a padronização dos dados para eliminar a diferença de escala entre as medianas de velocidade do vento e de temperatura do ar. Esse processo garante que ambas as variáveis possuam o mesmo peso, uma vez que elas apresentam magnitudes diferentes (HAIR JR. et al., 2014). Em outras palavras, a padronização remove o viés provocado pela diferença de escala das variáveis a partir da conversão de cada variável em escore (equação 1) e isso garante que ambas terão peso igual na análise de cluster (UNAL; KINDAP; KARACA, 2003).

$$Z = \frac{x - \mu}{\sigma} \quad \text{Equação 3.1}$$

onde  $x$  é a variável de interesse com média  $\mu$  e desvio padrão  $\sigma$ .

Após a padronização dos dados foi utilizado o algoritmo hierárquico aglomerativo para executar a tarefa de agrupamento. Isso significa que cada conjunto de variável V e T atribuído a uma estação meteorológica é tratado inicialmente como um cluster e estes são unidos uns aos outros iterativamente, até que todas as estações estejam sob um único agrupamento (LI; DE RIJKE, 2017). A escolha por este algoritmo em detrimento a outras técnicas não hierárquicas, como a k-means, se deu pelo fato de que esta última exige o conhecimento prévio do número de cluster a ser formado (HADZIMEJLIC; DONKO; HADZIMEJLIC, 2012). E de acordo com UNAL; KINDAP; KARACA (2003), este algoritmo hierárquico é ideal para análises exploratórias de dados, sendo este o objetivo da utilização da técnica de agrupamento.

Antes de iniciar o processo iterativo de agrupamento foi necessário calcular a distância métrica entre cada par de observações. Isto foi realizado utilizando-se uma medida de similaridade que indica o grau de correspondência entre cada conjunto de variável V e T. Em outras palavras, a homogeneidade dentro dos agrupamentos e a heterogeneidade entre eles é determinada pela medida de similaridade que é calculada por uma distância métrica (WILKS, 2011).

Uma vez determinada a similaridade entre os pares de observações (V e T), iniciou-se a ligação entre os clusters de maior similaridade (menor distância métrica), formando assim um

novo agrupamento. E esse processo iterativo terminou apenas quando a distância entre o último cluster e todos os agrupamentos já formados foi computada (HAIR JR. et al., 2014).

Entretanto, há na literatura diversas maneiras de calcular a distância métrica e o método de ligação para formar agrupamentos (LIU et al., 2013). E dependendo da escolha destes parâmetros pode-se obter resultados distintos de agrupamento para um mesmo conjunto de dados (HOX, 2010). Essa diferença pode ser evidenciada em um diagrama de relacionamento hierárquico chamado de dendrograma.

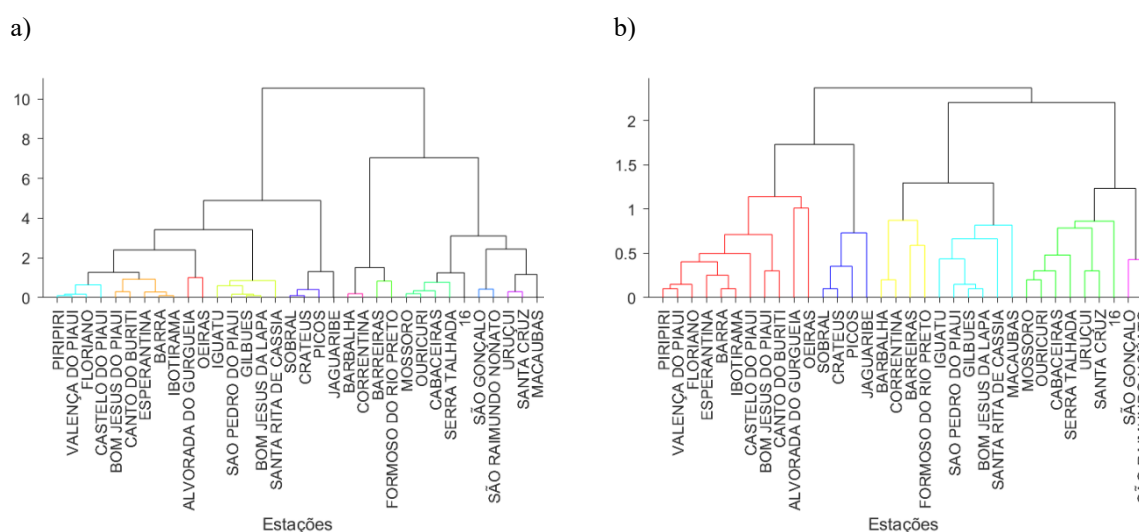
A figura 3.6 apresenta duas árvores hierárquicas construídas a partir do mesmo conjunto de dados presente na tabela 3.1, porém com matrizes de dissimilaridade calculadas com distâncias métricas e métodos de ligação diferentes. No eixo das ordenadas estão indicadas as medidas de similaridades e no eixo das abscissas, as estações meteorológicas que representam o conjunto de variáveis V e T. A união de duas estações (agrupamentos) é representada por uma barra horizontal e a posição dessa barra em relação ao eixo das ordenadas fornece a medida de similaridade entre ambas.

Para definir a quantidade de agrupamentos, é necessário efetuar um corte imaginário horizontal no dendrograma. Porém, para determinar a altura deste corte, é utilizado o coeficiente de inconsistência definido por JAIN; DUBES (1988). Este coeficiente compara a altura de cada nó (ponto de união de dois cluster) da árvore hierárquica com a altura média das outras ligações no mesmo nível de hierarquia. Desta forma, a união de clusters dissimilares possuem um alto valor de inconsistência enquanto ligações de agrupamentos com similaridade próxima possuem baixos coeficientes. Assim, todas as uniões de valor igual ou menor à máxima inconsistência é agrupada dentro de um cluster.

As cores presentes nos dendrogramas da figura 3.6 expressam os agrupamentos formados após o corte horizontal no limiar de inconsistência de 1,15. O que se observa é que a árvore hierárquica construída utilizando a distância métrica *cityblock* para calcular a similaridade e o método de ligação de *ward* para unir iterativamente cada cluster gerou 13 agrupamentos (figura 3.6a), contra 6 considerando a distância euclidiana e o método de pesos (figura 3.6b). Também é possível verificar na figura em questão que algumas estações são agrupadas entre si de maneira diferente dependendo dos parâmetros utilizados. Por exemplo, na figura 3.6a, a estação Macaúbas é considerada um cluster isolado enquanto na figura 3.6b, tal estação faz parte de um agrupamento com mais 5 estações. Ou seja, como exposto acima, para um mesmo conjunto de dados a análise de cluster é fortemente dependente do método de ligação e da distância métrica escolhida.

Assim, para determinar qual par de parâmetros (distância métrica e método de ligação) é mais consistente para o agrupamento hierárquico do conjunto de dados da tabela 3.1 foi calculado o coeficiente de correlação cofenética. Esse coeficiente compara os resultados do agrupamento para um mesmo conjunto de dados utilizando diferentes distâncias métricas e métodos de ligação (SOKAL; ROHLF, 1962). O valor próximo a 1 da correlação cofenética indica que os parâmetros utilizados para construir a árvore hierárquica representam bem a similaridade do conjunto de dados.

Figura 3.6 – Dendrograma formado utilizando a distância métrica (a) *cityblock* e método de ligação *ward* e (b) a distância euclidiana e o método de pesos.



Fonte: o autor.

Após construir 20 árvores hierárquicas a partir da combinação de pares de distância e de ligação distintos, verificou-se que a utilização da distância euclidiana com a ligação *average* apresentou a maior correlação cofenética (tabela 3.2) para o conjunto de dados em questão. Esta distância é a medida em linha reta entre dois conjuntos de variáveis  $V$  e  $T$ , enquanto o método *average* une agrupamentos em cada iteração considerando a distância mínima entre eles calculada a partir da média das distâncias entre todos os pares de observação (LI; DE RIJKE, 2017).

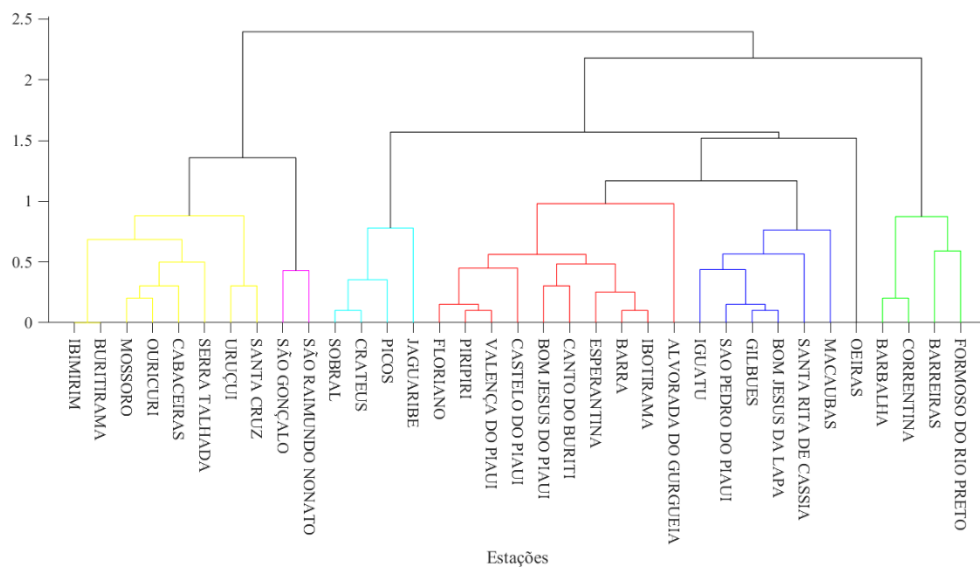
Por fim, como exposto anteriormente, o algoritmo hierárquico aglomerativo é finalizado quando todas as estações meteorológicas são unificadas em um grande cluster. E para concluir o processo de agrupamento, foi calculado o máximo coeficiente de inconsistência do dendrograma formado. Assim, a máxima similaridade entre as estações foi obtida ao nível de 1,1 da altura de similaridade da árvore hierárquica formada após o processo de agrupamento pelo método hierárquico.

Tabela 3.2 – Correlação cofenética dos pares de distância métrica e métodos de ligação utilizados para a construção de diferentes árvores hierárquicas utilizando o conjunto de variável V e T.

Distância métrica	Método de ligação	Correlação cofenética
euclidiana	ward	0,74
euclidiana	completo	0,67
euclidiana	simples	0,68
euclidiana	pesos	0,70
<b>euclidiana</b>	<b>média</b>	<b>0,78</b>
mahalanobis	ward	0,68
mahalanobis	completo	0,65
mahalanobis	simples	0,67
mahalanobis	pesos	0,66
mahalanobis	média	0,77
cityblock	ward	0,71
cityblock	completo	0,68
cityblock	simples	0,60
cityblock	pesos	0,69
cityblock	média	0,76
chebychev	ward	0,73
chebychev	completo	0,72
chebychev	simples	0,66
chebychev	pesos	0,70
chebychev	média	0,74

Fonte: o autor.

Figura 3.7 – Dendrograma construído utilizando a distância euclidiana como medida de similaridade e método de ligação *média* aplicado ao conjunto de variáveis das 35 estações meteorológicas presentes na tabela 3.1



\*parâmetro de corte de 1,13 da matriz de similaridade.

Fonte: o autor.



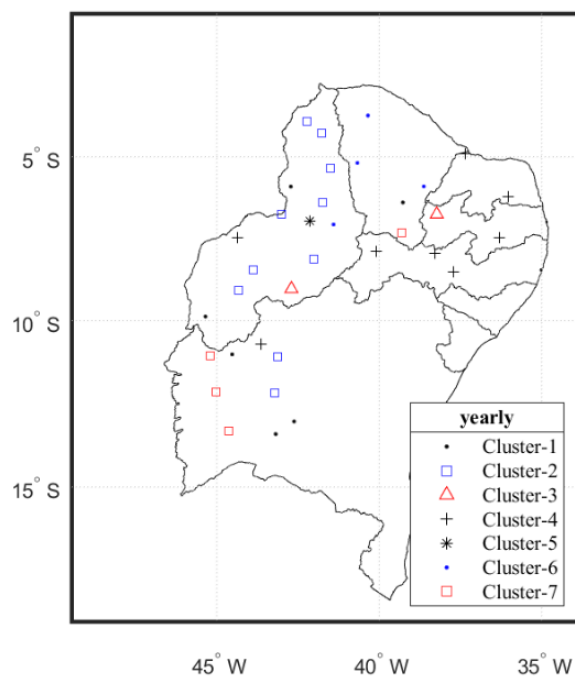
A tabela 3.3 sintetiza os sete agrupamentos formados a partir do parâmetro de corte aplicado ao dendrograma obtido pelo método hierárquico (Figura 3.7) e as medianas das temperaturas máximas diárias e das velocidades mínimas do vento. A distribuição espacial das estações agrupadas é apresentada na figura 3.8.

Tabela 3.3 – Agrupamentos formados por meio da análise de cluster considerando a similaridade da mediana das máximas temperaturas diárias do ar e da mediana das mínimas velocidades do vento.

Agrupamentos	Nº de estações	$T_{med}$	Desvio padrão	$V_{med}$	Desvio padrão
Cluster 1	6	33,0	0,25	0,5	0,08
Cluster 2	10	34,0	0,25	0,5	0,09
Cluster 3	2	33,4	0,00	0,9	0,07
Cluster 4	8	32,2	0,32	1,0	0,09
Cluster 5	1	34,7	0,00	0,3	0,00
Cluster 6	4	34,6	0,36	0,7	0,05
Cluster 7	4	31,9	0,34	0,4	0,10

Fonte: o autor.

Figura 3.8 – Distribuição espacial dos agrupamentos formados a partir do método hierárquico considerando 35 estações meteorológicas da rede INMET.



Fonte: o autor.

### 3.5 PREENCHIMENTO DE FALHAS

Ao contrário de séries sintéticas que fornecem um período completo de dados, a manipulação de séries temporais a partir de dados medidos por estações meteorológicas pode evidenciar que estas nem sempre estão completas. A ausência de dados na maioria das vezes ocorre por problemas de operação dos sensores durante a medição das variáveis meteorológicas ou pela remoção dos dados espúrios durante o processo de qualificação.

Assim, com o objetivo de preencher os buracos presentes nas séries temporais de velocidade do vento e temperatura do ar das 35 estações meteorológicas, foi avaliado o desempenho de 4 distintos métodos de preenchimento.

Para permitir a avaliação dos métodos, primeiramente, removeu-se 10% dos dados de velocidade do vento e temperatura do ar das séries temporais originais, os quais foram denominados de falhas artificiais. Estes buracos gerados artificialmente a partir de uma função randômica desenvolvida utilizando o software Matlab possuem tamanhos de 1 a 24 horas consecutivas e em locais aleatórios das séries temporais para o ano de 2017.

A escolha pelo ano de 2017 foi motivada por ter sido o último ano de dados completos do INMET presente no banco de dados modelado para o desenvolvimento da presente tese. Ou seja, comparando 2017 com os anos anteriores, é neste ano que há o maior número de estações em operação na região alvo deste trabalho.

As informações referentes ao tamanho e posição da falha, assim como os dados originais removidos artificialmente foram armazenadas para posterior avaliação de desempenho do preenchimento. A tabela 3.4 apresenta um exemplo hipotético das falhas geradas para uma série temporal de uma estação do INMET utilizada neste trabalho.

Tabela 3.4 – Exemplo hipotético das informações das falhas artificiais geradas para 10% dos dados de uma variável meteorológica qualquer pertencente a uma estação meteorológica.

Ano	Início da falha			Ano	Final da falha			Tamanho da falha
	Mês	Dia	Hora		Mês	Dia	Hora	
2017	1	28	10	2017	1	28	13	4
2017	8	13	22	2017	8	14	10	13
2017	3	11	18	2017	3	12	12	19
2017	8	13	8	2017	8	13	11	4
2017	6	15	20	2017	6	16	12	17
2017	9	10	20	2017	9	11	19	24
2017	12	25	13	2017	12	25	16	4
2017	7	5	23	2017	7	6	14	16

Continuação – Tabela 3.4 – Exemplo hipotético das informações das falhas artificiais geradas para 10% dos dados de uma variável meteorológica qualquer pertencente a uma estação meteorológica.

2017	1	21	0	2017	1	21	12	13
2017	2	19	1	2017	2	19	9	9
2017	10	14	16	2017	10	14	17	2
2017	3	13	19	2017	3	14	9	15
2017	12	30	13	2017	12	30	14	2
2017	2	20	15	2017	2	20	20	6
2017	8	24	3	2017	8	24	3	1
2017	3	12	17	2017	3	13	10	18
2017	8	7	1	2017	8	7	18	18
2017	7	27	11	2017	7	27	16	6
2017	7	8	8	2017	7	9	0	17
2017	1	9	2	2017	1	9	2	1
2017	5	18	11	2017	5	18	16	6
2017	4	26	4	2017	4	26	21	18
2017	5	12	4	2017	5	12	8	5
2017	7	29	6	2017	7	29	17	12
2017	12	21	18	2017	12	22	0	7
2017	7	26	20	2017	7	27	4	9
2017	4	24	19	2017	4	24	19	1
2017	9	1	17	2017	9	2	14	22
2017	1	1	20	2017	1	2	2	7
2017	12	30	6	2017	12	30	6	1
2017	9	25	8	2017	9	25	10	3
2017	8	20	0	2017	8	20	19	20
2017	3	3	17	2017	3	3	23	7
2017	1	2	18	2017	1	3	3	10
2017	3	16	17	2017	3	17	13	21
2017	9	12	15	2017	9	13	13	23
2017	4	28	19	2017	4	28	19	1
2017	2	21	17	2017	2	21	18	2
2017	5	20	21	2017	5	21	2	6
2017	12	8	16	2017	12	9	15	24
2017	1	15	18	2017	1	16	15	22
2017	12	5	13	2017	12	6	8	20
2017	5	8	20	2017	5	9	2	7
2017	3	10	17	2017	3	11	1	9
2017	9	1	13	2017	9	1	15	3
2017	6	8	4	2017	6	8	7	4
2017	2	1	19	2017	2	2	4	10
2017	8	4	6	2017	8	4	10	5
2017	6	19	7	2017	6	19	12	6
2017	1	20	2	2017	1	20	19	18
2017	9	7	9	2017	9	8	6	22
2017	9	18	1	2017	9	18	12	12
2017	1	10	7	2017	1	10	15	9
2017	6	27	15	2017	6	27	20	6

Continuação – Tabela 3.4 – Exemplo hipotético das informações das falhas artificiais geradas para 10% dos dados de uma variável meteorológica qualquer pertencente a uma estação meteorológica.

2017	6	9	15	2017	6	10	0	10
2017	6	10	13	2017	6	10	20	8
2017	6	13	1	2017	6	14	0	24
2017	6	18	15	2017	6	19	2	12
2017	1	8	20	2017	1	9	0	5
2017	2	7	5	2017	2	7	13	9
2017	12	2	20	2017	12	2	21	2
2017	12	29	2	2017	12	29	9	8
2017	3	25	0	2017	3	25	22	23
2017	9	20	2	2017	9	20	23	22
2017	2	19	17	2017	2	20	12	20
2017	9	6	11	2017	9	6	21	11
2017	1	25	17	2017	1	26	12	20
2017	5	25	13	2017	5	26	2	14
2017	5	6	7	2017	5	7	6	24
2017	5	20	16	2017	5	20	22	7
2017	12	21	14	2017	12	22	5	16
2017	5	13	9	2017	5	13	19	11
2017	1	22	19	2017	1	23	9	15
2017	5	3	8	2017	5	3	8	1
2017	4	15	6	2017	4	16	4	23
2017	10	19	18	2017	10	20	1	8

Fonte: o autor.

### 3.5.1 Imputação de dados

A imputação de dados para cada falha artificial gerada foi realizada por 4 maneiras distintas, aqui denominadas de métodos:

#### a) Interpolação linear ou média

Neste método o preenchimento dos buracos da série temporal é realizado a partir da interpolação linear se a falha tiver um tamanho de 1 hora ou com o valor médio entre o dado medido 24 horas anteriores à da falha e o posterior a falha para buracos de 2 a 24 horas.

#### b) Interpolação espacial

De acordo com BERNDT; HABERLANDT (2018), o uso de técnicas de interpolação espacial podem ser utilizadas com sucesso para estimar variáveis meteorológicas em locais onde não há dados medidos a partir de medidas próximas. Entretanto, estes autores afirmam que a confiabilidade da estimativa utilizando a interpolação espacial apresenta dependência

relacionada à densidade da rede de observação que é utilizada para a estimação, assim como à resolução temporal e à variabilidade espacial da variável a ser estimada.

Neste trabalho, decidiu-se pelo uso da interpolação ordinária de Kriging como um segundo método para preencher os valores removidos artificialmente das séries temporais originais. Alguns trabalhos, como WU; LI (2013), demonstram que este método é capaz de capturar a variabilidade espacial da temperatura do ar, embora o desempenho das estimativas apresente dependência com as estações do ano. Outros estudos também demonstraram o sucesso no uso da interpolação ordinária de Kriging para mapear a radiação solar global (TOVAR-PESCADOR et al., 2009) e para estimar a precipitação, a pressão atmosférica, a umidade relativa e outras variáveis devido à ocorrência de dados faltantes nas séries temporais (JEFFREY et al., 2001). Segundo LI; HEAP (2014), a interpolação utilizando krigagem ordinária é um método de estimativa geoestatística que utiliza a média ponderada dos vizinhos mais próximos para determinar a distribuição e a variabilidade espacial da variável de interesse com base em um modelo de variograma. Ou seja, o variograma quantifica a variabilidade espacial da variável a ser estimada.

Assim, a estimativa de uma determinada variável meteorológica  $U_0$  para um determinado ponto de interesse é uma combinação linear dos valores das estações vizinhas  $U_i$ :

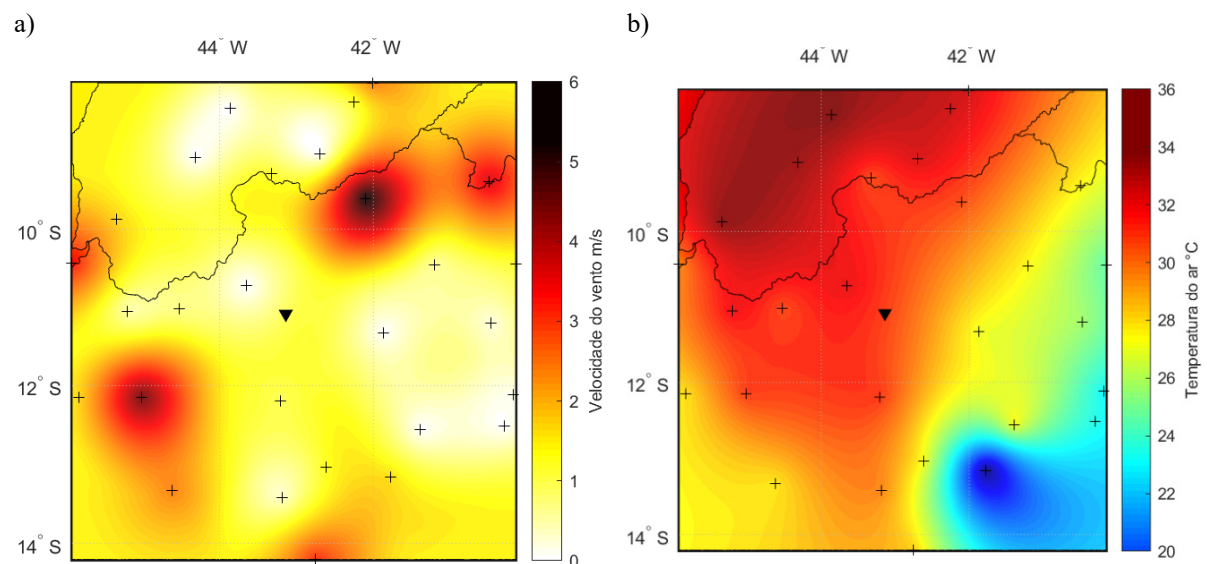
$$U_0 = \sum_{i=1}^{30} \lambda_i U_i \quad \text{Equação 3.2}$$

onde  $\lambda_i$  é o peso estimado para cada  $U_i$  a partir do sistema de equações de Kriging que usa a persistência espacial dos estimadores para realizar a estimação no ponto desejado. Essa persistência é obtida a partir do cálculo do variograma que é uma função linear de primeira ordem da distância entre o ponto a ser estimado e o estimador. Para calcular o variograma e, posteriormente, realizar a interpolação espacial foi definida a utilização máxima de 30 e mínima de 10 estações meteorológicas mais próximas à estação alvo. Em outras palavras, caso alguma das estações vizinhas selecionadas também não possua dado medido para o mesmo horário a ser estimado tal estação é descartada e o processo continua até que o mínimo de estações estimadoras seja igual a 10. Caso a quantidade de estações seja menor que 10, o valor atribuído ao buraco a ser preenchido receberá uma *string* NaN (Not a Number) para indicar um campo vazio. Adicionalmente, para realizar a interpolação espacial para cada estação alvo é necessário definir uma grade igualmente espaçada cujo limites neste trabalho foram definidos pelos mínimos e máximos de latitude e longitude das 30 estações vizinhas. A resolução utilizada para a grade foi de  $0,01^\circ$ , a qual representa uma melhor qualidade, porém com maior tempo de processamento computacional. A figura 3.9 apresenta a interpolação espacial realizada para as

0500 UTC do dia 28/10/2017 para (a) velocidade do vento e para as 1500 UTC do dia 22/08/2017 para (b) temperatura do ar. Tal interpolação utilizando a krigagem ordinária objetivou a estimativa das variáveis em questão para a estação Barra (Bahia), localizada no centro da grade (11,0° S e 43,1° O) a partir das 28 estações meteorológicas mais próximas. Ressalta-se que 2 estações foram eliminadas do preenchimento, pois não continham dados medidos para os horários em questão.

Assim, para cada horário a ser preenchido foi realizada uma interpolação para estimar o valor de velocidade do vento e temperatura do ar a partir de um máximo de 30 e um mínimo de 10 estações meteorológicas vizinhas.

Figura 3.9 – Domínio da grade e interpolação espacial utilizando o método ordinário kriging para estimar o dado de (a) velocidade do vento e (b) temperatura do ar para a estação meteorológica Barra (triângulo invertido) a partir de 28 estações meteorológicas mais próximas (sinal de positivo).



Fonte: o autor.

#### c) Imputação utilizando ERA5 – ECMWF

Utilizou-se os dados de reanálise do ERA5 descritos no item 3.2.2 como um terceiro método para preenchimento das falhas artificiais com tamanho de 1 a 24 horas consecutivas.

#### d) Downscaling dinâmico

Assim como na alínea anterior os dados obtidos após a integração do modelo atmosférico WRF também foram utilizados como um método adicional para preencher os dados removidos artificialmente de velocidade do vento e temperatura do ar. As configurações utilizadas para a realização do downscaling dinâmico estão descritas no item 3.2.2.

### 3.5.2 Avaliação de desempenho

Com o objetivo de avaliar qual dos 4 métodos supracitados apresentou melhor desempenho para preencher os diferentes tamanhos de buracos artificiais e, para posteriormente permitir o preenchimento das falhas encontradas nas séries temporais originais considerou-se o uso de três medidas de erros, uma categórica e uma de correlação. Ressalta-se sobretudo que esta avaliação de desempenho não foi realizada de maneira individualizada por estação. Ou seja, as medidas de erro, de correlação e categórica foram calculadas considerando-se o preenchimento das falhas artificiais de todas as estações presentes em cada cluster definido durante a análise de agrupamento. Isto porque, conforme discutido no item 3.4, cada agrupamento contempla estações meteorológicas cujo as medidas de velocidade do vento e temperatura do ar apresentam comportamento similar entre elas e dissimilar quando comparado com as de outros clusters. Em outras palavras, a avaliação de desempenho dos métodos utilizados para preencher os buracos gerados artificialmente foi realizada individualmente para cada um dos sete clusters definidos em 3.4.

De acordo com HALLAK; PEREIRA FILHO (2011), a diferença ( $d_i$ ) entre o valor estimado ( $y$ ) e o medido ( $x$ ) no ponto  $i$  do espaço bidimensional no tempo  $t$  é uma quantidade fundamental para o estudo de erros. Ou seja, neste trabalho quando  $d_i$  for igual a zero ou muito próximo a zero significa que houve sucesso em estimar o dado original removido artificialmente da série temporal. Por outro lado, valores de  $d_i$  distantes de zero demonstram baixo desempenho em estimar o dado original medido pela estação meteorológica. Assim, os índices utilizados neste trabalho serão apresentados em função desta diferença quando necessário.

A primeira medida de erro utilizada é o erro médio (ME) também denominado de *BIAS*. Essa medida indica se há erro sistemático (superestimativa ou subestimativa) dos métodos de preenchimento em estimar o valor original removido da série temporal.

$$ME = \frac{1}{N} \sum_{i=1}^N d_i \quad \text{Equação 3.3}$$

O segundo índice utilizado para avaliar o desempenho dos métodos de preenchimento é o erro médio absoluto (MAE). Essa medida é a média dos erros absolutos e elimina a possibilidade de valores individuais positivos e negativos de igual magnitude se anularem. Neste trabalho, o MAE também é utilizado para detectar se há algum método errando demasiadamente para estimar os dados removidos artificialmente.

$$MAE = \frac{1}{N} \sum_{i=1}^N |d_i| \quad \text{Equação 3.4}$$

Com o objetivo de apresentar os valores de erro na mesma unidade da variável analisada e para avaliar a presença de grandes erros individuais, também foi utilizada a raiz do erro quadrático médio (RMSE).

$$RMSE = \left[ \frac{1}{N} \sum_{i=1}^N (d_i)^2 \right]^{1/2} \quad \text{Equação 3.5}$$

O coeficiente de correlação de Pearson ( $r$ ) é um índice adimensional que mede o grau da correlação linear entre o valor medido e o estimado. Para uma correlação perfeita e positiva,  $r$  assume o valor de 1, indicando que o dado estimado está em fase com o medido. Por outro lado, quanto mais próximo  $r$  for de zero, maior será a defasagem entre o dado estimado e o medido. O índice  $r$  também pode assumir valor negativo, cujo valores próximos a -1 indicam que enquanto o valor do dado estimado aumenta, o do observado diminui ou vice-versa.

A última métrica de verificação é o cálculo de *scores* baseados em Tabela de Contingência. Em geral, os *scores* são calculados a partir da construção de uma tabela do tipo 2x2 considerando-se apenas a ocorrência e a não ocorrência de um determinado evento e a consequente estimativa de tal ocorrência ou não ocorrência (WILKS, 2011). Este fato permite discretizar as variáveis de interesse e, a partir de então realizar uma análise qualitativa. Também é possível predefinir limiares ou *range* de valores de modo a assumir que o valor estimado foi medido.

Neste trabalho, a ocorrência de um evento foi definida como cada dado de temperatura do ar e velocidade do vento removido artificialmente das séries temporais originais. Por outro lado, o valor estimado por cada método é considerado como ocorrido caso o  $d_i$  para a temperatura do ar (velocidade do vento) seja menor ou igual a 1° C (0,5 m/s). Esses *ranges* foram definidos para evitar penalização demasiada dos valores estimados pelos métodos de preenchimento. Ressalta-se ainda que os erros de 1° C para a temperatura do ar e 0,5 m/s para a velocidade do vento são aceitáveis no cálculo de ampacidade para linhas de transmissão, uma vez que se aproximam aos erros dos sensores utilizados nas medições das referidas variáveis atmosféricas.

Tabela 3.5 - Tabela de contingência hipotética para determinação da performance dos estimadores.

Valor estimado	Valor removido	
	Sim	Não
Sim	a	b
Não	c	d

Fonte: o autor.



A partir da tabela 3.5 nota-se que, para uma dada estação, “a” corresponde a quantidade de vezes em que o valor estimado por um determinado método foi igual ou próximo ao dado removido artificialmente considerando os limiares de 1°C e 0,5 m/s. As letras “b” e “d” da tabela 3.5 estão relacionadas a não ocorrência de um determinado evento. E para a presente tese tais letras serão desconsideradas, pois o dado removido sempre será considerado uma ocorrência a ser preenchida pelos métodos de estimação. E por fim, “c” refere-se ao erro na estimativa do valor original removido.

Portanto, o último índice de performance será derivado da tabela de contingência de cada cluster e variável preenchida. E devido à presença apenas de “a” e “c”, o indicador de sucesso será a Probabilidade de Detecção (POD) cujo valor expressa o percentual de acertos em estimar o dado removido da série original.

$$POD = \frac{a}{a+c} \times 100 \quad \text{Equação 3.6}$$

Adicionalmente, avaliou-se o desempenho dos métodos para preencher diferentes tamanhos de falhas (buracos consecutivos de 1 a 24 horas). Por este motivo, primeiramente realizou-se o teste estatístico de Shapiro-Wilk, que determina se a hipótese de nulidade de uma distribuição é aceita ou rejeitada diante da suposição de normalidade (ROYSTON, 1995). Após a hipótese ser rejeitada a um intervalo de confiança de 95% foram analisadas individualmente as propriedades da distribuição (histograma, curtose e assimetria) das probabilidades de detecção de cada método de imputação de dados. Essas análises foram realizadas com o objetivo de complementar o teste de Shapiro-Wilk no tocante à verificação da normalidade das distribuições.

Posterior a estas análises, foi calculada a equação polinomial por meio dos mínimos quadrados para cada estimador e, em seguida, ajustou-se uma reta de regressão ao conjunto de dados de POD médio para cada tamanho de falha artificial (1 a 24 horas) de cada método. E assim, esta análise de regressão indicou a tendência da probabilidade de detecção dos valores removidos artificialmente por meio do coeficiente angular de cada reta ajustada.



## 4 RESULTADOS

Nesta seção será apresentada a performance dos métodos utilizados para o preenchimento das falhas artificiais das séries temporais com vistas ao preenchimento dos buracos das séries temporais originais do INMET que foram utilizadas para atingir o objetivo proposto na presente tese.

Por fim, será apresentada a caracterização do regime de mínimos de vento medidos a 10 metros e dos máximos de temperatura do ar a 2 metros sobre o subsistema Nordeste do Sistema Interligado Nacional do setor elétrico. As escalas espacial e temporal foram consideradas para possibilitar a identificação dos sistemas atmosféricos que modulam as condições destes mínimos e máximos presentes nos clusters formados a partir da análise de agrupamentos.

A respeito da escala temporal considerou-se os seguintes períodos para análise e caracterização: verão, outono, inverno, primavera, manhã, tarde, noite e madrugada. Esta análise considerou as mínimas velocidades diárias do vento e as máximas temperaturas diárias do ar, assim como a variância destes dados ao longo das estações do ano utilizando o recurso gráfico *boxplot* para visualização.

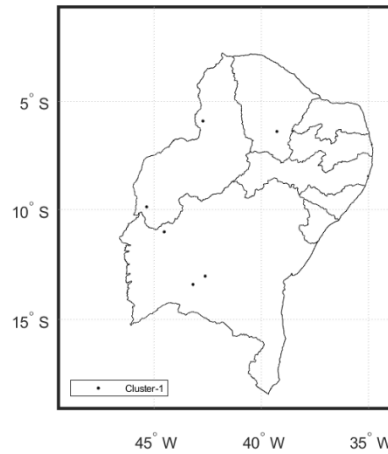
Este recurso gráfico permitiu apresentar de forma resumida a tendência central (mediana) de cada variável meteorológica considerada neste estudo assim como a dispersão dos dados em quartis, valores máximos e mínimos das séries temporais, intervalo de confiança e os *outliers*.

### 4.1 ANÁLISE DE PREENCHIMENTO DE FALHAS ARTIFICIAIS

#### 4.1.1 Cluster 1

O primeiro agrupamento denominado de cluster-1 é formado por 6 estações meteorológicas (figura 4.1) cuja mediana das medianas das mínimas velocidades diárias do vento ( $V_{med}$ ) e das máximas temperaturas diárias do ar ( $T_{med}$ ) é 0,5 m/s e 32,9 °C (tabela 4.1) com um desvio padrão de 0,1 e 0,25, respectivamente. Para este cluster, três estações (Iguatu, Santa Rita de Cássia e Bom Jesus da Lapa) possuem séries temporais de aproximadamente 11 anos (05/2007 a 03/2018) enquanto Macaúbas apresenta a menor série (22 meses de dados medidos). O cluster-1 é formado por estações localizadas a uma altitude média de 400 metros, em depressões e chapadas e, cuja vegetação predominante desse cluster é a caatinga e o cerrado.

Figura 4.1 – Localização das estações meteorológicas que formaram o cluster-1.



Fonte: o autor.

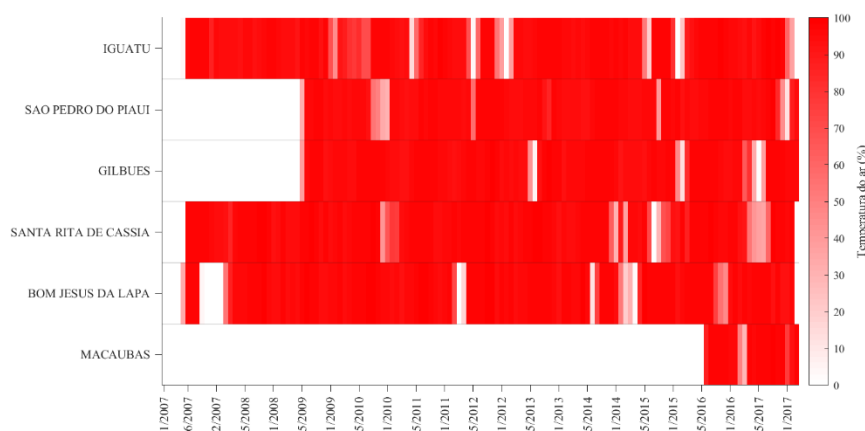
Tabela 4.1 – Estações meteorológicas agrupadas no cluster 1 e a mediana das máximas temperaturas diárias do ar ( $T_{med}$ ) e das mínimas velocidades diárias do vento ( $V_{med}$ ).

Latitude (°)	Longitude (°)	Altitude (m)	$T_{med}$ (°C)	$V_{med}$ (m/s)	Nome da estação	UF
-6,4	-39,3	233	33,0	0,6	Iguatu	CE
-5,9	-42,7	287	32,9	0,5	São Pedro do Piauí	PI
-9,9	-45,3	425	33,0	0,5	Gilbués	PI
-11,2	-44,5	450	32,9	0,4	Santa Rita de Cássia	BA
-13,4	-43,2	440	33,1	0,5	Bom Jesus da Lapa	BA
-13,0	-42,6	574	32,4	0,6	Macaúbas	BA
Mediana			32,9	0,5		

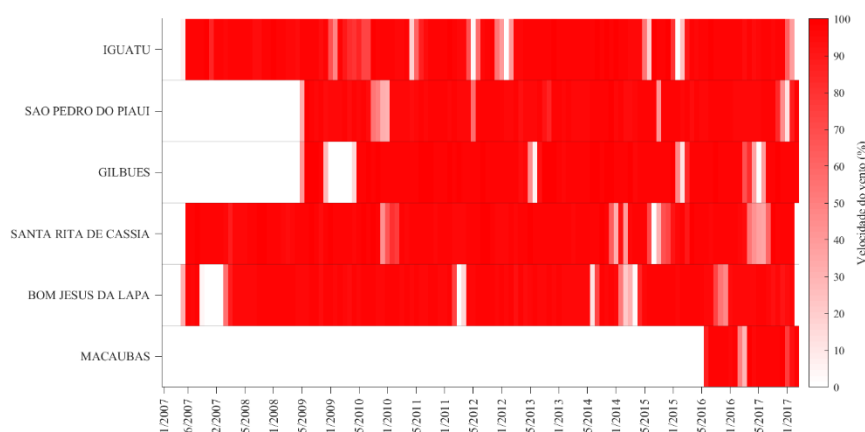
A figura 4.2 mostra as séries temporais da temperatura do ar e da velocidade do vento. No eixo das coordenadas são apresentadas as estações meteorológicas que compõem o cluster-1 enquanto no das abcissas o período da série temporal. Os dados disponíveis na escala horária ao longo dos meses e anos são representados pela barra horizontal na máxima tonalidade da cor vermelha. Quanto maior a quantidade de falhas horárias na série temporal menor é a tonalidade. Assim, a ausência de dados é representada pela cor branca.

Figura 4.2 – Séries temporais brutas da (a) temperatura do ar e (b) velocidade do vento das estações do INMET pertencentes ao cluster-1.

a)



b)



Fonte: o autor.

Com o objetivo de preencher as falhas com tamanho de 1 a 24 horas presentes nas séries temporais do cluster-1, inicialmente gerou-se 10% de buracos artificiais nos dados brutos de cada estação. Posteriormente, realizou-se o preenchimento das falhas geradas utilizando 4 métodos diferentes. E por fim, avaliou-se o desempenho dos métodos em preencher as falhas artificialmente criadas.

A figura 4.3 apresenta a média dos dados de temperatura do ar e velocidade do vento removidos artificialmente em cada hora de um dia (hora 1 a 24). Esta figura mostra também a média dos dados estimados nos mesmos horários para cada método considerado.

A partir da figura 4.3a foi possível verificar que ambos os métodos reproduziram com eficiência o ciclo diurno da temperatura do ar. Referente ao horário da máxima temperatura diária todos os métodos acertaram em estimá-lo para as 19 UTC mesmo que com subestimativas (WRF e kriging). O mesmo ocorreu para o horário da mínima, com exceção do modelo WRF que antecipou a menor temperatura para as 9 UTC, enquanto os demais acertaram em estimar

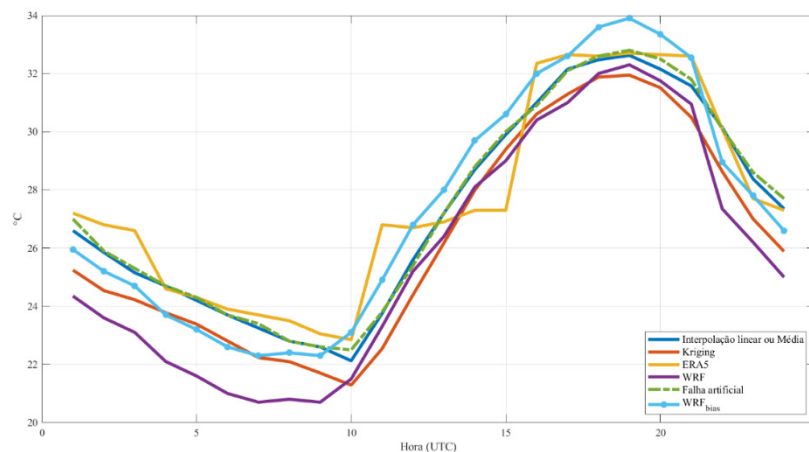
o mínimo de temperatura ocorrido as 10 UTC. De maneira geral, esse fato é evidenciado pelas altas correlações lineares apresentadas na tabela 4.2.

Para a velocidade do vento figura 4.3b os dados estimados também capturam a variabilidade horária desta variável na maior parte do dia. Ou seja, todos demonstraram habilidade em acompanhar a máxima elevação da velocidade e a consequente diminuição ao final do dia, embora em alguns horários, um ou outro método não tenha reproduzido o aumento ou diminuição da velocidade do vento.

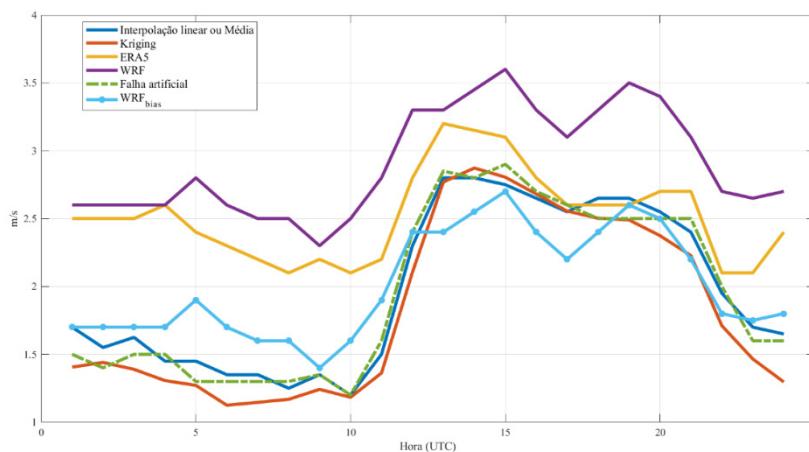
Porém, essas diferenças podem ser observadas na tabela 4.2 a qual apresenta correlações menores para todos os métodos quando comparadas à temperatura.

Figura 4.3 – Média dos valores removidos artificialmente (linha tracejada) e estimados por 4 diferentes métodos para (a) temperatura do ar e (b) velocidade do vento para o cluster-1.

a)



b)



Fonte: o autor.

A respeito dos erros sistemáticos das estimativas verifica-se na tabela 4.2 que o *bias* da reanálise ERA5 para a temperatura do ar é de 0,1. Esse valor indica que, assim como para a interpolação, o erro sistemático é baixo e, portanto, estes dois métodos possuem habilidade em

estimar os dados removidos artificialmente. No entanto, esta habilidade diminui para o ERA5 entre as 11 e 18 UTC (figura 4.3a) cujo período é o de maior aquecimento da atmosfera próxima a superfície devido a presença da irradiação solar. Para todas as demais imputações de dados notou-se a capacidade de reprodução do aumento de temperatura observado nos dados medidos.

Para as estimativas de velocidade do vento observou-se erros sistemáticos menores que 1 m/s ressaltando-se inclusive o *bias* igual zero calculado para a interpolação linear (tabela 4.2). Porém, embora este método não tenha apresentado erro sistemático para a velocidade do vento o MAE calculado é de 0,8 m/s (tabela 4.2).

Ao analisar o RMSE verificou-se que o WRF apresentou os maiores erros individuais na estimativa da temperatura do ar (3,1 °C) quando comparado com os demais métodos. O mesmo ocorreu para velocidade do vento, no entanto a maior diferença entre os erros foi de 0,6 m/s (WRF 1,6 contra 1 m/s da interpolação).

Por fim, com o intuito de avaliar se o modelo WRF pode ser utilizado para prever a temperatura do ar e a velocidade do vento sobre o cluster-1 removeu-se o *bias* ( $WRF_{bias}$ ) para os valores estimados por este modelo (figura 4.3a). A remoção do erro sistemático aproximou ambas as curvas do modelo regional à dos dados medidos com uma ligeira melhora para as métricas de verificação influenciadas pelo *bias* (tabela 4.2).

Tabela 4.2 – Métricas de verificação do desempenho do preenchimento de falhas geradas artificialmente para o cluster-1.

Cluster-1	<i>Bias</i>		MAE		r		RMSE		POD	
	T <sup>1</sup>	V <sup>2</sup>	T <sup>1</sup>	V <sup>2</sup>	T	V	T <sup>1</sup>	V <sup>2</sup>	T <sup>1</sup>	V <sup>2</sup>
Interpolação	-0,1	0,0	1,1	0,8	0,94	0,70	1,5	1,0	62,9	45,0
Kriging	-1,0	-0,2	1,9	0,9	0,87	0,50	2,5	1,3	33,5	35,5
ERA5	0,1	0,6	1,5	1,1	0,90	0,60	1,9	1,4	44,8	32,7
WRF	-1,6	0,9	2,4	1,3	0,84	0,60	3,1	1,6	31,4	26,6
WRF- <i>Bias</i> <sup>3</sup>	0,0	0,0	2,7	1,0	-	-	2,7	1,3	32,3	28,9

<sup>1</sup>Temperatura do ar em °C

<sup>2</sup>Velocidade do vento em m/s

<sup>3</sup> Índices após a remoção do *Bias*

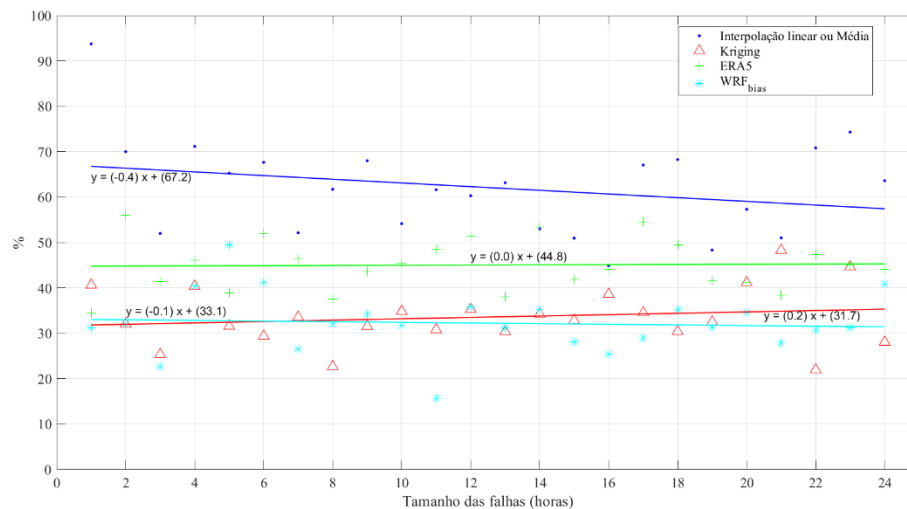
O preenchimento das falhas geradas artificialmente foi realizado de acordo com o tamanho dos buracos que foram removidos. Em outras palavras, verificou-se a habilidade dos métodos preencherem buracos individuais de 1 hora, 2 horas, 3 horas e assim por diante até 24 horas de falhas consecutivas. O objetivo desta análise foi avaliar se algum método performa melhor o preenchimento de acordo com o tamanho da falha. E desta maneira, possibilitar que pesquisadores e demais usuários das séries temporais do INMET possam preencher os buracos presentes nas séries originais de acordo com o tipo de falha.

Assim, de acordo com a figura 4.4a verificou-se para a temperatura do ar que, independentemente do tamanho das falhas, a interpolação apresentou maior probabilidade de detecção (estimativa) dos dados removidos artificialmente do que quando comparado com os demais métodos. No entanto, para a velocidade do vento (figura 4.4b), observou-se que apenas para falhas de 7, 20 e 21 horas consecutivas o método de kriging superou a interpolação linear no tocante a probabilidade de estimativa dos dados removidos.

A análise da figura 4.4 revelou também que para falhas de 2 horas a probabilidade de detecção utilizando a interpolação sofreu uma redução de 24% para a temperatura e 15% para o vento. Esse comportamento se deve ao fato de que para falhas entre 2 e 24 horas o dado imputado é o valor médio entre dado medido 24 horas antes e 24 horas depois da hora da falha ao invés da interpolação linear entre o dado medido na hora anterior e posterior para falhas com tamanho de 1 hora.

Figura 4.4 – Diagrama de dispersão e reta de regressão da Probabilidade de Detecção média dos valores médios removidos artificialmente das séries temporais do cluster-1 para (a) temperatura do ar e (b) velocidade do vento considerando diferentes tamanhos de falhas.

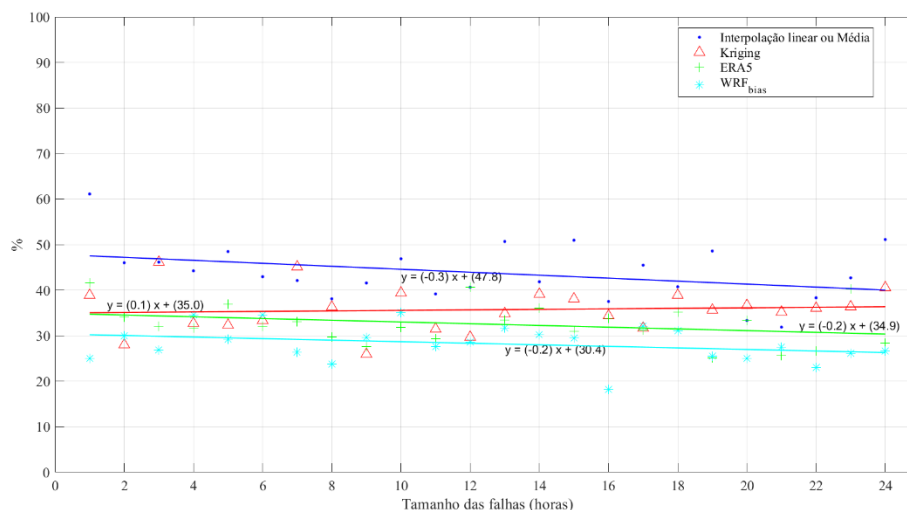
a)





Continuação - Figura 4.4 – Diagrama de dispersão e reta de regressão da Probabilidade de Detecção média dos valores médios removidos artificialmente das séries temporais do cluster-1 para (a) temperatura do ar e (b) velocidade do vento considerando diferentes tamanhos de falhas.

b)



Fonte: o autor.

A análise de regressão utilizada para permitir a avaliação do desempenho dos métodos no preenchimento de diferentes tamanhos de falhas indicou que para ambas as variáveis meteorológicas a tendência da probabilidade de detecção dos valores removidos artificialmente foi negativa e quase nula para todos os métodos (figura 4.4). A única exceção é observada para a interpolação espacial de kriging cujo coeficiente angular também é próximo de zero, porém positivo (0,1 tanto para temperatura quanto para velocidade do vento). Em outras palavras, a tendência quase nula indicada pelos coeficientes angulares demonstrou que a probabilidade de detecção, para qualquer método e variável a ser estimada, não sofreu alteração independentemente do número de falhas a ser preenchido.

Desta forma, a análise conjunta das métricas de verificação de desempenho do preenchimento de falhas para o cluster-1 revelou que para a temperatura do ar a interpolação performou melhor para todos os tamanhos de buracos. Entretanto, para a velocidade do vento verificou-se que o preenchimento de falhas consecutivas de 7, 20 e 21 horas utilizando o método de kriging apresentou maior probabilidade de estimar os dados removidos artificialmente. E para todos os demais tamanhos o uso da interpolação indicou maior probabilidade de detecção.

Assim, as séries temporais originais do cluster-1 foram preenchidas com o método de interpolação linear quando identificadas falhas de 1 a 24 horas para a temperatura do ar. Para a velocidade do vento utilizou-se o método geoestatístico de kriging para preencher buracos consecutivos de 7, 20 e 21 horas nas séries brutas do primeiro agrupamento e o de interpolação para os demais tamanhos.

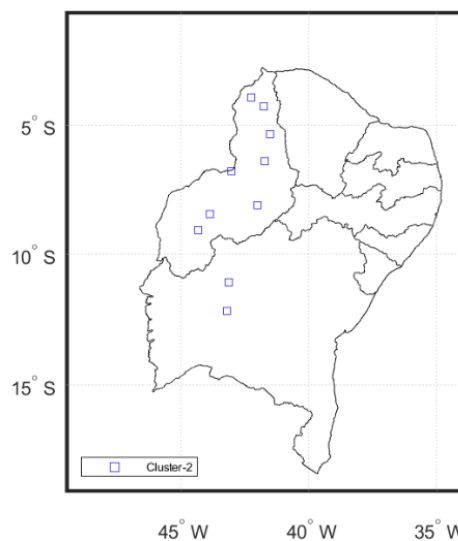
#### 4.1.2 Cluster 2

A maior parte das estações que formaram o cluster-2 (figura 4.5) estão situadas em três mesorregiões piauiense: três estão sobre o sudoeste (cerrado), enquanto uma está no sudeste, duas no centro-norte (meio norte) e outras duas no norte do estado (litoral). Estas cinco últimas estações meteorológicas estão sobre a caatinga.

Além das estações situadas no Piauí há outras duas estações que compuseram o segundo agrupamento e que estão localizadas na mesorregião baiana Vale São-Franciscano (Barra e Ibotirama) cujo bioma também é classificado como caatinga.

A respeito das variáveis meteorológicas em questão o desvio padrão entre as medianas das oito estações é de 0,1 m/s para a velocidade mínima diária do vento e de 0,2 °C para máxima temperatura diária do ar. E a mediana para todo o agrupamento é de 0,5 m/s e 33,9 °C (tabela 4.3).

Figura 4.5 – Localização das estações meteorológicas que formaram o cluster-2.



Fonte: o autor.

Tabela 4.3 – Estações meteorológicas agrupadas no cluster-2 e a mediana das máximas temperaturas diárias do ar ( $T_{med}$ ) e das mínimas velocidades diárias do vento ( $V_{med}$ ).

Latitude (°)	Longitude (°)	Altitude (m)	$T_{med}$ (°C)	$V_{med}$ (m/s)	Nome da estação	UF
-6,7	-43,0	123,27	34,1	0,5	Floriano	PI
-9,0	-44,3	415,0	34,0	0,6	Bom Jesus do Piauí	PI
-4,2	-41,7	161,2	34,2	0,5	Piripiri	PI
-8,4	-43,8	270,0	33,7	0,3	Alvorada do Gurgueia	PI
-3,9	-42,2	65,0	33,9	0,5	Esperantina	PI
-5,3	-41,5	286,0	34,1	0,4	Castelo do Piauí	PI

Continuação – Tabela 4.3 – Estações meteorológicas agrupadas no cluster-2 e a mediana das máximas temperaturas diárias do ar ( $T_{med}$ ) e das mínimas velocidades diárias do vento ( $V_{med}$ ).

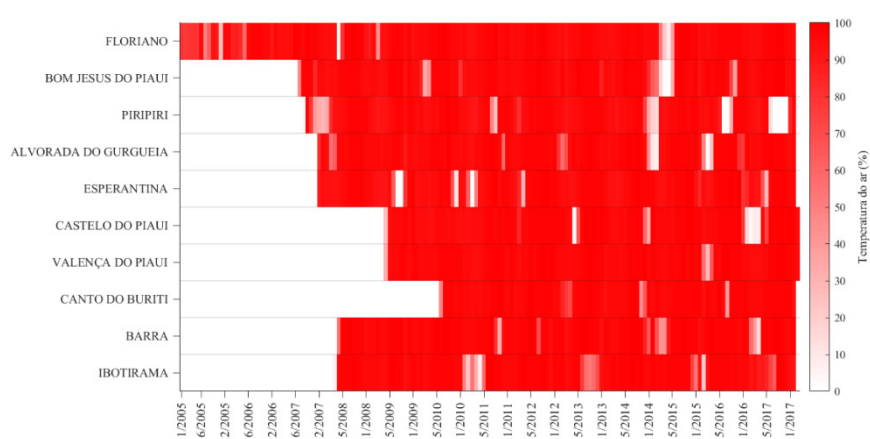
-6,3	-41,7	301,0	34,3	0,5	Valença do Piauí	PI
-8,1	-42,0	308,0	33,7	0,6	Canto do Buriti	PI
-11	-43,1	403,0	33,6	0,5	Barra	BA
-12,1	-43,2	430,0	33,7	0,5	Ibotirama	BA
	Mediana		33,9	0,5		

As séries temporais das estações do cluster-2 são apresentadas na figura 4.6. A partir desta figura, observou-se que a estação meteorológica localizada na cidade de Floriano (Piauí) possui a maior série temporal de dados medidos disponível neste cluster (13 anos) enquanto a menor, Canto do Buriti, tem aproximadamente 8 anos (05/2010 a 03/2018).

Verificou-se também que, por aproximadamente 2 anos (figura 4.6b), as séries temporais das estações Alvorada do Gurgueia (04/2010 a 10/2012) e Castelo do Piauí (10/2012 a 10/2014) não realizaram medidas da velocidade do vento. O motivo provavelmente está relacionado a problemas técnicos no anemômetro uma vez que há dados disponíveis para a temperatura do ar neste período. Outras estações como Floriano (05/2006 a 07/2006) e Piripiri (12/2016 a 11/2017) também ficaram longos períodos sem medidas de velocidade do vento.

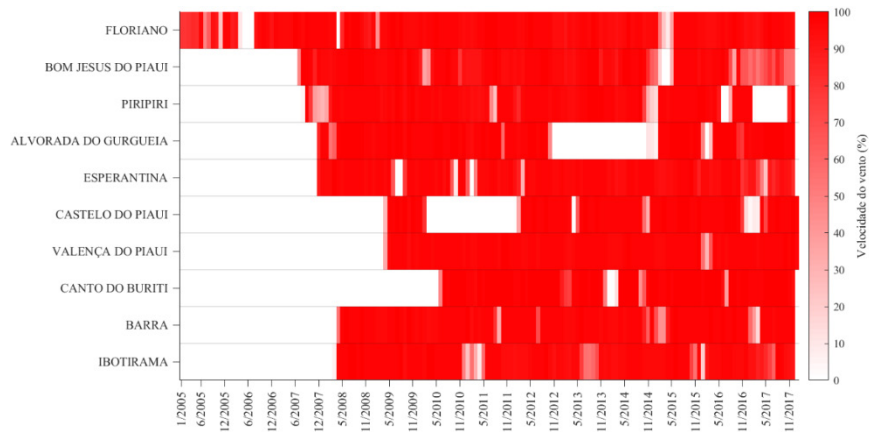
Figura 4.6 – Séries temporais brutas da (a) temperatura do ar e (b) velocidade do vento das estações do INMET pertencentes ao cluster-2.

a)



Continuação – Figura 4.6 – Séries temporais brutas da (a) temperatura do ar e (b) velocidade do vento das estações do INMET pertencentes ao cluster-2.

b)



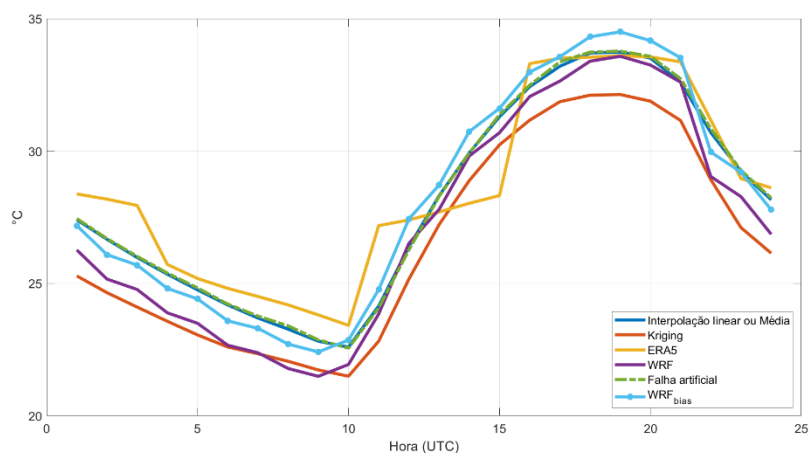
Fonte: o autor.

No tocante ao preenchimento das falhas de 1 a 24 horas das séries temporais do cluster-2, a partir da figura 4.7a, foi observado que ambos os métodos capturaram a variabilidade diurna da temperatura máxima diária do ar incluindo os horários de ocorrência dos valores mínimos e máximos. A correlação linear apresentada na tabela 4.4 evidenciou este fato assim como mostrou que o menor erro médio foi encontrado para o método da interpolação linear ( $-0,1\text{ }^{\circ}\text{C}$ ) seguido do ERA5 ( $0,3\text{ }^{\circ}\text{C}$ ).

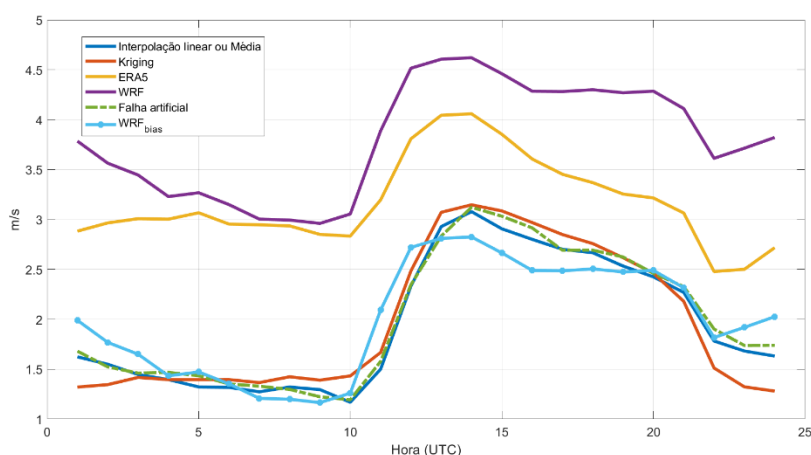
O preenchimento realizado a partir do downscaling dinâmico (WRF) também apresentou *bias* menor que  $1\text{ }^{\circ}\text{C}$ , embora negativo. Entretanto, a remoção deste erro sistemático não aumentou o desempenho do modelo de escala regional no preenchimento das falhas. Isto porque os outros índices estatísticos (RMSE e MAE) calculados para o WRF indicaram que o erro sistemático não é constante (tabela 4.4).

Figura 4.7 – Média dos valores removidos artificialmente (linha tracejada) e estimados por 4 diferentes métodos para (a) temperatura do ar e (b) velocidade do vento para o cluster-2.

a)



b)



Fonte: o autor.

Tabela 4.4 – Métricas de verificação do desempenho do preenchimento de falhas geradas artificialmente para o cluster-2.

Cluster-2	Bias		MAE		r		RMSE		POD	
	T <sup>1</sup>	V <sup>2</sup>	T <sup>1</sup>	V <sup>2</sup>	T	V	T <sup>1</sup>	V <sup>2</sup>	T <sup>1</sup>	V <sup>2</sup>
Interpolação	-0,1	-0,1	1,1	0,8	0,93	0,71	1,5	1,1	60,5	42,5
Kriging	-1,6	-0,1	2,2	1,1	0,86	0,48	2,8	1,5	30,8	26,2
ERA5	0,3	1,1	1,6	1,5	0,88	0,55	2,1	1,8	40,8	17,9
WRF	-0,9	1,8	2,2	2,0	0,81	0,50	2,9	2,5	31,8	13,1
WRF-Bias <sup>3</sup>	0,0	0,0	2,0	1,3	-	-	2,8	1,7	31,5	21,8

<sup>1</sup>Temperatura do ar em °C    <sup>2</sup>Velocidade do vento em m/s    <sup>3</sup>Índices após a remoção do Bias

Para a velocidade do vento (figura 4.7b) os modelos dinâmicos da atmosfera tiveram dificuldade em estimar os valores removidos artificialmente da série temporal enquanto os métodos estatísticos (interpolação e média e kriging) performaram com maior habilidade os preenchimentos. Os índices estatísticos apresentados na tabela 4.4 comprovaram a deficiência

dos modelos numéricos no preenchimento das falhas quando comparado aos outros métodos. Por exemplo, o *bias* calculado para os métodos de interpolação e kriging são de apenas -0,1 m/s enquanto para o ERA5 e WRF são de 1,1 e 1,8 m/s, respectivamente. O RMSE também indicou que ambos os modelos estimaram valores distantes da média dos dados removidos artificialmente. Entretanto, com a remoção do erro sistemático, observou-se que o  $WRF_{bias}$  reduziu a diferença entre os dados estimados e os medidos sobretudo durante o período de maior resfriamento da atmosfera próxima superfície (0 as 10 UTC). E para o período onde há o aumento da velocidade do vento devido aos processos relacionados ao aquecimento da atmosfera pela irradiância solar o  $WRF_{bias}$  subestimou a velocidade do vento.

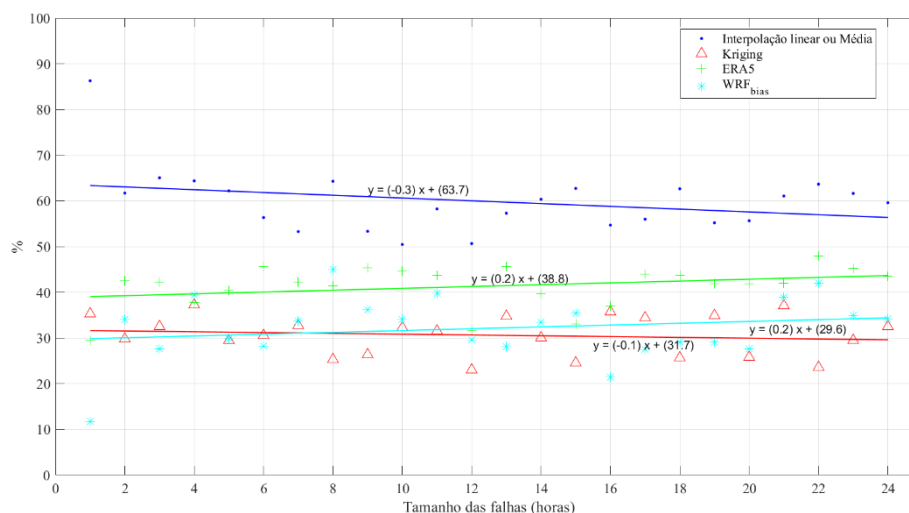
O baixo desempenho da reanálise do ERA5 e do downscaling dinâmico utilizando o WRF em estimar os valores de vento removidos artificialmente também estão refletidos no valor médio do índice de probabilidade de detecção calculado para o cluster-2 e apresentado na tabela 4.4.

Ao analisar o desempenho dos métodos em preencher os buracos criados artificialmente nas séries temporais do cluster-2 para tamanhos distintos de buracos foi observado que o método de interpolação apresentou as maiores probabilidades de detecção dos valores medidos para ambas variáveis meteorológicas em questão (figura 4.8).

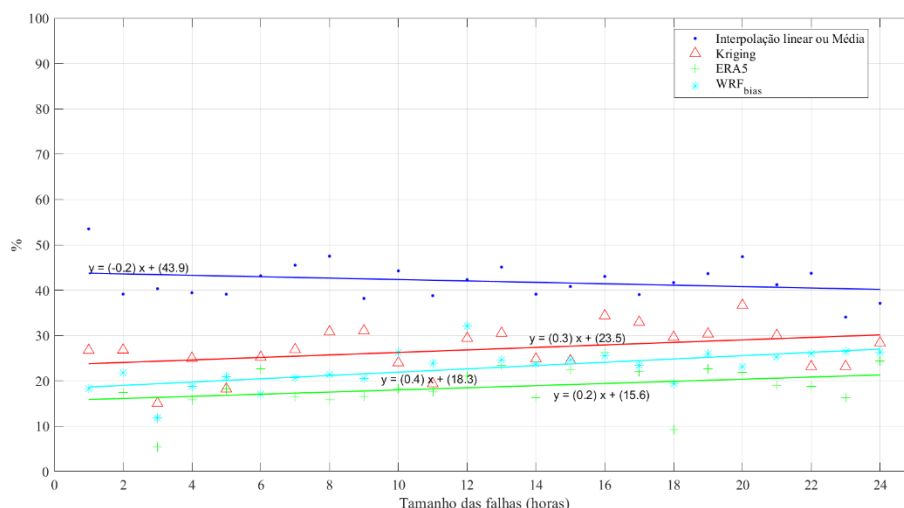
A análise da reta de regressão ajustada aos valores médios do índice de Probabilidade de Detecção para a temperatura do ar (figura 4.8a) demonstrou também que os modelos atmosféricos aumentaram a probabilidade de estimar os valores removidos de acordo com o aumento do número de falhas consecutivas enquanto os estatísticos diminuíram a probabilidade de acertos. Para a velocidade do vento (figura 4.8b) o único método que diminuiu a performance do preenchimento dos dados removidos com o aumento do tamanho de falhas foi o de interpolação. Entretanto, verificou-se que este método possui a maior probabilidade de estimar as falhas para buracos de 1 a 24 horas quando comparado com os demais métodos. Por fim, a análise da inclinação das retas da figura 4.8 e do coeficiente angular das equações indicou que embora tenha havido tendência no aumento ou diminuição da probabilidade de detecção esta foi próxima a zero.

Figura 4.8 – Diagrama de dispersão e reta de regressão da Probabilidade de Detecção média dos valores médios removidos artificialmente das séries temporais do cluster-2 para (a) temperatura do ar e (b) velocidade do vento considerando diferentes tamanhos de falhas.

a)



b)



Fonte: o autor.

Portanto, a partir da análise dos índices estatísticos apresentados na tabela 4.4 e do índice de probabilidade de detecção realizou-se o preenchimento dos buracos de 1 até 24 horas presentes nas séries temporais brutas do cluster-2 de ambas as variáveis utilizando a interpolação linear para falhas de 1 hora ou a média entre o valor medido 24 horas antes e 24 horas depois da falha.

### 4.1.3 Cluster 3

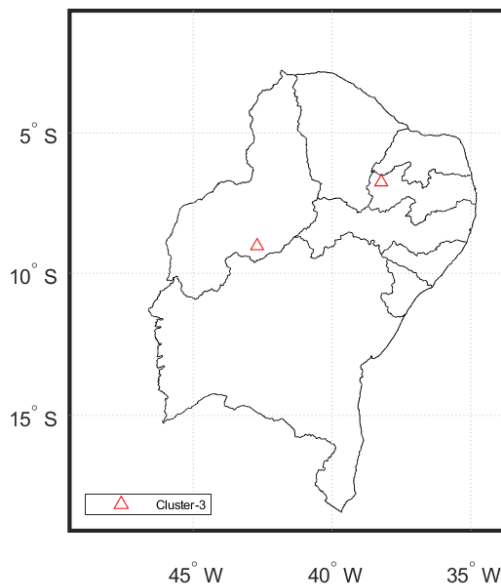
O terceiro agrupamento (figura 4.9) é formado por apenas duas estações meteorológicas distantes entre si por aproximadamente 540 km em linha reta. A primeira estação localiza-se

no sudoeste piauiense, a 402 metros de altitude, no município de São Raimundo Nonato (Piauí). Esta estação opera no sopé da Serra do Bom Jesus da Gurguéia e da Serra da Capivara cuja vegetação dominante é a caatinga.

Por outro lado, a segunda estação meteorológica associada a este cluster, São Gonçalo – distrito do município de Sousa, situa-se no sertão da Paraíba, na depressão Sertaneja do São Francisco, a 233 metros de altitude.

A mediana das máximas temperaturas diárias calculada para o cluster-3, formado por estas duas estações, foi de 33,4 °C e de 0,8 m/s para a mediana das mínimas velocidades diárias do vento (tabela 4.5).

Figura 4.9 – Localização das estações meteorológicas que formaram o cluster-3.



Fonte: o autor.

Tabela 4.5 – Estações meteorológicas agrupadas no cluster-3 e a mediana das máximas temperaturas diárias do ar ( $T_{med}$ ) e das mínimas velocidades diárias do vento ( $V_{med}$ ).

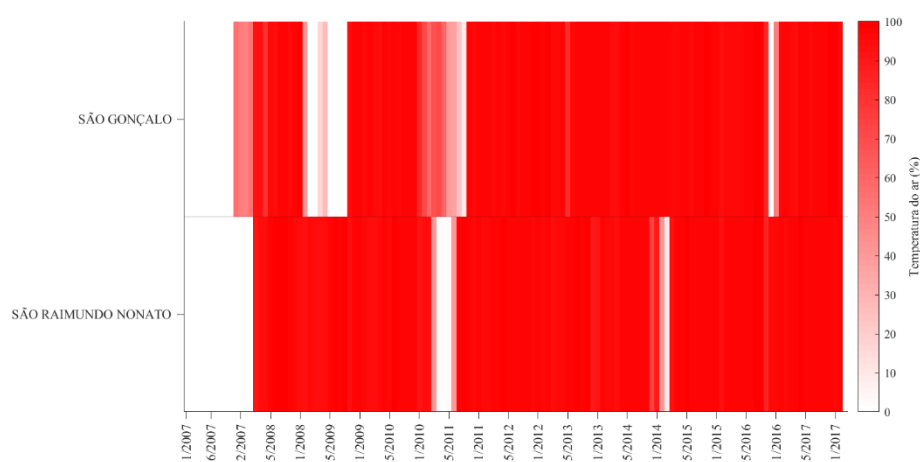
Latitude (°)	Longitude (°)	Altitude (m)	$T_{med}$ (°C)	$V_{med}$ (m/s)	Nome da estação	UF
-6,7	-38,2	233,0	33,4	0,9	São Gonçalo	PB
-9,0	-42,7	402,0	33,4	0,8	São Raimundo Nonato	PI
Mediana			33,4	0,85		



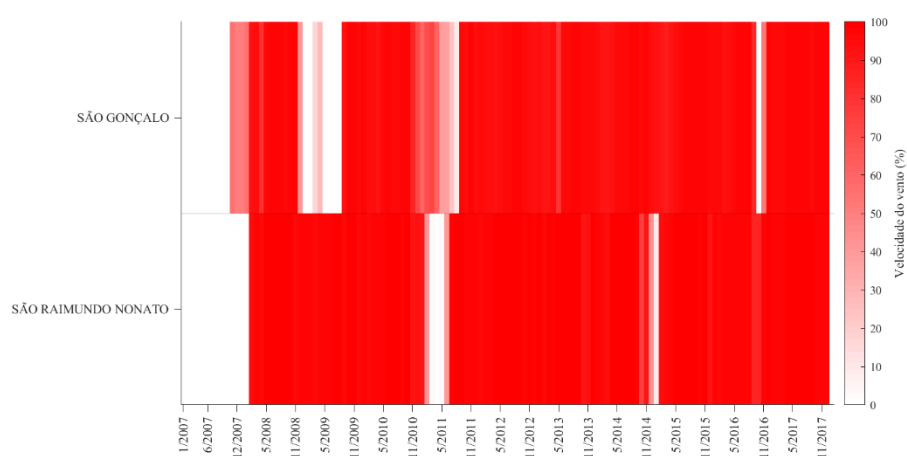
A análise da figura 4.10 revelou que a estação São Gonçalo e São Raimundo Nonato possuem mais de 10 anos de dados medidos disponíveis. Porém, tanto para a temperatura quanto para a velocidade do vento, verificou-se que há dois períodos maiores sem dados medidos na série da primeira estação (12/2008 a 04/2009 e 05/2009 a 09/2009) e um período para a segunda (02/2011 a 06/2011).

Figura 4.10 – Séries temporais brutas da (a) temperatura do ar e (b) velocidade do vento das estações do INMET pertencentes ao cluster-3.

a)



b)



Fonte: o autor.

Para realizar o preenchimento dos buracos de 1 até 24 horas identificados nas séries temporais do cluster-3 removeu-se artificialmente 10% dos dados presentes na figura 4.10. Após a remoção foi avaliado o desempenho do preenchimento destas falhas a partir de quatro métodos (figura 4.11).

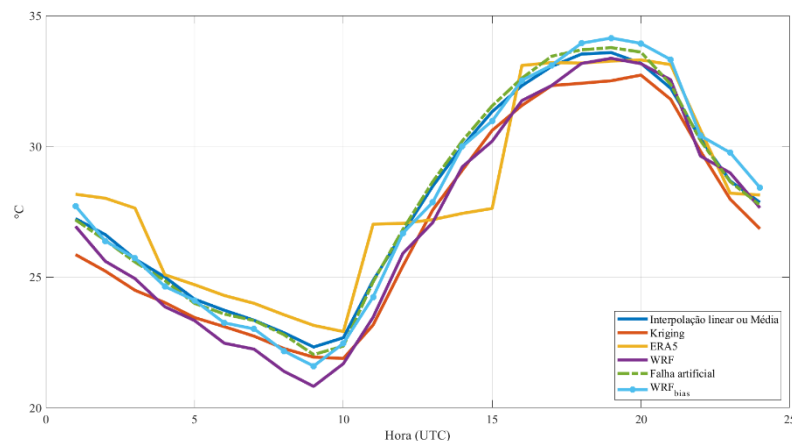
A figura 4.11a mostra que todos os métodos reproduziram a variabilidade dos valores médios removidos artificialmente da temperatura do ar. Entretanto, apenas o WRF e a interpolação estimaram com eficiência o horário de ocorrência da mínima e máxima temperatura do ar. Observou-se também que a reanálise do ERA5 durante o período de aquecimento da atmosfera próxima a superfície (entre 9 e 19 UTC) teve dificuldade em reproduzir o aumento gradativo da temperatura. Este fato pode estar associado a parametrização dos processos de radiação solar na superfície utilizado pelo modelo ERA5.

Para a velocidade do vento (figura 4.11b) a interpolação linear reproduziu a variabilidade desta variável com eficiência durante o todo o ciclo diurno. Por outro lado, o segundo método estatístico (Kriging) demonstrou dificuldade em capturar a variação do vento tanto no período diurno quanto noturno. O baixo desempenho para a velocidade do vento quando comparado ao da temperatura do ar pode estar associado a natureza estocástica da primeira variável que é estimada pelo método de Kriging a partir de dados também com padrão estocástico de estações vizinhas.

A respeito dos modelos numéricos da atmosfera, ambos superestimaram o vento durante o ciclo diurno, embora tenham capturado o aumento acentuado da velocidade entre as 10 e 13 UTC e, posteriormente a queda gradativa entre 13 e 24 UTC. Ainda a respeito do ERA5 e do WRF verificou-se que durante as primeiras horas do dia (0 as 12 UTC) o primeiro modelo tem um erro médio maior do que o WRF (1,4 contra 1,1 m/s). Porém, durante o restante do dia o desempenho dos modelos se inverteu. O WRF errou em média 1,6 m/s entre as 13 e 24 UTC contra 0,7 m/s do ERA5 (figura 4.11b).

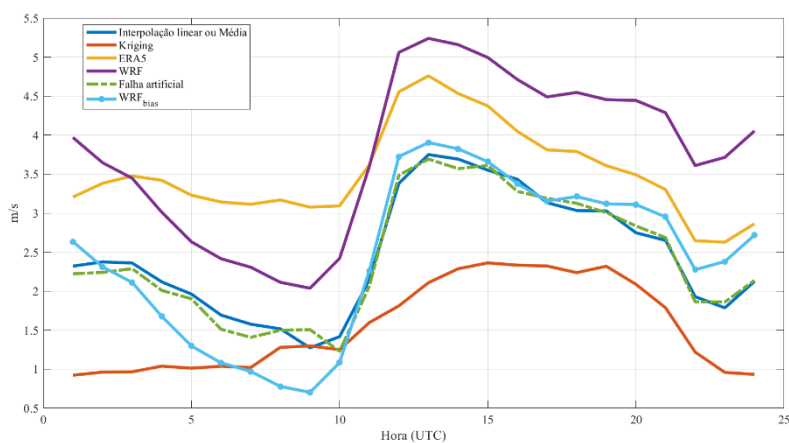
Figura 4.11 – Média dos valores removidos artificialmente (linha tracejada) e estimados por 4 diferentes métodos para (a) temperatura do ar e (b) velocidade do vento para o cluster-3.

a)



Continuação – Figura 4.11 – Média dos valores removidos artificialmente (linha tracejada) e estimados por 4 diferentes métodos para (a) temperatura do ar e (b) velocidade do vento para o cluster-3.

b)



Fonte: o autor.

A análise dos índices estatísticos demonstrou que a interpolação não possui erro sistemático para estimar a temperatura do ar e a velocidade do vento (tabela 4.6). Essa informação foi verificada também a partir figura 4.11 onde as diferenças entre os valores estimados e observados foram mínimas. Ainda a respeito do *bias* este índice indicou que o maior erro de estimativa da temperatura do ar foi de  $-0,8\text{ }^{\circ}\text{C}$  (Kriging e WRF) e para a velocidade do vento de  $1,3\text{ m/s}$  (WRF).

A habilidade de todos os métodos reproduzirem a variabilidade da temperatura do ar foi refletida na alta correlação linear apresentada na tabela 4.6. No entanto, este índice estatístico demonstrou que para a velocidade do vento os métodos utilizados para estimar as falhas artificiais tiveram dificuldade para reproduzir a variação diurna. O método de Kriging, por exemplo, possui a menor correlação ( $0,38$ ) evidenciando mais uma vez que este método geoestatístico não reproduz com habilidade a variação da velocidade do vento ao longo de um dia.

A análise do RMSE tanto para a temperatura do ar quanto para a velocidade do vento demonstrou que a interpolação linear apresentou os menores erros individuais de preenchimento ( $1,4\text{ }^{\circ}\text{C}$  e  $1,1\text{ m/s}$ ) enquanto os demais métodos registraram erros de  $2$  a  $2,5\text{ }^{\circ}\text{C}$  para a temperatura e  $1,5$  a  $2,5\text{ m/s}$  para a velocidade (tabela 4.6).

A remoção do erro sistemático do modelo WRF aumentou o desempenho deste modelo em estimar as falhas artificiais principalmente da velocidade do vento. Essa melhora é evidenciada com a redução do erro médio absoluto e do RMSE. Ressalta-se também que a partir da remoção do *bias* os valores médios estimados para a temperatura do ar pelo modelo WRF subestimaram os dados medidos em apenas  $0,2\text{ }^{\circ}\text{C}$  (contra  $1^{\circ}\text{C}$  sem a retirada do *bias*) durante

o período de aquecimento da atmosfera próxima a superfície (10 as 19 UTC). Para este mesmo período, a remoção do erro médio para a velocidade do vento demonstrou que o WRF superestima esta variável em média em apenas 0,1 m/s contra 1,4 m/s na opção sem a retirada do *bias*.

Este fato, demonstra que, na ausência de dados medidos na vizinhança de linhas de transmissão ou no tocante a previsão dessas variáveis sobre linhas de transmissão, o WRF pode ser utilizado para prover informações meteorológicas principalmente durante o período mais crítico para o carregamento de linhas de transmissão (maior demanda de carga e temperaturas do ar mais altas).

Tabela 4.6 – Métricas de verificação do desempenho do preenchimento de falhas geradas artificialmente para o cluster-3.

Cluster-3	<i>Bias</i>		MAE		r		RMSE		POD	
	T <sup>1</sup>	V <sup>2</sup>	T <sup>1</sup>	V <sup>2</sup>	T	V	T <sup>1</sup>	V <sup>2</sup>	T <sup>1</sup>	V <sup>2</sup>
Interpolação	0,0	0,0	1,0	0,9	0,95	0,64	1,4	1,1	60,8	41,1
Kriging	-0,8	-0,9	1,8	1,4	0,89	0,38	2,2	1,5	36,4	24,4
ERA5	0,1	1,1	1,6	1,3	0,89	0,66	2,0	1,8	43,9	25,3
WRF	-0,8	1,3	2,0	1,7	0,86	0,53	2,5	2,5	32,6	16,7
WRF- <i>Bias</i> <sup>3</sup>	0,0	0,0	1,9	1,2	-	-	2,4	1,6	34,5	27,2

<sup>1</sup>Temperatura do ar em °C

<sup>2</sup>Velocidade do vento em m/s

<sup>3</sup>Índices após a remoção do *Bias*

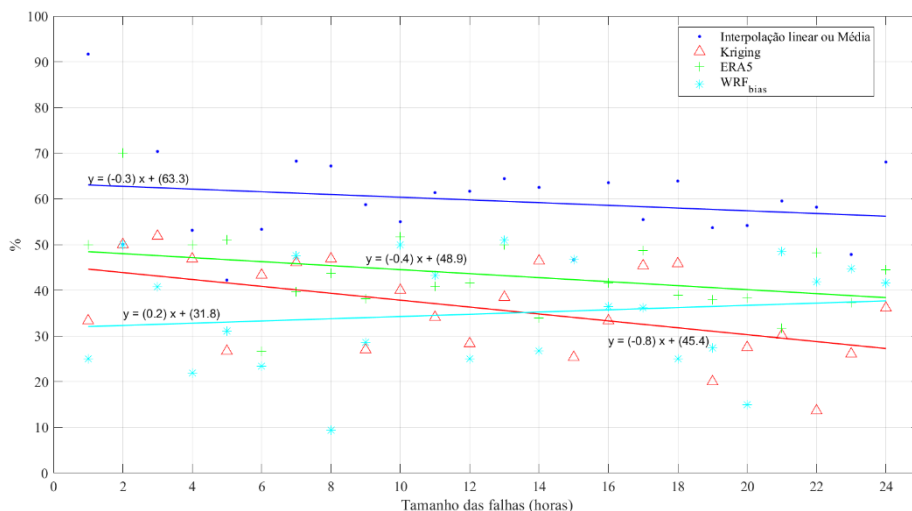
O índice de probabilidade de detecção utilizado para avaliar se os dados removidos foram estimados com sucesso pelos métodos considerados nesta tese demonstrou também que a interpolação linear apresentou o melhor desempenho para ambas as variáveis meteorológicas (tabela 4.6). A partir da tabela 4.6 observou-se também que, dentre os métodos em questão, o WRF após a remoção do erro médio tornou-se o segundo melhor método para estimar a velocidade do vento de acordo com o POD.

Para tamanhos distintos de falhas consecutivas foi observado que, com exceção de falhas com duas e cinco horas consecutivas para a temperatura do ar e de 13, 16 e 20 horas para a velocidade do vento, o método de interpolação apresentou a maior probabilidade de detecção do valor medido quando comparado aos demais métodos (figura 4.12).

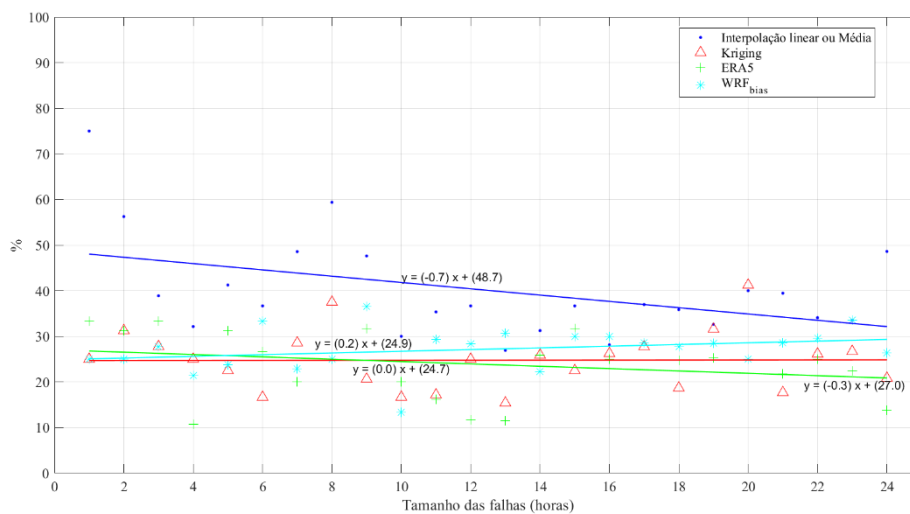
No entanto, assim como no cluster-1 e cluster-2 também foi verificado para o terceiro agrupamento que o método de interpolação apresentou um decréscimo acentuado na probabilidade de detecção entre buracos de apenas uma hora e os demais (figura 4.12). Como discutido anteriormente, isto ocorre porque para falhas de uma hora foi utilizada a interpolação linear entre a hora anterior e a posterior a falha enquanto para falhas de 2 a 24 horas utilizou-se a média do valor medido 24 horas antes e depois.

Figura 4.12 – Diagrama de dispersão e reta de regressão da Probabilidade de Detecção média dos valores médios removidos artificialmente das séries temporais do cluster-3 para (a) temperatura do ar e (b) velocidade do vento considerando diferentes tamanhos de falhas.

a)



b)



Fonte: o autor.

A análise da reta de regressão ajustada aos valores médios do índice de Probabilidade de Detecção demonstrou que apenas o WRF amentou a probabilidade de estimar os valores removidos artificialmente a partir do aumento do número de falhas consecutivas preenchidas, independentemente da variável meteorológica (figura 4.12).

Porém, embora o desempenho da interpolação tenha sido menor com o aumento do tamanho de falhas para ambas as variáveis, verificou-se que este método possui a maior

probabilidade de estimar as falhas para a maioria dos buracos de 1 a 24 horas das séries temporais quando comparado com os demais métodos.

Portanto, a partir da análise dos índices estatísticos apresentados na tabela 4.6 e do índice de probabilidade de detecção realizou-se o preenchimento das séries temporais brutas da temperatura do ar utilizando os dados da reanálise do ERA5 para buracos de 2 e 5 horas consecutivas. Para os demais tamanhos de falhas a imputação de dados foi realizada a partir do método de interpolação. A respeito das séries temporais brutas do cluster-3 o preenchimento de buracos com 13 falhas consecutivas foi realizado utilizando os dados do downscaling dinâmico (WRF sem o *bias*), de 20 horas consecutivas sem dados com o método de kriging e para os demais tamanhos de buracos o método de interpolação.

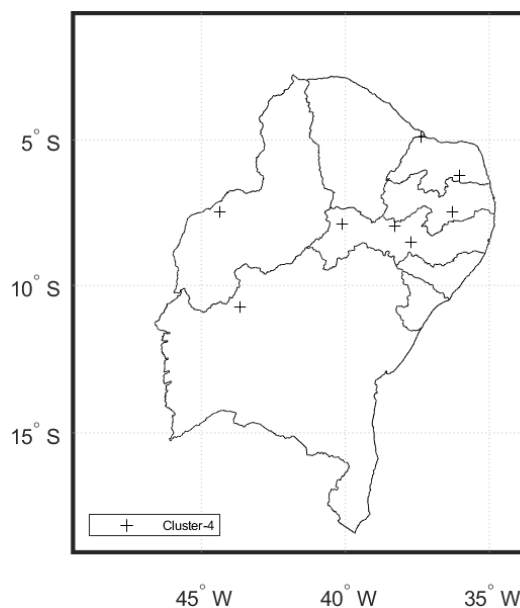
#### 4.1.4 Cluster 4

O quarto agrupamento é formado por oito estações meteorológicas (figura 4.13) das quais duas situam-se no estado do Rio Grande do Norte (Mossoró e Santa Cruz), uma no Piauí (Uruçuí), outra na Paraíba (Cabeceiras), uma na Bahia (Buritirama). As três últimas estações estão instaladas e operam em Pernambuco (Ibimirim, Serra Talhada e Ouricuri).

A temperatura e a velocidade do vento de algumas estações deste cluster são influenciadas também pela altitude e o relevo onde elas se encontram. Por exemplo, com exceção das estações localizadas no Oeste e Agreste Potiguar que estão a uma altitude de 36 e 237 metros, respectivamente, todas as demais estão acima de 430 metros de altitude (tabela 4.7). Adicionalmente, estações como a Santa Cruz está entre o Planalto da Borborema e a Serra dos Tabuleiros Costeiros. A Buritirama localizada na Bahia encontra-se sobre a Depressão do São Francisco. Outro exemplo é a estação instalada no Sertão Pernambucano, na cidade de Serra Talhada, a qual situa-se no Vale de Pajeú.

A respeito do bioma predominante a única estação do cluster-4 que não está sobre a caatinga é a Uruçuí, em operação no Sudoeste Piauiense, que pertence ao bioma do Cerrado.

Figura 4.13 - Localização das estações meteorológicas que formaram o cluster-4.



Fonte: o autor.

A análise exploratória das séries temporais que formaram o cluster-4 indicou que a mediana das máximas temperaturas diárias do ar e das mínimas velocidades diárias do vento para este cluster foi de 32,1 °C e de 1,0 m/s, respectivamente (tabela 4.7).

Tabela 4.7 - Estações meteorológicas agrupadas no cluster-4 e a mediana das máximas temperaturas diárias do ar ( $T_{med}$ ) e das mínimas velocidades diárias do vento ( $V_{med}$ ).

Latitude (°)	Longitude (°)	Altitude (m)	$T_{med}$ (°C)	$V_{med}$ (m/s)	Nome da estação	UF
-4,9	-37,3	36,0	32,2	1,0	Mossoró	RN
-6,2	-36,0	237,0	32,1	1,0	Santa Cruz	RN
-7,4	-40,1	464,0	32,4	0,8	Uruçuí	PI
-7,4	-36,2	436,0	31,8	1,0	Cabeceiras	PB
-10,7	-43,6	506,5	32,6	1,0	Buritirama	BA
-8,5	-37,7	448,0	32,6	1,0	Ibimirim	PE
-7,9	-38,2	461,0	31,8	0,9	Serra Talhada	PE
-7,8	-40,1	464,0	32,0	1,0	Ouricuri	PE
Mediana			32,1	1,0		

As séries temporais disponíveis para as análises referente ao cluster-4 são apresentadas na figura 4.14. A estação com maior número de dados disponíveis é a de Mossoró com aproximadamente 11 anos (06/2007 a 03/2018) enquanto a menor é a da Ouricuri e Santa Cruz com quase 8 anos (06/2010 a 03/2018).

A respeito dos buracos presentes nas séries observou-se que as estações de Mossoró, Santa Cruz, Uruçuí e Buritirama apresentaram mais falhas ao longo dos anos para a velocidade do vento do que para a temperatura do ar. Por exemplo, ao analisar os dados de velocidade do vento disponíveis para a estação Buritirama (figura 4.14b) observou-se que, principalmente a partir do final do ano de 2014, houve ausência de dados medidos por pequenos períodos não consecutivos até o final da série temporal. Entretanto, este fato não foi observado na série da temperatura do ar desta estação (figura 4.14a).

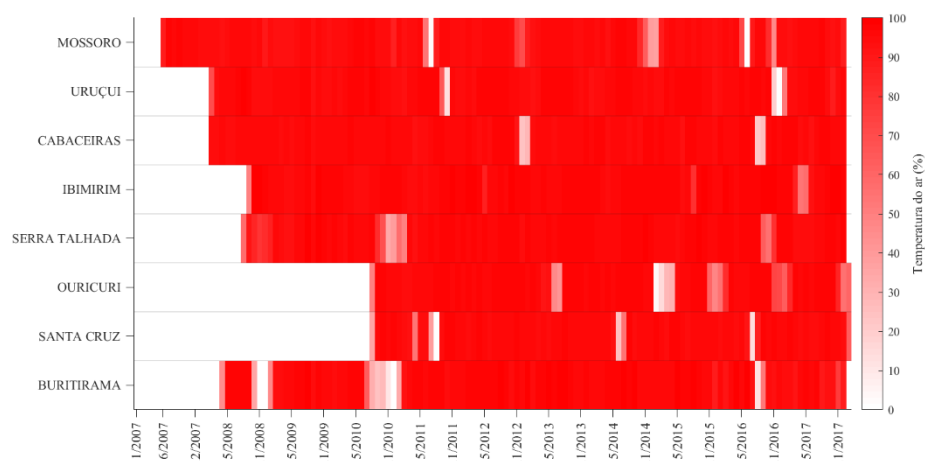
Diante desse fato, evidenciou-se que os distintos princípios de medida destas variáveis podem influenciar a quantidade de dados disponíveis. Em outras palavras, se o anemômetro utilizado para medir a velocidade do vento contiver partes móveis (tipo concha ou hélice) pode haver um maior número de falhas devido a problemas mecânicos do sensor quando comparado ao instrumento que mede a temperatura do ar o qual não possui partes móveis.

Por fim, não foi observado falhas consecutivas maiores que dois meses em nenhuma das séries temporais das estações meteorológicas presentes no cluster-4 (figura 4.14).

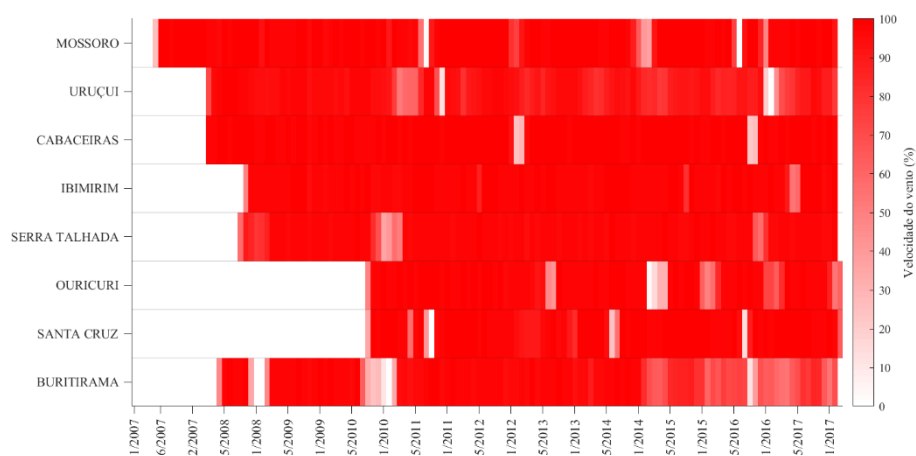


Figura 4.14 - Séries temporais brutas da (a) temperatura do ar e (b) velocidade do vento das estações do INMET pertencentes ao cluster-4.

a)



b)



Fonte: o autor.

O desempenho do preenchimento de falhas geradas artificialmente a partir dos métodos de imputação de dados definidos nesta tese é apresentado na figura 4.15. Após remover 10% dos dados brutos das séries temporais das estações do cluster-4 verificou-se que, assim como para os clusters anteriores, todos os métodos capturaram a variabilidade da temperatura do ar (figura 4.15a). O bom desempenho em estimar esta variabilidade foi confirmado pelos altos valores de correlação linear apresentados na tabela 4.8.

No tocante aos erros médios de estimativa, verificou-se que o modelo WRF foi o único que subestimou a temperatura durante todo o ciclo diurno. Esse fato foi evidenciado também pelo erro médio de  $-0,7\text{ }^{\circ}\text{C}$  encontrado para este modelo (tabela 4.8). O segundo método que

apresentou o maior *bias* foi o kriging (0,8 °C). Porém, ao contrário do WRF este método superestimou a temperatura do ar durante todos os horários do dia.

Ao analisar a figura 4.15a foi possível observar que apenas o modelo global teve dificuldade em reproduzir o acréscimo de temperatura durante o período de aquecimento da atmosfera (10 as 18 UTC). No entanto, as estimativas distantes dos valores médios removidos artificialmente se anularam no cálculo do *bias* e, por este motivo, o ERA5 apresentou um erro médio de apenas 0,2 °C. Porém, a análise complementar utilizando os índices MAE e RMSE (tabela 4.8) indicou a reanálise ERA5 não é muito superior ao WRF para estimar os dados removidos das séries temporais do cluster-4.

Ainda para a temperatura do ar ressalta-se a excelente performance da interpolação em estimar o ciclo diurno (figura 4.15a). Esse desempenho da interpolação refletiu-se no erro médio nulo e nos menores erros quando observado os demais índices presentes na tabela 4.8.

Por fim, a remoção do erro sistemático do modelo WRF demonstrou que, principalmente no período mais crítico para a operação de linhas de transmissão (período onde ocorre o aumento da temperatura do ar, 10 as 18 UTC) os valores estimados pelo modelo regional se ajustam perfeitamente ao dados observados (figura 4.15a). A partir da remoção do erro médio quase constante em todo o ciclo diurno, observou-se que as outras medidas de erro também diminuíram e se aproximaram das medidas de erro encontradas para a interpolação (tabela 4.8).

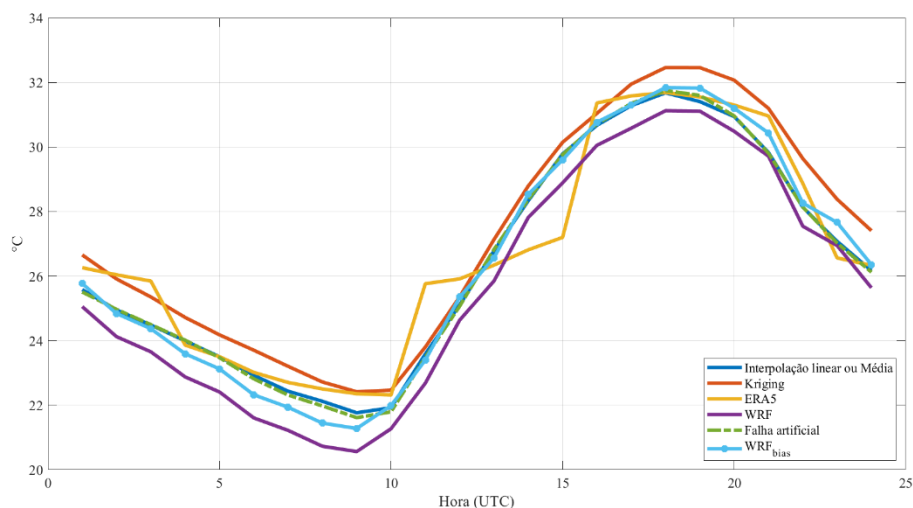
Para a velocidade do vento (figura 4.15b) a interpolação linear reproduziu a variabilidade desta variável com eficiência durante o todo o ciclo diurno. O método geoestatístico também capturou com sucesso a variação da velocidade do vento, embora entre 15 e 01 UTC tenha subestimado os dados medidos em até 0,7 m/s (23 UTC).

Por outro lado, os modelos dinâmicos ERA5 e WRF superestimaram esta variável em média em 1,6 m/s e 2,3 m/s, respectivamente (tabela 4.8). Observou-se também que com a remoção do erro médio do modelo WRF os índices MEA e RMSE diminuíram em 1,1 m/s e 1 m/s respectivamente. A análise da figura 4.15b demonstrou que o modelo WRF sem o *bias* estimou valores mais próximos aos dados medidos, principalmente, durante o período de aumento da velocidade do vento (10 e 13 UTC).

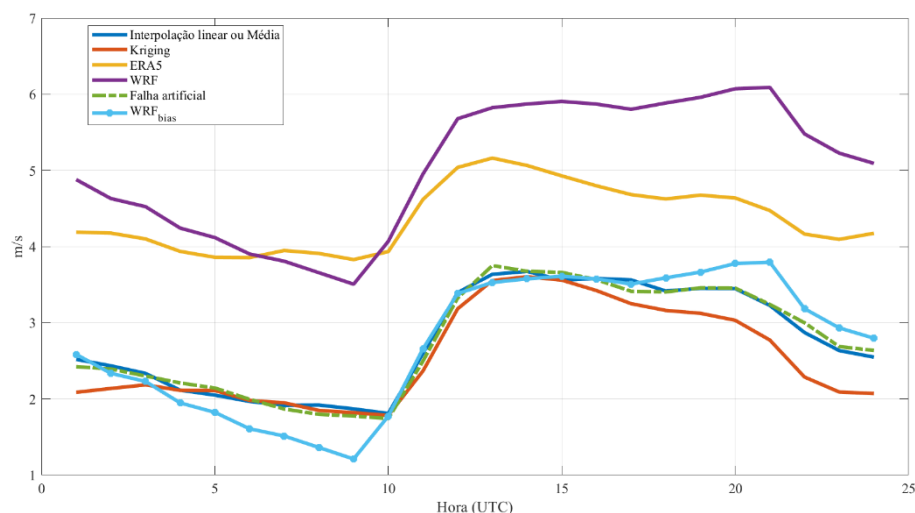
Por fim, ressalta-se que tanto para a temperatura do ar quanto para a velocidade do vento os índices estatísticos indicaram que a interpolação teve o melhor desempenho no preenchimento das falhas geradas artificialmente para o cluster-4.

Figura 4.15 - Média dos valores removidos artificialmente (linha tracejada) e estimados por 4 diferentes métodos para (a) temperatura do ar e (b) velocidade do vento para o cluster-4.

a)



b)



Fonte: o autor.

Tabela 4.8 - Métricas de verificação do desempenho do preenchimento de falhas geradas artificialmente para o cluster-4.

Cluster-4	Bias		MAE		r		RMSE		POD	
	T <sup>1</sup>	V <sup>2</sup>	T <sup>1</sup>	V <sup>2</sup>	T	V	T <sup>1</sup>	V <sup>2</sup>	T <sup>1</sup>	V <sup>2</sup>
Interpolação	0,0	0,0	0,9	0,8	0,94	0,76	1,3	1,1	67,7	43,8
Kriging	0,8	-0,2	2,0	1,2	0,82	0,48	2,6	1,6	32,2	27,7
ERA5	0,2	1,6	1,3	1,8	0,91	0,62	1,7	2,1	50,2	15,0
WRF	-0,7	2,3	1,5	2,5	0,91	0,51	1,9	2,9	44,4	8,6
WRF-Bias <sup>3</sup>	0,0	0,0	1,3	1,4	-	-	1,7	1,9	51,3	24,3

<sup>1</sup>Temperatura do ar em °C

<sup>2</sup>Velocidade do vento em m/s

<sup>3</sup>Índices após a remoção do Bias

A última avaliação do desempenho de preenchimento dos buracos artificiais considerou o índice de probabilidade de detecção. Este índice que, indica o sucesso dos métodos estimarem os dados removidos, demonstrou que a interpolação linear apresentou o melhor desempenho para ambas as variáveis meteorológicas (tabela 4.8). Para a temperatura do ar o POD médio do WRF após a remoção do *bias* aumentou de 44,4 para 51,3 % e para a velocidade do vento de 8,6% para 24,3%. Embora esses novos valores tenham superado o POD dos métodos de kriging e da reanálise do ERA5 no preenchimento dos dados o desempenho do modelo regional ainda foi baixo.

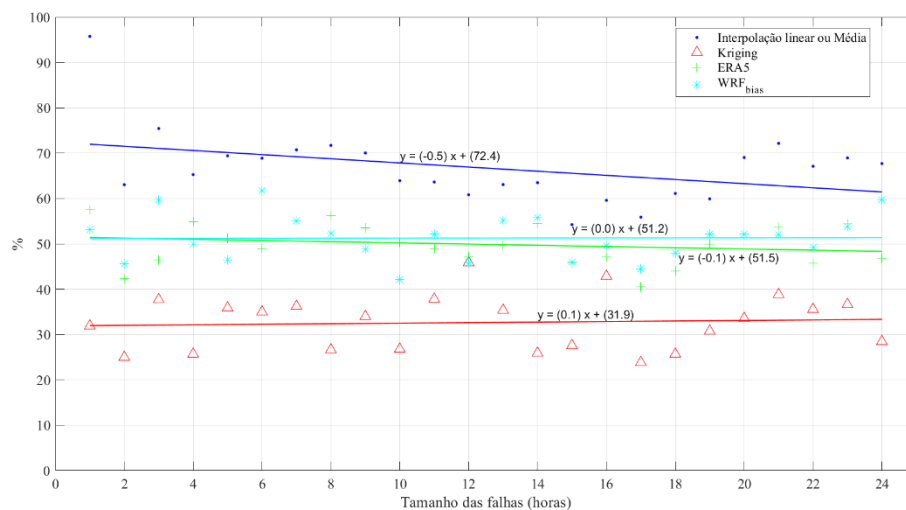
A análise do desempenho dos métodos para estimar diferentes tamanhos de buracos consecutivos (1 a 24 horas) indicou que para a temperatura do ar apenas o método de kriging aumentou a probabilidade de detecção dos valores removidos com o aumento do número de falhas (figura 4.16). A estimativa dos dados utilizando o método do downscaling dinâmico foi a única que não apresentou tendência para distintos tamanhos de falhas. Para os demais métodos, embora tenham indicado tendência positiva ou negativa, o coeficiente angular da reta de regressão demonstrou que a probabilidade de detecção varia apenas ligeiramente com o tamanho de buracos.

Para a velocidade do vento a probabilidade média horaria para estimar os valores removidos diminuiu para todos os métodos quando comparado com o desempenho para a temperatura (figura 4.16b). Porém, a interpolação foi o método que demonstrou ter o maior sucesso para detectar os valores removidos independentemente do tamanho de buracos consecutivos. Ou seja, para falhas de 1 hora até 24 horas a maior probabilidade de detecção sempre foi verificada para o método de interpolação.

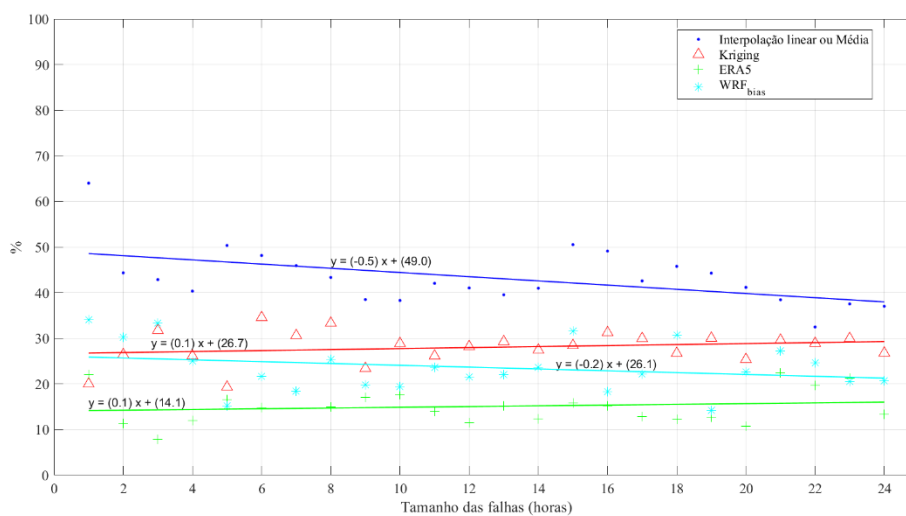
A inclinação da reta de regressão para a velocidade do vento (figura 4.16b) demonstrou que, assim como para a temperatura do ar, a tendência de aumento ou queda do sucesso de detecção é ligeiramente pequena independentemente do número consecutivo de falhas a ser preenchida. Ainda sobre a tendência, o menor coeficiente angular da reta de regressão para a velocidade do vento foi encontrado para o método de interpolação (-0,5). Esta tendência negativa pode ser justificada, pois para falhas de uma hora é utilizada a interpolação linear entre o dado medido uma hora antes e uma hora depois da falha a ser preenchida.

Figura 4.16 - Diagrama de dispersão e reta de regressão da Probabilidade de Detecção média dos valores médios removidos artificialmente das séries temporais do cluster-4 para (a) temperatura do ar e (b) velocidade do vento considerando diferentes tamanhos de falhas.

a)



b)



Fonte: o autor.

Portanto, a partir da análise dos índices estatísticos apresentados na tabela 4.8 e dos índices de probabilidade de detecção para a temperatura do ar e velocidade do vento realizou-se o preenchimento dos buracos de 1 até 24 horas das séries temporais brutas do cluster-4 utilizando a interpolação linear para falhas de 1 hora ou a média entre o valor medido 24 horas antes e 24 horas depois da falha.

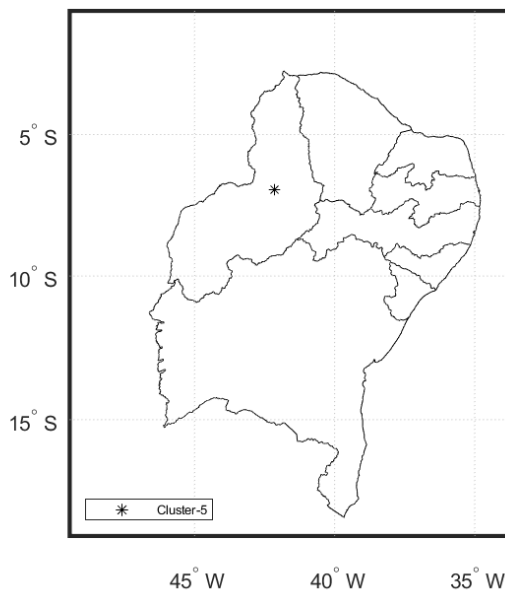
#### 4.1.5 Cluster 5

A análise de cluster indicou que a mediana das máximas temperaturas diárias do ar (34,7 °C) e das mínimas velocidades diárias do vento (0,3 m/s) da estação Oeiras são dissimilares a todas as outras medianas das estações consideradas no processo de agrupamento. Por este motivo, o cluster-5 (figura 4.17) foi formado por apenas esta estação que está localizada no vale do rio Canindé, na mesorregião Sudeste do Piauí.

Oeiras está situada sobre uma zona de transição entre a caatinga e o cerrado, a 156 metros de altitude (tabela 4.9) e possui precipitação anual acumulada de 800 mm sendo que maior parte deste volume está concentrado nos primeiros quatro meses do ano (INMET, 2009).

A partir destas características o cluster-5 apresentou as situações mais críticas para operação de linhas de transmissão, pois foi para este cluster que se encontrou a maior mediana das máximas temperaturas diárias do ar e a menor mediana para as mínimas velocidades diárias do vento.

Figura 4.17 - Localização da estação meteorológica que representou o cluster-5.



Fonte: o autor.

Tabela 4.9 - Estação meteorológica do cluster-5 e a mediana das máximas temperaturas diárias do ar ( $T_{med}$ ) e das mínimas velocidades diárias do vento ( $V_{med}$ ).

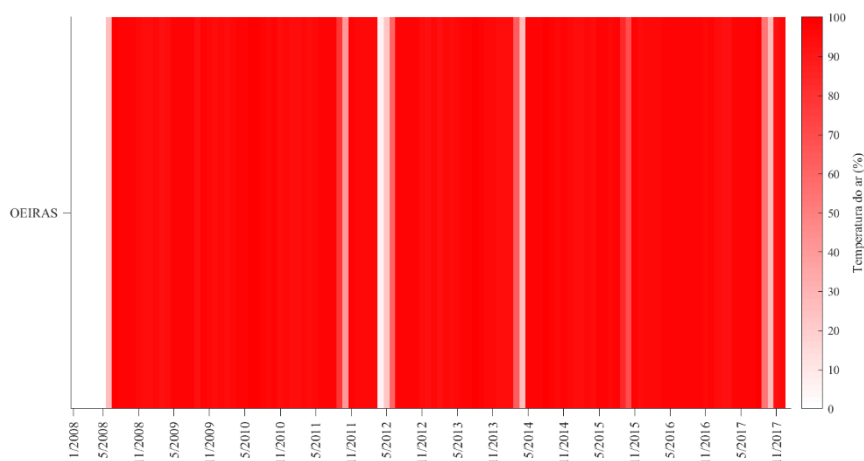
Latitude (°)	Longitude (°)	Altitude (m)	$T_{med}$ (°C)	$V_{med}$ (m/s)	Nome da estação	UF
-6,9	-42,1	156,0	34,7	0,3	Oeiras	PI

A série temporal disponível para as análises referente ao cluster-5 é apresentada na figura 4.18. A estação Oeiras possui aproximadamente 10 anos de dados medidos (05/2008 a 03/2018) e com poucas falhas ao longo da série de temperatura do ar e velocidade do vento.

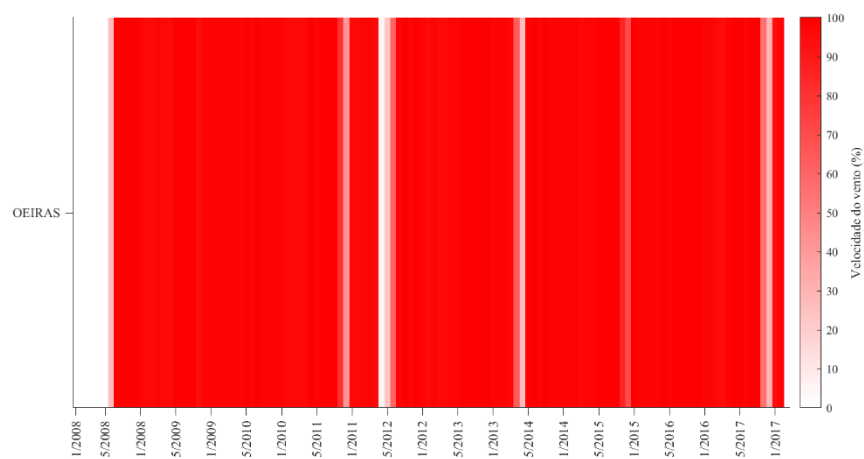
As maiores ausências de dados medidos na série temporal do cluster-5 foram observadas entre abril e maio do ano de 2012. Adicionalmente, verificou-se outras falhas com tamanhos menores, mas de ocorrência simultânea para ambas variáveis. Este fato sugere que a ausência de dados foi devido a operação da estação (provavelmente falta de bateria) ao invés de problemas individuais dos sensores de temperatura do ar e velocidade do vento.

Figura 4.18 - Série temporal bruta da (a) temperatura do ar e (b) velocidade do vento das estação do INMET pertencente ao cluster-5.

a)



b)



Fonte: o autor.

O desempenho do preenchimento de falhas geradas artificialmente a partir dos quatro métodos de imputação é apresentado na figura 4.19. Para a temperatura do ar, verificou-se que a variabilidade dos dados removidos foi reproduzida por todos os métodos estimadores (figura 4.19a). A alta correlação linear encontrada para estes métodos (tabela 4.10) também confirmou a boa performance na captura da variação da temperatura durante o ciclo diurno.

O menor valor médio temperaturas diárias removidas artificialmente foi observado as 10 UTC (figura 4.19a). E todos os métodos tiveram sucesso em estimar a menor temperatura diária para este mesmo horário. Entretanto, para o maior valor médio desta variável apenas o modelo WRF estimou corretamente o horário de ocorrência

No tocante aos erros médios de estimativa, verificou-se que a interpolação linear e o modelo WRF apresentaram os menores valores deste índice (0,1 °C e -0,5 °C, respectivamente) (tabela 4.10). Por outro lado, o maior erro médio foi encontrado para a ERA5 (1,1 °C). A presença de grandes erros médios individuais da reanálise do ERA5 observados na figura 4.19a refletiu-se no maior RMSE (2,8 °C) quando comparados ao dos demais métodos.

Assim como observado nos agrupamentos anteriores, observou-se também que a reanálise ERA5 teve dificuldade em reproduzir o acréscimo de temperatura durante o período de aquecimento da atmosfera (10 as 19 UTC). Neste período, a ERA5, diferente dos demais métodos, estimou a temperatura do ar semelhante a um padrão binário ou de escada. A elevação de temperatura não foi suave ou gradual conforme indicada pelos dados medidos (figura 4.19a). Ainda para a temperatura do ar, ressalta-se que, embora a interpolação tenha apresentado um *bias* próximo a zero (tabela 4.10), os índices MAE e RMSE indicaram que, em alguns horários, este método estimou valores distantes dos dados medidos.

Por fim, a remoção do erro médio do modelo WRF não demonstrou que, para o cluster-5, houve o aumento no desempenho em estimar as falhas artificiais. Este resultado pode estar associado ao fato de que o erro médio não foi constante ao longo da série.

Para a velocidade do vento (figura 4.19b) verificou-se que o erro médio da interpolação linear é nulo. Ao observar a figura 4.19b verificou-se também que ao logo do ciclo diurno este método não apresentou grandes erros individuais durante a estimativa da velocidade do vento. Por este motivo, os valores de MAE e RMSE indicados na tabela 4.10 são os menores em relação aos calculados para os demais métodos.

Para o cluster-5 as altas correlações lineares encontradas para todos os métodos indicaram que estes tiveram habilidade em capturar a variação da velocidade do vento ao longo do dia. Todos os métodos acertaram em estimar o horário de ocorrência do máximo valor da velocidade do vento ocorrido as 13 UTC (figura 4.19b). Entretanto, com exceção da



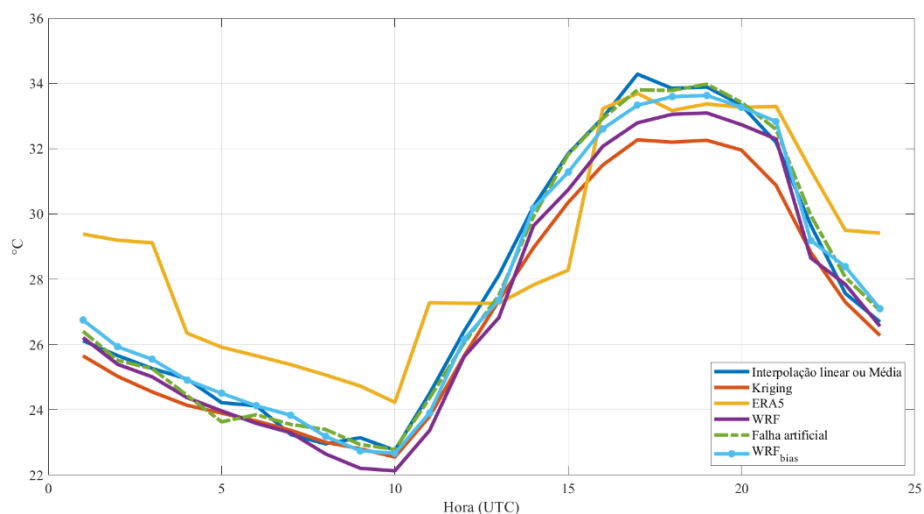
interpolação, os demais métodos apresentaram dificuldade em estimar corretamente os valores médios das falhas artificiais da velocidade do vento. Os modelos numéricos da atmosfera superestimaram os dados medidos em até 1,7 m/s enquanto o método geoestatístico subestimou em 0,8 m/s. O erro médio absoluto e a raiz do erro quadrático médio indicaram que os erros individuais foram grandes para estes três métodos supracitados.

Por fim, observou-se que devido ao WRF possuir um erro médio sistemático ao longo do ciclo diurno a remoção deste erro indicou que para o WRF sem o *bias*, os índices MEA e RMSE tiveram uma diminuição de 0,6 m/s e 0,7 m/s respectivamente (tabela 4.10).

Por fim, ressalta-se que tanto para a temperatura do ar quanto para a velocidade do vento os índices estatísticos indicaram que a interpolação teve o melhor desempenho no preenchimento das falhas geradas artificialmente para o cluster-5.

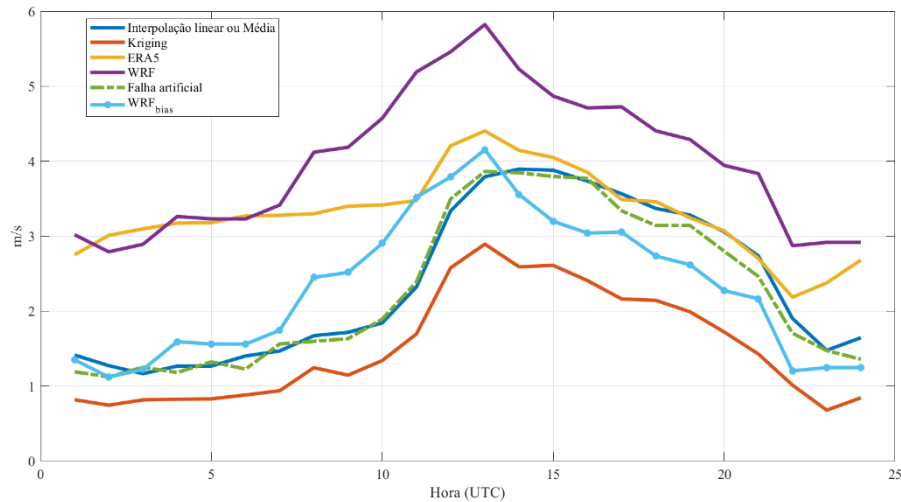
Figura 4.19 – Média dos valores removidos artificialmente (linha tracejada) e estimados por 4 diferentes métodos para (a) temperatura do ar e (b) velocidade do vento para o cluster-5.

a)



Continuação – Figura 4.20 – Média dos valores removidos artificialmente (linha tracejada) e estimados por 4 diferentes métodos para (a) temperatura do ar e (b) velocidade do vento para o cluster-5.

b)



Fonte: o autor.

Tabela 4.10 – Métricas de verificação do desempenho do preenchimento de falhas geradas artificialmente para o cluster-5.

Cluster-5	Bias		MAE		r		RMSE		POD	
	T <sup>1</sup>	V <sup>2</sup>	T <sup>1</sup>	V <sup>2</sup>	T	V	T <sup>1</sup>	V <sup>2</sup>	T <sup>1</sup>	V <sup>2</sup>
Interpolação	0,1	0,0	1,4	0,7	0,90	0,85	2,0	0,9	54,7	46,6
Kriging	-0,8	-0,8	1,6	1,1	0,92	0,70	2,0	1,5	36,4	35,3
ERA5	1,1	1,0	2,0	1,3	0,84	0,74	2,8	1,6	34,7	24,2
WRF	-0,5	1,7	2,0	1,8	0,85	0,73	2,5	2,2	26,0	18,2
WRF-Bias <sup>3</sup>	0,0	0,0	2,0	1,2	-	-	2,4	1,5	25,7	30,2

<sup>1</sup>Temperatura do ar em °C

<sup>2</sup>Velocidade do vento em m/s

<sup>3</sup> Índices após a remoção do *Bias*

A última avaliação do desempenho de preenchimento dos buracos artificiais considerou o sucesso dos métodos em estimar os dados removidos a partir da Probabilidade de Detecção. Este índice demonstrou que a interpolação linear apresentou o melhor desempenho em estimar os dados medidos para ambas as variáveis meteorológicas (tabela 4.10). Para a temperatura do ar, não houve aumento do POD após remover o *bias* do modelo WRF pelo fato do erro médio não ter sido constante para esta variável. Entretanto, para a velocidade do vento após a remoção do *bias* o POD médio do WRF aumentou de 18,2 para 30,2 %.

Ainda para a velocidade do vento, verificou-se a partir da tabela 4.10 que, com exceção do modelo WRF sem o *bias*, a probabilidade média horária para estimar os valores removidos diminuiu para todos os métodos quando comparado com o desempenho para a temperatura.

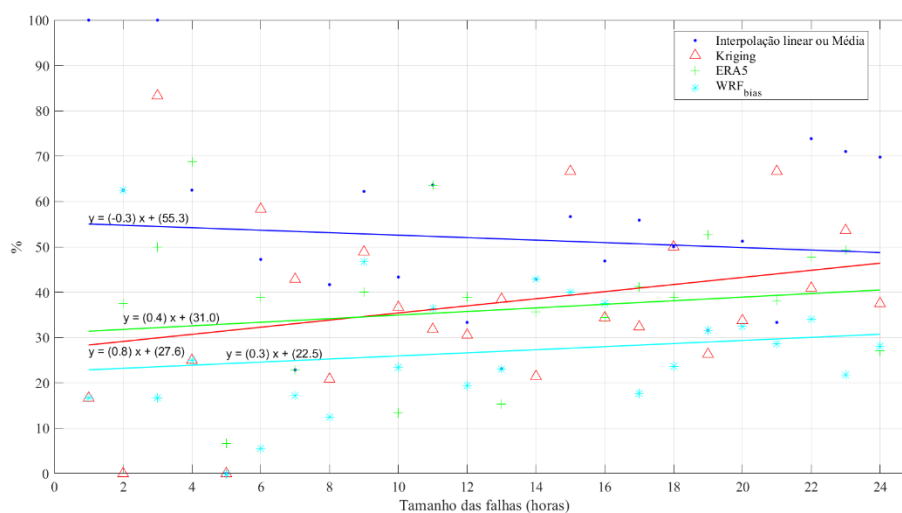
A última análise referente a probabilidade de detecção foi realizada utilizando a reta de regressão e a dispersão dos dados de POD para cada horário do ciclo diurno (figura 4.21). A inclinação da reta de regressão para a temperatura do ar (figura 4.21a) indicou que apenas a interpolação diminuiu a probabilidade de estimar os dados removidos artificialmente com o aumento do número de falhas. Também foi verificado para a temperatura do ar que para falhas de 1 e 3 horas a interpolação obteve 100% de sucesso para estimar os dados removidos. No entanto, com o aumento do número de buracos como por exemplo para falhas com tamanhos consecutivos de 4 a 7 horas, 12 a 15 horas, 19 e 21 horas, a probabilidade de detecção da interpolação foi superada por outros métodos de estimação (figura 4.21a).

Para a velocidade do vento (figura 4.21b) o único método que não apresentou tendência foi Kriging. Os demais métodos, embora tenham apresentando tendência positiva indicando que a probabilidade de acerto aumenta com o aumento do número de falhas, a inclinação da reta indicou que o índice POD teve um aumento ligeiramente pequeno.

A maior probabilidade de estimar os dados de velocidade do vento removidos artificialmente só não foi melhor para a interpolação em falhas com tamanho consecutivo de 2, 10, 13 e 15 horas (figura 4.21b).

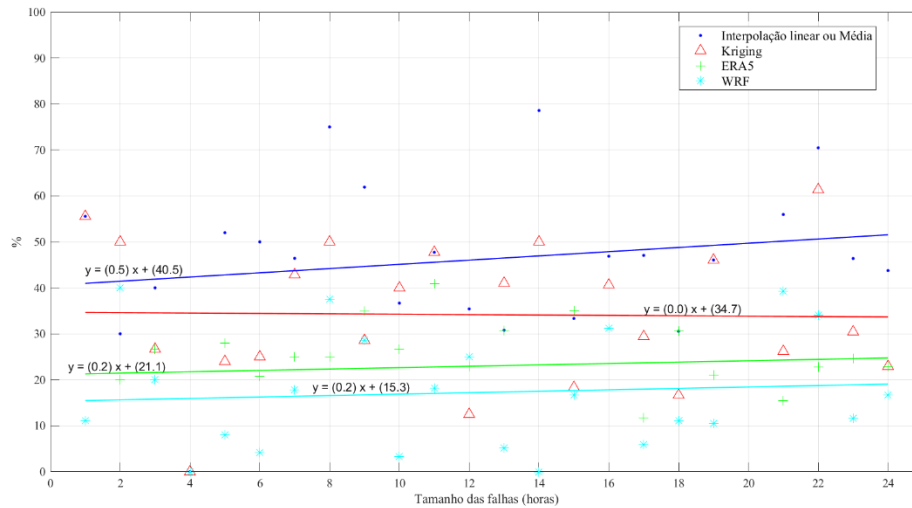
Figura 4.21 - Diagrama de dispersão e reta de regressão da Probabilidade de Detecção média dos valores médios removidos artificialmente da série temporal do cluster-5 para (a) temperatura do ar e (b) velocidade do vento considerando diferentes tamanhos de falhas.

a)



Continuação – Figura 4.22 - Diagrama de dispersão e reta de regressão da Probabilidade de Detecção média dos valores médios removidos artificialmente da série temporal do cluster-5 para (a) temperatura do ar e (b) velocidade do vento considerando diferentes tamanhos de falhas.

b)



Fonte: o autor.

Portanto, a partir da análise dos índices estatísticos apresentados na tabela 4.10 e dos índices de probabilidade de detecção para a temperatura do ar (figura 4.21a) realizou-se o preenchimento dos buracos de 1 até 24 horas da série temporal bruta do cluster-5 da seguinte maneira:

- Falhas de 4, 5, 12 e 19 horas consecutivas: ERA5;
- Falhas de 6, 7, 13, 15 e 21 horas consecutivas: kriging;
- Demais falhas: interpolação linear ou média;

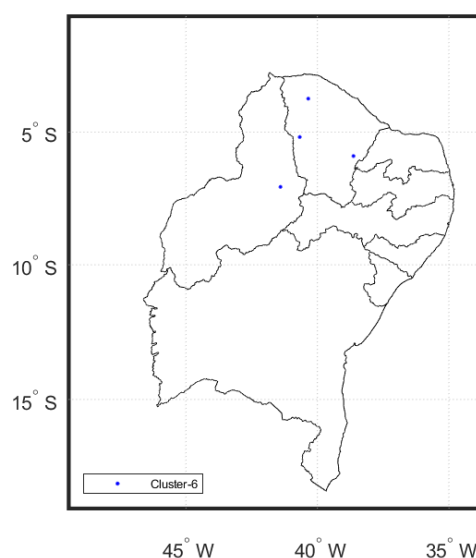
A análise semelhante foi realizada para a velocidade do vento. E, portanto, o preenchimento das falhas foi realizado utilizando o método de interpolação para todos os tamanhos com exceção de buracos com 2 e 10 horas consecutivas de falhas onde utilizou-se o método de kriging e de 6 horas consecutivas (WRF sem o *bias*).

#### 4.1.6 Cluster 6

O cluster-6 é formado por quatro estações meteorológicas (figura 4.23) das quais três situam-se no Estado do Ceará na seguintes mesorregiões: Noroeste cearense (Sobral), Sertões (Crateús) e Jaguaribe (Jaguaribe). As estações cearenses situam-se sobre a depressão sertaneja e o vale do Jaguaribe. Por outro lado, a quarta estação meteorológica está localizada no sudeste piauiense na cidade de Picos (Vale do Guaribas).

As estações do cluster-6 estão entre 100 e 300 metros de altitude (tabela 4.11) e pertencem ao clima semiárido – bioma caatinga. A mediana das temperaturas máximas deste cluster é de 34,6 °C e das velocidades mínimas do vento de 0,7 m/s.

Figura 4.23 – Localização das estações meteorológicas que formaram o cluster-6.



Fonte: o autor.

Tabela 4.11 – Estações meteorológicas agrupadas no cluster-6 e a mediana das máximas temperaturas diárias do ar ( $T_{med}$ ) e das mínimas velocidades diárias do vento ( $V_{med}$ ).

Latitude (°)	Longitude (°)	Altitude (m)	$T_{med}$ (°C)	$V_{med}$ (m/s)	Nome da estação	UF
-3,7	-40,3	109,6	34,4	0,7	Sobral	CE
-5,1	-40,6	291	34,3	0,7	Crateús	CE
-5,9	-38,6	184	35,1	0,8	Jaguaribe	CE
-7,0	-41,4	233	34,7	0,7	Picos	PI
Mediana			34,6	0,7		

As séries temporais disponíveis para as análises referente ao cluster-6 são apresentadas na figura 4.24. A estação com maior número de dados disponíveis é a de Sobral com pouco mais de 13 anos (01/2005 a 03/2018) enquanto as outras três séries possuem aproximadamente 10 anos de dados medidos (06/2008 a 03/2018).

Ao analisar as séries temporais de temperatura do ar (figura 4.24a) observou-se que há falhas com distintos tamanhos em todas as estações, porém para a estação Sobral verificou-se que houve um período de quase seis meses de ausência de dados (12/2014 a 05/2015). Para as

demais estações os maiores períodos sem medição de dados de temperatura do ar foram de aproximadamente 30 dias.

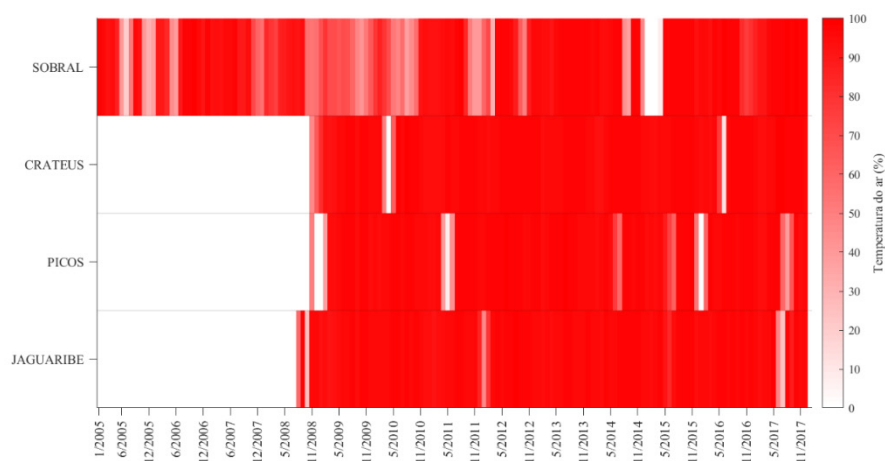
Para a série de velocidade do vento (figura 4.24b) da estação de Sobral não houve um período prolongado sem medições como aquele observado para a série temporal da temperatura. Porém, ainda foi verificado muitas falhas ao longo de toda a série temporal.

Ainda para a velocidade do vento, a análise da série temporal da estação de Picos indicou que esta estação permaneceu longos períodos sem realizar medições. Por exemplo, o maior período com ausência de dados medidos para a estação de Picos foi de dois anos (06/2010 a 06/2012). Outros períodos com até seis meses sem medições (02/2013 a 08/2013) também foram observados para esta estação.

Crateús e Jaguaribe também apresentaram mais buracos consecutivos nas séries temporais de velocidade do vento quando comparado as suas séries de temperatura do ar. E o fato de muitos buracos nas séries temporais da temperatura do ar e velocidade do vento não ocorrerem simultaneamente é um indicativo de que a maior parte das falhas se deve a problemas individuais do sensores de medida ao invés de falhas generalizadas da estação.

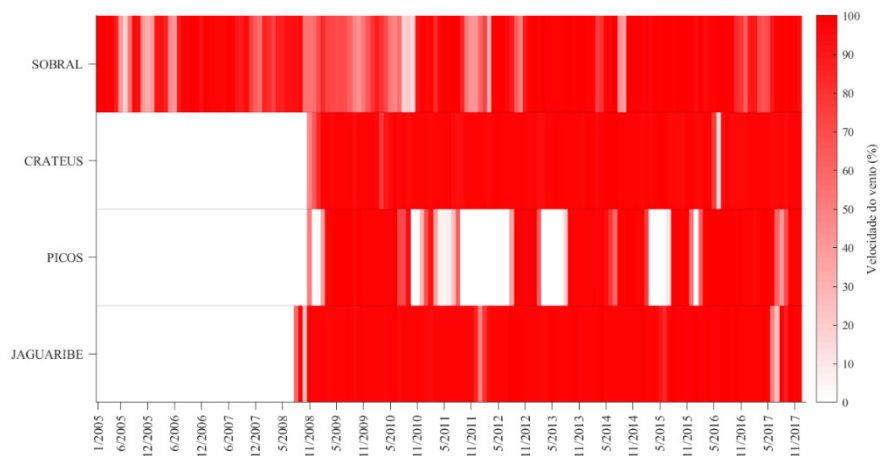
Figura 4.24 - Séries temporais brutas da (a) temperatura do ar e (b) velocidade do vento das estações do INMET pertencente ao cluster-6.

a)



Continuação – Figura 4.25 - Séries temporais brutas da (a) temperatura do ar e (b) velocidade do vento das estações do INMET pertencente ao cluster-6.

b)



Fonte: o autor.

O desempenho do preenchimento de falhas geradas artificialmente a partir dos métodos de imputação de dados é apresentado na figura 4.26. Após remover 10% dos dados brutos das séries temporais das estações do cluster-6 verificou-se que todos os métodos de imputação de dados capturaram a variabilidade da temperatura do ar ao longo do ciclo diurno (figura 4.26a). O desempenho em estimar esta variação da temperatura foi confirmado pelos altos valores de correlação linear apresentados na tabela 4.12.

Além de reproduzir com eficiência a variabilidade da temperatura o *bias* indicou que todos os métodos subestimam os dados medidos, porém com valores próximos a zero (tabela 4.12). A única exceção para esta afirmação é para o método de Kriging cujo erro médio foi de  $-1,1$  °C.

Ao analisar a figura 4.26a foi possível observar que, assim como para os clusters anteriores, apenas o modelo global teve dificuldade em reproduzir o acréscimo da temperatura do ar durante o período de aquecimento da atmosfera (10 as 19 UTC). Neste período a reanálise ERA5 apresentou um comportamento em forma de escada para reproduzir a elevação da temperatura do ciclo diurno. E, embora o *bias* da ERA5 tenha sido de apenas  $-0,1$  °C, os índices MAE e RMSE (tabela 4.12) indicaram que este método possui valores estimados distantes dos dados medidos.

Ainda para a temperatura do ar, verificou-se a alta performance da interpolação em estimar as falhas artificiais do ciclo diurno (figura 4.26a). Esse desempenho da interpolação

refletiu-se em medidas de erros menores quando comparado aos demais métodos de estimativa (tabela 4.12).

Por fim, a remoção do erro médio do modelo WRF não demonstrou uma diminuição significativa nos índices MAE e RMSE uma vez que o *bias* não foi constante para todo o período estimado.

Para a velocidade do vento (figura 4.26b), apenas o método de interpolação estimou valores em fase com a variação da velocidade do vento registrada ao longo do dia de acordo com a correlação linear (0,70). No entanto, todos métodos utilizados para estimar as falhas artificiais reproduziram com eficiência a elevação da velocidade do vento ocorrida a partir das 10 UTC. Entretanto, observou-se que a reanálise ERA5 estimou o decréscimo da velocidade a partir das 14 UTC enquanto os valores médios das falhas artificiais indicaram que este fato ocorreu apenas a partir das 21 UTC. E, de maneira abrupta, o método de kriging estimou a diminuição da velocidade do vento a partir das 19 UTC e o WRF após as 21 UTC.

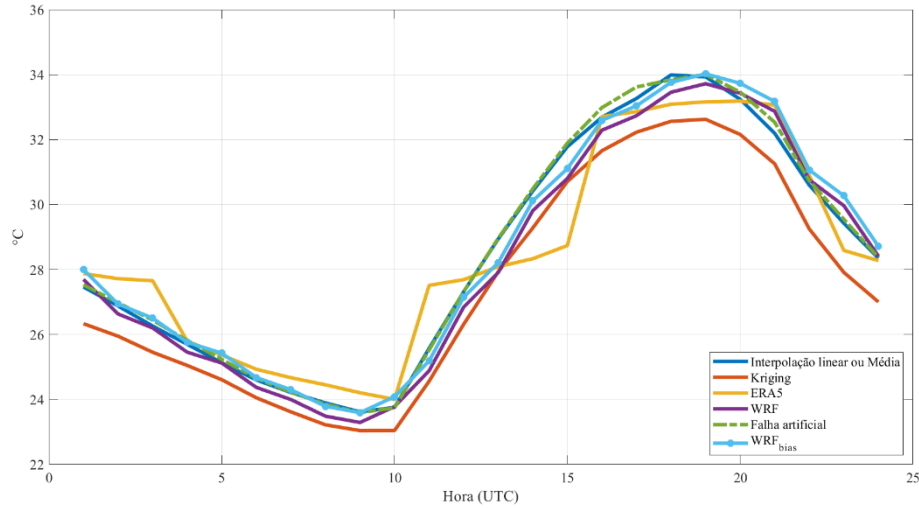
A respeito do erro médio para a velocidade do vento, verificou-se na tabela 4.12 que a reanálise ERA5 e o WRF possuem *bias* de 1 m/s e 1,4 m/s, respectivamente. A análise da figura 4.26b corrobora com o erro médio calculado para os modelos dinâmicos uma vez que foi possível observar estimativas distantes dos dados medidos.

Com o objetivo de corrigir as superestimativas do modelo WRF removeu-se o *bias* deste modelo e assim observou-se valores estimados pelo modelo regional mais próximos aos dados medidos (figura 4.26b). Porém, devido ao erro médio não ser constante para toda a série houve apenas uma ligeira diminuição nas medidas de erro contidas na tabela 4.12.

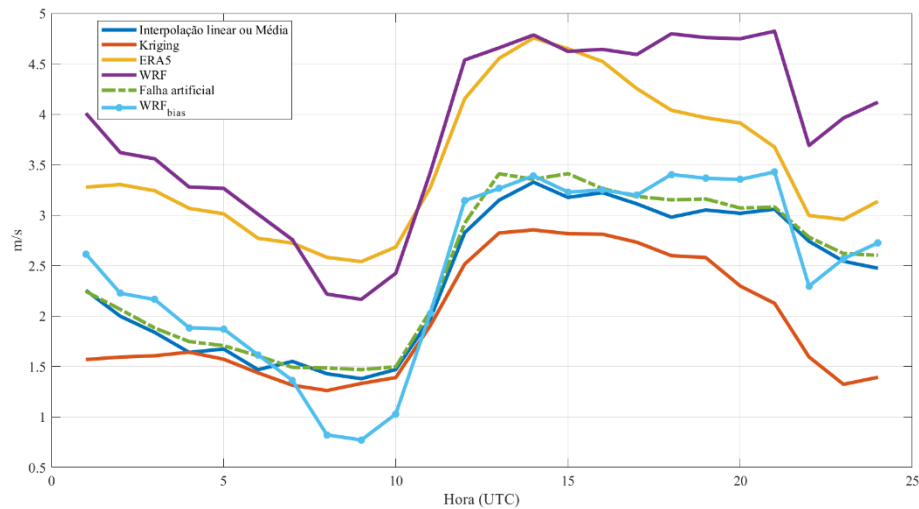


Figura 4.26 - Média dos valores removidos artificialmente (linha tracejada) e estimados por 4 diferentes métodos para (a) temperatura do ar e (b) velocidade do vento para o cluster-6.

a)



b)



Fonte: o autor.

Tabela 4.12 - Métricas de verificação do desempenho do preenchimento de falhas geradas artificialmente para o cluster-6.

Cluster-6	Bias		MAE		r		RMSE		POD	
	T <sup>1</sup>	V <sup>2</sup>	T <sup>1</sup>	V <sup>2</sup>	T	V	T <sup>1</sup>	V <sup>2</sup>	T <sup>1</sup>	V <sup>2</sup>
Interpolação	-0,1	-0,1	1,0	0,8	0,93	0,70	1,4	1,1	60,2	43,9
Kriging	-1,1	-0,5	1,6	1,2	0,89	0,42	2,1	1,5	36,5	26,3
ERA5	-0,1	1,0	1,3	1,4	0,90	0,55	1,7	1,8	49,0	23,2
WRF	-0,3	1,4	1,5	1,7	0,87	0,53	1,9	2,2	40,0	17,4
WRF-Bias <sup>3</sup>	0,0	0,0	1,4	1,3	-	-	1,9	1,7	39,8	26,9

<sup>1</sup>Temperatura do ar em °C

<sup>2</sup>Velocidade do vento em m/s

<sup>3</sup>Índices após a remoção do Bias

A última avaliação de desempenho dos métodos utilizados para o preenchimento dos buracos artificiais considerou o índice de probabilidade de detecção cuja porcentagem indicou o sucesso dos métodos em estimar os dados removidos artificialmente das séries temporais.

De acordo com o observado na figura 4.27a o método de interpolação apresentou a melhor performance na detecção dos dados removidos para todos os tamanhos consecutivos de falhas, com exceção para buracos de 24 horas para as quais a reanálise ERA5 demonstrou ter o maior desempenho entre todos os métodos analisados. A superioridade do índice POD para a temperatura do ar também é indicada pela média das probabilidades apresentada na tabela 4.12.

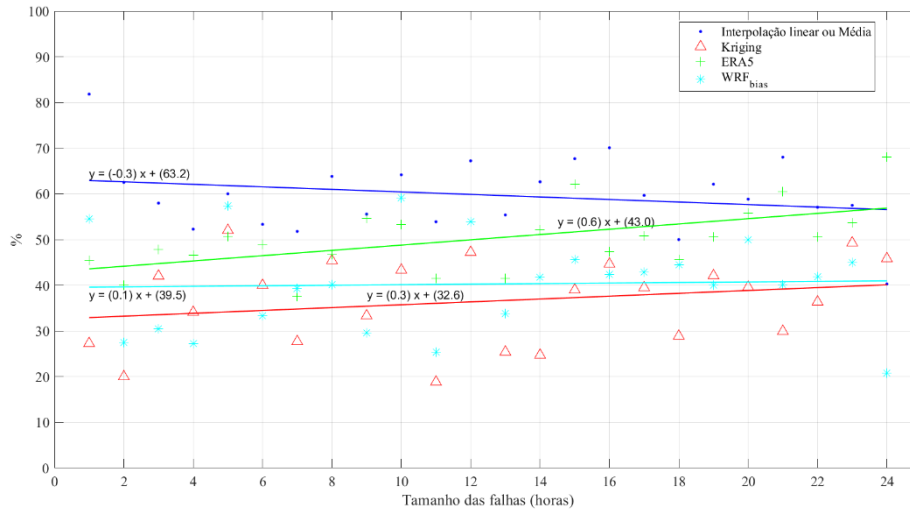
Ainda sobre a temperatura do ar, as retas de regressão presentes na figura 4.27 indicaram que apenas o método de interpolação apresentou tendência negativa para índice POD. Ou seja, a probabilidade de acerto das estimativas dos métodos kriging, ERA5 e WRF aumentaram de acordo com o aumento do tamanho de falhas consecutivas. No entanto, os coeficientes angulares das retas de regressão indicaram que esse aumento foi mínimo.

Para a velocidade do vento, a probabilidade de acerto da interpolação foi maior que os demais métodos em todos os tamanhos de falhas consecutivas (figura 4.27b). Este fato também foi evidenciado pela média do POD apresentada na tabela 4.12. Adicionalmente, a análise da inclinação das retas de regressão demonstrou que a tendência é quase nula para todos os métodos.

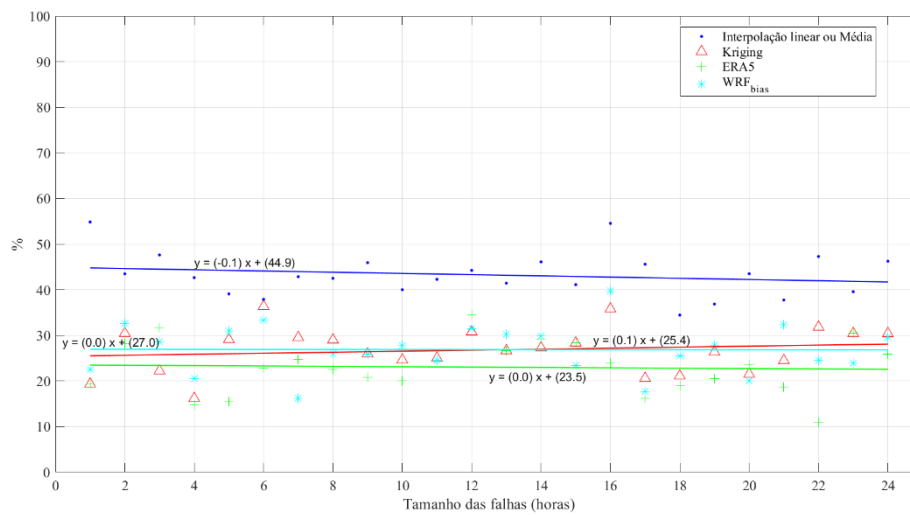
Por fim, após ter sido avaliado o desempenho dos métodos para preencher as falhas artificiais por meio dos índices estatísticos apresentados na tabela 4.12 e do índice de probabilidade de detecção para distintos tamanhos de buracos consecutivos realizou-se o preenchimento das falhas de 1 até 24 horas das séries temporais brutas do cluster-6 utilizando a interpolação linear para falhas de 1 hora ou a média entre o valor medido 24 horas antes e 24 horas depois da falha.

Figura 4.27 - Diagrama de dispersão e reta de regressão da Probabilidade de Detecção média dos valores médios removidos artificialmente das séries temporais do cluster-6 para (a) temperatura do ar e (b) velocidade do vento considerando diferentes tamanhos de falhas.

a)



b)



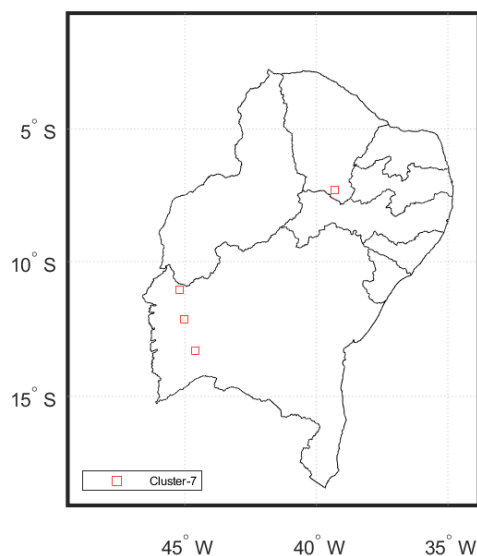
Fonte: o autor.

#### 4.1.7 Cluster 7

O último agrupamento constituído pelo processo de análise de cluster foi formado por três estações meteorológicas instaladas na Bahia e uma no Ceará (figura 4.28). A estação cearense localiza-se na cidade de Barbalha, região Sul do Estado e ao sopé da Chapada do Araripe. A vegetação é formada pelo cerrado, caatinga e resquícios de mata-atlântica e é essa diversidade que compõe a primeira Floresta Nacional do país.

As três estações baianas, Formosa do Rio Preto, Barreiras e Correntina, estão situadas no extremo oeste baiano e a uma altitude acima dos 470 metros e cuja vegetação predominante é o cerrado. O cluster-7 possui a menor mediana das máximas temperatura do ar ( $31,9\text{ }^{\circ}\text{C}$ ) quando comparada aos demais agrupamentos assim como, ao lado do cluster-5, a menor mediana das mínimas velocidades diárias do vento ( $0,3\text{ m/s}$ ) (tabela 4.13)

Figura 4.28 - Localização das estações meteorológicas que formaram o cluster-7.



Fonte: o autor.

Tabela 4.13 – Estações meteorológicas agrupadas no cluster-7 e a mediana das máximas temperaturas diárias do ar ( $T_{med}$ ) e das mínimas velocidades diárias do vento ( $V_{med}$ ).

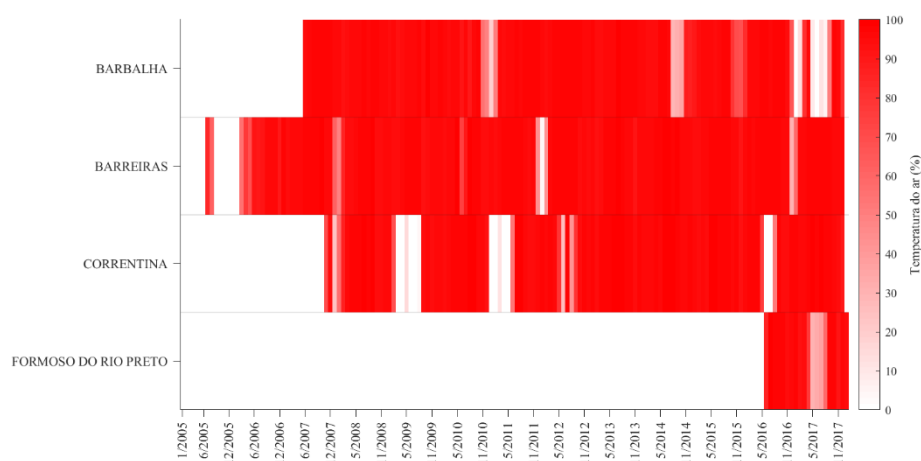
Latitude ( $^{\circ}$ )	Longitude ( $^{\circ}$ )	Altitude (m)	$T_{med}$ ( $^{\circ}\text{C}$ )	$V_{med}$ (m/s)	Nome da estação	UF
-7,3	-39,3	409,0	31,8	0,4	Barbalha	CE
-12,1	-45,0	470,3	32,0	0,2	Barreiras	BA
-13,3	-44,6	549,0	31,6	0,4	Correntina	BA
-11,0	-45,2	488,0	32,4	0,3	Formosa do Rio Preto	BA
Mediana			31,9	0,3		

A maior série temporal do cluster-7 está disponível para a estação Barreiras com aproximadamente 13 anos de dados medidos (06/2005 a 03/2018). No entanto, as séries temporais de temperatura do ar e velocidade do vento (figura 4.29) apresentaram períodos de até sete meses sem dados medidos. Por exemplo de 08/2005 a 03/2006 não houve dados medidos para a temperatura do ar da estação Barreiras e 03/2008 a 08/2009 para a temperatura do ar e velocidade do vento da estação Correntina. A estação cearense também permaneceu 2 meses sem coleta de dados de temperatura do ar (04/2017 a 06/2017).

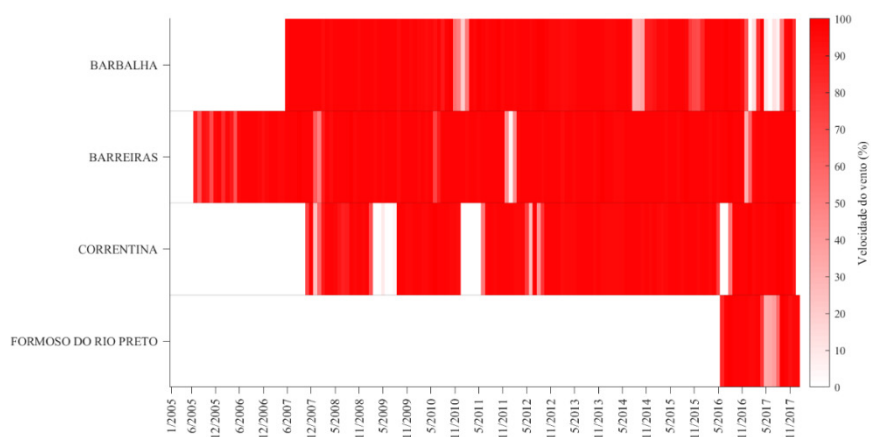
A menor série temporal disponível para a análise do cluster-7 é a da estação Formoso do Rio Preto com aproximadamente 20 meses de coleta de dados.

Figura 4.29 - Séries temporais brutas da (a) temperatura do ar e (b) velocidade do vento das estação do INMET pertencente ao cluster-7.

a)



b)



Fonte: o autor.

O desempenho do preenchimento de falhas geradas artificialmente a partir dos métodos de imputação de dados para o cluster-7 é apresentado na figura 4.30. Após remover 10% dos dados brutos das séries temporais observou-se que todos os métodos capturaram a variabilidade da temperatura do ar ao longo do ciclo diurno (figura 4.30a), embora a reanálise ERA5 tenha apresentado o mesmo comportamento já discutido nos clusters anteriores o qual demonstra grandes saltos entre 10 e 21 UTC para elevar a temperatura do ar.

O bom desempenho em estimar esta variação da temperatura foi confirmado pelos altos valores de correlação linear apresentados na tabela 4.14 (todos igual ou acima de 0,86). Nesta mesma tabela foi possível observar que a interpolação não apresenta erro sistemático, enquanto kriging e reanálise o *bias* foi de 0,4 °C. Por fim, apenas o WRF subestima os dados medidos, porém também com um valor próximo a zero (-0,6 °C).

A respeito das medidas de erro MAE e RMSE (tabela 4.14) foi possível concluir que o método kriging e o ERA5 apresentaram os maiores erros individuais em relação aos dados medidos para a temperatura do ar enquanto os menores foram observados para a interpolação linear ou média. Por este motivo, foi observado a excelente performance da interpolação em estimar as falhas artificiais do ciclo diurno (figura 4.30a).

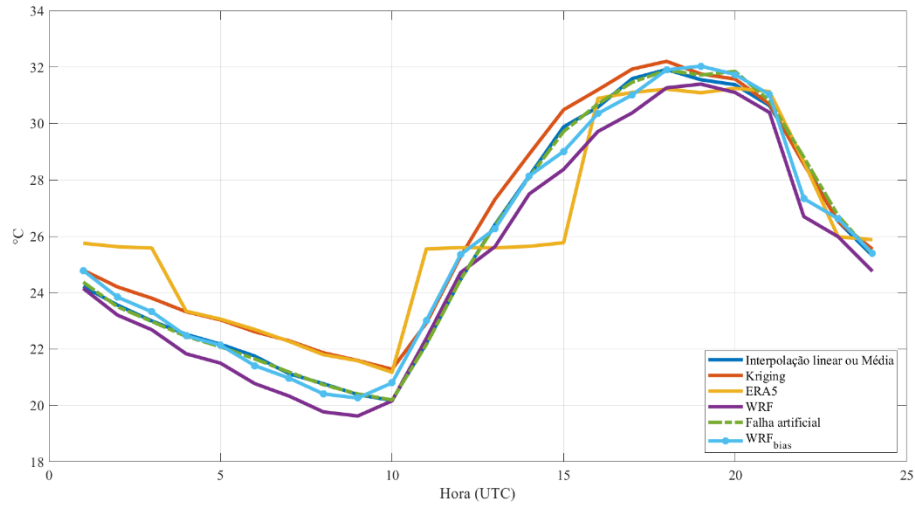
Por fim, a remoção do *bias* do modelo WRF não demonstrou uma diminuição significativa nas medidas de erro MAE e RMSE uma vez que o erro médio não foi constante para sobre todo o ciclo diurno. Para a velocidade do vento (figura 4.30b) apenas o método de interpolação estimou valores em fase com a variação da velocidade do vento registrada ao longo do dia. Esse fato foi evidenciado pela correlação linear apresentada na tabela 4.14 onde a maior correlação foi encontrada para a interpolação (0,70). Para a interpolação também o erro médio calculado foi nulo enquanto o maior *bias* foi encontrado para o WRF (1,3 m/s).

Embora todos os métodos utilizados para estimar as falhas artificiais tenham reproduzido com eficiência a fase de elevação da velocidade do vento ocorrida a partir das 10 UTC e a queda abrupta a partir das 21 UTC, o índice RMSE demonstrou que, com exceção da interpolação, os demais métodos apresentaram erros individuais acima de 1 m/s (tabela 4.14).

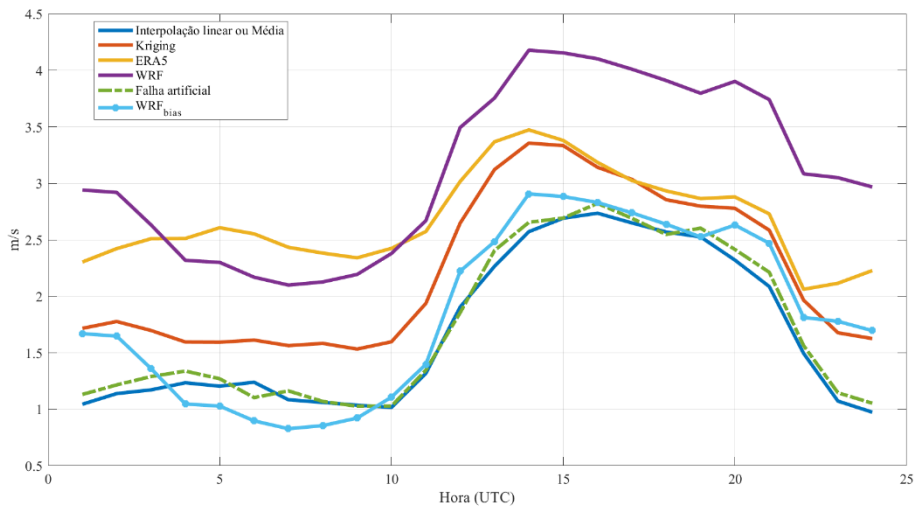
A análise da figura 4.26b também corrobora com as medidas de erro e correlação encontradas, principalmente para os modelos atmosféricos, uma vez que foi possível observar estimativas distantes dos dados medidos. A partir da remoção do *bias* do modelo WRF para a velocidade do vento observou-se estimativas mais próximas aos dados removidos artificialmente (figura 4.30). Com a retirada do erro sistemático constante do modelo regional observou-se valores mais próximos das falhas artificiais sobretudo entre 10 e 19 UTC. Esse fato resultou na redução de 0,5 m/s dos erros MAE e RMSE do WRF (tabela 4.14).

Figura 4.30 - Média dos valores removidos artificialmente (linha tracejada) e estimados por 4 diferentes métodos para (a) temperatura do ar e (b) velocidade do vento para o cluster-7.

a)



b)



Fonte: o autor.

Tabela 4.14 – Métricas de verificação do desempenho do preenchimento de falhas geradas artificialmente para o cluster-7.

Cluster-7	Bias		MAE		r		RMSE		POD	
	T <sup>1</sup>	V <sup>2</sup>	T <sup>1</sup>	V <sup>2</sup>	T	V	T <sup>1</sup>	V <sup>2</sup>	T <sup>1</sup>	V <sup>2</sup>
Interpolação	0,0	0,0	1,3	0,6	0,90	0,70	1,8	0,9	47,1	46,4
Kriging	0,4	0,4	2,0	0,8	0,86	0,61	2,6	1,1	30,5	34,2
ERA5	0,4	0,8	1,9	1,1	0,87	0,54	2,5	1,4	34,0	27,9
WRF	-0,6	1,3	1,8	1,5	0,88	0,55	2,3	1,8	30,3	16,8
WRF-Bias <sup>3</sup>	0,0	0,0	1,7	1,0	-	-	2,3	1,3	33,6	27,7

<sup>1</sup>Temperatura do ar em °C

<sup>2</sup>Velocidade do vento em m/s

<sup>3</sup>Índices após a remoção do *Bias*

Por fim, avaliou-se o preenchimento das falhas artificiais a partir do índice de probabilidade de detecção o qual indicou o sucesso dos métodos em estimar os dados removidos artificialmente das séries temporais. A figura 4.31 apresenta a probabilidade média de detecção (em porcentagem) de cada método para cada tamanho de falha consecutiva gerada artificialmente.

Para a temperatura do ar, o método de interpolação apresentou a maior probabilidade de estimar os dados removidos para todos os tamanhos de falhas com exceção de buracos com 9 horas, os quais foram melhor performado pelo método de Kriging (figura 4.31a). A superioridade do índice POD para a temperatura do ar também é indicada pela média das probabilidade sobre todos os tamanhos de falhas apresentada na tabela 4.14.

As retas de regressão presentes na figura 4.31a indicaram que apenas o método de kriging não apresentou tendência na probabilidade de detecção dos dados removidos artificialmente com o aumento do tamanho de falhas. No entanto, os demais métodos apresentarem tendência ligeiramente negativa.

Para a velocidade do vento a probabilidade de acerto da interpolação foi maior que os demais métodos em todos os tamanhos de falhas consecutivas, exceto para buracos de 2, 6, 8 e 18 horas, para os quais o método de kriging obteve maior sucesso (figura 4.31b). Adicionalmente, a análise da inclinação das retas de regressão demonstrou que a tendência da probabilidade de acerto conforme o aumento do tamanho de falhas é positiva, porém quase nula para o modelo WRF com o *bias* removido. Para os demais métodos a tendência é negativa sendo que a reta de regressão para reanálise ERA5 foi a que apresentou a maior inclinação.

Após a avaliação conjunta das métricas de verificação de desempenho contidas na tabela 4.14 e do índice de probabilidade de detecção das falhas artificiais para distintos tamanhos realizou-se o preenchimento das falhas de 1 até 24 horas das séries temporais originais do cluster-7. Para tanto, o preenchimento para a temperatura do ar foi realizado utilizando o

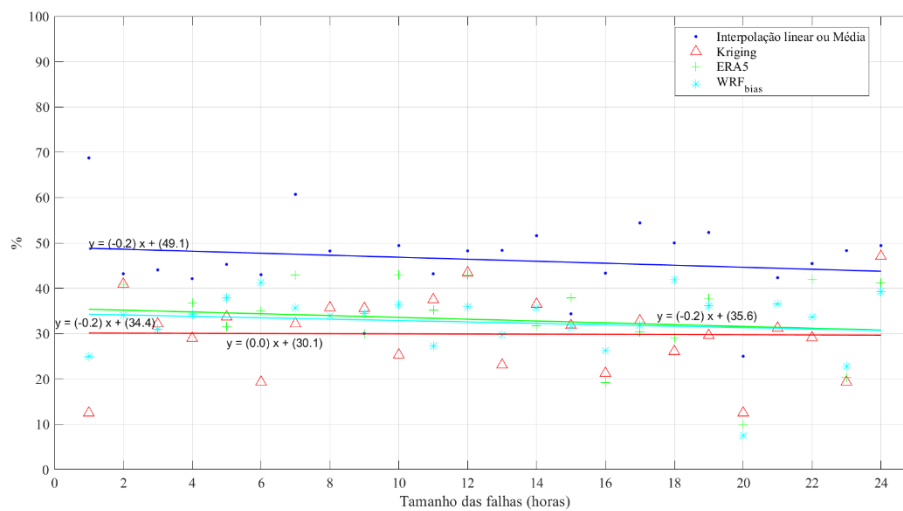


método de Kriging para buracos de 9 horas consecutivas, o da reanálise ERA5 para falhas de 15 horas e o método de interpolação para os demais tamanhos de buracos.

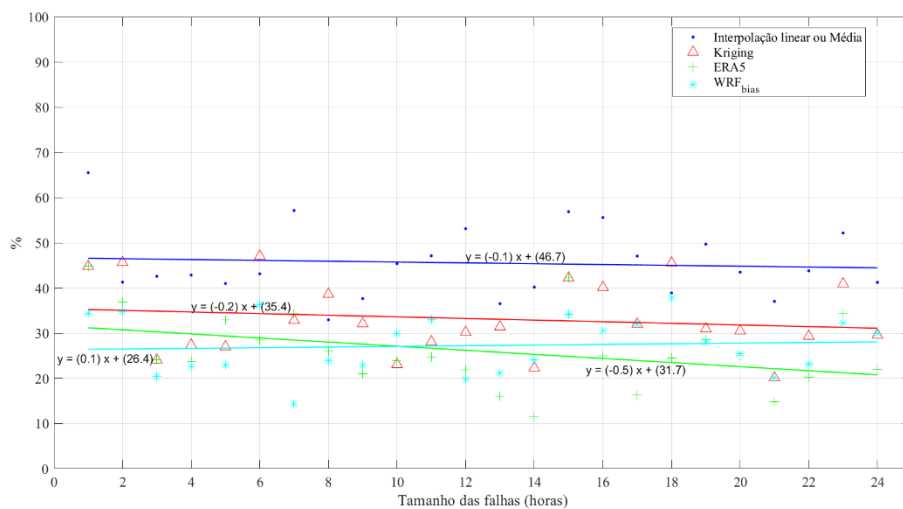
Para a velocidade do vento, as falhas consecutivas com tamanhos de 2, 6, 8 e 18 horas foram preenchidas utilizando o método geostatístico de Kriging. Os demais tamanhos de buracos foram preenchidos com o método de interpolação.

Figura 4.31 – Diagrama de dispersão e reta de regressão da Probabilidade de Detecção média dos valores médios removidos artificialmente das séries temporais do cluster-7 para (a) temperatura do ar e (b) velocidade do vento considerando diferentes tamanhos de falhas.

a)



b)



Fonte: o autor.

## 4.2 CARACTERIZAÇÃO DOS MÍNIMOS DE VENTO E MÁXIMOS DE TEMPERATURA DO AR

Após identificar as estações meteorológicas que indicaram as regiões com condições climatológicas adversas para o aumento de carga em linhas de transmissão e agrupá-las a partir da análise de cluster, caracterizou-se os regimes de velocidade do vento e temperatura do ar em diferentes escalas de tempo (estações do ano e ciclo diurno) para cada cluster.

A maior parte dos agrupamentos formados para realização da caracterização supracitada situam-se na sub-região nordestina denominada Sertão (28 estações). Há também estações localizadas no Meio-Norte (quatro) e no Agreste (três). Estas três sub-regiões pertencem a região de baixas latitudes onde a disponibilidade de radiação solar é intensa e pouco afetada pela variação do ciclo solar. Entretanto, essas sub-regiões sofrem influência de diferentes sistemas atmosféricos que determinam as condições de tempo e clima. Por exemplo, os sistemas convectivos de mesoescala, brisa marítima, terrestre e lacustre, vórtices ciclônicos de altos níveis, distúrbios ondulatórios de leste, sistemas frontais e zona de convergência intertropical modulam o regime de precipitação nessa região (CHAVES; CAVALCANTI, 2001; GOMES et al., 2019) e consequentemente os máximos de temperatura e mínimos de vento.

A análise do método gráfico *boxplot*<sup>1</sup> para a escala sazonal demonstrou que, com exceção do verão e outono do cluster-5, todas as medianas das máximas temperaturas diárias são estatisticamente diferentes entre si ao intervalo de confiança de 95% (figura 4.32). Essa variabilidade observada entre as estações do ano está associada aos máximos de precipitação que ocorrem entre dezembro e maio. E estes, por sua vez, estão associados a atividade convectiva da ZCIT (Zona de Convergência Intertropical) a qual tem sua posição mais ao sul neste período (RAO; DE LIMA; FRANCHITO, 1993).

Assim, o aumento da nebulosidade associada a atividade convectiva diminuiu a temperatura do ar e, consequentemente, as máximas temperaturas diárias em análise neste trabalho. Esse fato é evidenciado na figura 4.32 onde as menores temperaturas máximas ocorreram no período chuvoso (verão e outono) enquanto as máximas no período seco (inverno

---

<sup>1</sup> A linha horizontal tracejada representa o valor da temperatura do ar (31,5 °C) definida pela NBR 5422 para ser utilizada nos projetos de linhas de transmissão quando há ausência de dados medidos. A caixa (em azul) representa o intervalo interquartil (25% - 75%) e a linha vermelha dentro da caixa a mediana. As linhas verticais tracejadas (preta) se estendem do primeiro quartil até o valor mínimo dos dados e do terceiro quartil até o valor máximo. Ambos definidos como 1,5 vezes o intervalo interquartil da parte superior e inferior da caixa. Os valores discrepantes são apresentados com um sinal de positivo (vermelho). Os entalhes exibem a variabilidade da mediana entre as escalas de tempo e são utilizados para indicar, ao nível de significância de 5%, se as medianas são estatisticamente diferentes ou iguais de acordo com a sobreposição.

e primavera). Entretanto, apenas para o cluster-4 é possível observar que a menor mediana da temperatura máxima foi registrada no inverno (figura 4.32d). Isto porque, uma parte das estações que compõem este agrupamento está localizada no leste do Nordeste. E nesta região nordestina a estação chuvosa estende-se até agosto devido a interação com a ZCIT dos distúrbios de onda de leste, os quais são uma banda convectiva iniciada com o aumento da convecção no lado oeste da África Central e que se propaga para leste (GOMES et al., 2019).

Segundo COELHO; GAN; CONFORTE (2004) e SANTOS E SILVA et al. (2014) outros sistemas atmosféricos como frentes frias no sul do Nordeste e vórtices ciclônicos de altos níveis também contribuem para a variabilidade sazonal das máximas temperaturas. Isso porque, ambos sistemas supracitados também auxiliam na formação da convecção tropical (CHAVES; CAVALCANTI, 2001).

A respeito do limiar da temperatura máxima diária ( $31,5\text{ }^{\circ}\text{C}$ ) estabelecido na presente tese, observou-se na figura 4.32 que, independente da estação do ano, todas as medianas estiveram acima deste valor (linha horizontal tracejada). A única exceção foi para o cluster-4 ( $30,9\text{ }^{\circ}\text{C}$ ) no inverno (figura 4.32d) e cluster-7 ( $31,4\text{ }^{\circ}\text{C}$ ) no outono (figura 4.32g). Adicionalmente, ao excetuar-se os *outliers* da escala sazonal o menor limite inferior foi registrado no inverno ( $23,8\text{ }^{\circ}\text{C}$ ) para o cluster-4 (figura 4.32d) enquanto o maior na primavera ( $42,7\text{ }^{\circ}\text{C}$ ) para o cluster-2 (figura 4.32b).

Na escala diurna a evolução da temperatura do ar próxima a superfície é modulada pela radiação solar (DU et al., 2018). Isso porque o ciclo diurno da temperatura é influenciado pelo balanço de energia que envolve a radiação de onda curta e longa na superfície (DUTTON, 1990). Em outras palavras, nas primeiras horas do dia quando o sol está acima da linha do horizonte a superfície começa a ser aquecida devido a irradiância solar. Uma parte da energia absorvida pela superfície é emitida novamente para a atmosfera em forma de radiação de onda longa. E assim, o ar próximo a superfície é aquecido (ZHOU; WANG; MA, 2017). Esse saldo de radiação que aquece a atmosfera próxima à superfície é determinado pelo fluxo de calor sensível e latente. E, esses por sua vez, são diretamente influenciados pela cobertura vegetal e umidade disponível no solo (WATTERSON, 1997).

Desta forma, os biomas caatinga e cerrado, onde estão localizadas as estações meteorológicas que compõem os sete agrupamentos, se constituem em fatores fundamentais para realizar a alteração das propriedades radiativas da superfície como a emissividade e o albedo, as hídricas e térmicas do substrato e solo. E, por conseguinte, são determinantes para a variação da temperatura do ar próximo a superfície ao longo de um dia (FERREIRA; DE OLIVEIRA; SOARES, 2013).

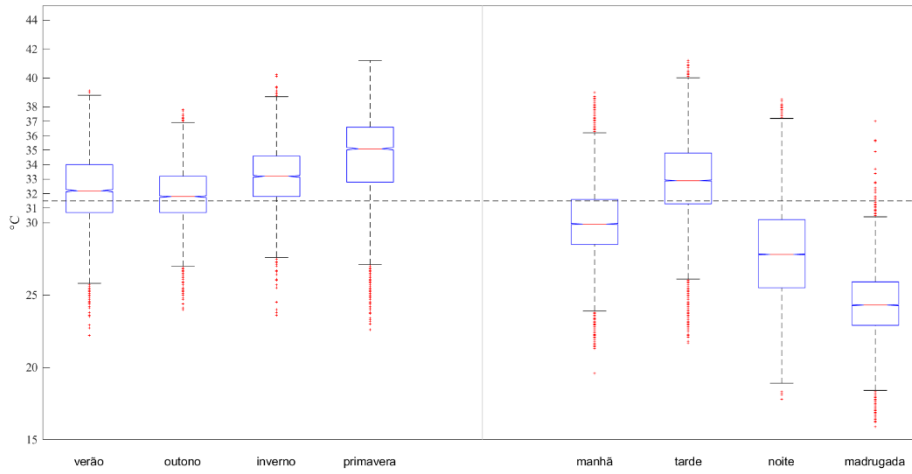
A partir da figura 4.32 verificou-se também que, com intervalo de confiança de 95%, no ciclo diurno as medianas das temperaturas máximas foram diferentes entre si em todos os clusters. Ainda sobre esta figura, para todos os agrupamentos no período de maior resfriamento da superfície (madrugada entre 4 as 9 UTC) registrou-se as menores medianas das temperaturas máximas devido à ausência de radiação solar no espectro de onda curta. Por outro lado, as maiores medianas das máximas temperaturas foram registradas no período em que o ângulo de elevação solar estava próximo ao máximo e, por conseguinte a irradiância ou o saldo de radiação na superfície era maior (tarde, entre 16 as 21 UTC).

No ciclo diurno as diferenças entre as maiores e menores medianas registradas no período da tarde e da madrugada, respectivamente, são em média na ordem de 9,2 °C. Por exemplo, o cluster-5 (figura 4.32e) formado por uma única estação meteorológica localizada no Piauí possui uma diferença de 10,3 °C entre o período de maior aquecimento e menor resfriamento da atmosfera próxima a superfície. A análise dos *boxplot* demonstrou também que, excetuando-se os *outliers*, o maior valor da temperatura máxima diária do ar encontrado dentro das séries temporais foi no período da tarde (41,7 °C) nos cluster-2 (figura 4.32b) e cluster-5 (figura 4.32e). Para este último, também foi registrado o maior valor da temperatura máxima para o período da madrugada (30,8 °C).

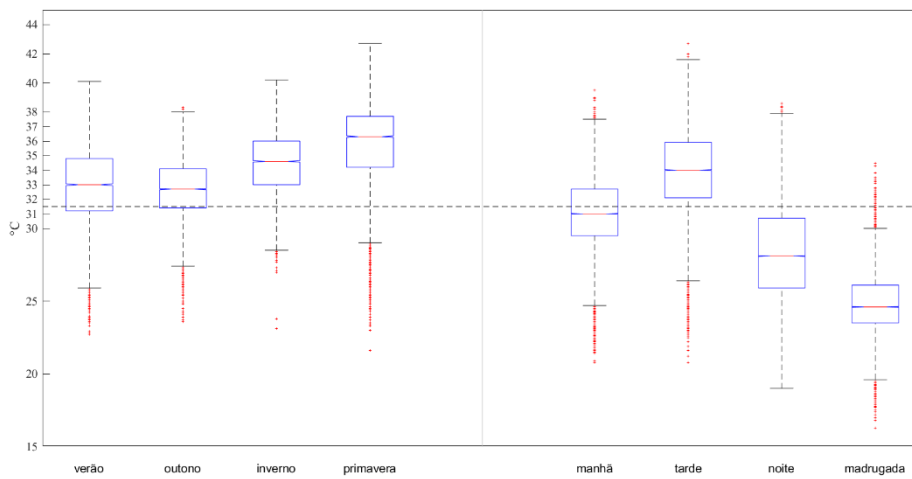
A respeito do limiar de temperatura do ar (31,5 °C) definido na presente tese observou-se que em todos os agrupamentos este valor foi ultrapassado apenas no período da tarde (figura 4.32). A maior mediana das temperaturas máximas diárias foi observada no cluster-5 (figura 4.32e) e cluster-6 (figura 4.32f) durante o período da tarde e cujo valor foi de 34,7 °C. Para esse período do dia a menor mediana dentre todos os agrupamentos (32,1 °C) foi registrada no cluster-7 (figura 4.32g) assim como a menor (22,7 °C) para o período da madrugada. Esse fato se deve a altitude (acima de 400 metros) e ao bioma (cerrado) onde estão localizadas as quatro estações que compõe o cluster-7. A altitude e a vegetação do cerrado na escala diurna afetam o saldo de radiação na superfície e conseqüentemente têm-se temperaturas máximas diárias do ar menores do que aquelas registradas sobre a caatinga.

Figura 4.32 – Distribuição de frequência das temperaturas máximas diárias na escala de tempo sazonal e diurna para o (a) cluster-1, (b) cluster-2, (c) cluster-3, (d) cluster-4, (e) cluster-5, (f) cluster-6 e (g) cluster-7.

a)



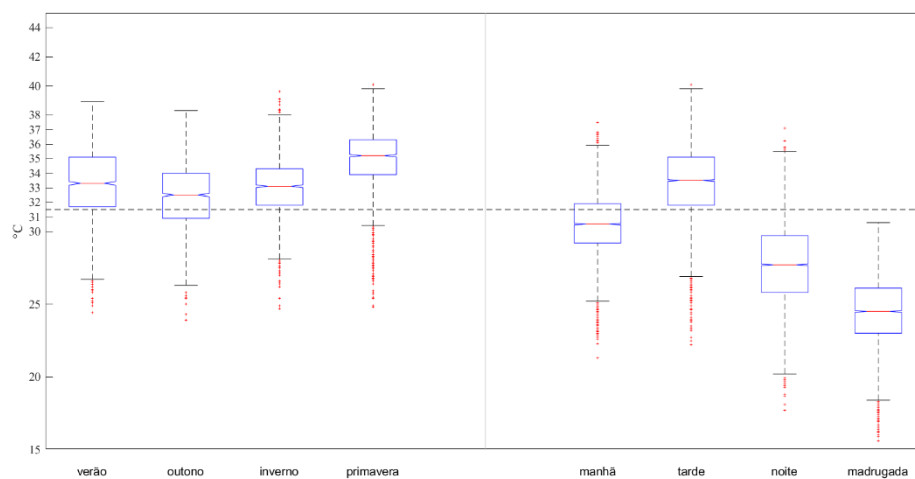
b)



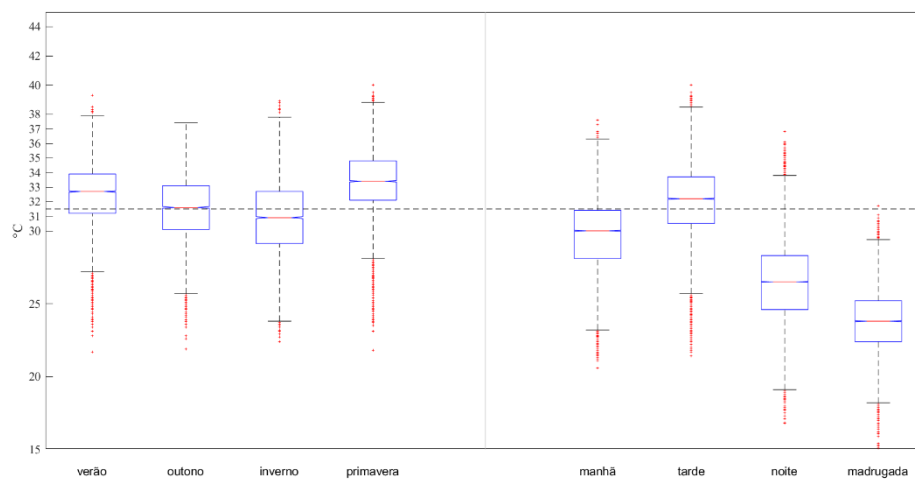
Fonte: o autor.

Continuação – Figura 4.30– Distribuição de frequência das temperaturas máximas diárias na escala de tempo sazonal e diurna para o (a) cluster-1, (b) cluster-2, (c) cluster-3, (d) cluster-4, (e) cluster-5, (f) cluster-6 e (g) cluster-7.

c)

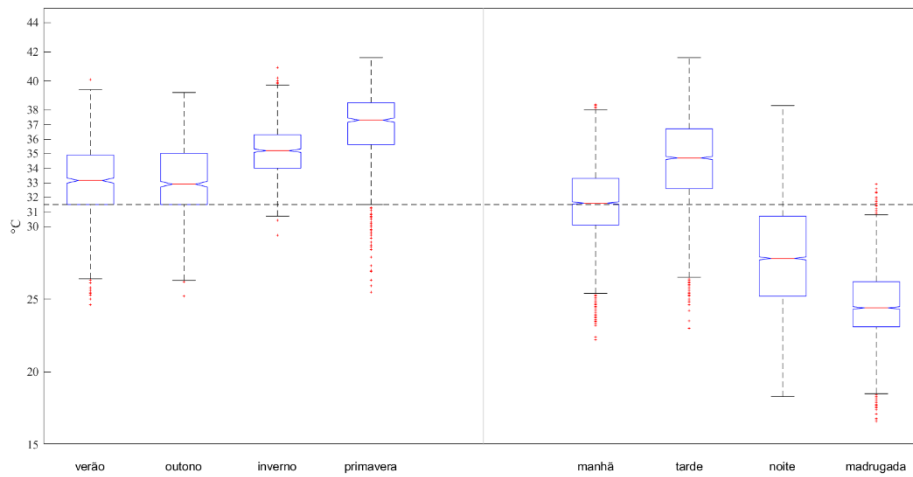


d)

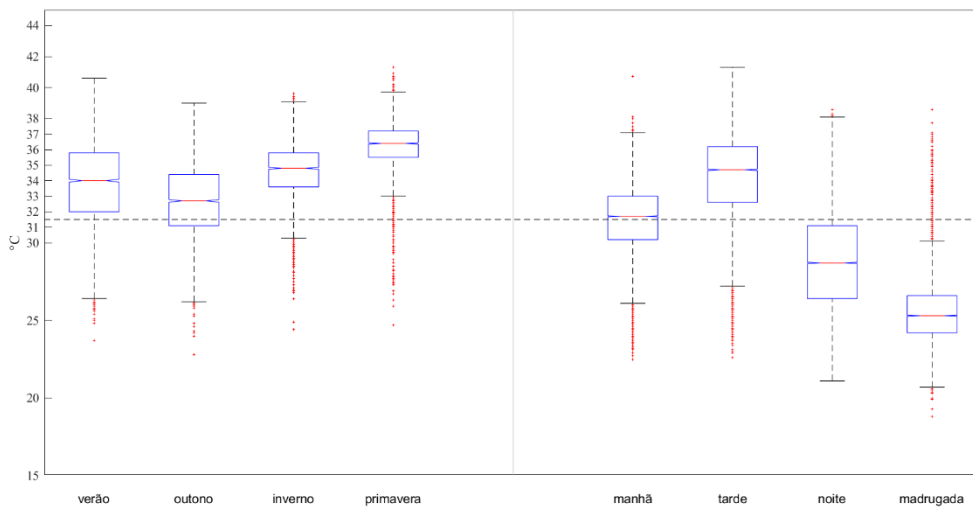


Continuação – Figura 4.30– Distribuição de frequência das temperaturas máximas diárias na escala de tempo sazonal e diurna para o (a) cluster-1, (b) cluster-2, (c) cluster-3, (d) cluster-4, (e) cluster-5, (f) cluster-6 e (g) cluster-7.

e)

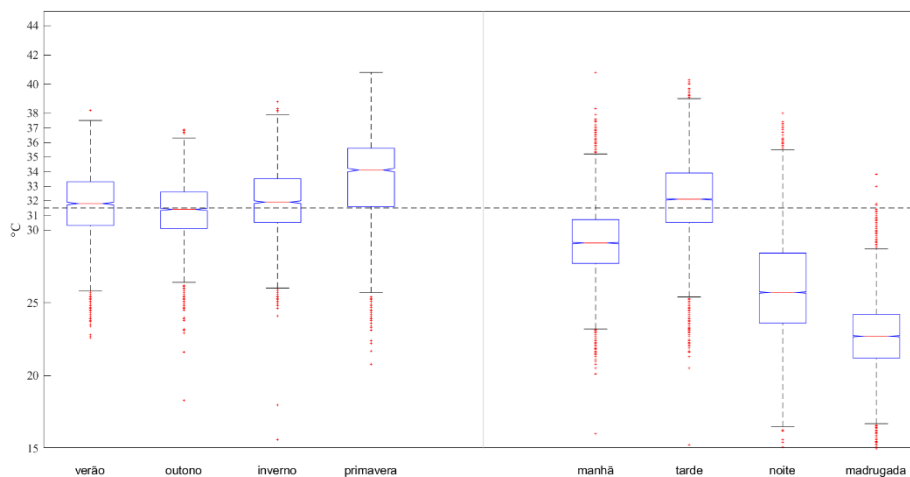


f)



Continuação -Figura 4.30– Distribuição de frequência das temperaturas máximas diárias na escala de tempo sazonal e diurna para o (a) cluster-1, (b) cluster-2, (c) cluster-3, (d) cluster-4, (e) cluster-5, (f) cluster-6 e (g) cluster-7.

g)



Fonte: o autor.

A análise dos *boxplot*<sup>2</sup> da figura 4.33 revelou que, com intervalo de confiança de 95%, as medianas das mínimas velocidades diárias do vento são iguais entre si para o verão e o outono em todos os agrupamentos com exceção do cluster-4 (figura 4.33d) e cluster-6 (figura 4.33f). Ao nível de significância de 5%, há diferença estatística entre as medianas das estações inverno e primavera de todos agrupamentos com exceção dos cluster-2 (figura 4.33b), cluster-3 (figura 4.33c) e cluster-5 (figura 4.33e).

A respeito do limiar de 1 m/s estabelecido pela NBR 5422 para ser utilizado em projetos de construção de linhas de transmissão na falta de dados medidos observou-se que, na escala sazonal e em todos os agrupamentos, a mediana das mínimas velocidades diárias do vento situaram-se sempre abaixo de 1 m/s (figura 4.33).

Essa análise demonstrou também que, para a mesma escala temporal, em estações do ano mais quentes a velocidade do vento é maior. Ou seja, a variação das mínimas velocidades

<sup>2</sup> A linha horizontal tracejada representa o valor da velocidade do vento (1 m/s) definida pela NBR 5422 para ser utilizada nos projetos de linhas de transmissão quando há ausência de dados medidos. A caixa (em azul) representa o intervalo interquartil (25% - 75%) e a linha vermelha dentro da caixa a mediana. As linhas verticais tracejadas (preta) se estendem do primeiro quartil até o valor mínimo dos dados e do terceiro quartil até o valor máximo. Ambos definidos como 1,5 o intervalo interquartil da parte superior e inferior da caixa. Os valores discrepantes são apresentados com um sinal de positivo (vermelho). Os entalhes exibem a variabilidade da mediana entre as escalas de tempo e são utilizados para indicar, ao nível de significância de 5%, se as medianas são estatisticamente diferentes ou iguais de acordo com a sobreposição.



diárias do vento ao longo das estações do ano é influenciada indiretamente pelo período chuvoso que normalmente ocorre entre o verão e o outono e o período seco entre o inverno e a primavera.

Em outras palavras, a atividade convectiva durante o período de maior precipitação diminui o gradiente de temperatura e como consequência o gradiente de pressão horizontal responsável pelo deslocamento de certa quantidade de massa de ar (vento) torna-se menor (LIRA et al., 2017). Por este motivo, observou-se em cinco dos sete clusters que as menores medianas das mínimas velocidades do vento ocorreram no verão e outono enquanto as maiores no inverno e na primavera (figura 4.33).

A exceção ocorreu para o cluster-3 (figura 4.33c) e cluster-7 (figura 4.33g) cuja sobreposição dos entalhes presentes nos *boxplot* demonstrou que não houve diferença estatística entre as medianas. Em outras palavras, não houve variabilidade sazonal das mínimas velocidades do vento para estes agrupamentos entre as estações. Ainda sobre o cluster-3, verificou-se a partir da figura 4.33c que a mediana para todas as estações do ano foi igual a 0 m/s e que não houve variabilidade das mínimas velocidades do vento durante o verão e o inverno.

A análise dos *boxplot* para todos os agrupamentos revelou também que, em média, a máxima diferença entre as medianas das mínimas velocidades na escala sazonal foi de 0,2 m/s. Ou seja, a variabilidade dos mínimos de vento entre uma estação do ano e outra é baixa. Adicionalmente, nesta mesma escala, observou-se que a maior variabilidade da velocidade mínima diária do vento ocorreu sempre no inverno. Para as demais estações do ano a caixa achatada indicou um menor desvio padrão entre os dados, por conseguinte, ventos calmos constantes na escala intra-sazonal. Em outras palavras, isso significa que durante a estação chuvosa as mínimas velocidades diárias possuem baixa variância em torno da mediana enquanto as maiores dispersões dos dados medidos encontram-se no inverno independentemente do agrupamento.

Ainda a respeito da variabilidade sazonal, apenas os *boxplot* dos clusters-4 (figura 4.33d) e cluster-6 (figura 4.33f) apresentaram dados com maior variância das mínimas velocidades diárias em torno da mediana em todas as estações do ano. Também foi nesses dois agrupamentos que as maiores medianas foram observadas na primavera ao contrário dos demais clusters cujo fato ocorreu no inverno. Entretanto, ressalta-se que essa diferença foi mínima: 0,3 m/s entre o inverno e a primavera do cluster-4 e 0,2 m/s do cluster-6.

Na análise do ciclo diurno das mínimas velocidades diárias do vento (figura 4.33) observou-se em todos os agrupamentos que as circulações locais induzidas termicamente

modulam esta variável. As circulações térmicas iniciam-se durante o período da manhã com o aquecimento da atmosfera próxima a superfície devido a radiação solar incidente. A partir deste aquecimento inicia-se a convecção térmica com a presença também da turbulência térmica e mecânica. Esta última devido a rugosidade aerodinâmica da superfície (PERRY; SCHOFIELD; JOUBERT, 1969). No período da tarde, onde tem-se o máximo aquecimento do ar adjacente a superfície, ocorre a intensificação do gradiente de pressão e, por conseguinte há a intensificação dos ventos próximos a superfície. Por outro lado, no período da noite os ventos enfraquecem devido ao esfriamento da atmosfera determinado pelo saldo de radiação. Durante a noite a única fonte de irradiação de calor para atmosfera próxima a superfície é devida a radiação de onda longa. Por fim, o máximo resfriamento do ar próximo a superfície ocorre durante a madrugada quando a radiação de onda longa é mínima.

Portanto, a variação das mínimas velocidades diárias do vento observadas na figura 4.33 é modulada na escala diurna pelo balanço de radiação na superfície. Assim, foi possível verificar para todos os clusters que as mínimas velocidades ocorreram sempre no período de maior resfriamento da atmosfera (madrugada) enquanto as maiores mínimas no período da tarde quando a atmosfera próxima a superfície estava mais aquecida. As medianas do período da manhã e da noite se posicionaram sempre entre as medianas da madrugada e tarde de tal modo a evidenciar a transição entre os mínimos de vento.

Outra circulação local e termicamente induzida que pode modular o ciclo diurno da velocidade do vento é brisa marítima e terrestre (FREITAS et al., 2007; OLIVEIRA; COSTA, 2011). Esta circulação é induzida pelo aquecimento diferencial que ocorre entre as superfícies heterogêneas oceano e continente. Neste trabalho, a influência mais pronunciada da brisa marítima e terrestre foi verificada sobre o cluster-4 (figura 4.33d) devido parte das estações meteorológicas que formaram este agrupamento situarem-se na costa leste do Nordeste (figura 3.8).

Adicionalmente, algumas estações do cluster-4 também sofrem influência da brisa lacustre. Esta circulação é semelhante a brisa marítima e terrestre, porém o aquecimento diferencial entre as superfícies heterogêneas ocorre devido a presença de grandes corpos d'água como rios e lagos (DE SOUSA MELO; DA SILVA ARAGÃO; CORREIA, 2014). Neste sentido, pode-se citar a presença do Açude do Saco próximo à estação Serra Talhada cujo maior comprimento e largura entre margens é de 4,5 km e 1,5 km, respectivamente. E a do Açude Poço da Cruz – maior reservatório de água de Pernambuco com uma bacia hidrográfica de 4.700 km<sup>2</sup>, localizado próximo à estação meteorológica Ibimirim.

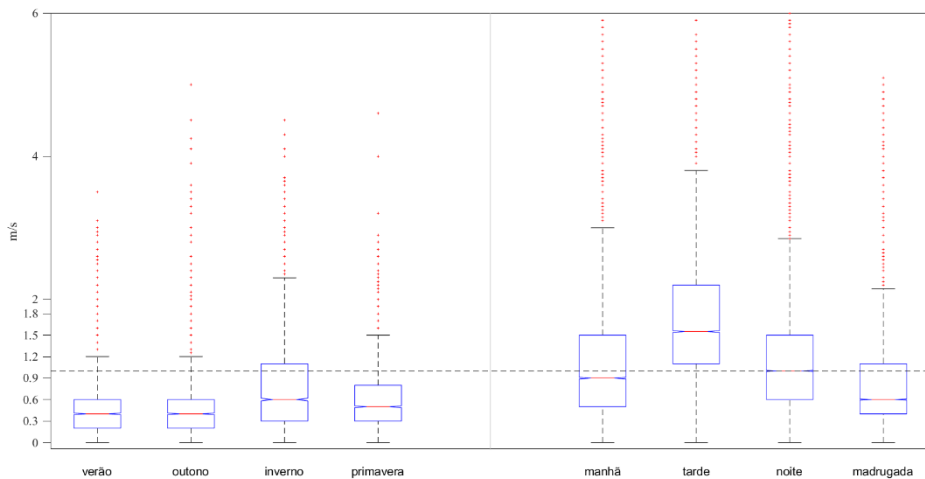
A intensidade das circulações de brisa é modulada principalmente pela radiação solar. Ou seja, a maior intensidade da brisa marítima ou lacustre ocorre durante o período da tarde, pois é quando há o máximo aquecimento diferencial entre as superfícies. Por outro lado, durante o período da noite e da madrugada estas circulações enfraquecem e mudam de sentido devido a inversão do gradiente de pressão. E então passam a ser denominadas de brisa terrestre.

Portanto, a interação entre as circulações térmicas e as de brisa elevaram as medianas das mínimas de vento do cluster-4 (Figura 4.33d) quando comparado aos demais agrupamentos. Esta comparação evidenciou também que em todos os períodos do dia, com exceção da madrugada, apenas as medianas do cluster-4 superaram o limiar de 1 m/s.

Por fim, observou-se ainda que para o ciclo diurno as maiores medianas dos mínimos de vento assim como a maior variabilidade e desvio padrão do conjunto de dados ocorreram para o cluster-4 quando comparado aos demais agrupamentos.

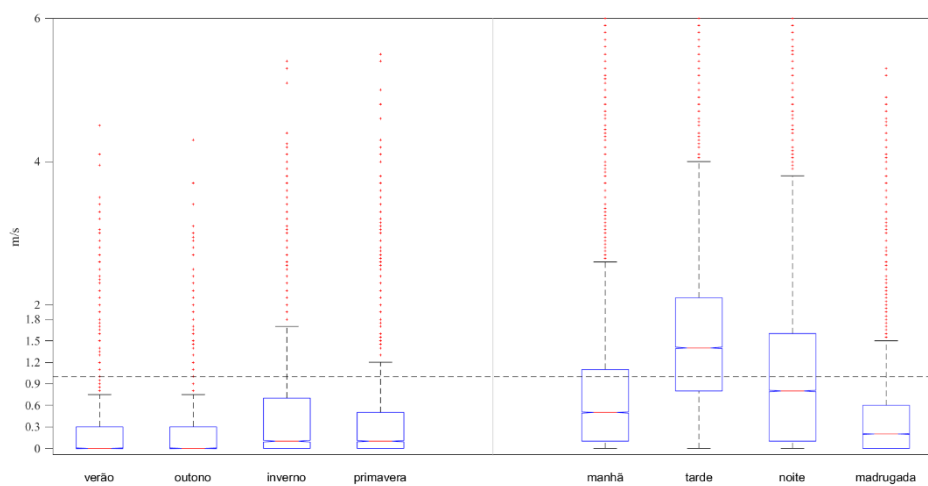
Figura 4.33 – Distribuição de frequência das mínimas velocidades diárias do vento na escala de tempo sazonal e diurna para o (a) cluster-1, (b) cluster-2, (c) cluster-3, (d) cluster-4, (e) cluster-5, (f) cluster-6 e (g) cluster-7.

a)

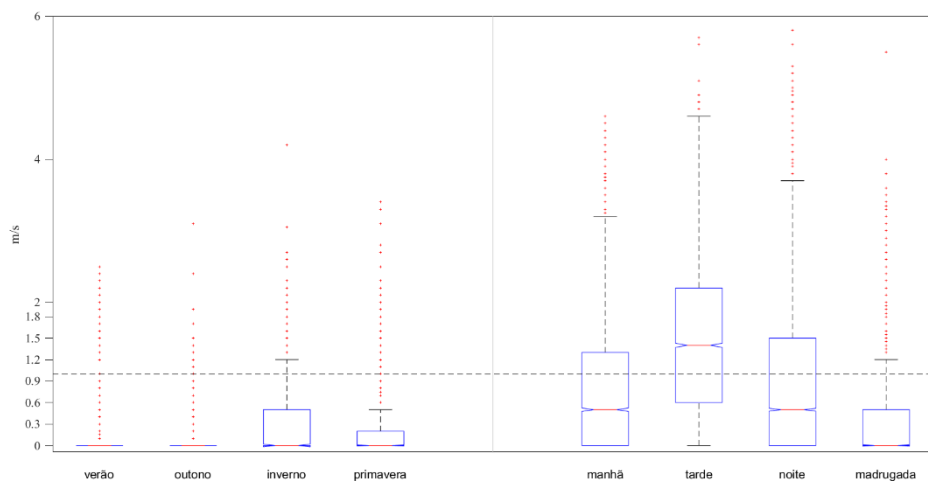


Continuação – Figura 4.31 – Distribuição de frequência das mínimas velocidades diárias do vento na escala de tempo sazonal e diurna para o (a) cluster-1, (b) cluster-2, (c) cluster-3, (d) cluster-4, (e) cluster-5, (f) cluster-6 e (g) cluster-7.

b)

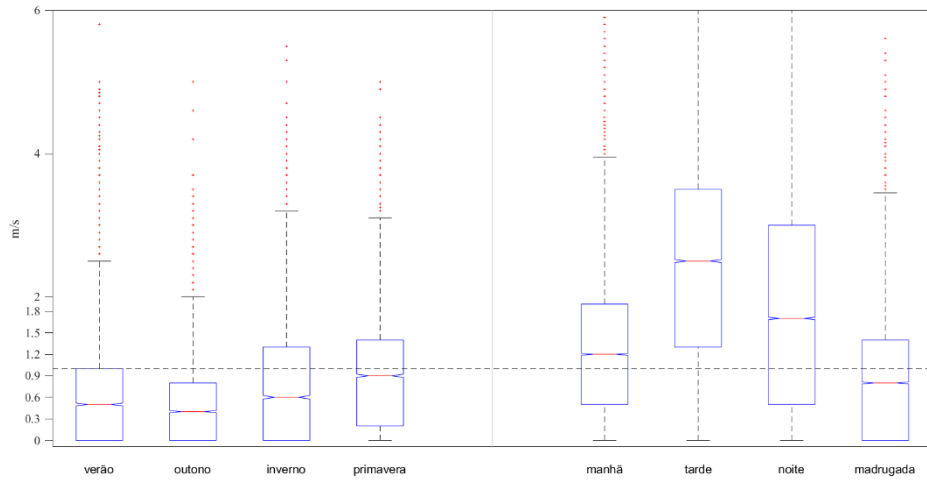


c)

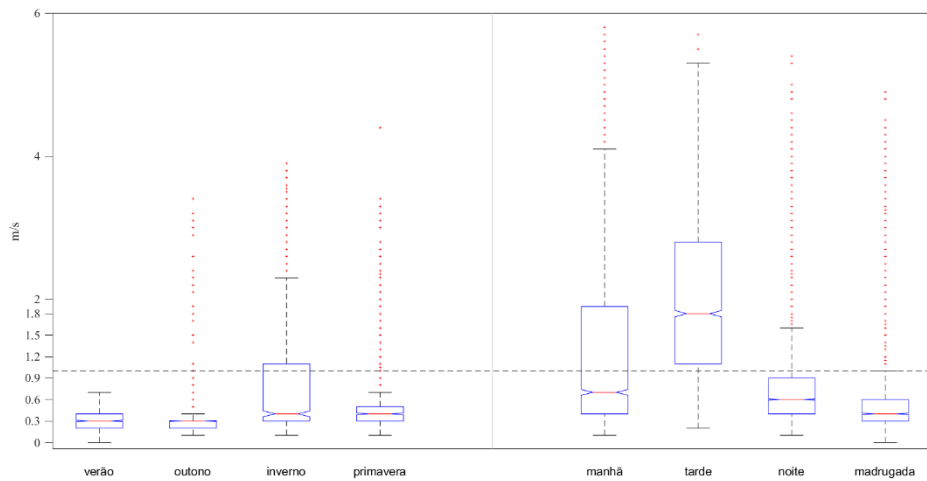


Continuação – Figura 4.31 – Distribuição de frequência das mínimas velocidades diárias do vento na escala de tempo sazonal e diurna para o (a) cluster-1, (b) cluster-2, (c) cluster-3, (d) cluster-4, (e) cluster-5, (f) cluster-6 e (g) cluster-7.

d)

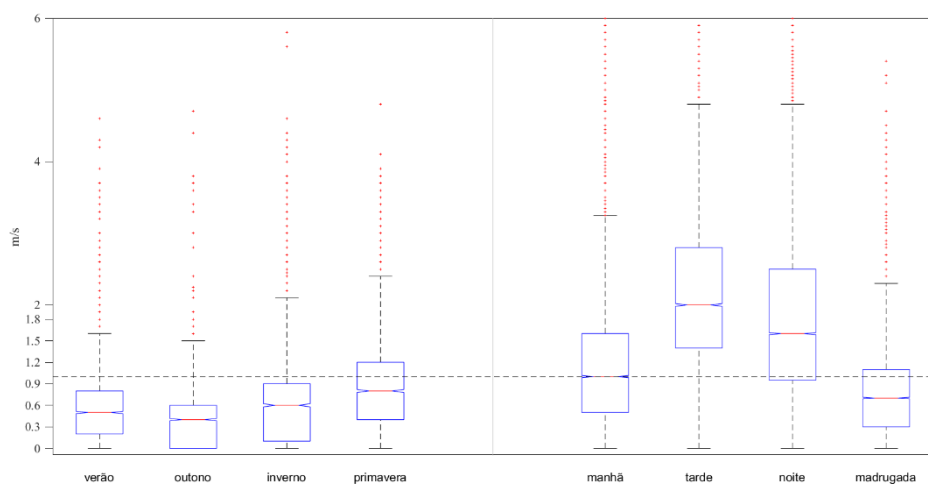


e)

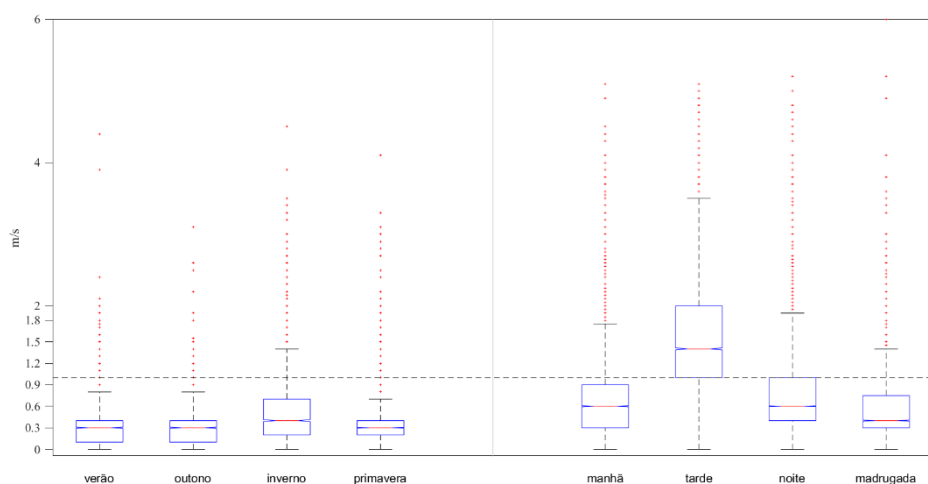


Continuação – Figura 4.31 – Distribuição de frequência das mínimas velocidades diárias do vento na escala de tempo sazonal e diurna para o (a) cluster-1, (b) cluster-2, (c) cluster-3, (d) cluster-4, (e) cluster-5, (f) cluster-6 e (g) cluster-7.

f)



g)



Fonte: o autor.

### 4.3 SÍNTESE DOS RESULTADOS

Nesta seção apresenta-se a síntese do desempenho dos quatro métodos utilizados para o preenchimento de falhas artificiais e da caracterização da variabilidade da temperatura máxima do ar e da velocidade mínima do vento na escala sazonal e diurna para os sete agrupamentos,

os quais indicaram as regiões meteorologicamente críticas para a transmissão de energia em linhas aéreas de transmissão.

A partir do diagrama de Taylor (figura 4.34), observou-se o desempenho de quatro métodos de imputação de dados (interpolação linear, kriging, ERA5 e WRF) no preenchimento de falhas geradas artificialmente. A interpolação linear apresentou os menores erros quando analisado o índice RMSE, assim como as maiores correlações lineares para ambas variáveis meteorológicas e independente do cluster. Ou seja, este método estimou dados mais próximos aos que foram removidos artificialmente.

Para a temperatura do ar (figura 4.34a), observou-se ainda que todos os métodos utilizados para o estimar os valores removidos apresentaram alta correlação linear (entre 0,80 e 0,95) e RMSE entre 1,5 °C e 3,0 °C. Adicionalmente, a posição da interpolação linear do cluster-3 no diagrama de Taylor é a mais próxima ao valor de referência (dado medido e removido artificialmente) quando comparado aos demais métodos por cluster. Ou seja, o melhor desempenho em estimar as falhas artificiais da temperatura do ar foi encontrado para a interpolação no cluster-3.

A análise do desempenho utilizando-se os métodos de modelagem dinâmica da atmosfera indicou que a reanálise ERA5 e o downscaling obtido por meio do modelo WRF para temperatura do ar possuem correlação linear entre 0,81 e 0,91 e RMSE entre 2 °C e 3°C. Adicionalmente, a posição dos marcadores de cada método no diagrama de Taylor indicou que, em geral, o WRF em todos os clusters apresenta um desvio padrão mais próximo ao valor de referência (medido) quando comparado ao ERA5 indicando baixa variabilidade das séries estimadas.

O diagrama de Taylor para a velocidade do vento (figura 4.34b) indicou que o modelo WRF apresentou os maiores erros de estimativa quando observado o índice RMSE. Com exceção do cluster-6 e cluster-7, o RMSE do método WRF esteve entre 1,5 e 2 m/s, enquanto esta medida de erro para os demais métodos concentrou-se entre 1 e 1,5 m/s.

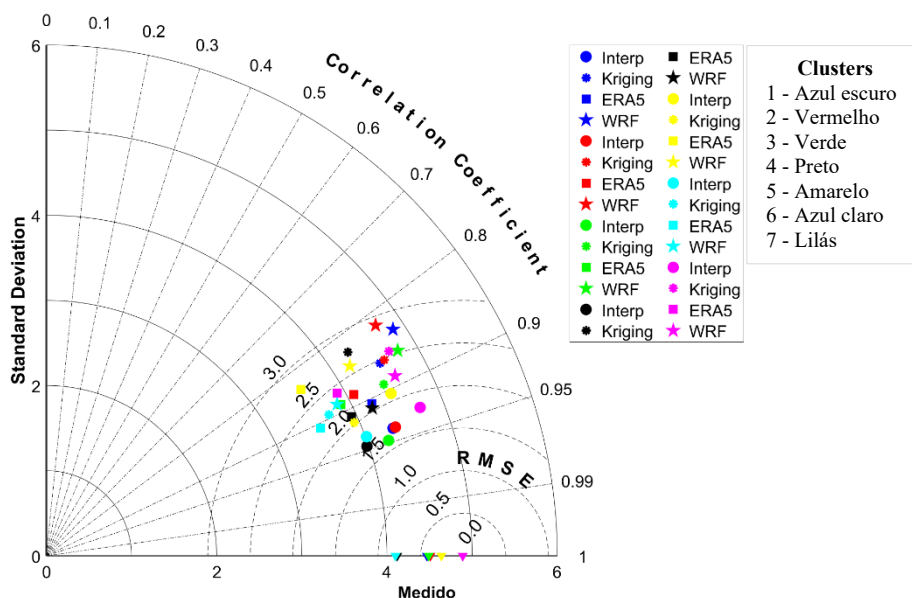
A respeito do coeficiente de correlação, ainda para a velocidade do vento, verificou-se que as estimativas utilizando-se a interpolação linear apresentaram em média valores entre 0,70 e 0,85 (figura 4.34b). No entanto, a correlação linear calculada para os demais métodos concentrou-se entre 0,50 e 0,70. A exceção foi observada apenas para o método de kriging, cuja maioria dos coeficientes de correlação estiveram entre 0,38 e 0,50.

A comparação do desempenho entre os modelos dinâmicos (ERA5 e WRF) indicou que para a velocidade do vento as estimativas utilizando a reanálise situaram-se mais próximas aos valores de referência (medidos) de todos os clusters. Esse resultado indicou que a reanálise do

ERA5 obteve maior habilidade em reproduzir as falhas geradas artificialmente nas séries temporais da velocidade do vento.

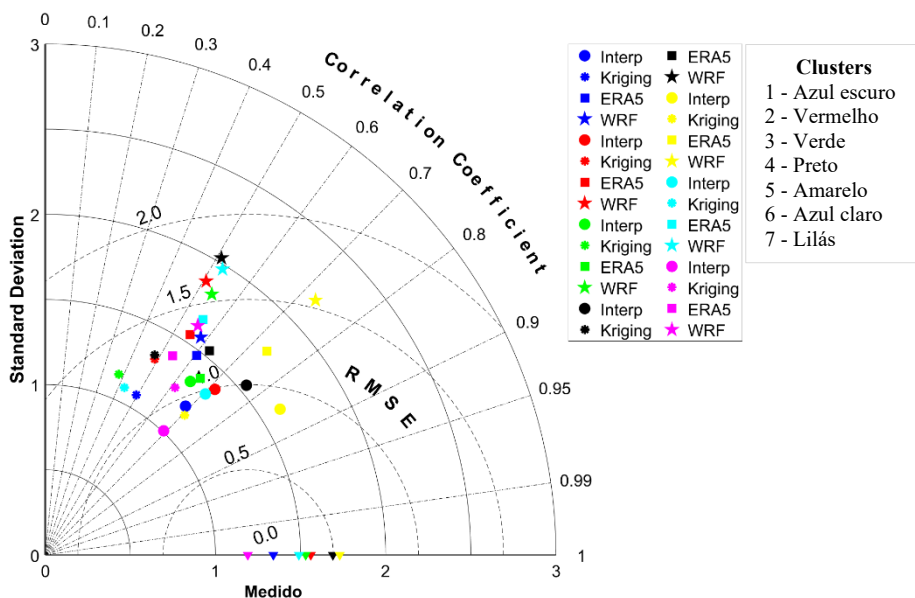
Figura 4.34 – Diagrama de Taylor para (a) temperatura do ar e (b) velocidade do vento considerando os quatro métodos de estimava para cada cluster e usando os valores removidos artificialmente (medido) como referência.

a)



Fonte:  
o  
autor.

b)



No intuito de verificar se o desempenho de cada método de imputação aumentaria ou diminuiria com o tamanho de falhas consecutivas a serem preenchidas realizou-se uma análise adicional, a qual considerou a probabilidade de sucesso de cada método em estimar os valores removidos artificialmente. A tabela 4.15 sumariza o desempenho de cada método de acordo com o índice de probabilidade de detecção (POD) indicando qual deles foi utilizado para



realizar o preenchimento das falhas encontradas nas séries temporais brutas da temperatura do ar e da velocidade do vento.

Tabela 4.15 - Métodos de preenchimento das séries temporais brutas da temperatura do ar e velocidade do vento de acordo com o índice de probabilidade de detecção.

Agrupamentos	Temperatura do ar				Velocidade do vento			
	Métodos				Métodos			
	1	2	3	4	1	2	3	4
Cluster-1	x				x	x		
Cluster-2	x				x			
Cluster-3	x		x		x	x		x
Cluster-4	x				x			
Cluster-5	x	x	x		x	x		x
Cluster-6	x		x		x			
Cluster-7	x	x	x		x	x		

1 - Interpolação    2 - Kriging    3 - ERA5    4 - WRF<sub>Bias</sub>

Fonte: o autor.

A presença de mais de um método por cluster para uma mesma variável indica que mais de um método foi utilizado para preencher as falhas artificiais. Por exemplo, os resultados demonstraram que para a temperatura do ar apenas a interpolação apresentou maior probabilidade independentemente do tamanho das falhas para o cluster-1, o cluster-2 e o cluster-4. Para a velocidade do vento o sucesso absoluto da interpolação se deu para o cluster-2, o cluster-4 e cluster-6. Entretanto, verificou-se para o cluster-3, cluster-5, cluster-6 e cluster-7 que a reanálise ERA5 apresentou maior probabilidade de acerto para as falhas da temperatura do ar de 2 e 5 horas; 4, 5, 12 e 19 horas; 24 horas; e 15 horas, respectivamente.

Adicionalmente, as estimativas que consideraram os valores medidos de temperatura do ar por estações vizinhas (método de Kriging), indicaram maior probabilidade de acerto para os cluster-5 e cluster-7 nos tamanhos de 6, 7, 13, 15 e 21 e 9 horas, respectivamente. Nota-se a partir da tabela 4.15 que, para a velocidade do vento, o método Kriging também apresentou maior probabilidade de estimar os valores removidos para o cluster-1, o cluster-3, o cluster-5 e o cluster-7. Para estes agrupamentos, a performance do método geoestatístico foi superior para o preenchimento de falhas de 20 horas; 2 horas; e 2, 6, 8 e 18 horas, respectivamente.

Na análise de evolução diurna dos valores médios removidos de temperatura do ar e velocidade do vento, os resultados demonstraram também que, considerando todos os agrupamentos, o modelo WRF apresenta um erro sistemático médio de  $-0,8$  °C para a temperatura do ar e de  $1,5$  m/s para a velocidade do vento. O menor erro médio encontrado para

os dados simulados de temperatura do ar, comparado aos valores removidos artificialmente, foi de  $-0,3\text{ }^{\circ}\text{C}$  (cluster-6) e de  $0,9\text{ m/s}$  para o vento (cluster-1).

No entanto, observou-se que para clusters cujo *bias* da velocidade do vento é constante ao longo da série temporal, a remoção deste erro durante a etapa de pós-processamento do modelo reduziu o erro médio no período de aquecimento da atmosfera próxima à superfície (10 às 19 UTC). Como exemplo, cita-se o resultado encontrado para o cluster-3, cujo *bias* existente entre o período das 10 às 19 UTC foi reduzido de  $1,4\text{ m/s}$  para apenas  $0,1\text{ m/s}$ .

Este resultado indicou que na ausência de dados medidos na vizinhança de linhas de transmissão, o WRF, após a remoção do erro médio sistemático, pode ser utilizado para prover informações meteorológicas, principalmente durante o período mais crítico para a operação de linhas de transmissão (maior demanda de carga e temperaturas do ar mais alta). Isto é justificado pelo fato de que é durante o período da tarde que está concentrado o maior consumo de energia e, conseqüentemente a maior demanda por geração e despacho de carga nas linhas de transmissão do SIN.

A caracterização do regime de mínimos de vento medidos a 10 metros e dos máximos de temperatura do ar a 2 metros sobre o subsistema Nordeste do SIN foi realizada considerando-se a frequência de ocorrência destas variáveis nas escalas de tempo sazonal e diurna e em cada agrupamento formado pela análise de cluster. Isso possibilita a identificação dos sistemas atmosféricos que influenciam as variáveis em questão.

Os resultados desta caracterização, obtidos a partir dos entalhes dos *boxplot*, demonstram que, de maneira geral, todas as medianas das temperaturas máximas são estatisticamente diferentes entre si dentro de cada agrupamento nas escalas analisadas (sazonal e diurna). Quando analisada a atuação dos sistemas atmosféricos sobre as regiões dos clusters, verificou-se que a variabilidade dos máximos de temperatura do ar de uma estação do ano para outra é influenciada pelos regimes de precipitação existentes na região estudada.

Por exemplo, os resultados mostraram que as menores medianas das temperaturas máximas ocorreram no verão e outono enquanto as maiores foram observadas na primavera. Este fato é devido aos máximos de precipitação registrados entre dezembro e maio sobre o subsistema Nordeste do SIN. Em outras palavras, o aumento da nebulosidade associada à atividade convectiva produzida por sistemas atmosféricos como ZCIT, VCAN, distúrbios de ondas de leste e até mesmo frentes frias, diminuiu a temperatura do ar e, conseqüentemente, as máximas temperaturas diárias para todos os agrupamentos.

Ainda para a escala sazonal, as medianas das temperaturas máximas situam-se sempre acima do limiar de  $31,5\text{ }^{\circ}\text{C}$  para todas as estações do ano com exceção do cluster-4, cuja menor

mediana situou-se no inverno. O motivo para tanto é que uma parte das estações que compuseram este agrupamento estão situadas no leste do Nordeste e nesta região o período chuvoso se estende até agosto devido a interação da ZCIT com os distúrbios de onda de leste. Portanto, as maiores medianas das temperaturas máximas do ar foram observadas sempre durante a primavera, assim como os extremos superiores.

Esta constatação indica que, analisando apenas a temperatura do ar, a primavera é a estação do ano de maior restrição para aumento de carregamento de linhas de transmissão do subsistema Nordeste do SIN, pois a ocorrência de altas temperaturas impactam negativamente no balanço de energia dos cabos condutores, diminuindo a distância de segurança entre cabo e solo.

Para todos os agrupamentos, observa-se por meio dos entalhes presentes nos *boxplot* que, na escala diurna, as medianas das temperaturas máximas são estatisticamente diferentes entre si independente do período do dia. Os resultados para esta escala demonstram que a variação das temperaturas máximas do ar em todos os clusters é fortemente influenciada pelo balanço de energia que ocorre na superfície.

Ou seja, no período de maior resfriamento da atmosfera próxima à superfície (madrugada entre 4 e 9 UTC) registraram-se as menores medianas das temperaturas máximas diárias devido à ausência de radiação solar no espectro de onda curta. Por outro lado, as maiores medianas são registradas no período em que o ângulo de elevação solar está próximo ao máximo e, por conseguinte a irradiância ou o saldo de radiação na superfície é maior (tarde, entre 16 e 21 UTC).

Outro resultado importante encontrado em todos os agrupamentos durante a análise do ciclo diurno da temperatura é que as temperaturas máximas do ar apenas ultrapassam o limiar de temperatura de 31,5 °C durante o período da tarde. Este fato sugere que projetos de linhas de transmissão que utilizaram a norma NBR5422 para estabelecer a temperatura máxima do ar na região de estudo poderiam flexibilizar seus limites de carregamento na escala diurna sem descuidar dos aspectos físicos e mecânicos da LT.

A respeito da caracterização dos regimes de vento verifica-se que a variação das mínimas velocidades na escala sazonal é influenciada indiretamente pelo período chuvoso. Por este motivo, observa-se que em cinco dos sete clusters, as menores medianas dessa variável meteorológica ocorrem no verão e outono enquanto as maiores no inverno e na primavera. A exceção foi encontrada apenas para os cluster-2 e cluster-3 (figura 4.33b,c), cuja medianas apresentaram o mesmo valor para as quatro estações do ano.

As análises das mínimas velocidades do vento na escala sazonal revelaram também que em média a máxima diferença entre as medianas é de 0,2 m/s. Este resultado indicou que a variabilidade dos mínimos de vento entre uma estação do ano e outra é baixa. Ou seja, a pequena variação da velocidade do vento demonstrou que há pouca possibilidade de explorar os limites de despacho de carga em LT considerando apenas essa variável na escala sazonal.

Como resultado adicional, observa-se que a maior variabilidade da velocidade mínima diária do vento ocorre sempre no inverno. Para as demais estações do ano o menor desvio padrão indicado pela largura das caixas dos *boxplot* indica ventos calmos constantes na escala sazonal (principalmente verão e outono). Isso significa que durante a estação chuvosa, as mínimas velocidades diárias possuem baixa variância em torno da mediana enquanto as maiores variações dos valores medidos encontram-se no inverno independentemente do agrupamento.

Por fim, para a escala diurna verifica-se que as maiores medianas das mínimas velocidades diárias do vento ocorreram sempre no período da tarde (entre 1,4 m/s e 2,8 m/s). Este fato se dá porque durante a tarde ocorre o maior aquecimento da atmosfera próxima a superfície e, por conseguinte o aumento da convecção térmica e da turbulência mecânica devido à rugosidade do terreno. Como resultado tem-se o aumento da intensidade do vento durante a tarde e o registro dos menores valores de vento durante o período da madrugada.



## 5 CONCLUSÕES

Por meio da aplicação de uma metodologia para identificar as regiões de calmaria de vento e temperaturas máximas do ar próximo à superfície estabeleceu-se a frequência de ocorrência na escala sazonal e diurna destas variáveis de acordo com os sistemas atmosféricos atuantes na região-alvo da presente tese.

Conforme já discutido neste trabalho, a NBR5422 estabelece que, em projetos de construção de linhas aéreas de transmissão sejam utilizados dados meteorológicos medidos para subsidiar o dimensionamento da linha afim de garantir a operação segura do ativo e do Sistema Interligado Nacional. Entretanto, esta mesma Norma define que na falta de dados medidos próximo à linha a velocidade do vento deve ser considerada igual a 1 m/s e a temperatura do ar igual à média da máxima temperatura diária do ar. Estes valores, segundo a Norma supracitada, são representativos das condições meteorológicas mais desfavoráveis para a transmissão de energia elétrica (baixa velocidade do vento e alta temperatura do ar).

Assim, com o intuito de estabelecer as regiões críticas para operação de linhas aéreas de transmissão investigou-se a frequência de ocorrência das mínimas velocidades do vento e máximas temperaturas do ar, com o objetivo de permitir que o setor elétrico avalie, posteriormente, o melhor aproveitamento das linhas existentes assim como direcione novos projetos de linhas com valores típicos de vento e temperatura próximos ao observado.

Para isso, selecionaram-se as estações meteorológicas do INMET instaladas na área de atuação do subsistema Nordeste do SIN e identificaram-se as estações que apresentaram as condições meteorológicas restritivas para despacho de carga em linhas aéreas: mínimas velocidades diárias do vento e máximas temperaturas diárias do ar. Para tanto, consideraram-se as estações com mediana destas variáveis igual ou menor a 1 m/s e maior ou igual a 31,5 °C.

Com o intuito de avaliar se as 35 estações meteorológicas selecionadas no passo anterior eram similares ou dissimilares no tocante às variáveis em estudo, aplicou-se a técnica de análise de cluster por meio do algoritmo hierárquico aglomerativo. Como resultado obtiveram-se sete agrupamentos, os quais indicaram as regiões de mínimos de vento e de máximos de temperatura do ar.

A análise exploratória inicial dos dados medidos pelas 35 estações meteorológicas indicou a ausência de valores medidos (buracos) com tamanhos distintos ao longo das séries temporais. Assim, o preenchimento destas falhas foi realizado após a avaliação de qual método de imputação de dados apresenta maior habilidade em estimar falhas geradas artificialmente

nas séries originais. Os métodos utilizados para este propósito foram: interpolação linear, kriging, ERA5 e WRF.

A interpolação linear foi o método com maior habilidade para estimar os dados das séries temporais independente da variável meteorológica avaliada e do cluster. Este método possui a vantagem da simples aplicação; no entanto, é dependente da existência de variáveis meteorológicas medidas para que possa ocorrer a estimativa.

A avaliação dos métodos de preenchimento que não fazem uso de dados medidos na vizinhança para estimar dados indica que, tanto a reanálise ERA5 quanto o downscaling obtido por meio do WRF, apresentam resultados similares para ambas variáveis em todos os clusters. Para a temperatura do ar, verifica-se que o WRF reproduz com habilidade a variabilidade desta variável durante o ciclo diurno com altos valores de correlação linear em todos os clusters.

Portanto, conclui-se que na ausência de dados medidos, o modelo WRF pode ser utilizado para estimar, na região do subsistema em questão, a temperatura do ar. Este resultado é de suma importância para o setor elétrico, principalmente para suportar a execução de novos projetos de linhas de transmissão e a recapacitação destas quando não houver disponibilidade de dados medidos por estações meteorológicas.

Para a velocidade do vento, no entanto, os dados obtidos por meio do downscaling dinâmico utilizando o WRF demonstram a dificuldade do modelo em reproduzir a variação diurna da velocidade do vento medida pelas estações. As medidas de erro indicam que este modelo apresentou grandes erros individuais para estimar os valores medidos. Porém, após a remoção do erro sistemático do WRF, as estimativas se aproximam dos dados medidos, principalmente entre as 10 e 13 UTC.

Diante do exposto, pode-se concluir que novos experimentos numéricos devem ser realizados utilizando outras análises numéricas globais como condição inicial e de fronteira, assim como é necessário aumentar a resolução horizontal da grade do modelo com o objetivo de obter resultados mais coerentes com os dados medidos de velocidade do vento no ciclo diurno.

Com relação a análise da frequência de ocorrência das máximas temperaturas do ar, conclui-se que as medianas se situam sempre acima do limiar de 31,5 °C na escala sazonal com exceção do inverno do cluster-4. Ainda para esta variável, a sazonalidade mais crítica para o despacho de energia deve ser considerada a primavera, onde são observadas as maiores temperaturas máximas, independente do agrupamento. Esse resultado vai de encontro ao esperado pelo setor elétrico, que adota o verão como a estação de máxima temperatura do ar e, por conseguinte, com maior restrição de capacidade operativa das LT na escala sazonal.

A respeito da escala temporal diurna, verifica-se que as medianas das máximas temperaturas estão sempre acima do limiar de 31,5 °C apenas durante o período da tarde.

Portanto, conclui-se que, excetuando-se o turno vespertino, poder-se-ia aumentar a capacidade de operação das linhas de transmissão da região-alvo da presente tese, quando considerado apenas os efeitos da temperatura do ar no balanço de energia do cabo condutor.

Para a velocidade do vento, observa-se que, além da baixa variabilidade desta variável entre as estações do ano, as medianas estão sempre abaixo do limiar de 1 m/s em todos os clusters, quando a variável é analisada na escala sazonal. A partir deste resultado, conclui-se que as medianas das mínimas velocidades do vento são menores do que o valor de 1 m/s estabelecido na NBR5422 como sendo o representativo da condição meteorológica do vento mais desfavorável para a transmissão de energia em linhas aéreas. Adicionalmente, conclui-se também que a baixa variabilidade desta variável observada na escala sazonal indica que a capacidade operativa entre as estações do ano não seria alterada se considerado apenas o regime de vento.

Para o ciclo diurno, os resultados indicam que em todos agrupamentos, as circulações locais induzidas termicamente modulam as mínimas velocidades do vento. Desta forma, observa-se para todos os clusters que as menores mínimas velocidades ocorrem sempre no período de maior resfriamento da atmosfera (madrugada), enquanto as maiores mínimas, durante a tarde (entre 1,4 m/s e 2,8 m/s), período em que a atmosfera próxima à superfície está mais aquecida. Portanto, para a velocidade do vento apenas, a escala diurna torna-se de maior importância para o despacho de energia sobre o subsistema Nordeste do SIN, pois é nesta escala que ocorrem as variações estatisticamente significativas para a operação de LT.

Os resultados expostos apontam que o preenchimento de séries temporais incompletas de temperatura do ar e velocidade do vento utilizando a interpolação linear possibilita a realização de estudos de capacidade a partir de valores mais fidedignos das condições críticas ao longo de uma linha de transmissão. Porém, os resultados sugerem ainda que o uso dos métodos de simulação dinâmica da atmosfera, também avaliados nesta tese, podem ser utilizados com eficiência para estimar a temperatura do ar quando não há dados medidos por estações meteorológicas próximas às linhas de transmissão. Por outro lado, os modelos dinâmicos estimaram a velocidade do vento com maiores erros no período estudado.

A caracterização da frequência de ocorrência das máximas temperaturas do ar demonstra que o período de maior restrição para o despacho de carga é, considerando apenas a temperatura do ar, na primavera e durante a tarde. No tocante à velocidade do vento, apenas o ciclo diurno apresenta variações significativas entre os períodos analisados. E a maior restrição



considerando esta variável ocorre para todos os períodos do ciclo diurno com exceção das tardes.

Desta forma, conclui-se que a capacidade operativa de linhas de transmissão pode possuir limites diferenciados na escala sazonal e diurna. Em adição, que se considera que os resultados encontrados neste trabalho podem contribuir para o desenvolvimento de projetos de novas linhas de transmissão e para o dimensionamento de novas capacidades operativas na escala sazonal e diurna em linhas de transmissão existentes sobre o subsistema Nordeste do SIN.

## 6 SUGESTÕES PARA TRABALHOS FUTUROS

A demanda por dados meteorológicos medidos é cada vez maior para suportar pesquisas endógenas e aplicadas à diversos setores da sociedade. No tocante ao setor elétrico, o uso de modelos numéricos, que representem adequadamente os processos atmosféricos incluindo os sistemas convectivos e processos radiativos, pode suportar pesquisas e desenvolvimento na área de energia por meio de séries sintéticas quando há ausência de dados medidos ou falta de séries temporais longas.

Portanto, sugerem-se estudos numéricos que contemplem:

- simulações com altíssima resolução horizontal, com espaçamentos de grade iguais ou menores que 3 km e com uso de parametrização de microfísica de nuvens em substituição às parametrizações de nuvens convectivas;
- o uso de modelos baseados em malhas não estruturadas de Voronoi, os quais permitem variar a resolução horizontal do modelo resolvendo as equações de movimento diretamente nas malhas não estruturadas, por exemplo o The Model for Prediction Across Scales (MPAS), desenvolvido pelo mesmo grupo de pesquisadores do modelo WRF. Com isso, pode-se aumentar a resolução horizontal sobre a região de interesse de maneira gradual sem a necessidade de realizar grandes saltos entre grades, como é verificado em modelos de grade com geometria retangular;
- o uso da técnica de ensemble que considera uma série de condições iniciais diferentes entre si com o modelo ligeiramente modificado no tocante a parametrizações físicas (esquema de nuvens, microfísica de nuvens e CLP). Isso porque, a reprodução do ciclo diurno da velocidade do vento é fortemente dependente do uso adequado das parametrizações dos processos físicos na atmosfera (ZHANG; ZHENG, 2004).

A partir dos resultados encontrados na presente tese sugere-se também que novas análises sejam realizadas considerando os limiares de vento e temperatura do ar definidos na Nota Técnica 94/2016 emitida pelo Operador Nacional do Sistema Elétrico (ONS, 2017) a qual define capacidades operativas sazonais nas escalas verão-dia, verão-noite, inverno-dia e inverno-noite.



## 7 REFERÊNCIAS

- ABNT. NBR 5422 - Projeto de Linhas Aéreas de Transmissão de Energia Elétrica - Procedimentos. 1985. p. 58.
- ALCANTARA, Clênia R.; SOUZA, Enio P. Teoria Para O Trabalho De Expansão Aplicada Às Brisas Do Nordeste Brasileiro. **Revista Brasileira de Meteorologia**, São José dos Campos, São Paulo, v. 24, n. 4, p. 423–435, 2009.
- ALVAREZ, David L. et al. An approach to dynamic line rating state estimation at thermal steady state using direct and indirect measurements. **Electric Power Systems Research**. v. 163, p. 599–611, 2018.
- ANEEL. **Acompanhamento diferenciado da expansão da transmissão**. Brasília, DF, Brasil, 2019.
- ANEEL, Agência Nacional de Energia Elétrica. **Atlas de Energia Elétrica do Brasil**. Brasília, DF, 2008.
- ANJOS, Rodnei Dias Dos. **Ampacidade de linhas de transmissão - Aspectos operativos**. 2013. Universidade Federal de Itajubá, Itajubá, MG, Brasil, 2013.
- ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR 5422: Projeto de linhas aéreas de transmissão de energia elétrica**. Rio de Janeiro, RJ, Brasil.
- ATLURI, Gowtham; KARPATNE, Anuj; KUMAR, Vipin. Spatio-Temporal Data Mining. **ACM Computing Surveys**, v. 51, n. 4, p. 1–41, 2018.
- BALDICK, Ross; NEILL, Richard P. O. Estimates of Comparative Costs for Uprating Transmission Capacity. **IEEE Transactions on Power Delivery**, v. 24, n. 2, p. 961–969, 2009.
- BANERJEE, Binayak. **Quantifying the Benefits of Dynamic Line Ratings in Congestion Management for Improved Wind Integration in Smart Grids**. 2015. Curtin University, Bentley, Austrália, 2015.
- BAUMBACH, Günter; VOGT, Ulrich. Experimental determination of the effect of mountain-valley breeze circulation on air pollution in the vicinity of Freiburg. **Atmospheric Environment**, v. 33, n. 24–25, p. 4019–4027, 1999.
- BEÇA, Juliana Maciel Maia et al. Estudo de caso de técnicas de repotencialização de uma linha de transmissão de 230 kV ( Recife II- Suape). In: **XIII ENCONTRO NACIONAL DOS ESTUDANTES DE ENGENHARIA ELÉTRICA**, Cabo de Santo Agostinho, PE, 2018.
- BERNDT, C.; HABERLANDT, U. Spatial interpolation of climate variables in Northern

- Germany—Influence of temporal resolution and network density. **Journal of Hydrology: Regional Studies**, v. 15, n. October 2017, p. 184–202, 2018. Disponível em: <<https://doi.org/10.1016/j.ejrh.2018.02.002>>
- BEZERRA, Francisco Diniz. Energia eólica no nordeste. **Caderno Setorial ETENE**, Fortaleza, CE, p. 1–20, 2019.
- BRAHMANANDA RAO, V.; DE LIMA, Marley C.; FRANCHITO, S. H. Seasonal and Interannual Variations of Rainfall over Eastern Northeast Brazil. **Journal of Climate**, v. 6, n. 9, p. 1754–1763, 1993. Disponível em: <<http://journals.ametsoc.org/doi/abs/10.1175/1520-0442%281993%29006%3C1754%3ASAIVOR%3E2.0.CO%3B2>>
- CARVALHO, Miguel Ângelo Vargas De; OYAMA, Marcos Daisuke. Variabilidade da largura e intensidade da Zona de Convergência Intertropical atlântica: aspectos observacionais. **Revista Brasileira de Meteorologia**, v. 28, n. 3, p. 305–316, 2013.
- CARVALHO, RENATTO VAZ. **Monitoramento térmico de linhas de transmissão**. 2017. Universidade Federal do Paraná, Curitiba, PR, 2017.
- CASTILHO, Renato Maioli. **Otimização do uso da linha de transmissão através do monitoramento em tempo real**. 2010. Universidade Estadual de Campinas, Campinas, SP, 2010.
- CAVASSIN, Ricardo S.; FERNANDES, Thelma S. Piazza. Uma abordagem multicritérios para recapacitação de linhas de transmissão. **Sba: Controle & Automação Sociedade Brasileira de Automatica**, v. 23, n. 6, p. 749–765, 2012.
- CHAKRABORTY, Sanjay; NAGWANI, N. K.; DEY, Lopamudra. Weather Forecasting using Incremental K-means Clustering. **International Journal of Biometrics and Bioinformatics**, 2014.
- CHAVES, Rosane Rodrigues; CAVALCANTI, Iracema Fonseca Albuquerque. Atmospheric Circulation Features Associated with Rainfall Variability over Southern Northeast Brazil. **Monthly Weather Review**, v. 129, n. 10, p. 2614–2626, 2001.
- CIGRE WG B2.12. **Technical Brochure 244: Conductors for the Uprating of Overhead Lines**Cigre. 2004.
- COELHO, Micheline De Souza; GAN, Manoel Alonso; CONFORTE, Jorge Conrado. Estudo da variabilidade da posição e da nebulosidade associada a ZCIT do Atlântico, durante a estação chuvosa de 1998 e 1999 no Nordeste do Brasil. **Revista Brasileira de Meteorologia**, São José dos Campos, SP, v. 19, n. 1, p. 23–34, 2004.
- COMPO, G. P. et al. The Twentieth Century Reanalysis Project. **Quarterly Journal of the Royal Meteorological Society**, v. 137, n. 654, p. 1–28, 2011.

- CORREIA, M. F.; DA SILVA DIAS, M. A. F.; DA SILVA ARAGÃO, M. R. Soil occupation and atmospheric variations over Sobradinho Lake area. Part one: an observational analysis. **Meteorology and Atmospheric Physics**, v. 94, n. 1–4, p. 103–113, 2006.
- COSTA, Gabriel Brito; LYRA, Roberto Fernando da Fonseca. Análise dos padrões de vento no Estado de Alagoas. **Revista Brasileira de Meteorologia**, v. 27, n. 1, p. 31–38, 2012.
- DACONTI, J. R.; LAWRY, D. C. Increasing power transfer capability of existing transmission lines. In: 2003 **IEEE PES TRANSMISSION AND DISTRIBUTION CONFERENCE AND EXPOSITION**. 2003.
- DANTAS, Marcelo Eduardo et al. Geodiversidade e análise da paisagem: uma abordagem teórico-metodológica geodiversity and landscape analysis: a theoretical and Methodological approach. **Terrae Didactica**, v. 11, n. 11, p. 4–13, 2015.
- DE FREITAS, Gilmar Fialho; ROMARCO DE OLIVEIRA, Marcelo Leles. Uma análise do programa Luz para Todos do Governo Federal. **Revista de Extensão e Estudos Rurais**, v. 6, n. 2, p. 143–155, 2017.
- DE SOUSA MELO, Ewerton Cleudson; DA SILVA ARAGÃO, Maria Regina; CORREIA, Magaly De Fatima. Regimes do vento à superfície na área de petrolina, submédio são francisco. **Revista Brasileira de Meteorologia**, v. 29, n. 2, p. 229–241, 2014.
- DIAZ, Henry F.; BRADLEL, Raymond S. **The Hadley Circulation: Present, Past and Future**. Dordrecht: Springer Netherlands, 2004. v. 21.
- DU, Jizeng et al. Diurnal Cycle of Surface Air Temperature within China in Current Reanalyses: Evaluation and Diagnostics. **Journal of Climate**, v. 31, n. 11, p. 4585–4603, 2018.
- DUPIN, Romain; KARINIOTAKIS, George; MICHIORRI, Andrea. Overhead lines Dynamic Line rating based on probabilistic day-ahead forecasting and risk assessment. **International Journal of Electrical Power & Energy Systems**, v. 110, n. June 2018, p. 565–578, 2019.
- DUTTON, Ellsworth G. Annual Forcing of the Surface Radiation Balance Diurnal Cycle Measured from a High Tower near Boulder, Colorado. **Journal of Climate**, v. 3, n. 12, p. 1400–1408, 1990.
- EMANUEL, Kerry A. Overview and Definition of Mesoscale Meteorology. In: **Mesoscale Meteorology and Forecasting**. Boston, MA: American Meteorological Society, 1986. p. 1–17.
- ENTSOE. **Dynamic Line Rating for overhead lines - V6**. Bruxelas, Bélgica, 2015.
- FAGHMOUS, James H.; KUMAR, Vipin. **Spatio-temporal Data Mining for Climate Data: Advances, Challenges, and Opportunities**. Data Mining and Knowledge Discovery for Big

Data. p. 83–116, 2014.

FANG, Shu et al. Boundaries and Characteristics of Arid Regions in Mountain Valleys in Southwestern China. **Mountain Research and Development**, v. 38, n. 1, p. 73, 2018.

FARIA, Bruno Lopes De. **Modelagem do potencial eólico do nordeste do brasil sob condições atuais e de aquecimento global: uma interface entre modelos numéricos computacionais de microescala e mesoescala**. 2010. Universidade Federal de Viçosa, Viçosa, MG, 2010.

FERREIRA, Antonio Geraldo; MELLO, Namir Giovanni da Silva. Principais sistemas atmosféricos atuantes sobre a Região Nordeste do Brasil e a influência dos oceanos Pacífico e Atlântico no clima da região. **Revista Brasileira de Climatologia**, v. 1, n. 1, p. 15–28, 2005.

FERREIRA, Mauricio Jonas; DE OLIVEIRA, Amauri Pereira; SOARES, Jacyra. Diurnal variation in stored energy flux in São Paulo city, Brazil. **Urban Climate**, [s. l.], v. 5, p. 36–51, 2013.

FIEBRICH, Christopher A. et al. Quality assurance procedures for mesoscale meteorological data. **Journal of Atmospheric and Oceanic Technology**, v. 27, n. 10, p. 1565–1582, 2010.

FOVELL, R. G.; FOVELL, M. Y. C. **Climate zones of the conterminous United States defined using cluster analysis** *Journal of Climate*, 1993.

FREITAS, Edmilson Dias De. **Circulações locais em São Paulo e sua influência sobre a dispersão de poluentes**. 2003. Universidade de São Paulo, São Paulo, SP, 2003.

FREITAS, Edmilson D. et al. Interactions of an urban heat island and sea-breeze circulations during winter over the metropolitan area of São Paulo, Brazil. **Boundary-Layer Meteorology**, v. 122, n. 1, p. 43–65, 2007.

FREITAS, Gisele; SILVEIRA, Suely de Fátima Ramos. Programa Luz Para Todos: Uma Representação Da Teoria Do Programa Por Meio Do Modelo Lógico Luz Para Todos Program: a Representation Theory From Across the Logic Model Program. **Planejamento e Políticas Públicas - PPP**, v. 45, 2015.

FRONTIN, Sergio De Oliveira et al. **Prospecção e hierarquização de inovações tecnológicas aplicadas a linhas de transmissão**. 1º Edição ed. Brasília, DF, Brazil: Goya Editora, 2010.

FURTADO, Ronan Gustavo Carvalho. **Métodos estatísticos aplicados ao cálculo da ampacidade e risco térmico de linhas aéreas de transmissão**. 2008. Universidade Federal de Juiz de Fora, Juiz de Fora, MG, 2008.

GIANNUZZI, G. M. et al. Overhead transmission lines dynamic line rating estimation in WAMS environments. In: 2015 **INTERNATIONAL CONFERENCE ON CLEAN**

**ELECTRICAL POWER**, 2015.

GOMES, Helber B. et al. Climatology of easterly wave disturbances over the tropical South Atlantic. **Climate Dynamics**, v. 53, n. 3–4, p. 1393–1411, 2019.

GOMES, Helber Barros. **Distúrbios ondulatórios de leste no nordeste brasileiro: climatologia e modelagem numérica**. 2012. Universidade de São Paulo, São Paulo, 2012.

GOVENDER, Paulene; BROOKS, Michael J.; MATTHEWS, Alan P. Cluster analysis for classification and forecasting of solar irradiance in Durban, South Africa. **Journal of Energy in Southern Africa**, v. 29, n. 2, p. 63–76, 2018.

HADZIMEJLIC, Nejra; DONKO, Dzenana; HADZIMEJLIC, Nijaz. Climate Data Analysis Using Clustering Data Mining Techniques. **Latest Trends in Applied Informatics and Computing Climate**, p. 94–101, 2012.

HAIR JR., Joseph F. et al. **Multivariate Data Analysis**. 7th. ed. Upper Saddle River, New Jersey: Print, 2014.

HALLAK, Ricardo; FILHO, Augusto José Pereira. Metodologia para análise de desempenho de simulações de sistemas convectivos na região metropolitana de São Paulo com o modelo ARPS: Sensibilidade a variações com os esquemas de advecção e assimilação de dados. **Revista Brasileira de Meteorologia**, São José dos Campos, SP, v. 26, n. 4, p. 591–608, 2011.

HIJMANS, Robert J. et al. Very high resolution interpolated climate surfaces for global land areas. **International Journal of Climatology**, v. 25, n. 15, p. 1965–1978, 2005.

HOX, Joop J. The Multilevel Approach to Meta-Analysis. In: **Multilevel analysis: Techniques and applications**. Second ed. New York, NY: George Marcoulides, 2010. p. 205.

INMET. **Normais climatológicas do Brasil, 1961-1990**. 2009, Brasília, DF, Brazil.

INMET. **Rede de Estações Meteorológicas Automáticas do INMET - NT 001/2001**. 2011, Brasília, DF, Brazil.

INMET. **Especificações técnicas dos sensores da MAWS301**. 2019, Brasília, DF, Brazil.

JAIN, Anil K.; DUBES, Richard C. **Algorithms for clustering data**. Englewood Cliffs, New Jersey: Prentice Hall, 1988.

JARDINETTI, Rosemary Boettger. **Repotencialização de linhas de transmissão localizadas em áreas de severas restrições ambientais**. 2014. Escola Politécnica da Universidade de São Paulo, São Paulo, SP, 2014.

JEFFREY, Stephen J. et al. Using spatial interpolation to construct a comprehensive archive of Australian climate data. **Environmental Modelling and Software**, v. 16, n. 4, p. 309–330,



2001.

KARIMI, Soheila; MUSILEK, Petr; KNIGHT, Andrew M. Dynamic thermal rating of transmission lines: A review. **Renewable and Sustainable Energy Reviews**, v. 91, n. April, p. 600–612, 2018.

KHAKI, Milad et al. Electric power system cost/loss optimization using Dynamic Thermal Rating and linear programming. In: 2010 **IEEE ELECTRICAL POWER & ENERGY CONFERENCE**. IEEE, 2010.

KÖPPEN, W.; GEIGER, R. **Das geographische System der Klimate**. Munique: Gebrüder Borntraeger, 1936.

KOUSKY, Vernon E. Frontal Influences on Northeast Brazil. **Monthly Weather Review**, v. 107, n. 9, p. 1140–1153, 1979.

LARRUSKAIN, D. M. et al. Power transmission capacity upgrade for overhead lines. **Renewable Energy and Power Quality Journal**, v. 1, n. 04, p. 221–227, 2006.

LI, Jin; HEAP, Andrew D. Spatial interpolation methods applied in the environmental sciences: A review. **Environmental Modelling & Software**, v. 53, p. 173–189, 2014.

LI, Ziming; DE RIJKE, Maarten. The Impact of Linkage Methods in Hierarchical Clustering for Active Learning to Rank. **Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval - SIGIR '17**, p. 941–944, 2017.

LIAO, Huanyue; LIU, Wei. Weather-based dynamic line rating of overhead transmission lines over Europe interconnected network. **IOP Conference Series: Materials Science and Engineering**, v. 382, n. 5, p. 052035, 2018.

LIRA, Marcos Antonio Tavares et al. Caracterização do Regime de Ventos no Piauí Para o Aproveitamento de Energia Eólica. **Revista Brasileira de Meteorologia**, v. 32, n. 1, p. 77–88, 2017.

LIU, Yanchi et al. Understanding and enhancement of internal clustering validation measures. **IEEE Transactions on Cybernetics**, v. 43, n. 3, p. 982–994, 2013.

MA, Nan et al. Design of Meteorological Information Display System Based on Data Warehouse. In: 2010 **INTERNATIONAL CONFERENCE ON MANAGEMENT AND SERVICE SCIENCE**. IEEE, 2010.

MANZANAS, R. et al. Dynamical and statistical downscaling of seasonal temperature forecasts in Europe: Added value for user applications. **Climate Services**, v. 9, p. 44–56, 2018.

MELO, Ewerton Cleudson de Sousa. **Simulação numérica das interações biosfera-**

- atmosfera em área de Caatinga: Uma análise da expansão agrícola em ambiente semiárido.** 2011. Universidade Federal de Campina Grande, Campina Grande, PB, 2011.
- MELO, Ewerton Cleudson de Sousa; CORREIA, Magaly de Fatima; SILVA, Maria Regina Da. Expansão da agricultura irrigada e mudanças nos Processos de interação superfície-atmosfera: um estudo numérico de impacto ambiental em áreas de Caatinga. **Revista Brasileira de Geografia Física**, v. 07, n. 05, p. 960–968, 2014.
- MISRA, Vasubandhu; KANAMITSU, Masao. Anomaly Nesting: A Methodology to Downscale Seasonal Climate Simulations from AGCMs. **Journal of Climate**, [s. l.], v. 17, n. 17, p. 3249–3262, 2004.
- MORAIS, Michelyne Duarte Coutinho De. **Vórtices ciclônicos de altos níveis que atuam no nordeste do Brasil: Estudo observacional e numérico.** 2016. INPE, São José dos Campos, SP, 2016.
- MOURA, Antonio Divino; TADEU, Lauro; FORTES, Guimarães. The Brazilian National Institute of Meteorology (INMET) and its contributions to agrometeorology. **Agrometeoros**, v. 24, n. 1, p. 15–27, 2016.
- NASCIMENTO, Carlos Alexandre Meireles. **Calculo da Ampacidade de Linhas Aereas de Energia Considerando a Influencia da Camada Limite da Atmosfera.** 2009. Universidade Federal de Minas Gerais, Belo Horizonte, MG, 2009.
- NETZEL, Pawel; STEPINSKI, Tomasz. On using a clustering approach for global climate classification. **Journal of Climate**, v. 29, n. 9, p. 3387–3401, 2016.
- NTULI, Mavula et al. Increasing the capacity of transmission lines via current uprating: An updated review of benefits, considerations and developments. In: 2016 **AUSTRALASIAN UNIVERSITIES POWER ENGINEERING CONFERENCE**. IEEE, 2016.
- OLIVEIRA, Juliana Lima; COSTA, Alexandre Araújo. Estudo De Variabilidade Do Vento Em Escala Sazonal Sobre O Nordeste. **Revista Brasileira de Meteorologia**, v. 26, n. 1, p. 53–66, 2011.
- ONS. **Metodologia para cálculo da capacidade sazonal de projeto de linhas de transmissão a serem licitadas.** Rio de Janeiro, RJ, Brasil.
- ONS. **Plano de ampliação e reforços nas instalações de transmissão do SIN - PAR 2019-2023.** Rio de Janeiro.
- ONS. **Plano de operação energética 2019-2023.** Rio de Janeiro, RJ, Brasil.
- ONS. **Visão do Sistema de Transmissão - Horizonte 2024.** 2020. Disponível em: <<http://www.ons.org.br/paginas/sobre-o-sin/mapas>>. Acesso em: 2 jan. 2020.
- ORLANSKI, Isidoro. A rational subdivision of scales for atmospheric processes. **Bulletin of**

- the **American Meteorological Society**, v. 56, n. 5, p. 527–530, 1975.
- PALMER, C. E. On high-level cyclones originating in the tropics. **Transactions, American Geophysical Union**, v. 32, n. 5, p. 683, 1951.
- PAREDAENS, Jan et al. **The Structure of the Relational Database Model**. Berlin, Heidelberg: Springer Berlin Heidelberg, 1989.
- PEREIRA, Heloisa Ramos; REBOITA, Michelle Simões; AMBRIZZI, Tércio. Características da Atmosfera na Primavera Austral Durante o El Niño de 2015/2016. **Revista Brasileira de Meteorologia**, v. 32, n. 2, p. 293–310, 2017.
- PERRY, A. E.; SCHOFIELD, W. H.; JOUBERT, P. N. Rough wall turbulent boundary layers. **Journal of Fluid Mechanics**, v. 37, n. 2, p. 383–413, 1969.
- QUADRO, Mario Francisco Leal De et al. Análise climatológica da precipitação e do transporte de umidade na região da ZCAS através da nova geração de reanálises. **Revista Brasileira de Meteorologia**, v. 27, n. 2, p. 152–162, 2012.
- RAMOS, Rodolpho Paes Leme. Precipitation characteristics in the Northeast Brazil dry region. **Journal of Geophysical Research**, v. 80, n. 12, p. 1665–1678, 1975.
- REBOITA, Michelle Simões et al. Regimes de precipitação na América do Sul: uma revisão bibliográfica. **Revista Brasileira de Meteorologia**, São José dos Campos, SP, v. 25, n. 2, p. 185–204, 2010.
- REBOITA, Michelle Simões; SANTOS, Isimar. INFLUÊNCIA DE ALGUNS PADRÕES DE TELECONEXÃO NA PRECIPITAÇÃO NO NORTE E NORDESTE DO BRASIL. **Revista Brasileira de Climatologia**, v. 15, n. 1935, p. 28–48, 2015.
- REBOUÇAS, Aldo da C. Água na região Nordeste: desperdício e escassez. **Estudos Avançados**, v. 11, n. 29, p. 127–154, 1997.
- REGIS, Vlândia Viana. **Reequilíbrio econômico-financeiro de contratos de concessão de serviços de transmissão: Estudo de casos**. 2017. FGV - Fundação Getúlio Vargas, 2017.
- ROYSTON, Patrick. Remark AS R94: A Remark on Algorithm AS 181: The W-test for Normality. **Applied Statistics**, v. 44, n. 4, p. 547, 1995.
- SANTANA, Lêda Valéria Ramos. **Análise do comportamento da velocidade do vento na região Nordeste do Brasil utilizando dados da ERA-40**. 2014. Universidade Rural de Pernambuco, Recife, PE, 2014.
- SANTOS E SILVA, C. M. et al. Dynamical downscaling of the precipitation in Northeast Brazil with a regional climate model during contrasting years. **Atmospheric Science Letters**, v. 15, n. 1, p. 50–57, 2014.
- SHARYGIN, E. Yu et al. Query compilation in PostgreSQL by specialization of the DBMS

- source code. **Programming and Computer Software**, v. 43, n. 6, p. 353–365, 2017.
- SHEN, Jinmei. Numerical Modelling of the Effects of Vegetation and Environmental Conditions on the Lake Breeze. **Boundary-Layer Meteorology**, v. 87, n. 3, p. 481–498, 1998.
- SILVA, Bruce Kelly N. et al. Rainfall-related natural disasters in the Northeast of Brazil as a response to ocean-atmosphere interaction. **Theoretical and Applied Climatology**, n. 2010, 2019.
- SILVA, Alexsandro Aleixo Pereira Da. **Modelagem para repotencialização de linhas de transmissão através da aplicação de condutores especiais**. 2009. Universidade Federal de Pernambuco, Recife, PE, 2009.
- SILVA, Bruce Francisco Pontes Da et al. Synoptic systems associated to heavy precipitation in the state of Alagoas. **Revista Brasileira de Meteorologia**, v. 26, n. 3, p. 323–338, 2011.
- SKAMAROCK, W. C. et al. A Description of the Advanced Research WRF Version 3. **Technical Report**, p. 113, 2008.
- SOARES, BRUNO FRANCO. **Análise da metodologia aplicada às tarifas de uso dos sistemas de distribuição para geradores - TUSDg**. 2009. Universidade de São Paulo, São Paulo, SP, 2009.
- SOKAL, Robert R.; ROHLF, F. James. THE COMPARISON OF DENDROGRAMS BY OBJECTIVE METHODS. **TAXON**, v. 11, n. 2, p. 33–40, 1962.
- SOUZA, Dayana Castilho. **Brisa Na Costa Do Norte E Nordeste Brasileiro**. 2016. INPE, São Paulo, SP, 2016.
- STULL, Roland B. **An Introduction to Boundary Layer Meteorology**. 5. ed. Dordrecht: Springer Netherlands, 1988.
- TAREK, Mostafa; BRISSETTE, François P.; ARSENAULT, Richard. Evaluation of the ERA5 reanalysis as a potential reference dataset for hydrological modeling over North-America. **Hydrology and Earth System Sciences Discussions**, p. 1–35, 2019.
- TOVAR-PESCADOR, Joaquin et al. A comparative study of ordinary and residual kriging techniques for mapping global solar radiation over southern Spain. **Agricultural and Forest Meteorology**. v. 149, p. 1343–1357, 2009.
- UNAL, Yurdanur; KINDAP, Tayfun; KARACA, Mehmet. Redefining the climate zones of Turkey using cluster analysis. **International Journal of Climatology**, v. 23, n. 9, p. 1045–1055, 2003.
- VIEIRA, Cicero Fernandes Almeida. **Sensibilidade do modelo WASP aos efeitos da estabilidade atmosférica no Nordeste do Brasil**. 2008. Universidade Federal do Ceará,

Fortaleza, CE, 2008.

WANG, Xianfeng et al. Wet periods in northeastern Brazil over the past 210 kyr linked to distant climate anomalies. **Nature**, v. 432, n. 7018, p. 740–743, 2004.

WATTERSON, I. G. The diurnal cycle of surface air temperature in simulated present and doubled CO<sub>2</sub> climates. **Climate Dynamics**, v. 13, n. 7–8, p. 533–545, 1997.

WILKS, Daniel S. **Statistical methods in the atmospheric sciences**. International Geophysics Series. 2011.

WU, Tingting; LI, Yingru. Spatial interpolation of temperature in the United States using residual kriging. **Applied Geography**, v. 44, p. 112–120, 2013.

XU, Zhongfeng; YANG, Zong-Liang. An Improved Dynamical Downscaling Method with GCM Bias Corrections and Its Validation with 30 Years of Climate Simulations. **Journal of Climate**, v. 25, n. 18, p. 6271–6286, 2012.

YIP, Tony et al. Dynamic Line Rating Protection for Wind Farm Connections. In: **CIGRE/IEEE PES JOINT SYMPOSIUM INTEGRATION OF WIDE-SCALE RENEWABLE RESOURCES INTO THE POWER DELIVERY SYSTEM**. 2009, Calgary, AB.

ZHANG, Da-Lin; ZHENG, Wei-Zhong. Diurnal Cycles of Surface Winds and Temperatures as Simulated by Five Boundary Layer Parameterizations. **Journal of Applied Meteorology**, v. 43, n. 1, p. 157–169, 2004.

ZHOU, Chunlüe; WANG, Kaicun; MA, Qian. Evaluation of Eight Current Reanalyses in Simulating Land Surface Temperature from 1979 to 2003 in China. **Journal of Climate**, v. 30, n. 18, p. 7379–7398, 2017.