

AMANDA CHOU PAZO BLANCO

Análise estatística da influência de anomalias de temperatura do Oceano Atlântico na predição da velocidade do vento em parques eólicos brasileiros

PROJETO DE GRADUAÇÃO EM ENGENHARIA DE PRODUÇÃO
APRESENTADO AO DEPARTAMENTO DE ENGENHARIA INDUSTRIAL
DA PUC-RIO, COMO PARTE DOS REQUISITOS PARA OBTENÇÃO
DO TÍTULO DE ENGENHEIRO DE PRODUÇÃO

Orientadora: Paula Medina Maçaira Louro
Coorientador: Fernando Luiz Cyrino Oliveira

Departamento de Engenharia Industrial
Rio de Janeiro, 12 de junho de 2024.

Agradecimentos

Agradeço à Deus por ter me proporcionado todas as condições para concluir essa jornada.

Aos meus pais, meu irmão e minha avó pelo amor, apoio e incentivo incondicionais.

Aos orientadores Paula Maçaira e Fernando Cyrino pela atenção dedicada durante a elaboração deste trabalho. A todos os professores da PUC-Rio que contribuíram para minha formação acadêmica e me inspiraram ao longo do curso. Um agradecimento especial ao professor Cristiano Fernandes, da disciplina “Análise de Séries Temporais”, por ensinar a teoria dos modelos ARIMA e a como implementá-los utilizando o *software* R. Sem suas aulas, concluir este trabalho teria sido muito mais difícil.

À pesquisadora titular do Instituto Nacional de Pesquisas Espaciais (INPE), Chou Sin Chan, por me apresentar os principais sistemas meteorológicos que atuam na região Nordeste do Brasil e por sugerir a utilização de um indicador de temperatura do Oceano Atlântico no modelo estatístico. Sua sugestão foi fundamental para a definição do tema deste trabalho.

Aos colegas do estágio na área de Pesquisa Quantitativa do BOCOM BBM por todos os ensinamentos, pelo apoio e por despertarem meu interesse por modelos estatísticos. Agradeço especialmente ao Daniel Reis, Thales Costa e João Luiz Neto.

Por fim, agradeço a todos os meus amigos principalmente Adailton Nascimento, Bruna Vallejos, Carolina Silva, Guilherme Ferrari, Laura Lopes, Leonardo Dias, Pedro Velozo, Rachel Szenberg, Simeon Georgiev e Sofia Nunes que foram muito importantes para mim ao longo dessa jornada e me deram forças para continuar nesta etapa final.

Consegui concluir esse ciclo graças a essa base forte que tive e tenho junto comigo e que contribuíram em muito para o meu desenvolvimento tanto pessoal quanto profissional.

Resumo

Prever a velocidade do vento de forma acurada é um desafio, uma vez que os ventos locais são influenciados por diferentes fenômenos atmosféricos de pequena, média e grande escala. As mudanças climáticas agravam esse desafio, podendo transformar um parque eólico atualmente produtivo em improdutivo no futuro. Então, para continuar impulsionando o crescimento da participação eólica na matriz energética brasileira, previsões mais precisas para estimar a geração energética são essenciais. Uma alternativa é considerar variáveis climáticas na modelagem. Nesse contexto, este estudo verificou inicialmente a correlação entre as séries de velocidade do vento de cada um dos parques eólicos — Alegria II, Praia Formosa e Ventos de São Januário 19, todos situados na região Nordeste do Brasil — e os índices *Tropical Northern Atlantic* (TNA) e *Tropical Southern Atlantic* (TSA), que monitoram as anomalias de temperatura no Oceano Atlântico. Os resultados mostraram que, para os parques Alegria II e Praia Formosa, o TSA defasado em um período apresentou uma alta correlação significativa. Para o parque Ventos de São Januário, a maior correlação foi observada com o TSA defasado em dois períodos. Em seguida, visando capturar os efeitos das mudanças do clima na previsão da velocidade do vento, utilizaram-se esses índices defasados como variável exógena no modelo SARIMAX. A aplicação do teste de Diebold-Mariano revelou que, para o ano de 2023, os modelos SARIMA e SARIMAX apresentaram capacidades preditivas similares na velocidade dos ventos do parque Ventos de São Januário 19, enquanto para os demais parques, o SARIMAX superou o SARIMA apenas no último período de dados, com previsões para os meses de outubro, novembro e dezembro de 2023.

Palavras-chave

Velocidade do Vento. Modelo SARIMAX. TNA. TSA. Nordeste do Brasil.

Abstract

Accurately predicting wind speed is challenging since local winds are influenced by atmospheric phenomena of small, medium, and large scales. Climate change exacerbates this challenge, potentially turning a currently productive wind farm into an unproductive one in the future. Therefore, to continue driving the growth of wind participation in the Brazilian energy matrix, more precise forecasts to estimate energy generation are essential. One alternative is to consider climatic variables in modeling. In this context, this study initially verified the correlation between the wind speed series of each of the wind farms—Alegria II, Praia Formosa, and Ventos de São Januário 19, all located in the Northeast region of Brazil—and the indices Tropical Northern Atlantic (TNA) and Tropical Southern Atlantic (TSA), which monitor temperature anomalies in the Atlantic Ocean. The results showed that, for the Alegria II and Praia Formosa farms, the TSA lagged by one period exhibited a highly significant correlation. For the Ventos de São Januário farm, the highest correlation was observed with the TSA lagged by two periods. Subsequently, aiming to capture the effects of climate changes on wind speed forecasting, these lagged indices were used as an exogenous variable in the SARIMAX model. The application of the Diebold-Mariano test revealed that, for the year 2023, the SARIMA and SARIMAX models showed similar predictive capabilities for the wind speeds of the Ventos de São Januário 19 farm, while for the other farms, SARIMAX outperformed SARIMA only in the last data period, with forecasts for the months of October, November, and December 2023.

Keywords

Wind speed. SARIMAX model. TNA. TSA. Brazilian northeast.

Sumário

1	Introdução	8
2	Referencial Teórico	10
2.1	Mudanças climáticas	10
2.2	Circulação Geral da Atmosfera	11
2.3	Regime de Ventos no Nordeste Brasileiro (NEB)	12
2.3.1	Localização	12
2.3.2	Principais características e sistemas meteorológicos atuantes	13
2.4	Padrões de Teleconexão	15
2.4.1	<i>Tropical Atlantic Dipole</i> (TAD)	16
2.5	Variáveis exógenas	18
2.5.1	Revisão da literatura	19
2.6	Modelos ARIMA	20
2.6.1	Modelo Autorregressivo (AR)	20
2.6.2	Modelo Médias Móveis (MA)	21
2.6.3	Modelo Autorregressivo Médias Móveis (ARMA)	21
2.6.4	Modelo Autorregressivo Integrado de Médias Móveis (ARIMA)	22
2.6.5	Modelo Sazonal Autorregressivo Integrado de Médias Móveis (SARIMA)	23
2.6.6	Modelo Sazonal Autorregressivo Integrado de Médias Móveis com Variáveis Exógenas (SARIMAX)	24
3	Metodologia	26
3.1	Pré-processamento	26
3.2	Análise descritiva dos dados	26
3.2.1	Velocidade do vento	26
3.2.2	Indicadores TNA e TSA	28
3.2.3	Relação entre as velocidades do vento e os indicadores TNA e TSA	30
3.3	Metodologia Box & Jenkins	32
4	Resultados e Discussões	34
4.1	Identificação das ordens do modelo e estimação dos parâmetros	35
4.1.1	Parque eólico Alegria II	37
4.1.2	Parque eólico Praia Formosa	38

4.1.3	Parque eólico Ventos de São Januário 19	39
4.2	Diagnósticos de adequação do modelo	40
4.2.1	Teste Ljung-Box	40
4.2.2	Teste Jarque-Bera	41
4.3	Comparação da performance preditiva dos modelos	42
5	Conclusões	46

Lista de Figuras

2.1	Escala de confiança para avaliar os impactos observados. Fonte: (IPCC, 2023)	10
2.2	Localização do Nordeste do Brasil e dos três parques eólicos selecionados. .	13
2.3	Complementaridade do vento e da precipitação no parque eólico Praia Formosa.	14
2.4	Áreas de cálculo para os índices TNA e TSA. Fonte: (WMO, 2024)	18
3.1	Séries temporais de velocidade do vento por parque eólico em frequência mensal.	27
3.2	Função de autocorrelação das séries temporais de velocidade do vento por parque eólico.	28
3.3	Gráfico de área das séries temporais dos indicadores TNA e TSA	29
3.4	Função de autocorrelação das séries dos indicadores TNA e TSA.	30
3.5	Função de correlação cruzada velocidades do vento <i>versus</i> índices.	31
3.6	Etapas da metodologia Box & Jenkins	33
4.1	Resíduos do modelo SARIMA(1,0,1)x(0,1,1) + lag1 TSA ajustado	38
4.2	Estimativa e significância estatística de cada coeficiente do modelo SARIMA(1,0,1)x(0,1,1) + lag1 TSA.	38
4.3	Resíduos do modelo SARIMA(1,0,1)x(0,1,1) + lag1 TSA ajustado.	39
4.4	Estimativa e significância estatística de cada coeficiente do modelo SARIMA(1,0,1)x(0,1,1) + lag1 TSA.	39
4.5	Resíduos do modelo SARIMA(0,0,0)x(0,1,1) + lag2 TSA ajustado.	40
4.6	Estimativa e significância estatística de cada coeficiente do modelo SARIMA(0,0,0)x(0,1,1) + lag2 TSA.	40

Lista de Tabelas

3.1	Estatísticas descritivas das séries de velocidade do vento por parque eólico.	27
3.2	Variável exógena, defasagem selecionada e valor da correlação.	32
4.1	Janelas de ajuste e previsão.	34
4.2	Resultados do teste ADF-II.	36
4.3	Valores AICc para os modelos SARIMAX aplicados às séries de velocidade do vento com a adição da série TSA defasada em um período. Modelo selecionado de forma automática: SARIMA(1,0,1) x (2,1,0) + lag1 TSA. .	37
4.4	Valores AICc para os modelos SARIMAX aplicados às séries de velocidade do vento com a adição da série TSA defasada em um período. Modelo selecionado de forma automática: SARIMA(1,0,1) x (2,1,0) + lag1 TSA. .	38

4.5	Valores AICc para os modelos SARIMAX aplicados às séries de velocidade do vento com a adição da série TSA defasada em dois períodos. Modelo selecionado de forma automática: SARIMA(0,0,0) x (2,1,0) + lag2 TSA. .	39
4.6	Resultado do teste de Ljung-Box.	41
4.7	Resultados do teste de Jarque-Bera.	42
4.8	Parque eólico Alegria II: performance dos modelos.	43
4.9	Parque eólico Praia Formosa: performance dos modelos.	43
4.10	Parque eólico Ventos de São Januário 19: performance dos modelos.	44
4.11	Resultados do teste Diebold-Mariano.	45

1. Introdução

Com as mudanças no clima, eventos climáticos extremos estão mais frequentes e intensos (IPCC, 2023). Recordes de ondas de calor, chuvas torrenciais e secas severas são alguns exemplos. Em resposta, diversos países têm se mobilizado para conter o aumento da temperatura média global em 1,5°C até o final do século XXI e lidar com seus impactos (ONU, 2015). Uma das medidas de redução de emissão de gases de efeito estufa envolve expandir o uso de fontes renováveis na matriz energética (IRENA, 2023). Nesse contexto, o investimento em fontes, como a eólica, tem aumentado significativamente no mundo, em especial, no Brasil.

Mais de 80% da matriz energética brasileira é composta por fontes renováveis. Desse total, as hidrelétricas predominam, contribuindo com 47% (ONS, 2024). Em seguida, vêm os parques eólicos, com uma participação de 14%. Esta fonte é a que cresce de forma mais acelerada no país. De 2012 a 2022, sua capacidade instalada passou de 2,5 GW para 25,63 GW, um aumento de dez vezes (ABEEÓLICA, 2022a). Desde 2021, o país ocupa o sexto lugar no mundo em termos de capacidade instalada (GWEC, 2023), com projeções indicando um alcance de 44,78 GW até 2028. Esse crescimento é liderado pela região Nordeste do Brasil (NEB) cujo fator de capacidade, quase o dobro da média mundial, é atribuído aos ventos consistentes, com intensidade adequada e velocidade e direção sem mudanças bruscas (ABEEÓLICA, 2022b).

O restante da matriz provém principalmente de termelétricas, que contribuem com 19% (ONS, 2024). Essas usinas são acionadas para complementar a demanda interna de energia, especialmente durante períodos de estiagem (ONS, 2020). Logo, para reduzir a dependência de usinas movidas a combustíveis fósseis, é necessário ampliar e/ou otimizar o uso de fontes renováveis. Adicionalmente, existe uma característica complementar entre a geração hidrelétrica e eólica no NEB que evidencia ainda mais a importância da fonte eólica. Contudo, essa fonte de energia está sujeita a incertezas ligadas às condições meteorológicas, fato que é intensificado pelas mudanças nos padrões climáticos (ONS, 2020; NOAA, 2020). Portanto, a incorporação de variáveis climáticas na modelagem da velocidade do vento pode ser uma alternativa para reduzir incertezas na geração energética prevista como apontam alguns estudos (MAÇAIRA, 2018; NETO et al., 2021).

Uma baixa precisão na estimativa da energia eólica a ser gerada compromete a estabilidade, a confiabilidade e o planejamento operacional do fornecimento de energia elétrica. Nesse sentido, com o crescimento da participação eólica na matriz energética brasileira, o modelo GEVAZP, utilizado para gerar uma série sintética multivariada de vazões para reservatórios, foi estendido para gerar uma série sintética multivariada mensal de vento,

conforme proposto por Maceira et al. (MACEIRA et al., 2021). Esses cenários de velocidades de vento alimentam o modelo principal, NEWAVE, que busca otimizar o uso dos recursos hídricos e das demais energias renováveis, minimizando os custos relacionados à geração térmica (CEPEL, 2024b).

A informação macroclimática é considerada apenas na geração de cenários hidrológicos que capturam a influência do El Niño por meio do índice *Oceanic Niño Index* (ONI) (CEPEL, 2024a). No entanto, outros índices relativos à temperatura dos oceanos poderiam ser incorporados em cenários de vento. Segundo o NOAA, a relação entre o oceano e os eventos climáticos extremos é complexa e requer um entendimento mais profundo para melhorar as previsões desses eventos. Por exemplo, foi identificado nos Estados Unidos da América (EUA) que a ocorrência de ondas de calor está relacionada às variações no padrão global de monções (NOAA, 2024b). O surgimento de projetos de implementação de eólicas offshore no Brasil (PETROBRAS, 2023) também ressalta a necessidade de analisar a conexão oceano-clima. Contudo, pesquisas que conectam a previsão da velocidade do vento a variáveis que refletem essas variações são escassas.

Nesse sentido, o presente estudo propõe modelar as velocidades de vento de três parques eólicos localizados na região Nordeste do Brasil por meio do modelo *Seasonal AutoRegressive Integrated Moving Average with exogenous variables* (SARIMAX), isto é, do modelo SARIMA com a adição de uma variável explicativa para melhor estimar os valores da velocidade de vento e, conseqüentemente, da geração eólica. Dada a maior proximidade da costa nordestina com o Oceano Atlântico, dois índices de variação de temperatura deste oceano foram testados: *Tropical Northern Atlantic Index* (TNA) e *Tropical Southern Atlantic Index* (TSA). O índice que apresenta maior significância estatística com a série de velocidades do vento é utilizado como variável exógena, a fim de capturar os efeitos das mudanças climáticas. Desse modo, busca-se responder as seguintes perguntas de pesquisa:

- Fenômenos de larga escala oriundos da variação de temperatura do Oceano Atlântico têm alta correlação com as séries de velocidade de vento no Nordeste do Brasil?
- A utilização do TNA ou TSA como variável exógena no modelo SARIMAX torna a estimação de séries de velocidades de vento mais precisa?

A estrutura do estudo está organizado em cinco seções. Após a Introdução, o Referencial Teórico apresenta o cenário atual das mudanças climáticas e desenvolve a base conceitual para a compreensão do regime de ventos no NEB, assim como as anomalias do Oceano Atlântico e suas interações com o clima. Introduz também o conceito de variáveis exógenas, apresentando uma breve revisão da literatura sobre modelos que as incorporam, e fornece a formulação matemática do modelo empregado. A terceira seção trata da Metodologia, que inclui uma análise descritiva dos dados e o método seguido. A quarta seção descreve os resultados obtidos com o ajuste do modelo e, por fim, são fornecidas as conclusões.

2. Referencial Teórico

2.1 Mudanças climáticas

Não há dúvida de que as atividades humanas, principalmente por meio das emissões de gases de efeito estufa (GEE), causaram o aquecimento da atmosfera, do oceano e da terra. Comparando o período de 2011 a 2020 com o de 1850 a 1900, houve um aumento de 1,09°C na temperatura média global da superfície. Outro dado relevante é que essa temperatura subiu mais rapidamente entre 1970 e 2020 do que em qualquer outro período de 50 anos nos últimos 2000 anos. Consequentemente, extremos climáticos e meteorológicos têm afetado todas as regiões do mundo, provocando impactos adversos na segurança alimentar e hídrica, na saúde humana e nas economias e sociedades, além de perdas e danos irreversíveis à natureza e às pessoas (IPCC, 2023).

Para classificar o grau de certeza dos impactos observados, o relatório do *Intergovernmental Panel on Climate Change* (IPCC) utiliza uma escala que varia de “média confiança” até “fato”, conforme mostrado na figura 2.1, onde “média confiança” representa o menor nível de certeza e “fato”, o maior.

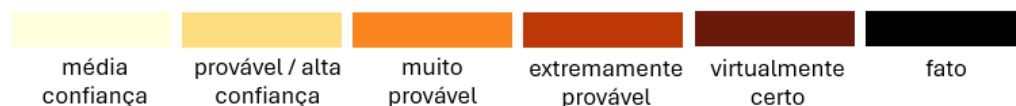


Figura 2.1: Escala de confiança para avaliar os impactos observados. Fonte: (IPCC, 2023)

Dentre os dados levantados por esse relatório, relativos aos oceanos, concluiu-se que é “virtualmente certo” que a superfície do oceano (0-700m) aqueceu desde a década de 1970. Com “alta confiança”, afirma-se que o aquecimento oceânico foi responsável por 91% do aquecimento no sistema climático, sendo o aquecimento terrestre, a perda de gelo e o aquecimento atmosférico correspondendo a cerca de 5%, 3% e 1%, respectivamente. Quanto à circulação atmosférica e o clima, é “virtualmente certo” que extremos de calor tornaram-se mais frequentes e mais intensos na maioria das áreas terrestres desde a década de 1950, enquanto extremos de frio tornaram-se menos frequentes e menos severos. Com “alta confiança”, observou-se também um aumento na frequência e intensidade dos eventos de precipitação em várias regiões do globo desde 1950.

A perspectiva é que o aquecimento global continue aumentando no curto prazo (2021-2040) principalmente em decorrência do crescimento das emissões acumuladas de CO₂ em quase todos os cenários considerados. Acredita-se com “alta confiança” que essa continuidade

intensificará ainda mais o ciclo global da água, provocando, por exemplo, estações e eventos de tempo e clima bastante úmidos e secos. Projeta-se também que cada região experimente mais mudanças simultâneas e diversas de fatores relacionados ao clima, agravando os impactos. Outra alteração prevista, mas com “média confiança”, é a intensificação de ciclones tropicais e/ou tempestades extratropicais.

Dessa forma, reduzir as perdas e danos previstos para seres humanos e ecossistemas implica limitar o aquecimento global a 1,5° até o final do século XXI, o que requer emissões líquidas zero de CO₂. Para atingir essa meta, ações globais rápidas e profundas que reduzam essas emissões são essenciais. Um dos principais tratados internacionais que busca alcançar esse objetivo é o Acordo de Paris, adotado em 2015 pelos países signatários da *United Nations Framework Convention on Climate Change* (UNFCCC). Neste acordo, as nações estabelecem as próprias metas para reduzir a emissão de gases de GEE, chamadas de *Nationally Determined Contribution* (NDC). A NDC brasileira, revisada em 2023, determina que o Brasil deve diminuir suas emissões em 48% até 2025 e 53% até 2030, comparado aos níveis de 2005. Além disso, o país reafirmou seu compromisso de alcançar emissões líquidas neutras até 2050.

Em meio à transição de combustíveis fósseis para fontes de energia de muito baixo ou zero carbono, o Brasil ocupa uma posição privilegiada graças à sua matriz energética predominantemente limpa, com uma forte presença de hidrelétricas. Contudo, os impactos das mudanças climáticas no sistema hidrológico, que afetam diretamente a geração hidrelétrica, destacam a necessidade de continuar diversificando a matriz com outras fontes renováveis, como a eólica (GOV.BRAZIL, 2023).

No entanto, a produção de energia por meio dessa fonte enfrenta desafios. Ao mesmo tempo em que ela contribui para a mitigação das mudanças climáticas, também é influenciada por essas mudanças. Portanto, é fundamental que as estratégias de expansão eólica no Brasil também incluam estudos relacionados à variabilidade do clima, de modo a manter a eficiência e sustentabilidade dos parques eólicos no futuro.

2.2 Circulação Geral da Atmosfera

A superfície da Terra é aquecida pelo Sol de maneira desigual ao longo do dia e entre as estações do ano. A diferença de temperatura e consequente mudança na densidade e pressão do ar desenvolve um gradiente de pressão horizontal, que é a principal causa do movimento do ar de áreas de alta pressão para áreas de baixa pressão (BARRY; CHORLEY, 2012). Portanto, o vento pode ser definido como o ar em movimento tentando equalizar as diferenças de pressão.

Os ventos variam em diferentes escalas espaciais e temporais, sendo classificados como macro (global), meso (regional) e micro (local) (ORLANSKI, 1976).

- As circulações em macroescala possuem dimensões maiores que 100 km e podem durar de um dia a meses. Elas são formadas principalmente pelo aquecimento desigual entre o Equador e os Polos e pelo movimento de rotação da Terra. Também desempenham um papel importante na definição das características climáticas e sazonais de diferentes regiões do planeta. Furacões, ciclones e anticiclones estão incluídos nesta categoria.
- As circulações em mesoescala variam de 1 a 100 km e têm duração de 1 hora a 1 dia. Nesses níveis, a topografia influencia os ventos. Exemplos incluem brisas marítimas e terrestres, Sistemas Convectivos de Mesoescala (SCM) e Linhas de Instabilidade (LI), que serão explicados na Seção 2.3.
- As circulações em microescala variam de 1 metro a 1 km e duram de segundos a minutos. Fatores como a superfície do solo, a cobertura vegetal e o relevo afetam significativamente essas circulações. Alguns exemplos são os movimentos turbulentos e as rajadas de vento (Atlas Eólico RN, 2022).

Apesar da categorização em escalas, os fenômenos atmosféricos transcorrem simultaneamente, interagindo e impactando uns aos outros. Os ventos próximos à superfície, que são os de interesse para a geração de energia, resultam da sobreposição das circulações atmosféricas nessas diferentes escalas espaço-temporal (Atlas Eólico BA, 2013).

2.3 Regime de Ventos no Nordeste Brasileiro (NEB)

Os padrões de ventos em determinada região são influenciados pelos sistemas meteorológicos atuantes, bem como pelas condições de TSM no Pacífico Equatorial e no Atlântico Tropical. Consequentemente, eles apresentam um forte ciclo diurno, além de variações sazonais e interanuais (RAO; LIMA; FRANCHITO, 1993).

Neste contexto, a seção é dividida em duas partes. Primeiro, descreve-se a localização da região em análise e, em seguida, abordam-se os principais sistemas meteorológicos que afetam o regime de ventos.

2.3.1 Localização

O NEB está localizado aproximadamente entre as latitudes 1°S e 18°S e longitudes 35°W e 47°W, abaixo da linha do Equador e a leste da Floresta Amazônica. A região é composta pelos estados de Alagoas, Bahia, Ceará, Maranhão, Paraíba, Pernambuco, Piauí, Rio Grande do Norte e Sergipe, totalizando uma área de 1.552.175 km² (IBGE, 2022). Além disso, é banhada pelo Oceano Atlântico tanto ao norte quanto à leste.

Os três parques eólicos em estudo são: Praia Formosa, Alegria II e Ventos de São Januário 19. O parque eólico Praia Formosa está localizado no município de Camocim, no estado do Ceará (CE). Alegria II situa-se em Guamaré, Rio Grande do Norte (RN). Ventos de

São Januário localiza-se em Morro do Chapéu, na Bahia (BA). A seguir, na figura 2.2, é possível observar a localização da região como um todo e dos três parques eólicos.

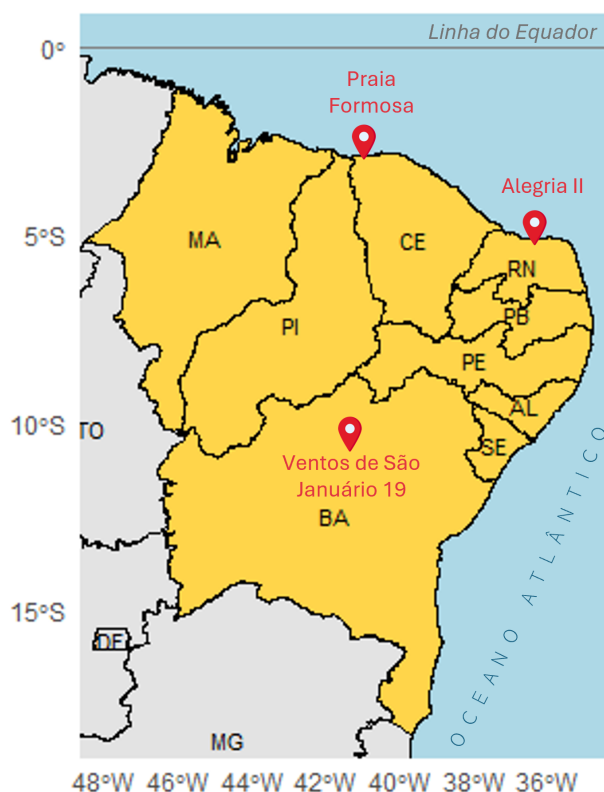


Figura 2.2: Localização do Nordeste do Brasil e dos três parques eólicos selecionados.

2.3.2 Principais características e sistemas meteorológicos atuantes

O NEB apresenta uma maior incidência de ventos mais intensos durante o inverno e a primavera do Hemisfério Sul (HS), de junho a dezembro. Por esse motivo, ocorre um aumento no fator de capacidade, que corresponde ao percentual de geração de uma usina eólica em relação a sua capacidade total instalada. Ao contrário, durante o verão, de dezembro a março, esse fator diminui. A redução é ainda mais acentuada no outono, de março a junho. Durante esse período de ventos mais fracos, inicia-se a estação chuvosa (ONS, 2020).

Diversos estudos sobre o potencial energético da região exploram essa complementaridade sazonal dos padrões hidrológicos e eólicos (MARINHO; AQUINO, 2009; WITZLER, 2014; AMARANTE et al., 2001). Essa característica levou ao uso combinado de parques eólicos e usinas hidrelétricas no processo de geração de energia. Em 2021, por exemplo, durante o agravamento da crise hídrica que atingiu todo o país, principalmente devido ao fenômeno La Niña, os parques eólicos do NEB geraram energia suficiente para abastecer toda a

região e até exportar excedentes para os subsistemas de outras regiões do país (ONS, 2022).

A fim de ilustrar esse comportamento, as séries de vento do parque eólico de Praia Formosa e as séries de precipitação total de 1º de janeiro de 2023 a 31 de dezembro de 2023 foram plotadas na figura 2.3 abaixo. Os dados dessas séries foram coletados do banco de dados de reanálise MERRA-2, disponibilizado pela NASA (NASA, 2024).

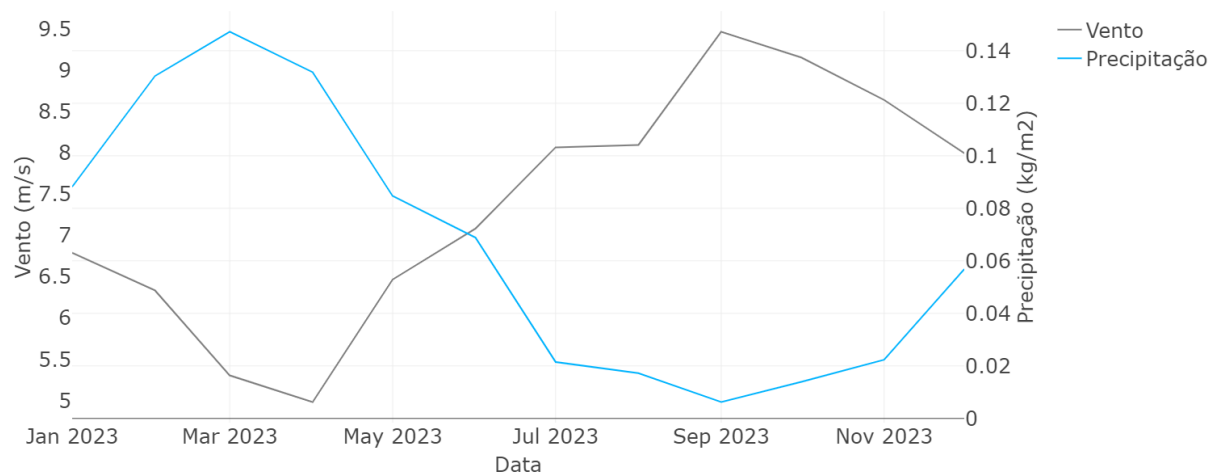


Figura 2.3: Complementaridade do vento e da precipitação no parque eólico Praia Formosa.

Em termos meteorológicos, a presença dessa complementaridade pode ser explicada em grande parte pela dinâmica da Zona de Convergência Intertropical (ZCIT), uma banda de nuvens que circunda o globo terrestre na região equatorial e que é considerada o principal sistema meteorológico atuante sobre o tempo e o clima em diversas áreas localizadas nos trópicos, em especial, no litoral norte do NEB (INMET, 2023). A ZCIT desloca-se de sua posição mais ao norte, onde permanece principalmente entre agosto e setembro, para sua posição mais ao sul, onde se mantém principalmente entre março e abril (UVO, 1989). Essa migração sazonal, junto com os fatores que influenciam o fortalecimento ou enfraquecimento dos ventos alísios, desempenha um papel fundamental na determinação da estação chuvosa/seca na região (CAVALCANTI et al., 2009).

Outros importantes sistemas meteorológicos atuam no regime de ventos da região, interagindo com a ZCIT e entre si. Dentre os sistemas de macroescala destacam-se a Alta Subtropical do Atlântico Sul (ASAS), os Vórtices Ciclônicos de Altos Níveis (VCAN) e os Distúrbios Ondulatórios de Leste (DOL). A ASAS é um sistema de alta pressão semipermanente no oceano Atlântico Sul, associada à convergência de ventos em superfície e que contribui diretamente com a manutenção da ZCIT. Quanto maior sua proximidade com o continente sul-americano, mais intenso são os ventos que chegam no litoral do

NEB. Essa maior proximidade com a América do Sul ocorre nos meses de inverno do HS (MOSCATI; GAN, 2007).

Os VCAN são sistemas ciclônicos com núcleo mais frio do que a periferia que se desenvolvem na troposfera. Devido a essa característica de núcleo mais frio, a precipitação é observada apenas nas áreas ao redor do sistema. Consequentemente, os VCAN conseguem interferir na variabilidade sazonal do vento no NEB (FEDOROVA et al., 2016). Os vórtices ciclônicos de origem tropical formam-se nos meses de primavera (KOUSKY; GAN, 1981). Já nos meses de outono e especialmente no inverno, os Distúrbios Ondulatórios de Leste (DOL), também referido como ondas de leste, perturbam o padrão de vento e precipitação na costa leste do NEB (CHAN, 1990). Os DOL originam-se na costa oeste do continente africano e seguem o fluxo dos ventos alísios, contribuindo significativamente para os totais anuais pluviométricos da costa leste do NEB.

O principais sistemas de mesoescala que atuam na região são os Sistemas Convectivos de Mesoescala (SCM) e Linhas de Instabilidade (LI) (OLIVEIRA, 2012), responsáveis por ventos fortes em superfície e tempestades. As brisas marítimas e terrestres também são fenômenos frequentemente observados. A brisa marítima origina-se durante o dia quando a superfície terrestre se aquece mais do que o oceano. O gradiente térmico e consequente gradiente de pressão em níveis acima da superfície impulsionam o movimento do ar do mar para a terra. A brisa terrestre, por sua vez, acontece no período noturno, quando verifica-se a inversão do sentido da circulação. Neste caso, o ar próximo a superfície desloca-se da terra para o mar (Atlas Eólico RN, 2022).

Já a variabilidade interanual da ZCIT está ligada à variação no seu posicionamento típico em decorrência da atuação de padrões de teleconexão, que serão aprofundados na Seção 2.4 a seguir. Quando a ZCIT posiciona-se mais ao norte da sua localização média típica durante as estações de verão e outono, o ano é mais seco. Em contrapartida, quando atinge posições mais ao sul nesse mesmo período, o ano é mais chuvoso (CAVALCANTI et al., 2009).

2.4 Padrões de Teleconexão

Fontes oceânicas de calor perturbam a atmosfera, que por sua vez dispersa ondas na tentativa de retornar ao seu estado inicial. Essas ondas afetam o clima e o tempo em regiões distantes e sua dinâmica estabelece padrões de teleconexão (REBOITA et al., 2021). Esses padrões podem durar semanas e meses, mas quando perduram por anos consecutivos, revelam uma parte importante da variabilidade interanual e interdecenal da circulação atmosférica. Segundo Cavalcanti et al (2009), eles podem explicar como anomalias que se desenvolvem em uma região específica são associadas a anomalias ocorridas em outro tempo ou espaço (CAVALCANTI et al., 2009).

Um dos mais importantes e mais estudados padrões de teleconexão é o fenômeno El Niño–Southern Oscillation (ENSO) pela sua capacidade de alterar a circulação atmosférica global, afetando a temperatura e a precipitação em várias regiões do mundo. Este fenômeno climático natural de grande escala envolve flutuações periódicas na temperatura da superfície do mar (TSM) no Oceano Pacífico Equatorial Central e Leste, bem como na pressão atmosférica acima desta região. O ENSO manifesta-se em ciclos irregulares que variam de 2 a 7 anos e apresenta três fases distintas: El Niño, La Niña e uma fase neutra (NWS, 2023).

A fase El Niño refere-se ao aquecimento acima da média da superfície do mar. Em geral, quanto mais alta for a temperatura oceânica, mais forte é o El Niño. Enquanto a fase La Niña é exatamente o oposto, tratando-se de um resfriamento abaixo da média. Nesse caso, quanto mais baixa for a temperatura oceânica, mais forte é o La Niña. Já a fase neutra ocorre quando nenhum destes eventos está presente ou quando um deles parece estar atuando, mas não apresenta uma resposta atmosférica correspondente, ou vice-versa (NOAA, 2024a). No Brasil, por exemplo, eventos de El Niño impactam o clima, levando a chuvas acima do normal na região Sul e chuvas abaixo do normal na região Nordeste. Em períodos de La Niña, ocorre o contrário.

No entanto, além das anomalias do Oceano Pacífico, as anomalias de TSM do Oceano Atlântico também influenciam a variabilidade de precipitação e de vento no Nordeste do Brasil (NEB) (CAVALCANTI et al., 2009). Inúmeros estudos apontam o gradiente anômalo de TSM entre o Atlântico Tropical Norte e o Atlântico Tropical Sul como o principal fator que influencia a estação chuvosa nessa região (HASTENRATH; HELLER, 1977; MARKHAM; MCLAIN, 1977; FOLLAND; PALMER; PARKER, 1986).

2.4.1 *Tropical Atlantic Dipole (TAD)*

O padrão de teleconexão *Tropical Atlantic Dipole* (TAD) refere-se à configuração frequente de um dos principais modos climáticos da Variabilidade do Atlântico Tropical (VAT), denominado *Atlantic Meridional Mode* (AMM). Esse modo manifesta-se através de uma zona de águas mais quentes no Atlântico Tropical Norte e uma zona complementar de águas mais frias no Atlântico Tropical Sul, ou vice-versa. Em outras palavras, o AMM caracteriza-se por apresentar anomalias de TSM antissimétricas em relação ao Equador, estabelecendo, assim, um gradiente meridional de temperatura inter-hemisférico sobre o Atlântico Equatorial (REBOITA et al., 2021).

Devido à falta de consenso sobre sua verdadeira natureza, o TAD também é conhecido como *decadal tropical Atlantic SST variability* (National Academies of Sciences, Engineering, and Medicine, 1998). No entanto, Ruiz-Barradas et al. (2000) observaram que, considerando um sistema acoplado de variáveis atmosféricas e oceânicas, há evidências da

presença desse padrão dipolar. Eles descobriram, por exemplo, que uma anomalia positiva de temperatura da superfície do mar (TSM) no Hemisfério Norte está ligada a uma anomalia de estresse de vento em direção ao norte e uma subsequente circulação ciclônica (anti-horária) nas regiões subtropicais (RUIZ-BARRADAS; CARTON; NIGAM, 2000).

Outros estudos também identificaram a influência das fases do TAD na atmosfera. A fase quente ou positiva do TAD no Atlântico Tropical Norte é marcada por anomalias positivas de TSM, enfraquecimento dos ventos alísios de nordeste e anomalias negativas de pressão ao nível do mar. Simultaneamente, no Atlântico Tropical Sul, observam-se anomalias negativas de TSM, intensificação dos ventos alísios de sudeste e anomalias positivas de pressão ao nível do mar. De forma inversa, a fase fria ou negativa do TAD apresenta anomalias positivas de TSM no Atlântico Tropical Sul e negativas no Atlântico Tropical Norte (MOURA; SHUKLA, 1981; NOBRE; SHUKLA, 1996; UVO et al., 1998; REBOITA et al., 2021).

Em relação ao impactos das fases do TAD especificamente no clima do NEB, Nobre et al. (1996), Hastenrath et al. (2006), Foltz et al. (2019) mostraram que anomalias positivas de TSM no Atlântico Tropical Sul (fase negativa do TAD) estão relacionadas com o aumento da evaporação no oceano, o que intensifica o fluxo de ventos alísios do sul e faz o fluxo de umidade convergir, aumentando a precipitação. Enquanto anomalias negativas (fase positiva do TAD), contribuem para condições mais secas por promover o deslocamento da ZCIT para o norte (NOBRE; SHUKLA, 1996; HASTENRATH, 2006; FOLTZ; BRANDT; RICHTER, 2019).

Hastenrath (1978) também observou anomalias positivas de TSM no Pacífico Leste (HASTENRATH; LAMB, 1978). Desde então a variabilidade das TSMs no Atlântico Tropical tem sido relacionada ao fenômeno ENSO. Pezzi e Cavalcanti (2001) verificaram que eventos de El Niño, quando associados a condições positivas do TAD, resultaram em um déficit de precipitação no NEB. Esses mesmos eventos quando ligados a condições negativas do TAD levaram a uma redução das chuvas na região, ao passo que foram registradas anomalias positivas de precipitação no norte do NEB. Por outro lado, eventos de La Niña associados à fase positiva do TAD resultaram em condições mais secas na região. Quando associados à fase negativa do TAD, observou-se uma quantidade excessiva de chuvas (PEZZI; CAVALCANTI, 2001).

Nesse contexto, para o monitoramento dessas anomalias dipolares de TSM, Enfield et al. (1999) definiram dois índices: o *Tropical Northern Atlantic* (TNA) e o *Tropical Southern Atlantic* (TSA). O índice TNA é determinado pela média das TSMs na região retangular entre 5°N a 25°N e 55°W a 15°W. Já o TSA corresponde à média das TSMs na área de 20°S a 0° a 30°W a 10°E, conforme mostra a figura 2.4 abaixo. Os autores identificaram que, embora configurações não-dipolares sejam mais frequentes, essa relação dipolar foi

observada em períodos de 8 a 12 anos para as estações de inverno e primavera do Hemisfério Norte, e de 2 a 3 anos para as estações de verão e outono desse mesmo hemisfério (ENFIELD et al., 1999).

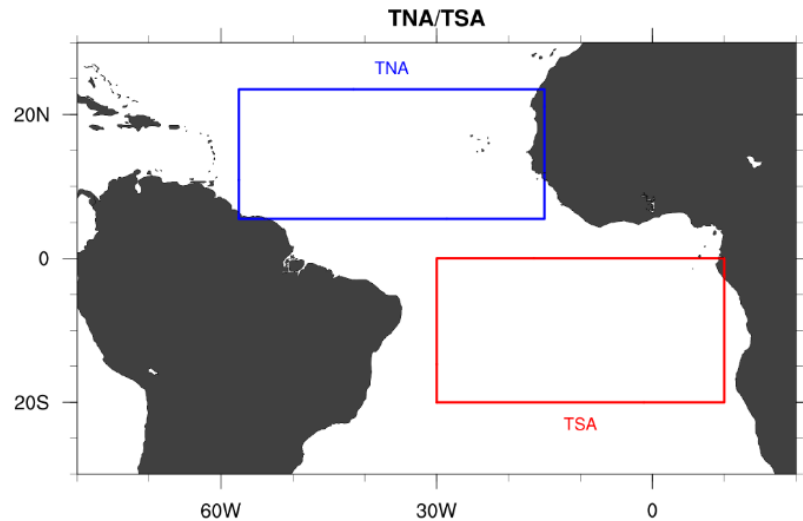


Figura 2.4: Áreas de cálculo para os índices TNA e TSA. Fonte: (WMO, 2024)

2.5 Variáveis exógenas

Segundo James H. Stock e Mark W. Watson (2020), uma variável exógena pode ser definida como uma variável que não é correlacionada com o termo de erro, enquanto uma variável endógena é correlacionada. Em termos mais simples, uma variável exógena é determinada fora do modelo, enquanto uma variável endógena é determinada dentro do modelo (STOCK; WATSON, 2020).

De acordo com Rob J. Hyndman e George Athanasopoulos (2021), a inclusão de uma ou mais variáveis exógenas em um modelo pode explicar parte de variações históricas e pode levar a previsões mais precisas. Por exemplo, pode-se desejar prever mudanças no consumo pessoal (série z_t) com base em variações de renda (variável exógena x) (HYNDMAN; ATHANASOPOULOS, 2021). Nesse contexto, o presente estudo propõe a adição dos indicadores, TNA ou TSA, como variável exógena na tentativa de gerar de forma mais precisa cenários sintéticos de velocidade do vento.

Desde 2021, a modelagem da incerteza do vento no Brasil é realizada por meio do ajuste de um modelo autorregressivo periódico de ordem p , PAR(p), às velocidades médias mensais do vento de Parques Eólicos Equivalentes (PEEs), que são agrupamentos de parques eólicos com padrões de vento semelhantes. Os resíduos do modelo são ajustados a distribuições de Weibull e transformados em espacialmente correlacionados para preservar as dependências espaciais entre as séries de velocidades de vento e as séries de afluências

(MACEIRA et al., 2021). Na geração de cenários sintéticos de ventos especificamente, variáveis exógenas não são incluídas.

Portanto, para um melhor embasamento deste estudo, foi realizada uma breve revisão da literatura sobre modelos que incorporam variáveis climáticas.

2.5.1 Revisão da literatura

Lima et al. (2014) propuseram um modelo Linear Dinâmico (DLM) que inclui componentes sazonais e autorregressivos, assim como variáveis climáticas como precipitação e o indicador Niño 3.4 do El Niño, para prever as vazões em reservatórios brasileiros por um período de cinco anos. Os resultados mostraram que a inclusão do indicador Niño 3.4 não melhorou as previsões, enquanto o DLM com dados de precipitação mostrou-se superior ao modelo usado no Brasil (LIMA; POPOVA; DAMIEN, 2014). Com base neste estudo, Maçaira (2018) identificou que os modelos Periódico Aditivo Generalizado (PGAM), utilizado para relacionar as variáveis climáticas e hidrológicas de forma não-linear, e Periódico Autoregressivo com diferentes indicadores do El Niño (PARX) superaram a performance dos modelos propostos por Lima et al. em todas as bacias e o modelo vigente (Periódico Autoregressivo) PAR (MAÇAIRA, 2018).

Cadenas et al. (2015) utilizaram o modelo *Nonlinear Autoregressive with Exogenous inputs* (NARX) para prever a velocidade do vento com 1 h de antecedência na cidade de La Mata, no México. Eles identificaram, por meio do teste de Granger, que informações passadas de radiação solar possuem valor preditivo significativo para a velocidade do vento. Assim, a radiação solar foi incorporada como variável exógena no modelo. Comparativamente, o modelo NARX mostrou-se superior, superando o modelo NAR em 4% e o de persistência em 11% em termos de desempenho (CADENAS et al., 2017). De Mattos Neto et al., ao realizarem a previsão horária para a série de velocidades do vento para o estado do Recife, por exemplo, através de modelos lineares e não-lineares, observaram que o modelo de *Machine Learning, Support Vector Regression* (SVR), obteve o melhor valor de MSE, enquanto o melhor MAPE foi alcançado pelos modelos ARIMAX e SARIMAX.

Este estudo utiliza o modelo SARIMAX para prever as velocidades do vento dos parques eólicos Alegria II (RN), Praia Formosa (CE) e Ventos de São Januário 19 (BA), incorporando um dos índices de TSM como variável exógena, abordado na Seção 2.4.1. Embora modelos lineares não capturem a complexidade do vento, o SARIMAX foi escolhido por considerar valores passados da série, ser simples e eficaz na modelagem de dados sazonais, além de permitir a inclusão de variáveis exógenas.

2.6 Modelos ARIMA

Os modelos da classe ARIMA, desenvolvidos por Box& Jenkins, são os mais utilizados dentre os modelos univariados, isto é, modelos que consideram apenas o histórico da própria série de tempo (BOX et al., 2015). Alguns dos requisitos essenciais para a eficácia desses modelos incluem a estacionariedade da série e resíduos que sejam independentes e normalmente distribuídos. Para guiar a seleção do mais adequado, os autores propuseram uma metodologia iterativa que será detalhada no Capítulo 3.

A seguir, são apresentados os modelos da classe ARIMA necessários para entender a formulação do modelo SARIMAX.

2.6.1 Modelo Autorregressivo (AR)

O modelo AR(p), ou Modelo Autorregressivo de ordem p, estabelece que o valor atual de uma série temporal, denotado por z_t , é determinado através de uma relação linear com seus valores passados, acrescido de um choque estocástico aditivo, também conhecido como termo de erro, a_t (BOX et al., 2015). A descrição do modelo é dada por:

$$z_t = \phi_1 z_{t-1} + \phi_2 z_{t-2} + \dots + \phi_p z_{t-p} + a_t \quad (2.1)$$

Onde:

- p é a ordem do modelo;
- ϕ_i são os parâmetros de peso, $i = 1, \dots, p$;
- a_t é ruído branco, ou seja, uma sequência de variáveis aleatórias não correlacionadas, com média zero e variância constante:

$$E[a_t] = 0 \quad (2.2)$$

$$var[a_t] = \sigma_a^2 \quad (2.3)$$

Utilizando o operador de atraso B, a equação pode ser reformulada do seguinte modo:

$$\phi(B)z_t = a_t \quad (2.4)$$

Onde:

- $B^k z_t = z_{t-k} \quad \forall k \in \{1, 2, \dots, p\}$;
- $\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$ é o operador autorregressivo de ordem p .

É importante destacar que, para um modelo autorregressivo ser eficaz na previsão, ele deve atender à condição de estacionariedade. Esta condição estipula que as raízes do polinômio

característico $\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p = 0$ devem estar localizadas fora do círculo unitário no plano complexo, ou seja, todas as raízes do polinômio devem ter módulos maiores que 1.

2.6.2 Modelo Médias Móveis (MA)

O modelo MA(q), ou Modelo de Médias Móveis de ordem q , é formado por uma combinação linear que inclui o termo de erro atual, a_t , e uma soma ponderada dos q valores passados deste termo de erro. Em outras palavras, o modelo, descrito na equação 2.5, sugere que cada valor observado na série temporal é o resultado de um choque aleatório atual mais a influência acumulada dos choques de até q períodos anteriores.

$$z_t = a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2} + \dots + \theta_q a_{t-q} \quad (2.5)$$

Onde:

- q é a ordem do modelo;
- θ_i são os parâmetros de peso, $i = 1, \dots, q$;
- a_t é ruído branco, com média zero e variância constante.

Utilizando o operador de defasagem B , a equação pode ser expressa da seguinte forma:

$$z_t = \theta(B) a_t \quad (2.6)$$

Onde:

- $\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q$ é o operador de médias móveis de ordem q .

Analogamente ao modelo AR, para um modelo MA ser capaz de representar adequadamente um distúrbio aleatório corrente como função do valor corrente e dos valores passados da série e dos parâmetros do modelo, ele deve satisfazer a condição de invertibilidade. Esta condição exige que as raízes do polinômio $1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q = 0$ tenham módulos maiores que 1.

2.6.3 Modelo Autorregressivo Médias Móveis (ARMA)

O modelo ARMA(p, q) é a combinação dos dois modelos anteriores, descrevendo a série z_t tanto por meio de seus valores passados quanto por meio dos termos de erro atual e passados. A estrutura do modelo é expressa na equação 2.7 a seguir.

$$z_t = \phi_1 z_{t-1} + \phi_2 z_{t-2} + \dots + \phi_p z_{t-p} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q} \quad (2.7)$$

Onde:

- $\phi(B)$ é o operador autorregressivo de ordem p , que satisfaz a condição de estacionariedade;
- $\theta(B)$ é o operador de médias móveis de ordem q , que satisfaz a condição de invertibilidade;
- a_t é ruído branco.

Utilizando o operador de defasagem B , obtém-se a equação 2.8 abaixo.

$$(1 - \phi_1 B - \phi_2 B^2 + \dots + \phi_p B^p)z_t = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q)a_t \quad (2.8)$$

ou

$$\phi(B)z_t = \theta(B)a_t \quad (2.9)$$

Para este modelo e seus derivados, tanto a condição de estacionariedade associada ao componente AR, quanto a condição de invertibilidade associada ao componente MA devem ser satisfeitas.

2.6.4 Modelo Autorregressivo Integrado de Médias Móveis (ARIMA)

O modelo ARIMA(p,d,q) é adequado para descrever séries não estacionárias que apresentem homogeneidade, ou seja, séries com variação no nível e na tendência, mas que quando submetidas a um número finito de diferenciações, tornam-se estacionárias.

Este modelo é uma extensão do modelo linear estacionário ARMA(p,q), expresso na equação 2.10.

$$\phi(B)W_t = \theta(B)a_t, \quad (2.10)$$

Se W_t é estacionária e igual à diferença de z_t , ou seja, $\nabla^d z_t$, então z_t é a integral de W_t . Por esse motivo, o modelo descrito na equação 2.11 abaixo chama-se Modelo Autorregressivo Integrado de Médias Móveis.

$$\phi(B)\nabla^d z_t = \theta(B)a_t, \quad (2.11)$$

Onde:

- $\phi(B)$ é o operador autorregressivo de ordem p ;
- $\theta(B)$ é o operador de médias móveis de ordem q ;
- $\nabla = 1 - B$ é o operador diferença, tal que:

$$z_t - z_{t-1} = z_t - Bz_t = (1 - B)z_t = \nabla z_t \quad (2.12)$$

- d é a ordem de integração, que corresponde ao número de diferenças;

- a_t é ruído branco.

2.6.5 Modelo Sazonal Autorregressivo Integrado de Médias Móveis (SARIMA)

Algumas séries temporais apresentam correlação significativa em lags sazonais, isto é, múltiplos do período s . Por exemplo, em dados mensais, altas correlações estão presentes nos lags 12, 24, 36, etc. Nesse caso, deve-se considerar uma sazonalidade estocástica (MORETTIN; TOLOI, 1981).

A sazonalidade sugere que uma observação z_t correspondente a um determinado mês, Agosto, por exemplo, está relacionada ao mês de Agosto de anos anteriores. Assim sendo, para relacionar os meses de Agosto, considere modelar uma série com frequência mensal, isto é, com periodicidade $s = 12$, por meio de um ARIMA sazonal, conforme a equação 2.13 a seguir.

$$\Phi(B^{12})\nabla_{12}^D z_t = \Theta(B^{12})\alpha_t \quad (2.13)$$

Onde:

- $\Phi(B^{12}) = 1 - \Phi_1 B^{12} - \Phi_2 B^{24} - \dots - \Phi_P B^{12P}$ é o operador autorregressivo sazonal de ordem P ;
- $\Theta(B^{12}) = 1 - \Theta_1 B^{12} - \Theta_2 B^{24} - \dots - \Theta_Q B^{12Q}$ é o operador autorregressivo sazonal de ordem Q ;
- ∇_{12}^D é o operador diferença sazonal, tal que:

$$\nabla_{12}^D = (1 - B^{12})^D \quad (2.14)$$

sendo D , o número de diferenças sazonais.

Para relacionar os meses de Julho, o modelo seria similar à equação 2.13:

$$\Phi(B^{12})\nabla_{12}^d z_{t-1} = \Theta(B^{12})\alpha_{t-1} \quad (2.15)$$

Entretanto, os componentes de erro α_t , α_{t-1} e, assim por diante, não seriam ruído branco, ou seja, possuiriam correlação. Nesse caso, o mês de Agosto, além de ser relacionado com Agosto de anos anteriores, também é relacionado com Julho, Junho, e outros meses.

Então, dado o ARIMA usual, definido na equação 2.11, onde a_t é ruído branco e substituindo-a na equação 2.13, obtém-se o Modelo Sazonal Multiplicativo de ordem $(p, d, q) \times (P, D, Q)_s$, mais conhecido como SARIMA(p, d, q) $\times$ (P, D, Q) $_s$ (BOX et al., 2015). Este é matematicamente expresso da seguinte forma:

$$\phi(B)\Phi(B^s)\nabla^d\nabla_s^D z_t = \theta(B)\Theta(B^s)\alpha_t \quad (2.16)$$

Onde:

- s é a periodicidade da sazonalidade;
- $\phi(B)$ é o operador autorregressivo de ordem p ;
- $\theta(B)$ é o operador de médias móveis de ordem q ;
- $\Phi(B^s)$ é o operador autorregressivo sazonal de ordem P ;
- $\Theta(B^s)$ é o operador de médias móveis sazonal de ordem Q ;
- ∇^d é o operador diferença não-sazonal, tal que:

$$\nabla^d = (1 - B)^d \quad (2.17)$$

sendo d o número de diferenças não-sazonais.

- ∇_s^D é o operador diferença sazonal, tal que:

$$\nabla_s^D = (1 - B)^D \quad (2.18)$$

sendo D o número de diferenças sazonais

- a_t é ruído branco.

O modelo estabelece que se deve tomar d diferenças simples e D sazonais da série z_t para que o processo $W_t = \nabla^d\nabla_s^D z_t$ seja estacionário.

2.6.6 Modelo Sazonal Autorregressivo Integrado de Médias Móveis com Variáveis Exógenas (SARIMAX)

O modelo SARIMAX é simplesmente o modelo SARIMA, expresso na equação 2.13, com a incorporação de uma ou mais variáveis exógenas ¹ x_t multiplicadas por seus coeficientes β , conforme mostra a equação 2.19. Diferentemente de uma regressão, este coeficiente não representa o efeito sobre a série z_t quando x_t aumenta em uma unidade. A desvantagem é que a interpretação desse coeficiente é pouco intuitiva, uma vez que ele só pode ser interpretado condicionalmente ao valor das defasagens de z_t (HYNDMAN, 2010).

$$\phi(B)\Phi(B^s)\nabla^d\nabla_s^D z_t = \theta(B)\Theta(B^s)\alpha_t + \sum_{i=1}^k \beta_i x_{it} \quad (2.19)$$

Onde:

- x_{it} é a variável exógena i no tempo t , com i variando de 1 até k ;

¹As variáveis exógenas, x , também são chamadas de regressores, variáveis independentes, explicativas ou preditoras. Enquanto a variável de previsão, z_t , também é denominada regressando, variável dependente ou variável explicada.

- β é o coeficiente da variável exógena;
- s é a periodicidade da sazonalidade;
- $\phi(B)$ é o operador autorregressivo de ordem p ;
- $\theta(B)$ é o operador de médias móveis de ordem q ;
- $\Phi(B^s)$ é o operador autorregressivo sazonal de ordem P ;
- $\Theta(B^s)$ é o operador de médias móveis sazonal de ordem Q ;
- ∇^d é o operador diferença não-sazonal;
- ∇_s^D é o operador diferença sazonal;
- a_t é ruído branco.

Dada a dificuldade de interpretar a estrutura acima, Rob J. Hyndman sugeriu representar o modelo SARIMAX como um modelo de regressão com erros SARIMA, apresentado na equação 2.20 abaixo. Nesse caso, os erros da regressão possuem autocorrelação.

$$z_t = \sum_{i=1}^k \beta_i x_{it} + \eta_t \quad (2.20)$$

$$\begin{aligned} \eta_t = & \phi_1 \eta_{t-1} + \phi_2 \eta_{t-2} + \cdots + \phi_p \eta_{t-p} + \Phi_1 \eta_{t-s} + \Phi_2 \eta_{t-2s} + \cdots + \Phi_P \eta_{t-Ps} \\ & - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \cdots - \theta_q a_{t-q} - \Theta_1 a_{t-s} - \Theta_2 a_{t-2s} - \cdots - \Theta_Q a_{t-Qs} \\ & + a_t, \end{aligned} \quad (2.21)$$

Observe que a equação 2.21 possui dois termos de erro, o erro do modelo de regressão η_t e o erro do modelo SARIMA a_t . Assume-se que apenas os erros do modelo SARIMA são ruído branco.

Assim, usando o operador de defasagem B , o modelo SARIMAX pode ser descrito como:

$$z_t = \sum_{i=1}^k \beta_i x_{it} + \frac{\theta(B)\Theta(B^s)}{\phi(B)\Phi(B^s)\nabla^d\nabla_s^D} a_t \quad (2.22)$$

Vale ressaltar que, antes de estimar a regressão com erros SARIMA, é preciso verificar se z_t é estacionária e as variáveis exógenas x_t também. Caso não sejam, os coeficientes estimados não serão consistentes.

3. Metodologia

Neste capítulo, realiza-se a análise descritiva das séries temporais de velocidade do vento, das variáveis climáticas TNA e TSA, bem como a relação entre as séries e as variáveis. Em seguida, detalha-se a metodologia proposta por Box & Jenkins para a construção de modelos ARIMA. As análises presentes neste capítulo e no Capítulo 4 foram efetuadas com o auxílio do *software* R.

3.1 Pré-processamento

Para fins de análise, o ideal seria ter tanto os dados de produção eólica quanto de velocidade do vento dos três parques eólicos selecionados: Alegria II, Praia Formosa e Ventos de São Januário 19. No entanto, por questões de confidencialidade, esses dados não são divulgados. Apenas a partir de 2018 a Câmara de Comercialização de Energia Elétrica (CCEE) começou a disponibilizar dados verificados de produção eólica em base horária e o Operador Nacional do Sistema Elétrico (ONS) passou a disponibilizar valores previstos de produção eólica e velocidade do vento em base semi-horária (MACEIRA et al., 2021). Todavia, esse histórico é insuficiente para a modelagem.

Então, os dados de velocidade do vento de cada localidade foram extraídos do banco de dados de reanálise MERRA-2, disponibilizado pela NASA (NASA, 2024). As séries temporais de velocidade dos ventos apresentam frequência mensal e abrangem o período de janeiro de 1980 até dezembro de 2023. Os dados são relativos a uma altura de 100m, que é a altura média dos aerogeradores (ENEL, 2023), e a unidade de medida é em metros por segundo (m/s).

Já os dados referentes aos índices TNA e TSA foram coletados do banco de dados de índices climáticos, disponibilizado pela *National Oceanic and Atmospheric Administration* (NOAA). Esses dados também possuem frequência mensal e englobam o período de janeiro de 1980 até dezembro de 2023. A unidade de medida é em graus Celsius (°C).

3.2 Análise descritiva dos dados

3.2.1 Velocidade do vento

As séries históricas de velocidade do vento dos parques eólicos apresentam um comportamento sazonal bem marcado, caracterizado pela ocorrência de um vale e um pico ao longo de cada ano, como pode ser observado na figura 3.1. Verifica-se também que os períodos de maior e menor intensidade na velocidade do vento coincidem nos três parques, sendo

mais intensos principalmente entre os meses de junho e dezembro e menos intensos entre março e abril, conforme exposto na seção de caracterização do regime de ventos no NEB.

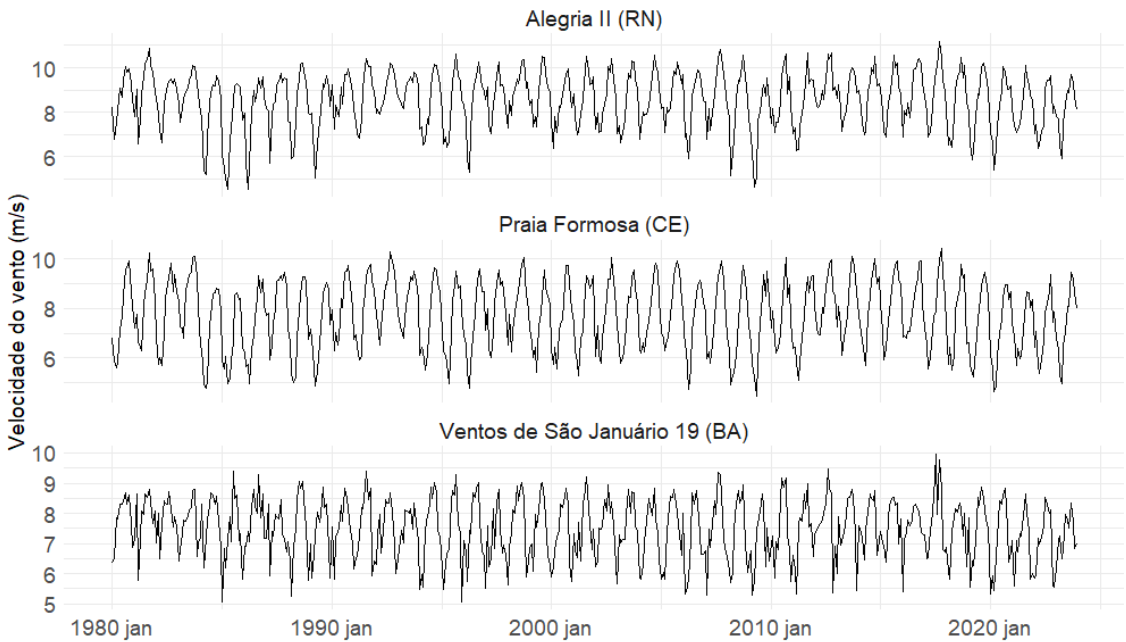


Figura 3.1: Séries temporais de velocidade do vento por parque eólico em frequência mensal.

Em relação às estatísticas descritivas, observa-se na tabela 3.1 que o parque eólico Alegria II apresenta a maior média de velocidade do vento, enquanto Ventos de São Januário 19, a menor. Nos três parques, nota-se que as medianas estão próximas às respectivas médias, concentrando-se mais à direita da média. Quanto ao coeficiente de variação, Ventos de São Januário 19 tem o menor valor quando comparado aos demais, o que indica que as velocidades do vento variam menos nessa localidade. Considerando as medidas de assimetria (S) e curtose (K), que refletem o padrão de distribuição dos dados, $S < 0$ para os três casos revela assimetrias com cauda à esquerda, confirmando a maior concentração de massa à direita da média. O $K < 3$ em Praia Formosa e Ventos de São Januário 19 sugere que a função de distribuição é platicúrtica, ou seja, mais “achatada” que a distribuição normal. Enquanto $K \approx 3$ em Alegria II, mostra que a função de distribuição é mesocúrtica, apresentando a curtose de uma distribuição normal padrão.

Tabela 3.1: Estatísticas descritivas das séries de velocidade do vento por parque eólico.

Parque eólico	Média	Mediana	Desvio Padrão	Coeficiente de Variação	Assimetria	Curtose
Alegria II	8,55	8,77	1,26	0,148	-0,631	3,03
Praia Formosa	7,72	7,84	1,40	0,181	-0,204	1,98
Ventos de São Januário 19	7,49	7,54	0,990	0,132	-0,305	2,37

As características de não-estacionariedade e sazonalidade das séries de velocidade do vento também podem ser observadas nos gráficos da função de autocorrelação (FAC) mostrados na figura 3.2. O decaimento da FAC como uma senoide amortecida é típico de dados sazonais, assim como a presença de picos significativos nos lags múltiplos de 12.

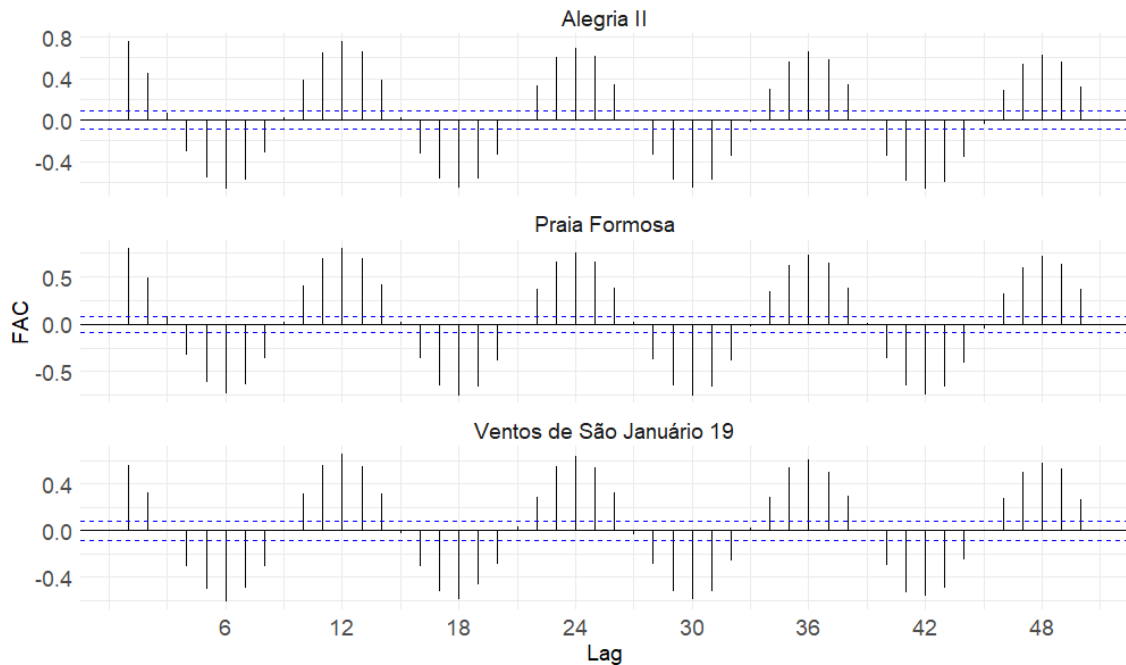


Figura 3.2: Função de autocorrelação das séries temporais de velocidade do vento por parque eólico.

3.2.2 Indicadores TNA e TSA

Ao analisarem o gradiente de temperatura do oceano Atlântico tropical, (ENFIELD et al., 1999) não identificaram um padrão dipolo dominante no geral, apenas em períodos específicos, como os meses de inverno-primavera e verão-outono boreal. Essa observação não significa, todavia, que flutuações significativas nos gradientes meridionais de TSM do Atlântico sejam inexistentes, nem que essas flutuações sejam irrelevantes para o clima do Atlântico tropical e das regiões terrestres adjacentes. Segundo os autores, as distribuições dipolares ficam mais evidentes em estudos que correlacionam as anomalias de temperatura do oceano às anomalias de variáveis atmosféricas sensíveis à posição da ZCIT como, por exemplo, precipitação e vento.

Eles observaram que configurações não-dipolares são mais comuns e o gradiente de temperatura entre o Atlântico Norte e Sul também pode ser importante para entender o clima. Outra informação relevante levantada é que valores opostos de temperatura não indicam necessariamente a existência de verdadeiras configurações dipolares. Logo, considerando esse contexto mais a dificuldade de encontrar dados sobre os períodos em

que o padrão de teleconexão TAD esteve atuando, não foi possível realizar uma análise mais detalhada além das características básicas relativas ao comportamento das séries de TNA e TSA.

Na figura 3.3, são apresentados os gráficos de área das séries temporais dos dois indicadores. A área em vermelho representa os valores positivos de temperatura, enquanto a área em azul indica os valores negativos.

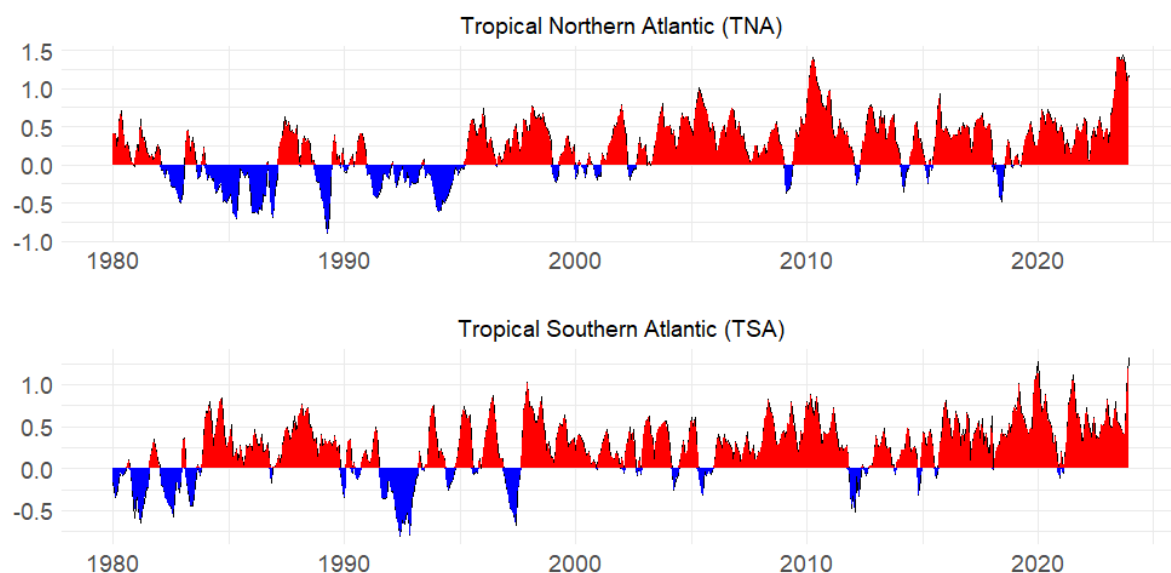


Figura 3.3: Gráfico de área das séries temporais dos indicadores TNA e TSA

Observa-se uma tendência de aumento da temperatura tanto no Atlântico Sul quanto no Norte. Além disso, confirma-se que os padrões não-dipolares ocorrem de forma frequente. Com o aquecimento progressivo dos oceanos, conclui-se também que é essencial continuar monitorando o padrão de teleconexão TAD e cada um desses índices.

Nos gráficos das FACs das séries de TNA e TSA, apresentados na figura 3.4 abaixo, nota-se que as autocorrelações começam altas para os primeiros lags, decaem rapidamente até o lag 7 aproximadamente e, depois, oscilam de forma significativa em lags mais altos. Em outras palavras, verifica-se uma persistência das autocorrelações. No entanto, essas oscilações não indicam um padrão cíclico ou sazonal, dado que não aparecem em intervalos regulares ou em lags múltiplos de 12.

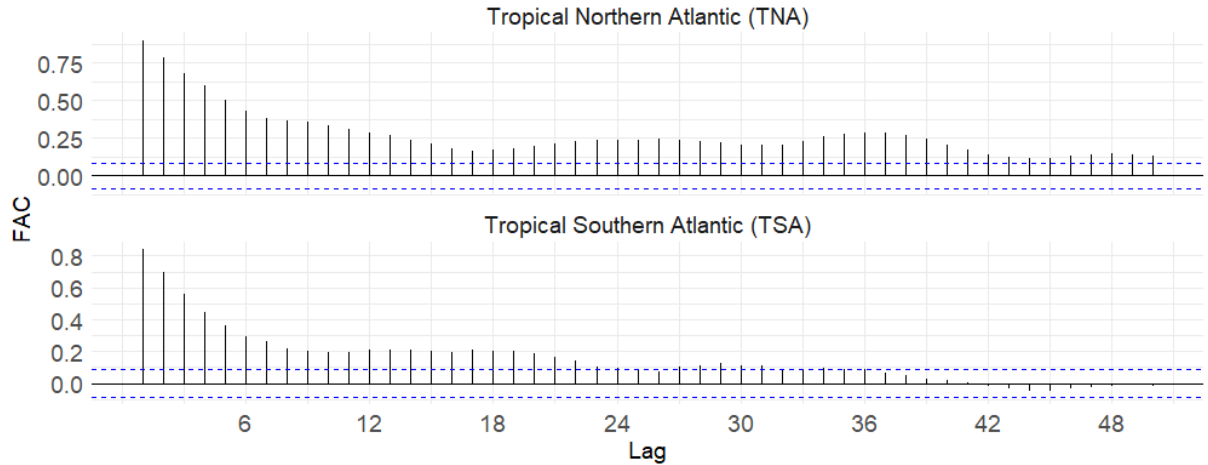


Figura 3.4: Função de autocorrelação das séries dos indicadores TNA e TSA.

3.2.3 Relação entre as velocidades do vento e os indicadores TNA e TSA

Para verificar a correlação entre uma série univariada x_{t+h} e os valores passados de uma outra série univariada y_t , utilizou-se o gráfico da Função de Correlação Cruzada (FCC). No eixo y , os valores variam entre -1 e 1, onde 1 representa a correlação positiva perfeita, -1 a correlação negativa perfeita e 0 a ausência de correlação. No eixo x , os lags de interesse são somente os negativos, pois busca-se analisar a correlação dos valores passados de x em y , ou seja, a correlação dos valores passados de cada variável climática nas séries de velocidade do vento.

Os gráficos da FCC para cada série de velocidade do vento e cada índice encontram-se na figura 3.5. Observa-se que há correlações significativas entre as velocidades do vento de cada parque eólico e cada índice, com valores máximos em torno de $|0,20|$. Em todos os casos, as correlações são significativas em períodos múltiplos de 6 (t , $2t$, $3t$, etc.) e meses ao redor ($t-1$, $t-2$, $2t-1$, $2t-2$, $3t-1$, $3t-2$, etc.). Além disso, as velocidades do vento de todos os parques apresentam correlações significativas com o índice TSA principalmente no tempo corrente t e em até três períodos anteriores ($t-1$, $t-2$ e $t-3$).

Nota-se também que as correlações oscilam entre valores positivos e negativos. O sinal positivo da correlação indica que as séries movem-se na mesma direção. Por exemplo, no parque eólico Alegria II, a correlação de +0,19 em um lag -1 indica que um aumento no índice TNA um período antes está associado a um aumento da velocidade do vento no tempo corrente (t). Enquanto o sinal negativo indica que as séries movem-se em direções opostas. Por exemplo, a correlação de -0,15 em um lag -6 significa que um aumento no índice TNA 6 meses antes está associado a uma diminuição na velocidade do vento no tempo t .

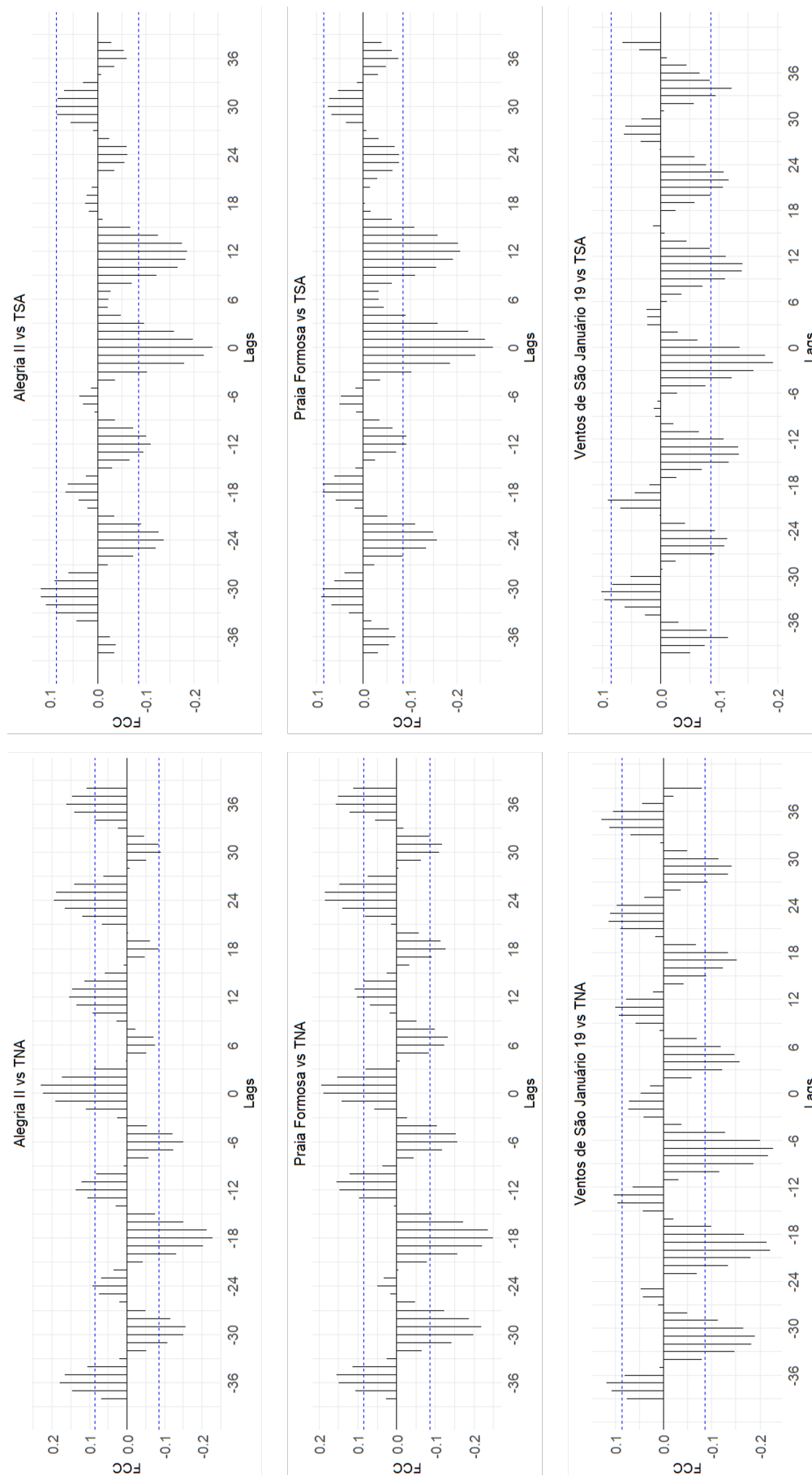


Figura 3.5: Função de correlação cruzada velocidades do vento *versus* índices.

Por simplificação, adotou-se um critério específico: em vez de testar vários lags e permitir que o modelo identificasse o mais significativo, optou-se por selecionar o lag da variável exógena que, além de apresentar correlação significativa, também estivesse mais próximo ao tempo corrente (t). Portanto, a tabela 3.2 abaixo apresenta a variável selecionada para a modelagem das séries de vento de cada parque eólico, a respectiva defasagem e o valor da correlação cruzada.

Tabela 3.2: Variável exógena, defasagem selecionada e valor da correlação.

Parque eólico	Variável exógena	Lag	Correlação
Alegria II	TSA	-1	-0,22
Praia Formosa	TSA	-1	-0,24
Ventos de São Januário 19	TSA	-2	-0.19

3.3 Metodologia Box & Jenkins

Com base no procedimento iterativo proposto por Box & Jenkins para a construção de modelos (BOX et al., 2015), esta análise seguiu cinco etapas:

1. Verifica-se a necessidade de realizar uma transformação Box-Cox para estabilizar a variância e de tratar outliers para não distorcer os resultados;
2. Realiza-se o teste de raiz unitária como o teste *Augmented Dickey Fuller* (ADF) para investigar a estacionariedade da série e, se necessário, aplicam-se diferenciações. Identificam-se as ordens do modelo por meio da função de autocorrelação (FAC), função de autocorrelação parcial (FACP) e critérios de informação. Estes realizam um *trade-off* entre a qualidade do ajuste do modelo testado e a sua complexidade, revelada no número de parâmetros. Os critérios de informação mais utilizados são o *Bayesian Information Criterion* (BIC), o *Akaike Information Criterion* (AIC) e o AIC corrigido (AICc). Hyndman e Athanasopoulos recomendam a utilização do AICc (HYNDMAN; ATHANASOPOULOS, 2021);
3. Estima-se os parâmetros do modelo identificado na etapa 2. Na função ARIMA() do pacote *fable*, essa estimação é realizada por máxima verossimilhança (MLE). Nessa técnica, encontram-se valores que maximizam a probabilidade de obter os dados observados (HYNDMAN; ATHANASOPOULOS, 2021);
4. Aplicam-se testes para verificar se o modelo é adequado para descrever a série temporal. Dentre os testes mais utilizados, destacam-se o teste Jarque-Bera para investigar a normalidade dos resíduos e o teste Ljung-Box para verificar a presença de autocorrelação nos resíduos.
5. Caso o modelo seja adequado, realiza-se a previsão.

Caso o modelo não seja adequado, reinicia-se o procedimento.

A figura 3.6 ilustra essas etapas. As três primeiras são descritas na Seção 4.1, a de diagnósticos é detalhada na Seção 4.2 e os resultados da previsão são apresentados na Seção 4.3.

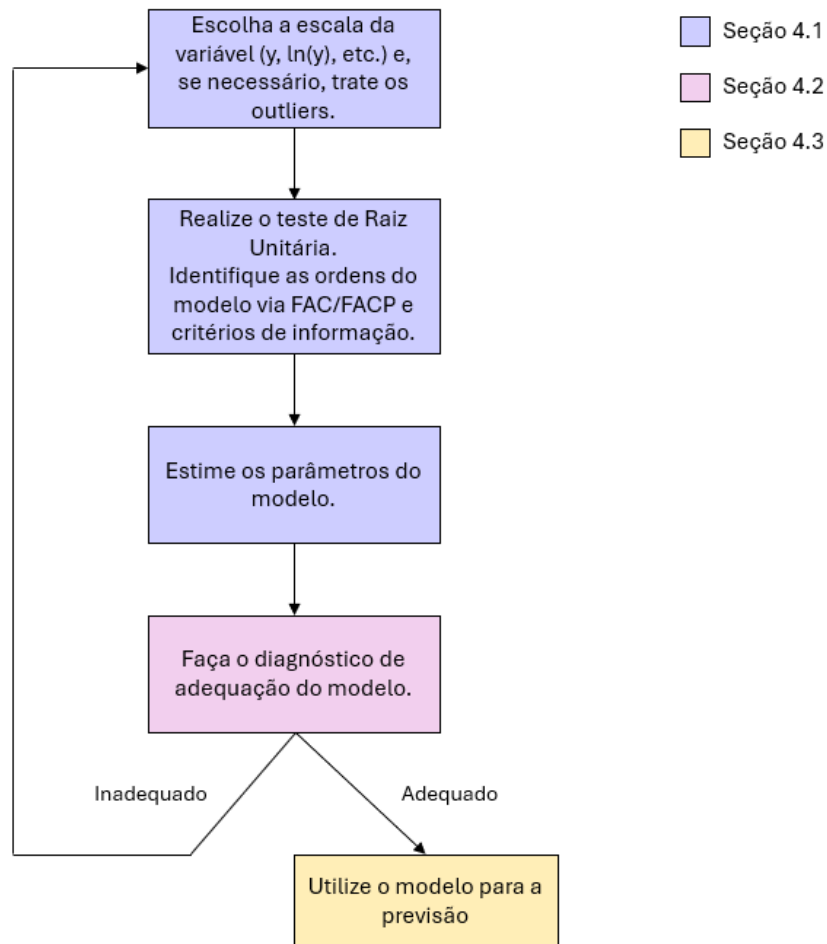


Figura 3.6: Etapas da metodologia Box & Jenkins

4. Resultados e Discussões

Neste capítulo, o modelo SARIMAX mais adequado para cada parque eólico foi selecionado dentre um conjunto de candidatos, seguindo a metodologia descrita no Capítulo 3. Em seguida, a performance dos modelos na previsão dos cenários de velocidade do vento foi verificada por meio de métricas de erro e, por fim, os modelos com e sem a variável exógena foram comparados entre si e com o modelo *Seasonal Naïve*.

Inicialmente, a base de dados apresentada na Seção 3.1 foi separada em 4 janelas de treinamento e teste, descritas na tabela 4.1.

Tabela 4.1: Janelas de ajuste e previsão.

Janela	In-sample	Out-of-sample
1	jan/1980-dez/2022	jan/2023-mar/2023
2	jan/1980-mar/2023	abr/2023-jun/2023
3	jan/1980-jun/2023	jul/2023-set/2023
4	jan/1980-set/2023	out/2023-dez/2023

A janela 1 foi utilizada para definir as ordens do modelo, isto é, o modelo selecionado no período *in-sample* da primeira janela foi aplicado às demais e utilizado para a previsão. O seguinte passo-a-passo foi realizado:

1. Ajustar a série de ventos para o período *in-sample*
2. Simular o período *out-of-sample*
3. Comparar os valores previstos (média dos cenários) com o observado;
4. Guardar os erros obtidos.

O modelo SARIMAX requer a especificação da variável climática para o período *out-of-sample*. Portanto, as previsões dos índices TNA e TSA para um período de três meses foram coletados no *site* da Agência Meteorológica do Japão (Japan Meteorological Agency, 2023). Devido à disponibilidade limitada de previsões para as variáveis exógenas, que compreende apenas um período de três meses, os dados foram divididos em quatro janelas trimestrais para ser possível realizar a previsão para todo o ano de 2023.

Os erros entre o valores previstos e observados foram obtidos para cada uma das quatro janelas com o objetivo de verificar de forma mais robusta a capacidade preditiva dos modelos. Métricas de erro muito utilizadas em estudos que preveem a velocidade do vento são o MAE, RMSE e o MAPE (SOMAN et al., 2010). Para esta análise, escolheram-se

as métricas *Mean absolute percentage error* (MAPE) por permitir a comparação entre diferentes modelos e *Mean Absolute Error* (MAE) pela facilidade de interpretação. Elas são expressas nas equações 4.1 e 4.2 abaixo.

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \quad (4.1)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (4.2)$$

Como modelo *benchmark*, optou-se pelo *Seasonal Naïve*, devido à forte sazonalidade presente nas séries de velocidades do vento. Neste caso, a previsão é igual ao último valor observado no mesmo período sazonal. Por exemplo, para dados mensais, a previsão para março deste ano é igual ao valor de março do ano passado (HYNDMAN; ATHANASOPOULOS, 2021). Caso o modelo proposto não alcance o nível mínimo de desempenho de previsão, ele pode ser substituído pelo *benchmark* (GILLILAND; SGLAVO, 2010).

Dado que os padrões sazonais das séries de velocidade do vento são relativamente estáveis ao longo do tempo, o Seasonal Naïve é um forte concorrente para realizar previsões. É importante observar também que as previsões do modelo SARIMAX estão sujeitas aos erros de previsão associados à variável exógena, TSA, o que pode afetar a precisão global do modelo.

4.1 Identificação das ordens do modelo e estimação dos parâmetros

Na inspeção visual, realizada na Seção 3.2, observou-se a não-estacionariedade das séries de velocidade de vento e possivelmente das séries dos dois índices. Então, para investigar a estacionariedade de maneira formal, aplicou-se o teste de Raiz Unitária (RU) ADF-II sobre as séries z_t (DICKEY; FULLER, 1979). Este teste estabelece que:

- H_0 : A série temporal possui uma raiz unitária $\leftrightarrow$ A série é não estacionária ($\Phi = 1 \leftrightarrow \alpha = 0$).
- H_1 : A série temporal não possui uma raiz unitária $\leftrightarrow$ A série é estacionária ($|\Phi| < 1 \leftrightarrow \alpha < 0$).

A estatística de teste é dada por:

$$t = \frac{\hat{\alpha} - 0}{SE(\hat{\alpha})} \quad (4.3)$$

Onde $\hat{\alpha}$ representa a estimativa do coeficiente e $SE(\hat{\alpha})$, o erro padrão dessa estimativa.

A tabela 4.2 abaixo apresenta os resultados.

Tabela 4.2: Resultados do teste ADF-II.

Séries	Valor da estatística de teste	Valor crítico ($\alpha = 5\%$)
Velocidades do vento	-2,76	-2,86
TNA	-2,46	-1,95
TSA	-2,15	-1,95

Observa-se que as séries de velocidade do vento dos três parques eólicos são não estacionárias, pois o valor da estatística de teste de -2,76 é maior que o valor crítico de -2,86 para um nível de significância de 5%. Enquanto os índices TNA e TSA são estacionários, visto que $-2,46 < -1,95$ e $-2,15 < -1,95$, respectivamente. Para tornar séries de velocidade do vento estacionárias, é preciso diferenciá-las. O gráfico da Função de Autocorrelação (FAC), apresentado na figura 3.2, mostra a necessidade de realizar pelo menos uma diferença sazonal.

Quanto à identificação das ordens do modelo pela análise da FAC e FACP, trata-se de uma tarefa complexa nesse caso, pois ambas as funções decaem como uma senoide amortecida. Portanto, foi utilizada a função ARIMA() do pacote *fable* para a seleção automática de um modelo, que serviu como referência para sugerir outros candidatos. As ordens do modelo foram definidas por meio do critério de informação, AICc, expresso na equação 4.4.

$$AICc = AIC + \frac{2(p + q + k + 1)(p + q + k + 2)}{T - p - q - k - 2} \quad (4.4)$$

Onde o termo $(p + q + k + 1)$ corresponde ao número de parâmetros do modelo, $k = 1$ se $c \neq 0$ e $k = 0$ se $c = 0$, e T é o tamanho da amostra.

Esse critério é uma correção do AIC, descrito na equação 4.5. Ele é utilizado quando o tamanho de amostra é pequeno ou quando o número de parâmetros ajustados é uma fração moderada a grande do tamanho da amostra. À medida que o número de observações aumenta, o AICc converge para o AIC (HURVICH; TSAI, 1989).

$$AIC = -2 \log(L) + 2(p + q + k + 1) \quad (4.5)$$

Onde L é a verossimilhança dos dados.

Para cada parque eólico, os resultados desta etapa da metodologia Box & Jenkins são apresentados a seguir. De forma geral, realizou-se:

- Ajuste de modelos candidatos às séries de velocidade do vento, com a série defasada da variável exógena incorporada:**

Considerando possíveis modelos com diferentes ordens p , q , P , Q , o selecionado foi

o primeiro de cada uma das tabelas 4.3, 4.4, 4.5, por apresentar o menor valor de AICc.

2. Análise visual dos resíduos:

- **Gráfico dos resíduos:** Não é evidente a presença de volatilidade ou aumento da amplitude dos resíduos ao longo do tempo. É preciso aplicar testes para sanar a dúvida.
- **Histograma:** Observa-se uma forma semelhante à distribuição normal, sugerindo que os resíduos podem ser normalmente distribuídos.
- **FAC:** Verifica-se que o modelo escolhido parece satisfatório para prever a série, uma vez que a FAC dos resíduos se comporta praticamente como um ruído branco, ou seja, quase todos os lags >0 são estatisticamente nulos.

3. Avaliação da significância estatística dos coeficientes estimados:

Nas figuras 4.2, 4.4, 4.6, observa-se que todos os coeficientes são estaticamente significantes, inclusive os coeficientes relacionados às variáveis exógenas, pois $p\text{-valor} < \alpha = 0,05$, onde α é o nível de significância.

4.1.1 Parque eólico Alegria II

Tabela 4.3: Valores AICc para os modelos SARIMAX aplicados às séries de velocidade do vento com a adição da série TSA defasada em um período. Modelo selecionado de forma automática: SARIMA(1,0,1) x (2,1,0) + lag1 TSA.

Modelo	AICc
SARIMA(1,0,1) x (0,1,1) + lag1 TSA	930
SARIMA(1,0,1) x (3,1,0) + lag1 TSA	1029
SARIMA(1,0,1) x (2,1,0) + lag1 TSA	1044

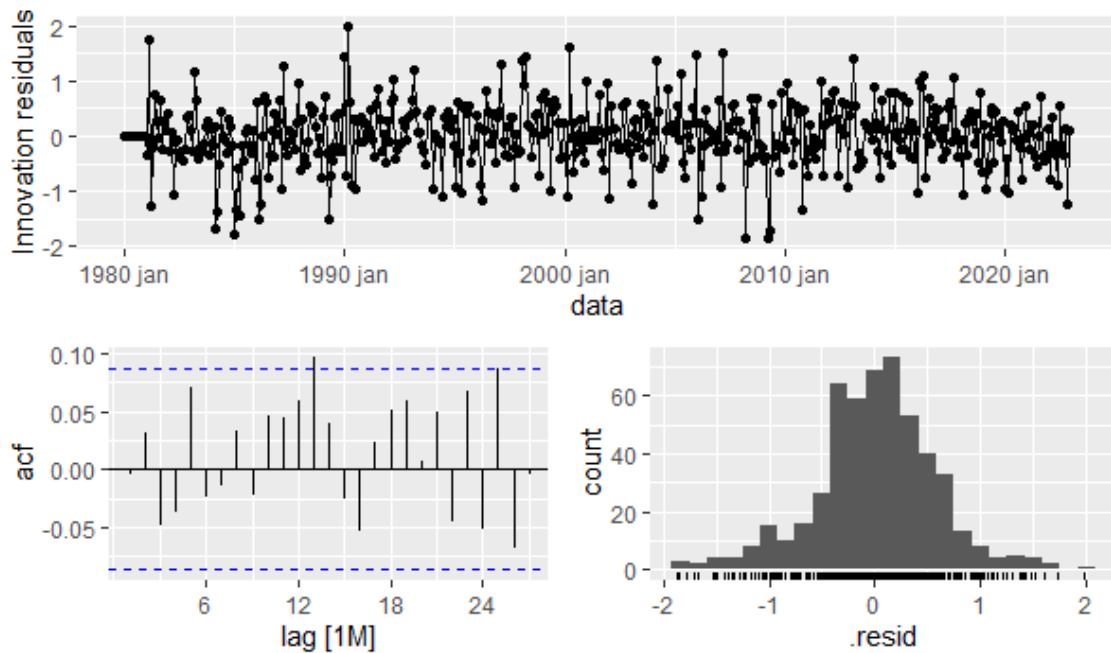


Figura 4.1: Resíduos do modelo SARIMA(1,0,1)x(0,1,1) + lag1 TSA ajustado

.model	term	estimate	std.error	statistic	p.value
<chr>	<chr>	<dbl>	<dbl>	<dbl>	<dbl>
sarima101011_tsa	ar1	0.701	0.0641	10.9	4.27e-25
sarima101011_tsa	ma1	-0.295	0.0844	-3.50	5.09e-4
sarima101011_tsa	sma1	-1.00	0.157	-6.37	4.36e-10
sarima101011_tsa	tlsa_lag1	-0.251	0.127	-1.97	4.91e-2

Figura 4.2: Estimativa e significância estatística de cada coeficiente do modelo SARIMA(1,0,1)x(0,1,1) + lag1 TSA.

4.1.2 Parque eólico Praia Formosa

Tabela 4.4: Valores AICc para os modelos SARIMAX aplicados às séries de velocidade do vento com a adição da série TSA defasada em um período. Modelo selecionado de forma automática: SARIMA(1,0,1) x (2,1,0) + lag1 TSA.

Modelo	AICc
SARIMA(1,0,1) x (0,1,1) + lag1 TSA	778
SARIMA(1,0,1) x (1,1,1) + lag1 TSA	779
SARIMA(1,0,1) x (2,1,0) + lag1 TSA	897

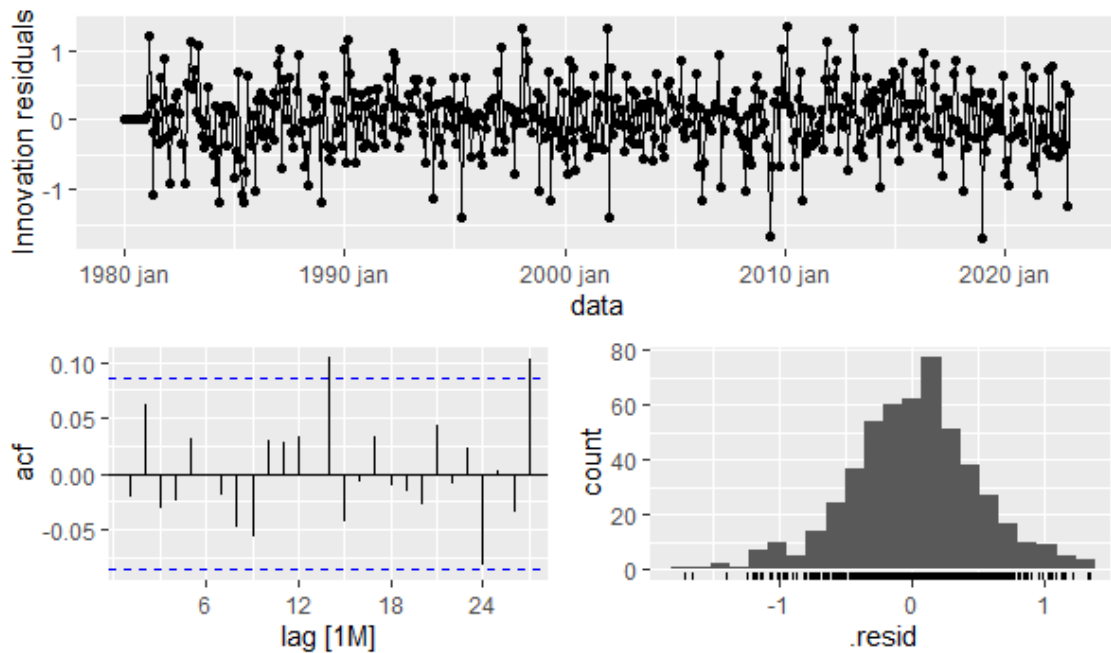


Figura 4.3: Resíduos do modelo SARIMA(1,0,1)x(0,1,1) + lag1 TSA ajustado.

.model	term	estimate	std.error	statistic	p.value
<chr>	<chr>	<dbl>	<dbl>	<dbl>	<dbl>
sarima101011_tsa	ar1	0.836	0.0417	20.1	3.09e-66
sarima101011_tsa	ma1	-0.492	0.0643	-7.65	9.95e-14
sarima101011_tsa	sma1	-1.00	0.0527	-19.0	6.10e-61
sarima101011_tsa	tsta_lag1	-0.474	0.109	-4.35	1.64e- 5

Figura 4.4: Estimativa e significância estatística de cada coeficiente do modelo SARIMA(1,0,1)x(0,1,1) + lag1 TSA.

4.1.3 Parque eólico Ventos de São Januário 19

Tabela 4.5: Valores AICc para os modelos SARIMAX aplicados às séries de velocidade do vento com a adição da série TSA defasada em dois períodos. Modelo selecionado de forma automática: SARIMA(0,0,0) x (2,1,0) + lag2 TSA.

Modelo	AICc
SARIMA(0,0,0) x (0,1,1) + lag2 TSA	966
SARIMA(0,0,0) x (1,1,1) + lag2 TSA	967
SARIMA(0,0,0) x (2,1,0) + lag2 TSA	1051

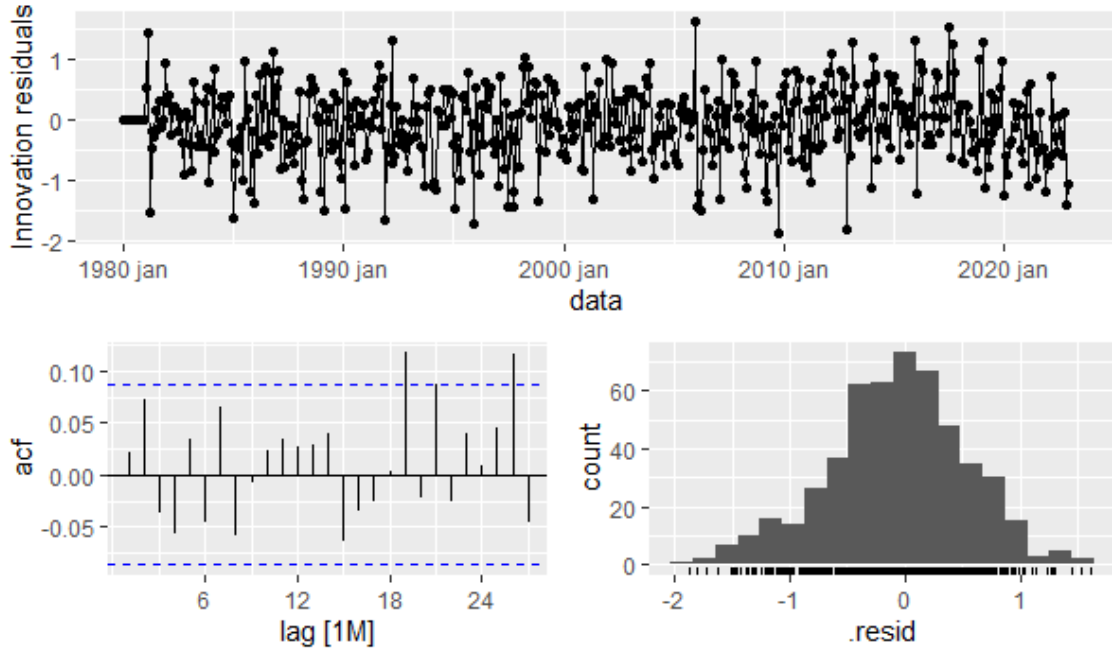


Figura 4.5: Resíduos do modelo SARIMA(0,0,0)x(0,1,1) + lag2 TSA ajustado.

	.model	term	estimate	std.error	statistic	p.value
	<chr>	<chr>	<dbl>	<dbl>	<dbl>	<dbl>
1	sarima000011_tsa	sma1	-0.938	0.0287	-32.7	1.85e-126
2	sarima000011_tsa	tsa_lag2	-0.164	0.0810	-2.03	4.30e- 2

Figura 4.6: Estimativa e significância estatística de cada coeficiente do modelo SARIMA(0,0,0)x(0,1,1) + lag2 TSA.

4.2 Diagnósticos de adequação do modelo

Nesta etapa, aplicaram-se testes nos resíduos do modelo com o objetivo de verificar estaticamente se os modelos selecionados são adequados para a previsão. Foi utilizado um nível de significância $\alpha = 0,05$ para investigar a presença de autocorrelação e normalidade.

Esses testes servem como uma confirmação formal das observações visuais realizadas na etapa de identificação das ordens do modelo.

4.2.1 Teste Ljung-Box

O teste de Ljung-Box investiga a autocorrelação conjunta dos resíduos, estabelecendo:

- $H_0 : \rho(1) = \rho(2) = \dots = \rho(12) = 0$
- $H_1 : \text{pelo menos uma das } \rho(k)\text{'s} \neq 0$

A estatística de teste é dada por:

$$LB(m) = T(T+2) \sum_{k=1}^m \frac{\hat{\rho}^2(k)}{(T-k)} \sim \chi^2(v) \quad (4.6)$$

Onde T é o tamanho da amostra, $\hat{\rho}_k$ é o coeficiente de autocorrelação amostral na defasagem k , m é o número total de lags considerados e $\chi^2(v)$ a Chi-Quadrada com v graus de liberdade.

Após a aplicação do teste, considerando 24 lags, foram obtidos os resultados apresentados na tabela, 4.2.1.

Tabela 4.6: Resultado do teste de Ljung-Box.

Parque eólico	Modelo	p-valor
Alegria II	SARIMA(0,0,0) x (0,1,1) + lag1 TSA	0,157
Praia Formosa	SARIMA(1,0,1) x (0,1,1) + lag1 TSA	0,147
Ventos de São Januário 19	SARIMA(1,0,1) x (0,1,1) + lag2 TSA	0,448

Como $p\text{-valor} > \alpha$ nos três casos, aceita-se a hipótese nula de que todas as autocorrelações são estatisticamente iguais a zero, ou seja, de que os resíduos são descorrelatados.

4.2.2 Teste Jarque-Bera

O Teste de Jarque-Bera verifica a normalidade dos resíduos, tendo como base duas características da distribuição normal (JARQUE; BERA, 1980):

1. É simétrica (assimetria = 0)
2. A velocidade de decaimento das caudas é igual (curtose = 3)

Este teste estabelece que:

- $H_0 : S = 0$ e $K = 3$
- $H_1 : S = 0$ e $K \neq 3$ ou $S \neq 0$ e $K = 3$ ou $S \neq 0$ e $K \neq 3$

A estatística de teste é dada por:

$$JB = \frac{T}{6}(\hat{S} - 0)^2 + \left(\frac{T}{24}\right)(\hat{K} - 3)^2 \sim \chi^2(2) \quad (4.7)$$

Onde T é o tamanho da amostra, $\hat{S}$ é a assimetria estimada dos dados, $\hat{K}$ é a curtose estimada dos dados e $\chi^2(v)$ é a Chi-Quadrada com 2 graus de liberdade.

Os resultados da aplicação do teste podem ser observados na tabela 4.2.2.

Tabela 4.7: Resultados do teste de Jarque-Bera.

Parque eólico	Modelo	p-valor	K	S
Alegria II	SARIMA(0,0,0) x (0,1,1) + lag1 TSA	2,53e-06	4,00	-0,207
Praia Formosa	SARIMA(1,0,1) x (0,1,1) + lag1 TSA	3,69e-03	3,60	-0,189
Ventos de São Januário 19	SARIMA(1,0,1) x (0,1,1) + lag2 TSA	9,07e-02	3,10	-0,238

Para resíduos do modelo ajustado à série de velocidades do vento de Ventos de São Januário 19, como $p\text{-valor}=0,0907 > \alpha = 0,05$, aceita-se a hipótese nula de normalidade, $K \approx 3$ e $S \approx 0$. Nos demais casos, rejeita-se a hipótese nula. Uma alternativa para resolver essa não normalidade é tratar os *outliers*. No entanto, dado que não se sabe a causa dos *outliers*, podendo estar relacionada a ocorrência de outros fenômenos atmosféricos, optou-se por não removê-los da amostra.

4.3 Comparação da performance preditiva dos modelos

Nesta etapa, o modelo selecionado para descrever as séries de velocidade do vento para cada parque eólico foi comparado com o modelo equivalente sem a variável e com o modelo *Seasonal Naïve*.

Para o parque eólico Alegria II, conforme mostrado na tabela 4.3, o modelo que incorpora a variável exógena superou os demais nas primeiras duas janelas, de acordo com o MAPE. Na janela 1, esse modelo foi mais preciso do que o modelo sem variável exógena e melhor do que o modelo *benchmark*. O mesmo acontece na janela 2. Por outro lado, o modelo sem a variável exógena obteve desempenho superior ao modelo com a variável nas janelas 3 e 4. Contudo, nestas últimas janelas, ambos os modelos obtiveram performance inferior quando comparados ao *Seasonal Naïve*. Nota-se, todavia, que este modelo apresenta variações consideráveis nos seus erros absolutos (MAE) ao longo dos diferentes períodos do ano. Portanto, nenhum modelo mostrou-se consistentemente superior em todas as janelas.

Uma observação é que o modelo SARIMAX apresentou melhores resultados nas janelas 1 e 2, que incluem os meses de janeiro à junho do ano de 2023, exatamente o período caracterizado pela menor incidência de ventos. Contudo, uma limitação deste estudo é que não foram realizadas previsões para janelas adicionais que incluíssem outros anos, o que seria importante para verificar se esse resultado é recorrente.

Tabela 4.8: Parque eólico Alegria II: performance dos modelos.

	SARIMA(1,0,1)x(0,1,1) + lag1 TSA		SARIMA(1,0,1)x(0,1,1)		Seasonal Naïve	
Janela	MAE (m/s)	MAPE (%)	MAE	MAPE	MAE	MAPE
1	0,261	3,53	0,280	3,78	0,361	4,75
2	0,345	5,63	0,355	5,75	0,661	9,10
3	0,413	4,57	0,386	4,28	0,302	3,32
4	0,460	5,40	0,438	5,16	0,401	4,59

Para o parque eólico Praia Formosa, o modelo *Seasonal Naïve* apresentou pior desempenho em todas as janelas com base no MAPE, como observado na tabela 4.3. Comparando o modelo SARIMAX com o SARIMA, o SARIMAX obteve uma melhoria nas janela 1 e 3. Enquanto o SARIMA obteve uma melhoria nas janela 2 e 4. Portanto, pode-se considerar um empate. Vale destacar que, na janela 4, tanto o SARIMAX quanto o SARIMA geraram previsões mais precisas, com MAE em torno de 0,1.

Tabela 4.9: Parque eólico Praia Formosa: performance dos modelos.

	SARIMA(1,0,1)x(0,1,1) + lag1 TSA		SARIMA(1,0,1)x(0,1,1)		Seasonal Naïve	
Janela	MAE (m/s)	MAPE (%)	MAE	MAPE	MAE	MAPE
1	0,208	3,73	0,238	4,15	0,498	8,12
2	0,306	5,45	0,280	4,98	0,456	7,42
3	0,328	3,94	0,338	4,05	0,393	4,37
4	0,160	1,85	0,131	1,54	0,599	5,99

Para o parque eólico Ventos de São Januário 19, o modelo *Seasonal Naïve* mostrou-se superior nas três primeiras janelas com base no MAPE, como pode ser verificado na tabela 4.3. Nota-se que as velocidades do vento nesse parque variam menos do que nos outros, como mostrado na análise descritiva pelo coeficiente de variação. Na janela 4, no entanto, o *Seasonal Naïve* obteve um MAPE de 14,7%, cerca de três vezes maior do que o MAPE dos demais modelos. Nesta última janela, o destaque foi para o modelo SARIMA por apresentar o menor MAPE.

Nesse caso, modelar a série via SARIMAX, não se mostrou mais eficaz na maioria das janelas, com exceção da última. A adição do índice TSA defasado em dois períodos também não gerou melhores resultados que o modelo *benchmark* na maioria das janelas, portanto essa variável exógena parece não ser a mais adequada para explicar as variações da série de velocidade de vento para o ano de 2023. Outros fenômenos podem afetar o comportamento dos ventos de forma mais acentuada nessa localidade. Além disso, esse parque não está localizado no litoral como os demais, então a influência do relevo também

pode ser significativa. Logo, é preciso realizar uma análise meteorológica e geográfica mais detalhada para esclarecer essas questões.

Tabela 4.10: Parque eólico Ventos de São Januário 19: performance dos modelos.

	SARIMA(0,0,0)x(0,1,1) + lag2 TSA		SARIMA(0,0,0)x(0,1,1)		Seasonal Naïve	
Janela	MAE (m/s)	MAPE (%)	MAE	MAPE	MAE	MAPE
1	0,530	7,68	0,572	8,27	0,509	7,42
2	0,451	6,65	0,441	6,51	0,418	5,99
3	0,633	8,25	0,649	8,45	0,635	8,19
4	0,247	3,42	0,241	3,34	1,03	14,7

Em síntese, a inclusão do índice TSA pode melhorar o desempenho do modelo para alguns períodos do ano de 2023 nos parques eólicos Alegria II e Praia Formosa. No entanto, os valores de MAE obtidos foram muito próximos, sugerindo que as melhorias na precisão das previsões podem não ser estatisticamente significantes. Em outras palavras, os resultados sugerem que as previsões dos modelos SARIMAX e SARIMA são comparativamente similares, uma vez que as diferenças nos erros de previsão entre eles foram relativamente pequenas. Para chegar a uma conclusão mais robusta, devem ser realizadas validações dos modelos para mais janelas e, inclusive, para mais parques eólicos.

Nesse sentido, aplicou-se o teste Diebold-Mariano para avaliar se esses dois modelos de previsão diferem significativamente em termos de precisão (DIEBOLD; MARIANO, 1995; HARVEY; LEYBOURNE; NEWBOLD, 1997).

Este teste estabelece que:

- $H_0 : \mathbb{E}[d_t] = 0 \forall t$ (As duas previsões têm a mesma acurácia)
- $H_1 : \mathbb{E}[d_t] \neq 0$ (As duas previsões têm diferentes níveis de acurácia)

A estatística de teste é dado por:

$$DM = \frac{\bar{d}}{\sqrt{\frac{\hat{\sigma}_d^2}{T}}} \quad (4.8)$$

Onde $\bar{d}$ é a média das diferenças dos erros de previsão, $\hat{\sigma}_d^2$ é a estimativa da variância dessas diferenças, e T é o tamanho da amostra.

Os resultados são apresentados na tabela 4.11. Considerando um intervalo de valores críticos entre $[-1,96, 1,96]$, referente à $\alpha = 0,05$, a diferença entre os modelos SARIMA e SARIMAX é estatisticamente insignificante para todas as janelas do parque Ventos de São Januário 19, pois os valores da estatística de teste caíram dentro do intervalo. Portanto, a

hipótese nula é aceita. Para os parques Alegria II e Praia Formosa, esse resultado também é identificado nas três primeiras janelas, mas na quarta janela verificou-se significância estatística. Como o valor da estatística de teste é positiva, o modelo SARIMAX tem uma acurácia superior ao modelo SARIMA, embora o MAE e o MAPE tenham mostrado o oposto, isto é, o SARIMA performando melhor do que o SARIMAX.

Tabela 4.11: Resultados do teste Diebold-Mariano.

Parque eólico	Janela	Valor da estatística de teste
Alegria II	1	-0,807
	2	1,28
	3	1,25
	4	5,11
Praia Formosa	1	0,354
	2	0,280
	3	0,504
	4	2,40
Ventos de São Januário 19	1	-2,19
	2	1,97
	3	-2,89
	4	-0,253

5. Conclusões

O objetivo inicial deste estudo foi verificar a existência de correlação entre cada série de velocidade do vento e cada índice de anomalia de temperatura do Oceano Atlântico. O segundo objetivo foi avaliar se a incorporação do TNA e TSA no modelo torna as previsões das velocidades do vento mais acuradas.

Para alcançar o primeiro objetivo, foram analisados os gráficos da função de correlação cruzada entre as séries de velocidade do vento e os índices de anomalia. Todas as análises revelaram correlações significativas, tanto positivas quanto negativas, em lags múltiplos de seis meses e períodos ao redor. As séries de velocidades do vento com o índice TSA, especificamente, mostraram uma correlação significativa no tempo corrente e nos três períodos anteriores. Por essa razão, o TSA defasado em um período foi escolhido como variável exógena para a modelagem das velocidades do vento dos parques eólicos Alegria II (RN) e Praia Formosa (CE). O TSA defasado em dois períodos foi selecionado para o parque Ventos de São Januário 19.

Em relação ao segundo objetivo, na primeira etapa, testou-se a estacionariedade das séries de velocidade e, posteriormente, dos dois índices. O teste ADF-II indicou que as séries de velocidade são não estacionárias, enquanto as séries dos índices são estacionárias. Na segunda etapa, para identificar as ordens do modelo, sugeriu-se um conjunto de candidatos com base no modelo indicado pela função ARIMA() do pacote *fable*. O modelo selecionado foi o que apresentou menor AICc. Para os parques Alegria II e Praia Formosa, foi o SARIMA(1,0,1)x(0,1,1) com lag 1 do TSA. Já para o parque Ventos de São Januário foi SARIMA(0,0,0)x(0,1,1) com lag 2 do TSA. Verificando os coeficientes estimados, todos mostraram-se estatisticamente significantes.

Na terceira etapa, o teste de Ljung-Box confirmou que os resíduos são ruído branco. O teste de Jarque-Bera mostrou que apenas os resíduos do modelo ajustado à série de velocidades do vento do parque Ventos de São Januário 19 seguem uma distribuição normal. Para os demais parques, os outliers não foram tratados, visto que não é possível removê-los de forma arbitrária. Por fim, na quarta etapa, foi realizada uma comparação da performance preditiva dos modelos para o ano de 2023 com base no MAPE, principalmente. No parque Ventos de São Januário 19, o modelo *Seasonal Naïve* superou os outros nas três primeiras janelas temporais. Na última, o SARIMA foi mais eficaz. Já o SARIMAX não gerou previsões mais precisas, sugerindo que, nesse caso, a adição do índice TSA defasado não é eficaz. Além disso, pelo teste de Diebold-Mariano, a diferença entre os modelos SARIMA e SARIMAX é estatisticamente insignificante para todas as janelas. Nos outros parques,

segundo esse mesmo teste, a incorporação do índice TSA gerou previsões mais acuradas apenas na última janela de dados, que compreende todas as observações disponíveis.

Como próximos passos para a continuação deste trabalho, é essencial aprofundar a análise, incluindo mais parques eólicos e janelas temporais de treinamento e teste com o objetivo de verificar a consistência dos resultados obtidos para o ano de 2023, isto é, determinar se os resultados observados podem ser generalizados para outros período e locais no NEB. Também é importante explorar de forma mais detalhada a relação desses resultados com os sistemas meteorológicos atuantes. Além disso, pode-se investigar o potencial de outros índices de anomalia do Oceano Atlântico, além do TNA e TSA e/ou adicionar outras variáveis climáticas no modelo.

Referências Bibliográficas

ABEEÓLICA. *Boletim Anual 2022*. 2022. <<https://abeeolica.org.br/wp-content/uploads/2023/06/Boletim-de-Geracao-Eolica-2022.pdf>>. Acessado em: 01/02/2024.

ABEEÓLICA. *The Sector - Wind power development in Brazil*. 2022. <<https://abeeolica.org.br/en/wind-energy/the-sector/>>. Acessado em: 01/02/2024.

AMARANTE, O. A. C. do et al. Wind / hydro complementary seasonal regimes in brazil. *Dewi Magazin*, v. 19, p. 79–86, 2001.

Atlas Eólico BA. *Atlas eólico Bahia*. 2013. <<http://www.infraestrutura.ba.gov.br/arquivos/File/publicacoes/atlaseolicobahia2013.pdf>>. Acessado em: 15/02/2024.

Atlas Eólico RN. *Atlas eólico e solar do Estado do Rio Grande do Norte*. 2022. <<http://www.adcon.rn.gov.br/ACERVO/sedec/DOC/DOC000000000298278.PDF>>. Acessado em: 15/02/2024.

BARRY, R. G.; CHORLEY, R. J. *Atmosfera, Tempo e Clima*. Porto Alegre: Bookman, 2012.

BOX, G. E. P. et al. *Time Series Analysis: Forecasting and Control*. 5. ed. Hoboken, NJ: Wiley, 2015.

CADENAS, E. et al. Wind speed forecasting using the narx model, case: La mata, oaxaca, méxico. *Neural Computing and Applications*, 2017.

CAVALCANTI, I. F. de A. et al. *Tempo e Clima no Brasil*. São Paulo: Oficina de Textos, 2009.

CEPEL. *GEVAZP: Modelo de Geração de Cenários de Energias e Vazões*. 2024. <<https://www.cepel.br/linhas-de-pesquisa/gevazp-saibamais/>>. Acessado em: 04/04/2024.

CEPEL. *Planejamento da Operação, Despacho e Formação de Preço*. 2024. <https://see.cepel.br/manual/libs/latest/modelos_otimizacao_energetica/modelos_otimizacao_energetica.html#uso-oficial-dos-modelos-no-setor-eletrico-brasileiro>. Acessado em: 10/03/2024.

CHAN, C. S. *Análise de Distúrbios Ondulatórios de Leste sobre o Oceano Atlântico Equatorial Sul*. 134 p. Dissertação (Dissertação (Mestrado em Meteorologia)) — Instituto Nacional de Pesquisas Espaciais (INPE), 1990. INPE 5222 - TDL/437.

DICKEY, D. A.; FULLER, W. A. Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, v. 74, n. 366, p. 427–431, 1979.

DIEBOLD, F. X.; MARIANO, R. S. Comparing predictive accuracy. *Journal of Business and Economic Statistics*, v. 13, p. 253–263, 1995.

ENEL. *Hélice eólica*. 2023. <<https://www.enelgreenpower.com/pt/learning-hub/energias-renoveveis/energia-eolica/helice-eolica>>. Acessado em: 09/05/2024.

ENFIELD, D. B. et al. How ubiquitous is the dipole relationship in tropical atlantic sea surface temperatures? *Journal of Geophysical Research*, v. 104, p. 7841–7848, 1999.

FEDOROVA, N. et al. Middle tropospheric cyclonic vortex in northeastern brazil and the tropical atlantic. *Pure and Applied Geophysics*, v. 174, n. 1, p. 397–411, 2016.

FOLLAND, C. K.; PALMER, T. N.; PARKER, D. E. Sahel rainfall and worldwide sea temperatures, 1901-85. *Nature*, v. 320, p. 602–607, 1986.

FOLTZ, G. R.; BRANDT, P.; RICHTER, I. The tropical atlantic observing system. *Frontiers in Marine Science*, v. 6, p. 206, 2019.

GILLILAND, M.; SGLAVO, U. Worst practices in business forecasting. *Analytics*, p. 12–17, 2010.

GOV.BRAZIL. *Nationally Determined Contribution (NDC) to the Paris Agreement under the UNFCCC*. 2023. <<https://unfccc.int/sites/default/files/NDC/2023-11/Brazil%20First%20NDC%202023%20adjustment.pdf>>. Acessado em: 09/05/2024.

GWEC. *Global Wind Report 2023*. 2023. <https://gwec.net/wp-content/uploads/2023/03/GWR-2023_interactive.pdf>. Acessado em: 01/02/2024.

HARVEY, D.; LEYBOURNE, S.; NEWBOLD, P. Testing the equality of prediction mean squared errors. *International Journal of Forecasting*, v. 13, n. 2, p. 281–291, 1997.

HASTENRATH, S. Circulation and teleconnection mechanisms of northeast brazil droughts. *Progress in Oceanography*, v. 70, p. 407–415, 2006.

HASTENRATH, S.; HELLER, L. Dynamics of climatic hazards in northeast brazil. *Quarterly Journal of the Royal Meteorological Society*, v. 103, p. 77–92, 1977.

HASTENRATH, S.; LAMB, P. On the dynamics and climatology of surface flow over the equatorial oceans. v. 30, n. 5, p. 436–448, 1978.

HURVICH, C. M.; TSAI, C.-L. Regression and time series model selection in small samples. *Biometrika*, v. 76, n. 2, p. 297–307, 1989.

HYNDMAN, R. J. *The ARIMAX model muddle*. 2010. Disponível em: <<https://robjhyndman.com/hyndsight/arimax/>>.

HYNDMAN, R. J.; ATHANASOPOULOS, G. *Forecasting: Principles and Practice*. 3rd. ed. OTexts, 2021. Disponível em: <<https://otexts.com/fpp3/>>.

IBGE. *Planilha áreas das Regiões Brasileiras 2022*. 2022. <<https://www.ibge.gov.br/geociencias/organizacao-do-territorio/estrutura-territorial/15761-areas-dos-municipios.html?edicao=36514&t=acesso-ao-produto>>. Acessado em: 20/02/2024.

INMET. *O que é Zona de Convergência Intertropical? Entenda agora!* 2023. <<https://portal.inmet.gov.br/noticias/zona-de-convergencia-intertropical-zcit>>. Acessado em: 03/04/2024.

- IPCC. *Summary for Policymakers. In: Climate Change 2023: Synthesis Report*. Geneva, Switzerland: IPCC, 2023. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, H. Lee and J. Romero (eds.)]. Pp. 1–34, doi: 10.59327/IPCC/AR6-9789291691647.001.
- IRENA. *NDCs and renewable energy targets in 2023: Tripling renewable power by 2030*. 2023. <<https://www.irena.org/Publications/2023/Dec/NDCs-and-renewable-energy-targets-in-2023-Tripling-renewable-power-by-2030>>. Acessado em: 15/04/2024.
- Japan Meteorological Agency. *3-Month Forecast of Climate Indices*. 2023. <https://ds.data.jma.go.jp/tcc/tcc/products/model/indices/3-mon/indices1/shisu_forecast.php>. Acessado em: 20/05/2023.
- JARQUE, C. M.; BERA, A. K. Efficient tests for normality, homoscedasticity and serial independence of regression residuals. *Economics Letters*, v. 6, n. 3, p. 255–259, 1980. ISSN 0165-1765. Disponível em: <<https://www.sciencedirect.com/science/article/pii/0165176580900245>>.
- KOUSKY, V. E.; GAN, M. A. Upper tropospheric cyclonic vortices in the tropical south atlantic. *Tellus*, v. 36, n. 6, p. 538–551, 1981.
- LIMA, L. M. M.; POPOVA, E.; DAMIEN, P. Modeling and forecasting of brazilian reservoir inflows via dynamic linear models. *International Journal of Forecasting*, Elsevier, v. 30, n. 3, p. 464–476, 2014.
- MAÇAIRA, P. M. *Modelos de Simulação Estocástica de Cenários de Vazão com Incorporação de Variáveis Climáticas*. Dissertação (Tese de Mestrado) — Pontifícia Universidade Católica do Rio de Janeiro, 2018.
- MACEIRA, M. E. P. et al. *Uma abordagem para a representação das incertezas da fonte de geração eólica no planejamento da operação de longo e médio prazos - modelo NEWAVE - relatório final*. [S.l.], 2021.
- MARINHO, M. H. d. N.; AQUINO, R. R. B. d. Oferta de energia através da complementariedade sazonal hidro-eólica. *Revista PCH Notícias & SHP NEWS*, v. 11, n. 40, p. 36–42, 2009.
- MARKHAM, C. G.; MCLAIN, D. R. Sea surface temperature related to rain in ceará, north-eastern brazil. *Nature*, v. 265, p. 320–323, 1977.
- MORETTIN, P. A.; TOLOI, C. M. d. C. *Modelos para previsão de séries temporais*. [S.l.]: Instituto de Matemática Pura e Aplicada, 1981.
- MOSCATI, M. C.; GAN, M. A. Rainfall variability in the rainy season of semiarid zone of northeast brazil (neb) and its relation to wind regime. *International Journal of Climatology: A Journal of the Royal Meteorological Society*, v. 27, n. 4, p. 493–512, 2007.
- MOURA, A. D.; SHUKLA, J. On the dynamics of droughts in northeast brazil: observations, theory and numerical experiments with a general circulation model. *Journal of Atmospheric Sciences*, v. 38, p. 2653–2675, 1981.

- NASA. *The second Modern-Era Retrospective analysis for Research and Applications*. 2024. Acessado em: 01/02/2024. Disponível em: <https://disc.gsfc.nasa.gov/datasets/M2TMNXFLX_5.12.4/summary>.
- National Academies of Sciences, Engineering, and Medicine. *Decade-to-Century-Scale Climate Variability and Change: A Science Strategy*. Washington, DC: The National Academies Press, 1998. Disponível em: <<https://doi.org/10.17226/6129>>.
- NETO, P. S. de M. et al. An adaptive hybrid system using deep learning for wind speed forecasting. *Elsevier*, v. 581, p. 495–514, 2021.
- NOAA. *El Niño Southern Oscillation in a Changing Climate*. 2020. <https://www.pmel.noaa.gov/news-story/el_nino_southern_oscillation_changing_climate>. Acessado em: 02/02/2024.
- NOAA. *El Niño/Southern Oscillation - Technical Discussion*. 2024. <<https://www.ncei.noaa.gov/access/monitoring/enso/technical-discussion>>. Acessado em: 23/03/2024.
- NOAA. *The Ocean and Extreme Weather*. 2024. <<https://www.aoml.noaa.gov/extreme-weather/>>. Acessado em: 04/04/2024.
- NOBRE, P.; SHUKLA, J. Variations of sea surface temperature, wind stress, and rainfall over the tropical atlantic and south america. *Journal of Climate*, v. 9, p. 2464–2479, 1996.
- NWS. *What is El Niño-Southern Oscillation (ENSO)?* 2023. <<https://www.weather.gov/mhx/ensowhat>>. Acessado em: 25/03/2024.
- OLIVEIRA, F. P. F. *Fatores associados a iniciacao de linhas de instabilidade na regioao do centro de lancamentos de Alcantara no mes julho*. 102 p. Dissertação (Dissertacao (Mestrado em Meteorologia)) — Instituto Nacional de Pesquisas Espaciais (INPE), 2012.
- ONS. *O planejamento da operação energética no Sistema Interligado Nacional*. 2020. <<https://www.ons.org.br/AcervoDigitalDocumentosEPublicacoes/Livros-da-Diretoria-de-Planejamento-na-biblioteca-digital-ONS/LIVRO-O-PLANEJAMENTO-DA-OPERA%C3%87%C3%83O-ENERG%C3%89TICA-NO-SISTEMA%20INTERLIGADO-NACIONAL.pdf>>. Acessado em: 05/02/2024.
- ONS. *Flexibilização dos limites de intercâmbio gerou economia de R\$ 11 bilhões para o país durante a crise hídrica*. 2022. <[https://www.ons.org.br/Paginas/Noticias/20220706-Flexibiliza%C3%A7%C3%A3o-dos-limites-de-interc%C3%A2mbio-gerou-economia-de-R\\$-11-bilh%C3%B5es-para-o-Pa%C3%ADs-durante-a-cri-se-h%C3%ADrica.aspx](https://www.ons.org.br/Paginas/Noticias/20220706-Flexibiliza%C3%A7%C3%A3o-dos-limites-de-interc%C3%A2mbio-gerou-economia-de-R$-11-bilh%C3%B5es-para-o-Pa%C3%ADs-durante-a-cri-se-h%C3%ADrica.aspx)>. Acessado em: 15/03/2024.
- ONS. *Matriz de Energia Elétrica do SIN*. 2024. <<https://www.ons.org.br/paginas/sobre-o-sin/o-sistema-em-numeros>>. Acessado em: 01/03/2024.
- ONU. *The Paris Agreement*. 2015. <<https://www.un.org/en/climatechange/paris-agreement>>. Acessado em: 13/04/2024.

ORLANSKI, I. A simple boundary condition for unbounded hyperbolic flows. *Journal of Computational Physics*, v. 21, p. 251–269, 1976.

PETROBRAS. *Energia eólica offshore no Brasil: conheça nossos projetos*. 2023. <<https://nossaenergia.petrobras.com.br/w/transicao-energetica/energia-eolica-offshore/acordo-equinor>>. Acessado em: 20/03/2024.

PEZZI, L. P.; CAVALCANTI, I. F. A. The relative importance of enso and tropical atlantic sea surface temperature anomalies for seasonal precipitation over south america: a numerical study. *Climate Dynamics*, v. 17, p. 205–212, 2001.

RAO, V. B.; LIMA, M. C. de; FRANCHITO, S. H. Seasonal and interannual variations of rainfall over eastern northeast brazil. *Journal of Climate*, v. 6, p. 1754–1763, 1993.

REBOITA, M. S. et al. Impacts of teleconnection patterns on south america climate. 2021.

RUIZ-BARRADAS, A.; CARTON, J. A.; NIGAM, S. Structure of interannual-to-decadal climate variability in the tropical atlantic sector. *Journal of Climate*, v. 13, n. 18, p. 3285–3297, 2000.

SOMAN, S. S. et al. A review of wind power and wind speed forecasting methods with different time horizons. In: . [S.l.: s.n.], 2010.

STOCK, J. H.; WATSON, M. W. *Introduction to Econometrics*. Fourth. [S.l.]: Pearson Education Limited, 2020.

UVO, C. B. et al. The relationships between tropical pacific and atlantic sst and northeast brazil monthly precipitation. *Journal of Climate*, v. 11, p. 551–562, 1998.

UVO, C. R. B. *A Zona de Convergência Intertropical (ZCIT) e sua relação com a precipitação da Região Norte do Nordeste Brasileiro*. Dissertação (Tese de Mestrado) — Instituto de Pesquisas Espaciais - INPE, 1989.

WITZLER, L. T. *Metodologia para reconstrução de séries históricas de vento e geração eólica visando a análise da complementariedade energética no Sistema Interligado Nacional*. Dissertação (Tese de Mestrado) — Universidade de São Paulo, 2014.

WMO. *Climate Indices*. 2024. Acessado em: 20/05/2024. Disponível em: <<https://www.wmolc.org/contents/index/Climate+Indices>>.