

UNIVERSIDADE FEDERAL DE SANTA CATARINA  
CENTRO TECNOLÓGICO  
DEPARTAMENTO DE ENGENHARIA ELÉTRICA E ELETRÔNICA  
CURSO DE GRADUAÇÃO EM ENGENHARIA ELÉTRICA

Lucas Roberto Dagort

**Estimação de Erros de Previsão Eólica do Sistema Interligado Nacional  
com Modelo Periódico Multivariável e Autorregressivo**

Florianópolis

2022

Lucas Roberto Dagort

**Estimação de Erros de Previsão Eólica do Sistema Interligado Nacional  
com Modelo Periódico Multivariável e Autorregressivo**

Trabalho de Conclusão do Curso de Graduação em Engenharia Elétrica do Centro Tecnológico da Universidade Federal de Santa Catarina como requisito para a obtenção do título de Bacharel em Engenharia Elétrica.

Orientador: Prof. Erlon Cristian Finardi, Dr. Eng.

Florianópolis

2022

Ficha de identificação da obra elaborada pelo autor,  
através do Programa de Geração Automática da Biblioteca Universitária da UFSC.

Dagort, Lucas Roberto  
Estimação de Erros de Previsão Eólica do Sistema  
Interligado Nacional com Modelo Periódico Multivariável e  
Autorregressivo / Lucas Roberto Dagort ; orientador, Erlon  
Cristian Finardi , 2022.  
99 p.

Trabalho de Conclusão de Curso (graduação) -  
Universidade Federal de Santa Catarina, Centro Tecnológico,  
Graduação em Engenharia Elétrica, Florianópolis, 2022.

Inclui referências.

1. Engenharia Elétrica. 2. Engenharia Elétrica. 3.  
Sistema de Energia Elétrica. 4. Previsão de desvios geração  
eólica. 5. Geração de Cenários. I. , Erlon Cristian Finardi.  
II. Universidade Federal de Santa Catarina. Graduação em  
Engenharia Elétrica. III. Título.

Lucas Roberto Dagort

**Estimação de Erros de Previsão Eólica do Sistema Interligado Nacional  
com Modelo Periódico Multivariável e Autorregressivo**

Este Trabalho Conclusão de Curso foi julgado adequado para obtenção do Título de “Bacharel em Engenharia Elétrica” e aceito, em sua forma final, pelo Curso de Graduação em Engenharia Elétrica.

Florianópolis, 28 de julho de 2022.



Documento assinado digitalmente  
**Miguel Moreto**  
Data: 02/08/2022 12:21:52-0300  
CPF: 948.850.100-63  
Verifique as assinaturas em <https://v.ufsc.br>

---

**Prof. Miguel Moreto, Dr.**  
Coordenador do Curso de Graduação em Engenharia Elétrica

**Banca Examinadora:**



Documento assinado digitalmente  
**Erlon Cristian Finardi**  
Data: 01/08/2022 15:22:53-0300  
CPF: 020.364.749-18  
Verifique as assinaturas em <https://v.ufsc.br>

---

**Prof. Erlon Cristian Finardi, Dr.**  
Orientador  
Universidade Federal de Santa Catarina



Documento assinado digitalmente  
**ROILAN HERNANDEZ VALDES**  
Data: 03/08/2022 09:39:16-0300  
CPF: 877.156.510-87  
Verifique as assinaturas em <https://v.ufsc.br>

---

**Roilan Hernandez, Me**  
Avaliador  
ENGIE



Documento assinado digitalmente  
**PAULO VITOR LARROYD**  
Data: 01/08/2022 15:34:36-0300  
CPF: 051.844.979-35  
Verifique as assinaturas em <https://v.ufsc.br>

---

**Eng. Eletric. Paulo Vitor Larroyd, Dr.**  
Avaliador  
Norus Tecnologia

Experiência não é o que acontece com um homem; é o que um homem faz com o que lhe acontece.

## **AGRADECIMENTOS**

Gostaria de agradecer à minha família pelo apoio incondicional durante toda a minha vida. Agradeço, especialmente, aos meus pais, Dilso e Eliete Dagort por me incentivarem e me ajudarem a tomar as decisões mais importantes da minha vida em particular durante o meu período de estudo na UFSC. Sou grato, igualmente, pelos conselhos que me deram que foram e que serão de grande ajuda em minha vida.

Sou grato a todos os professores que tive ao longo do curso que me guiaram e que compartilharam o seu conhecimento indispensável para a minha formação. Agradeço também aos meus colegas de curso que estiveram ao meu lado durante o curso e que ainda estão presentes em minha vida.

Aos meus colegas da Norus Tecnologia, empresa onde trabalho, gostaria de agradecer pela ajuda, amizade e companheirismo. Sou grato, em especial aos integrantes do grupo em que trabalho, Paulo Vitor Larroyd, Gabriel Teixeira e Felipe Béltran Rodriguez. Suas contribuições e orientações foram essenciais para o desenvolvimento deste trabalho e seus conselhos essenciais para a minha vida profissional.

Por fim, os meus orientadores, professor Erlon Finardi, agradeço pela orientação neste trabalho, pelo apoio e pelo aprendizado que as discussões sobre o assunto me proporcionaram durante o desenvolvimento deste trabalho.

## RESUMO

No contexto da transição energética, muitas nações ao redor do mundo estão investindo na integração de fontes de energia renováveis, tais como a eólica e a solar, com o intuito de reduzir as emissões de gás de efeito estufa. Um dos problemas associados à estas fontes é a natureza intermitente e o grau significativo de incerteza. Neste contexto, o operador do sistema deve considerar adequadamente estas incertezas para planejar a operação, mantendo a segurança do sistema elétrico. Uma estratégia amplamente utilizada pelos operadores de sistemas elétricos é empregar previsões de geração. Contudo, no planejamento da operação para o dia seguinte não é possível obter previsões perfeitas da geração de renováveis. Neste contexto, este trabalho tem o objetivo de propor um modelo capaz de gerar cenários a partir das previsões de geração eólica empregados pelo Operador Nacional do Sistema Elétrico (ONS) no programa de despacho de curtíssimo prazo DESSEM. Para tanto, uma metodologia baseada no modelo Autorregressivo Periódico (PAR) com matriz de carga nos resíduos é proposta. Para parametrizar este modelo é necessário utilizar os dados de previsão de geração dos arquivos de entrada do DESSEM e os dados de geração observada do histórico fornecidos pelo ONS. A ideia é gerar desvios de geração eólica a partir do uso de uma série temporal de previsão de geração escolhida pelo usuário. Por sua vez, o modelo é responsável por replicar as propriedades estatísticas observadas nos dados de treinamento na série escolhida pelo usuário. A qualidade do modelo parametrizado é verificada por meio de duas etapas de validação. A primeira consiste em usar uma curva de previsão de geração retirada da base de dados de treinamento, enquanto a segunda consiste em usar uma curva de previsão externa. De modo geral, os resultados mostram que o modelo é capaz de gerar cenários de desvios coerentes que reduzem os erros de previsão de geração eólica. Neste sentido, o modelo pode ser empregado para quantificar as incertezas de fonte eólica, tornando a operação do sistema mais segura. Destaca-se que a metodologia proposta pode ser aplicada para reduzir as incertezas em outros contextos tais como a de geração solar e das afluências hidrológicas.

**Palavras-chave:** Transição energética. Incertezas. Geração Eólica. Desvios de Geração. Modelo PAR com matriz de carga nos resíduos. Modelo DESSEM.

## ABSTRACT

In the context of the energy transition, many nations worldwide are investing in integrating renewable energy sources, such as wind and solar, into their electricity mix to reduce greenhouse gas emissions. One of the challenges associated with these sources is their intermittency and their significant degree of uncertainty. Therefore, system operators have to deal with these uncertainties to plan the operation while maintaining the security of the electrical power system. One of the solutions widely used by power system operators is forecast the power generated. Nonetheless, it is impossible to obtain perfect forecasts regarding power generation from renewable sources for the day-ahead horizon. This work proposes a model capable of generating forecast deviation scenarios from the wind generation forecasts employed by ONS in the short-term unit commitment model called DESSEM. To achieve this goal, a methodology based on the Periodic Autoregressive Model (PAR) with load matrix in the residuals is proposed. The model is parametrized by using the generation forecast data from the input files of the DESSEM model and the observed wind power generation data from the history provided by ONS. The idea of the model is to generate wind generation deviations from a wind generation forecast series chosen by the user. Then, the model applies the statistical properties observed in the training data in the time series chosen by the user. The quality of the model is verified through two validation steps. In the first step, the model is validated using a wind generation forecast series chosen from the training database. In the second step, a time series unknown to the model is used. The results show that the model could generate coherent forecast deviation scenarios that reduce wind generation forecast errors. In this sense, the model can be employed to quantify wind source uncertainties, making the system operation safer. The proposed methodology can also reduce uncertainties in other cases, such as solar generation and hydrological inflows.

**Keywords:** Energy Transition. Uncertainties. Wind Power Generation, Forecast deviation. Periodic Autoregressive Model (PAR) with load matrix in the residuals. DESSEM model.

## LISTA DE FIGURAS

Figura 1 – Comparação da geração eólica prevista (escala de cinza) e observada (azul) para o SIN entre os dias 2 e 8/1/2021 .....	18
Figura 2 – Comparação dos desvios de previsão para o SIN para o período entre janeiro de 2020 e abril de 2022 de acordo com o horizonte de previsão. ....	19
Figura 3 – Comparação entre as previsões de geração uma hora à frente e a geração verificada do sistema de energia Bonneville (Estados Unidos). ....	25
Figura 4 – Histograma dos desvios de previsão um dia à frente para o sistema alemão em 2010. Média, desvio padrão, assimetria e curtose iguais a $\mu = 0.0092$ , $\sigma = 0.045$ , $\gamma = -0.2891$ e $\kappa = 3.5896$ respectivamente. ....	26
Figura 5 – a) Função de densidade de Probabilidade, b) <i>Boxplots</i> dos desvios horário de geração eólica entre 2010 e 2013 por cluster. ....	28
Figura 6 – Histogramas dos desvios normalizados e respectivos parâmetros das distribuições obtidas de cada modelo de previsão de geração eólica. ....	29
Figura 7 – Exemplo de uma distribuição normal truncada.....	29
Figura 8 – Densidade de probabilidade do erro e simulação da geração eólica. ....	30
Figura 9 – Comparação dos histogramas e função de densidade acumulada da WPFE estimada pela distribuição Normal, Beta e Lévy- $\alpha$ estável.....	31
Figura 10 – Distribuição normalizada dos desvios observados, Cauchy, Weibull e Beta. ....	32
Figura 11 – Distribuição dos desvios de previsão normalizados. (a) Curvas quantílicas sobre a distribuição dos desvios. (b) Função de densidade de probabilidade.....	33
Figura 12 – Conjunto de 25 Cenários de Previsão de Geração Eólica para múltiplas fazendas eólicas na Dinamarca para um horizonte de tempo de 43h. ....	34
Figura 13 – Comparação entre a geração eólica real, a prevista e os 10 cenários de previsão. ....	34
Figura 14– Cenários de Previsão de Geração Eólica. (a) Baixa geração. (b) Alta geração. ....	35
Figura 15 – Gráfico acumulado da distribuição condicional dos desvios de previsão para diferentes níveis de previsão de geração. ....	35
Figura 16 – Evolução da capacidade eólica instalada no Brasil.....	36
Figura 17– Curva Vento Potência estimada para o trigésimo período do dia. ....	38
Figura 18 – Metodologia para previsão do vento e geração eólica do ONS. ....	39
Figura 19 – Gradação linear de pesos para a suavização dos desvios por intervalo. ....	40

Figura 20 – Exemplo da distribuição dos desvios brutos do subsistema Nordeste (NE) e Sul (S) entre janeiro 2017 e novembro 2018. ....	43
Figura 21 – Desvio da Geração Eólica entre setembro de 2020 e agosto de 2021 para o SIN e subsistemas Nordeste, Norte e Sul. ....	43
Figura 22 – EOL-OBS horária para o sistema brasileiro no dia 01/02/2021 .....	46
Figura 23 – Esquemático de execução do modelo DESSEM.....	47
Figura 24 – Exemplo de Arquivo com a EOL-PREV no modelo DESSEM .....	48
Figura 25 – Exemplo cálculo do erro de previsão para EOL-OBS e da EOL-PREV com diferentes bases temporais. ....	49
Figura 26 – Exemplo dos desvios observados referentes ao mês de janeiro de 2021 para o ponto de conexão ACARAU II – 69 kV.....	51
Figura 27– Exemplo de cenário de geração eólica e desvios previstos .....	52
Figura 28 – Metodologia geral de geração de cenários .....	52
Figura 29 – Função de Autocorrelação obtido para o histórico de geração prevista de janeiro de 2021 para o ponto de conexão ACARAU II – 69 kV.....	54
Figura 30 – Função de Autocorrelação obtido para os desvios de previsão de janeiro de 2021 para o ponto de conexão ACARAU II – 69 kV.....	54
Figura 31 - Média móvel (curva em preto) e média global (curva em azul) antes (a) e depois (b) do processo de estacionarização para o ponto de conexão Grupo João Piauú .....	63
Figura 32- Distribuição das amostras antes (a) e depois (b) e a comparação da distribuição cumulativa (c) antes e depois da estacionarização para o ponto de conexão Grupo João Piauú .....	64
Figura 33 - Boxplots das amostras por hora para o ponto de conexão Grupo Ibiapina para o período entre 2 de janeiro a 31 de janeiro de 2021 com e sem truncamento. ....	66
Figura 34 – Mapa de calor correlação espacial entre os pontos de conexão agrupados referente à hora 0 de janeiro de 2021 dos dados de treinamento .....	67
Figura 35 – Mapa de calor correlação espacial entre os pontos de conexão agrupados referente à hora 0 de janeiro de 2021 da série sintética de desvios .....	68
Figura 36 – Localização dos pontos de conexão agrupados no mapa do SIN.....	69
Figura 37- Comparação da distribuição dos desvios, geração prevista e observada para o ponto de conexão Grupo Formoso Bonfim para o período entre 23 de janeiro a 30 de janeiro de 2021 .....	70

Figura 38- Comparação da distribuição dos desvios, geração prevista e observada para o ponto de conexão Grupo Quinta para o período entre 23 de janeiro a 30 de janeiro de 2021 .....	70
Figura 39- Comparação da distribuição dos desvios, geração prevista e observada para o ponto de conexão Grupo Ceará Extremoz para o período de 2 a 9 de janeiro de 2021 .....	71
Figura 40 - Boxplots das amostras por hora para o ponto de conexão Grupo Ibiapina para o período entre 2 de janeiro a 31 de janeiro de 2022 após truncamento .....	73
Figura 41- Comparação da distribuição dos desvios, geração prevista e observada para o ponto de conexão Grupo Formoso Bonfim para o período entre 23 de janeiro a 30 de janeiro de 2022 .....	73
Figura 42- Comparação da distribuição dos desvios, geração prevista e observada para o ponto de conexão Grupo Quinta para o período entre 23 de janeiro a 30 de janeiro de 2022.....	74
Figura 43- Comparação da distribuição dos desvios, geração prevista e observada para o ponto de conexão Grupo Ceará Extremoz para o período entre 2 de janeiro a 9 de janeiro de 2021 .	74
Figura B.1 – Variação Sazonal da temperatura mínima diária entre 1971 e 1973 do Central Park- Nova York.....	87
Figura B.2 – Exemplo de séries temporais estacionárias e não estacionárias .....	88
Figura C.1 – Afluência média mensal do rio Saugeen (Canadá) entre janeiro de 1967 à dezembro de 1976.....	90

## LISTA DE TABELAS

Tabela 1 – Definição dos horizontes de tempo dos métodos de previsão e suas aplicações na operação dos sistemas de energia. ....	22
Tabela 2 – Comparação dos desvios de geração eólica em diferentes países e horizontes de tempo. ....	25
Tabela 3 – Métodos Paramétricos de Determinação da WPFE.....	26
Tabela 4 – Erro de previsão de geração eólica uma hora à frente para o CAISO em 2006 expresso em fração da capacidade instalada por estação do ano. ....	30
Tabela 5 – Medidas normalizadas dos momentos estatísticos do erro. ....	31
Tabela 6 – Métodos Não-Paramétricos de Determinação da WPFE.....	32
Tabela 7 – Desvios de previsão de fonte eólica no Nordeste entre outubro 2017 a setembro 2018. ....	42
Tabela 8 – Desvios de previsão de fonte eólica no Sul entre outubro 2017 a setembro 2018. ....	42
Tabela 9 – Características da estrutura dos dados da EOL-OBS e da EOL-PREV.....	49
Tabela 10– Características da estrutura dos dados da EOL-OBS e da EOL-PREV.....	50
Tabela A.1 - Comparação dos modelos determinísticos de previsão de ventos e geração eólica	85
Tabela A.2 - Comparação dos modelos probabilísticos de previsão de ventos e geração eólica	86
Tabela E.1 Lista dos pontos de conexão da base de dados antes e depois do reagrupamento ..	98

## LISTA DE ABREVIATURAS E SIGLAS

Abeeólica	Associação Brasileira de Energia Eólica
CCEE	Câmara de Comercialização de Energia Elétrica
CEPEL	Centro de Pesquisas de Energia Elétrica
COPT	Tabela de Probabilidade de Interrupção da Capacidade de Geração - <i>Capacity Outage Probability Table</i>
CPTEC	Centro de Previsão de Tempo e Estudos Climáticos
ECMWF	Centro Europeu de Previsão de Tempo para médio Prazo
EOL-OBS	Geração Eólica Observada
EOL-PREV	Geração Eólica Prevista
EPE	Empresa de Pesquisa Energética
FERV	Fonte de Energia Renovável Variável
NOAA	Administração Oceânica e Atmosférica Nacional
ONS	Operador Nacional do Sistema Elétrico
PAR	Modelo Autorregressivo Periódico - <i>Periodic Autoregression Model</i>
PARMA	Modelo Autorregressivo de Médias Móveis - <i>Periodic Autoregression Moving Averages Model</i>
PEACF	Função de Autocorrelação Periódica - <i>Periodic Autocorrelation Function</i>
PEACVF	Função de Autocovariância Periódica - <i>Periodic Autocovariance Function</i>
PEPACF	Função de Autocorrelação Periódica Parcial - <i>Partial Periodic Autocorrelation Function</i>
PDO	Programação Diária da Operação
SIN	Sistema Interligado Nacional
TND	Distribuição Gaussiana Truncada - <i>Truncated Normal Distribution</i>
WPFE	Erro de Previsão de Geração Eólica - <i>Wind Power Forecasting Error</i>
WPF	Previsão de Geração Eólica - <i>Wind Power Forecasting</i>

## LISTA DE SÍMBOLOS

MW	Megawatt
$x_{d,h}$	Valor da grandeza $x$ a ser determinado para o dia $d$ e hora $h$
$\mu x_h$	Média dos registros $x$ na hora $h$
$\sigma \eta_h$	Desvio padrão dos registros $x$ na hora $h$
$\sigma \eta_h^r$	Desvio padrão dos resíduos do modelo PAR para a hora $h$
$\gamma_0^{(h)}$	Variância dos registros $x$ na hora $h$
$\rho_k^{(h)}$	Função de Autocorrelação periódica para a hora $h$ com lag $k$
$\gamma_k^{(h)}$	Função de Autocovariância periódica para a hora $h$ com lag $k$
$np^{(h)}$	Ordem máximo do modelo PAR referente à hora $h$
$\phi_p^{(h)}$	Coefficiente autorregressivo do modelo PAR na hora $h$ de ordem máxima
$nr_{PC}$	Número de pontos de conexão do SIN contendo usinas eólicas
$\omega$	Ruído branco aleatório com distribuição normal de média zero e variância unitária
$\omega_h$	Vetor de ruídos brancos independentes de dimensão $(nr_{PC}, 1)$ na hora $h$
$\omega_{c,h}$	Vetor de ruídos brancos correlacionados de dimensão $(nr_{PC}, 1)$ na hora $h$
$RC_h$	Matriz de carga triangular inferior de dimensão $(nr_{PC}, nr_{PC})$
$\rho RC_h$	Matriz de correlação para o período de hora $h$ de dimensão $(nr_{PC}, nr_{PC})$
$E[Y]$	Valor esperado de $Y$

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO .....</b>	<b>17</b>
1.1	CONTEXTUALIZAÇÃO E MOTIVAÇÃO .....	19
1.2	OBJETIVOS .....	20
1.2.1	<b>Objetivos Gerais.....</b>	<b>20</b>
1.2.2	<b>Objetivos Específicos .....</b>	<b>21</b>
1.3	ESTRUTURA DO TRABALHO .....	21
<b>2</b>	<b>REVISÃO DO ESTADO DA ARTE.....</b>	<b>22</b>
2.1	MÉTODOS DE PREVISÃO DE GERAÇÃO EÓLICA .....	23
2.1.1	<b>Métodos de Previsão Determinísticos .....</b>	<b>23</b>
2.1.2	<b>Métodos de Previsão Probabilísticos.....</b>	<b>23</b>
2.2	MÉTODOS DE DETERMINAÇÃO DOS DESVIOS DE PREVISÃO DE GERAÇÃO EÓLICA .....	24
2.2.1	<b>Desvios de Previsão Eólica: Experiência Internacional .....</b>	<b>24</b>
2.2.2	<b>Métodos de Previsão de Erro Paramétricos .....</b>	<b>26</b>
2.2.3	<b>Métodos de Previsão de Desvios Não Paramétricos .....</b>	<b>32</b>
2.3	CASO BRASILEIRO: MÉTODOS DE PREVISÃO DA GERAÇÃO EÓLICA PARA O SIN .....	36
2.3.1	<b>Preparação dos Dados .....</b>	<b>37</b>
2.3.2	<b>Previsão da Geração Eólica .....</b>	<b>38</b>
<b>3</b>	<b>METODOLOGIA PARA A PARAMETRIZAÇÃO DO MODELO DE GERAÇÃO DE CENÁRIOS .....</b>	<b>45</b>
3.1	DESCRIÇÃO E PROCESSAMENTO DOS DADOS .....	45
3.2	PROCESSO DE CÁLCULO DOS DESVIOS DE PREVISÃO .....	50
3.3	DESCRIÇÃO DO MODELO DE GERAÇÃO DE CENÁRIOS DE DESVIOS .....	51
3.3.1	<b>Descrição do Algoritmo de Parametrização do modelo PAR.....</b>	<b>55</b>
<b>4</b>	<b>RESULTADOS .....</b>	<b>62</b>

4.1	BASE DE DADOS .....	62
4.1.1	<b>Tratamento da Base de Dados .....</b>	<b>62</b>
4.2	RESULTADOS DAS SIMULAÇÕES .....	65
4.2.1	<b>Base de dados de Treinamento .....</b>	<b>65</b>
4.2.2	<b>Base de dados de Validação Final .....</b>	<b>72</b>
4.3	CONSIDERAÇÕES FINAIS .....	74
5	<b>CONCLUSÕES.....</b>	<b>77</b>
	<b>APÊNDICE A - Modelos determinísticos e probabilísticos de previsão de geração eólica .....</b>	<b>85</b>
	<b>APÊNDICE B - Requisitos para parametrização do modelo PAR.....</b>	<b>87</b>
	<b>APÊNDICE C - Modelo Periódico Autorregressivo (PAR).....</b>	<b>90</b>
	<b>APÊNDICE D - Integração da Correlação Espacial .....</b>	<b>96</b>
	<b>APÊNDICE E - Lista de Pontos de Conexão .....</b>	<b>98</b>

## 1 INTRODUÇÃO

Atualmente, a tendência de descarbonização dos sistemas elétricos motivada pelas mudanças climáticas incentivou uma ampla disseminação de Fontes de Energia Renovável Variáveis (FERV) nos sistemas elétricos ao redor do mundo. Além do fator socioambiental, nos últimos anos, essas fontes têm apresentado preços cada vez mais competitivos, dominando o panorama de novos projetos de geração ao redor de todo o globo. Nos últimos 20 anos, experimentou-se um aumento de 4.652%<sup>1</sup> da produção mundial de geração eólica (OUR WORLD IN DATA, 2022a) e, atualmente, em países como Alemanha, Espanha e a República da Irlanda, mais de 20% da geração provêm da fonte eólica. Por outro lado, no Brasil, cerca 11% é proveniente da geração eólica (OUR WORLD IN DATA, 2022b).

Não obstante, suas vantagens socioambientais e econômicas, a forte disseminação das FERVs traz consigo novos desafios à operação dos sistemas elétricos devido às incertezas associadas a esse tipo de fonte e à alta variabilidade na geração de eletricidade (HOLTTINEN, 2005). Conseqüentemente, os impactos energéticos e elétricos vão desde o pré-despacho, programação intra-diária até o planejamento em horizontes maiores da operação do sistema. Assim, para que a ampla implantação de FERVs seja bem-sucedida e ao mesmo tempo atender as demandas da sociedade e do operador do sistema, é necessário um amplo entendimento dos desafios atuais e futuros decorrentes da alta penetração de FERVs em toda cadeia de operação do sistema.

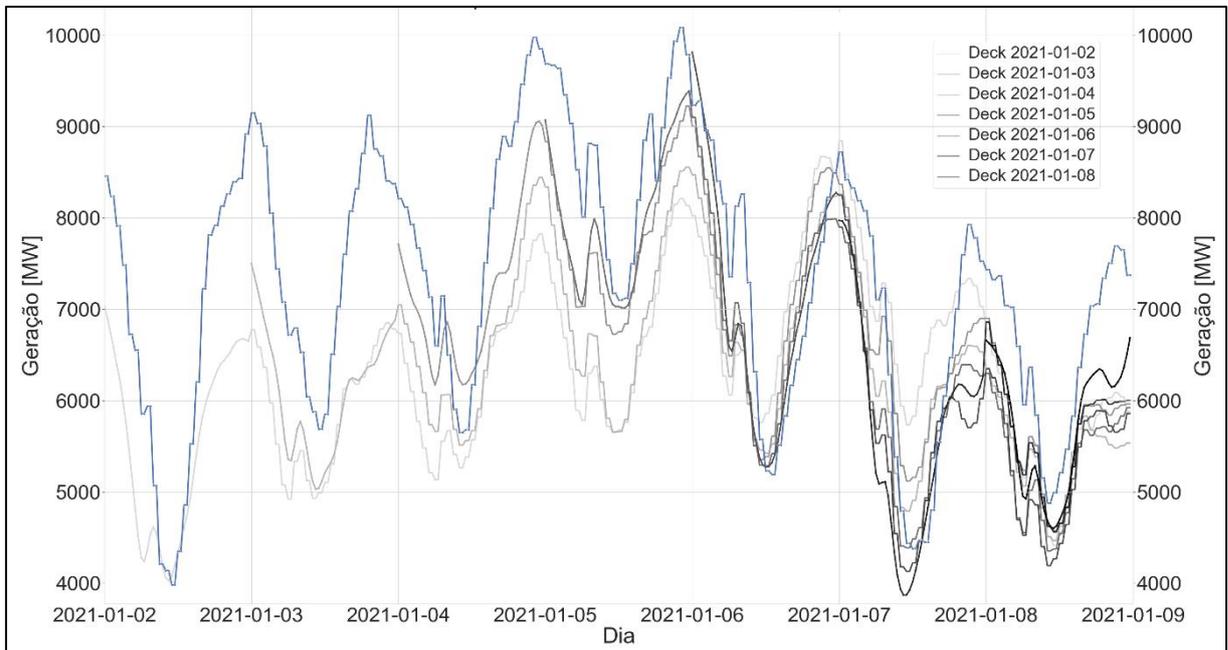
Especificamente, a alta penetração das FERVs em um sistema de potência aumenta a natureza estocástica da matriz elétrica devido ao caráter intermitente destas fontes de energia (GONZÁLEZ-APARICIO; ZUCKER, 2015). Apesar de dispor na atualidade de métodos sofisticados de previsão, diferenças significativas ainda são encontradas entre a geração prevista e a realizada, sendo mais frequentes em sistemas com uma alta penetração das FERVs. Assim, a tomada de decisão operativa deve considerar a incerteza na disponibilidade prevista dos recursos energéticos para garantir os requisitos de segurança e confiabilidade do sistema. De forma ilustrativa, a Figura 1 apresenta uma comparação da geração eólica observada e prevista para a semana operativa entre os dias 02 e 08/01/2021 no sistema brasileiro. Note que a geração prevista apresenta uma atualização a cada dia da semana (previsão 1 dia à frente) até o final da

---

<sup>1</sup> A geração mundial de energia eólica saltou de 38 TWh em 2001 para 1813 TWh em 2021.

semana<sup>2</sup>. Neste exemplo, a máxima amplitude dos desvios das previsões feitas para o dia seguinte<sup>3</sup> são de 623 MW (sobre estimação) para o deck do dia 3/1/2021 e de -2760 MW (subestimação) para o deck do dia 4/1/2021.

Figura 1 – Comparação da geração eólica prevista (escala de cinza) e observada (azul) para o SIN entre os dias 2 e 8/1/2021



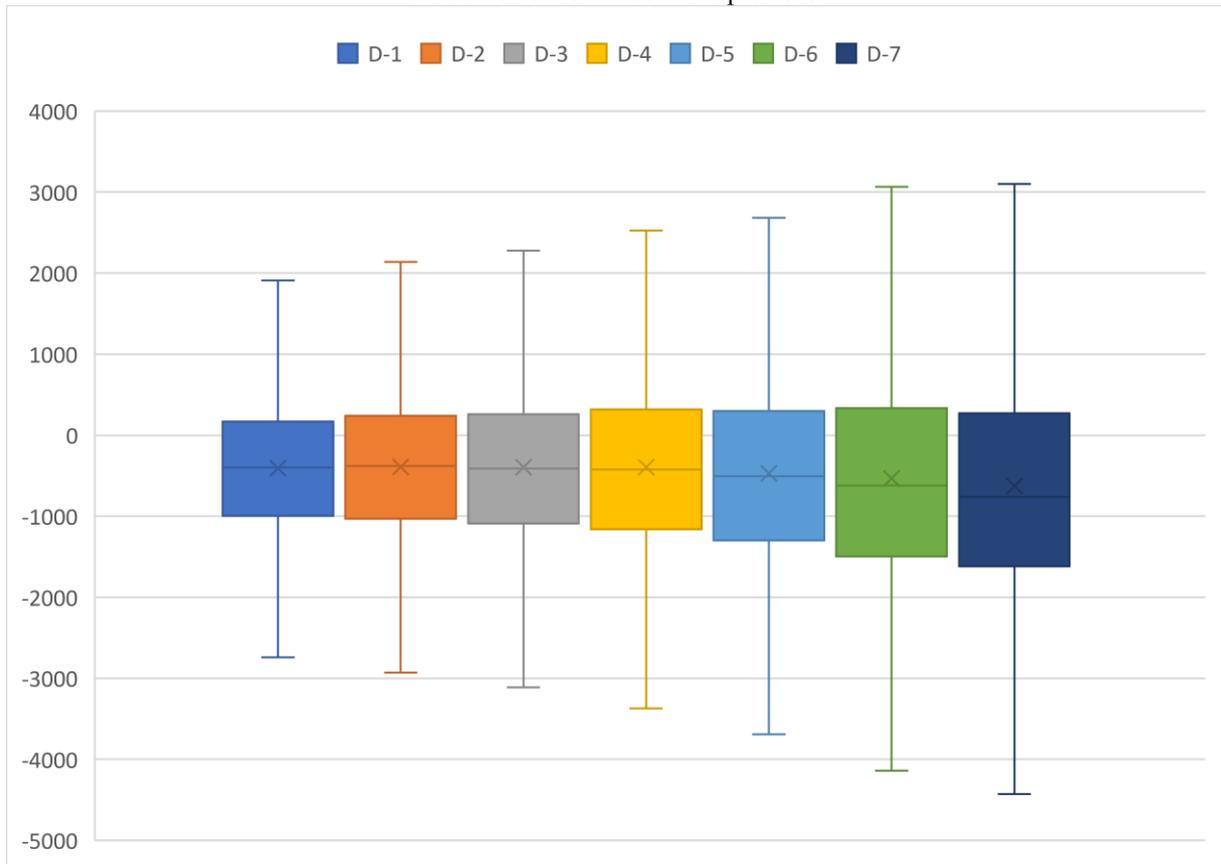
Fonte: do autor.

Percebe-se a partir da Figura 1 que cada atualização da previsão prevê até o final da semana operativa. i.e., a previsão de geração eólica do deck 2021-01-03 possui valores previstos até a meia noite do dia 8/1/2021. Um dos fatos que deve ser notado é que a magnitude dos desvios de previsão varia de acordo com o horizonte previsto. Isto é, as previsões feitas no dia D-1 para o dia D tendem a ter menores desvios se comparados às previsões feitas, por exemplo, no D-2 (dois dias antes da operação). A Figura 2 mostra que as previsões geradas no D-1 tendem a ser mais próximas da geração observada visto que a dispersão dos desvios é menor.

<sup>2</sup> O Operador Nacional do Sistema Elétrico (ONS) define como uma semana operativa o período de sete dias entre as meias noites de duas sextas-feiras consecutivas.

<sup>3</sup> Previsões feitas no dia que antecede o dia da operação (D-1).

Figura 2 – Comparação dos desvios de previsão para o SIN para o período entre janeiro de 2020 e abril de 2022 de acordo com o horizonte de previsão.



Fonte: do autor.

Ressalta-se que os dados relativos às previsões de geração eólica mostrados na Figura 1 foram obtidos a partir do conjunto de dados de entrada diários (decks) do modelo computacional DESSEM<sup>4</sup>. Já os dados de geração observada foram obtidos por meio do histórico de geração fornecido pelo Operador Nacional do Sistema Elétrico (ONS) (ONS, 2022a). Por fim, os desvios mostrados na Figura 2 foram calculados a partir da diferença da geração prevista e a observada.

### 1.1 CONTEXTUALIZAÇÃO E MOTIVAÇÃO

No Brasil, o ONS determina as diretrizes da operação para o dia seguinte com base na resolução de um problema de *unit commitment* denominado programação diária da operação (PDO). Para formular o problema da PDO, são requeridas previsões de um dia à frente de variáveis aleatórias como a afluência nas hidrelétricas, a geração eólica, a geração solar e o comportamento da carga. Além disso, na modelagem do problema, esses dados são considerados determinísticos,

<sup>4</sup> O modelo DESSEM determina o despacho de geração de potência ativa para a programação diária de operação.

ou seja, o despacho das usinas é feito com base em apenas um único cenário. Contudo, não é possível se obter previsões perfeitas, já que há uma aleatoriedade significativa nestas variáveis. Neste contexto, é imprescindível caracterizar os desvios de previsão<sup>5</sup> das FERVs para que as margens de segurança operacional sejam incorporadas no processo de tomada de decisão, e, conseqüentemente mitigar os riscos associados à falta ou excesso de geração. A grande questão é se as margens operativas definidas com horas ou até dias de antecedência serão suficientes para evitar desvios significativos entre a previsão e os valores observados.

Uma das abordagens para responder esta questão é a caracterização dos desvios de previsão das variáveis com maior grau de incerteza. Assim, este trabalho propõe um foco nos desvios de previsão da geração eólica pelas seguintes razões:

- 1) a significativa porcentagem de capacidade instalada da geração eólica em respeito à capacidade total do sistema e perspectivas de um alto crescimento a médio prazo (12.3% - 21.6 GW em 2022 e 13.9% - 28.7 GW em 2026) (ONS, 2022b) e;
- 2) a previsão da geração eólica apresenta um grau de incerteza maior quando comparada com a previsão da geração solar, carga e aflúências (GONZÁLEZ-APARICIO; ZUCKER, 2015) em horizontes de curtíssimo prazo, i.e., 1 dia.

## 1.2 OBJETIVOS

Para o presente trabalho, identificam-se os objetivos gerais e específicos descritos nas seções a seguir.

### 1.2.1 Objetivos Gerais

Este trabalho tem como objetivo desenvolver um modelo capaz de gerar cenários de desvios de geração eólica a partir dos históricos de previsão do ONS e de geração verificada do SIN. Salienta-se que este modelo será empregado para gerar cenários de desvios para o horizonte da PDO.

---

<sup>5</sup> Por exemplo, identificar a magnitude, a probabilidade de ocorrência, a sazonalidade, a periodicidade e as correlações temporais e espaciais dos erros.

### 1.2.2 Objetivos Específicos

Os objetivos específicos deste trabalho são:

- 1) Identificação dos modelos empregados internacionalmente e nacionalmente para a integração das incertezas de geração eólica a partir da revisão do estado de arte;
- 2) Determinação dos desvios de previsão de geração eólica do SIN;
- 3) Análise das séries temporais de desvios de previsão e de geração eólica e identificação de propriedade estatísticas;
- 4) Parametrização dos modelos de geração de cenários compatíveis com a PDO;
- 5) Validação do modelo parametrizado.

### 1.3 ESTRUTURA DO TRABALHO

Este trabalho está organizado em cinco capítulos. No Capítulo 2 é apresentado uma revisão do estado da arte sobre os modelos de previsão de geração eólica e de determinação dos desvios de previsão empregados internacionalmente. Este capítulo apresenta, igualmente, um panorama a respeito dos modelos empregados pelo ONS para prever a geração eólica para o problema da PDO. Em seguida, o Capítulo 3 descreve a metodologia empregada no trabalho para parametrizar os modelos de geração de cenários de desvios. Neste capítulo, são descritos os dados empregados nos modelos, assim como as questões relacionadas aos tratamentos necessários para obter uma base de dados consistente. O Capítulo 4 apresenta os resultados obtidos por meio da aplicação da metodologia descrita no Capítulo 3. Finalmente, as conclusões são descritas no Capítulo 5, juntamente com as discussões acerca das limitações dos modelos utilizados e sugestões para trabalhos futuros.

## 2 REVISÃO DO ESTADO DA ARTE

Um dos maiores desafios para a integração em grande escala de fontes de energia eólica é relacionado ao perfil intermitente dos ventos. Para uma melhor representação da real capacidade de geração destas fontes, verifica-se a necessidade de modelos de previsão de geração eólica com níveis de confiança adequados. Nos últimos anos, diversos modelos de previsão de vento e de geração têm sido desenvolvidos e ainda seguem sendo aprimorados. Uma compilação do estado da arte sobre estes métodos entre os anos 2000 e 2020 pode ser revisada em BAZIONIS; GEORGILAKIS (2021a).

Os modelos de previsão podem ser classificados em dois grandes grupos: modelos determinísticos e probabilísticos (BAZIONIS; GEORGILAKIS, 2021a). Para um determinado horizonte de previsão, os modelos probabilísticos fornecem intervalos de confiança da geração prevista, enquanto os métodos determinísticos fornecem uma estimativa única sem intervalo de confiança. Neste capítulo são introduzidos os diferentes métodos de previsão de vento e de geração eólica encontrados na literatura. Além disso, apresentam-se experiências internacionais a respeito da caracterização dos desvios de previsão de geração eólica. Para concluir, a metodologia empregada atualmente pelo ONS para a previsão da geração eólica é discutida.

Antes de começar a descrever o estado da arte é necessário realizar seguintes definições a respeito ao horizonte de previsão de geração eólica. Este aspecto é de extrema relevância visto que quanto maior o horizonte de tempo da previsão, maior será o nível de incertezas. De forma esquemática, a Tabela 1 apresenta uma definição de diferentes horizontes de previsão conforme o intervalo de tempo abrangente, além de possíveis aplicações em sistemas de energia elétrica. É importante notar que a definição dos horizontes de tempo mostrado pela tabela abaixo é apenas indicativa e pode mudar de acordo com os sistemas e mercados de energia.

Tabela 1 – Definição dos horizontes de tempo dos métodos de previsão e suas aplicações na operação dos sistemas de energia.

<b>Horizonte de previsão</b>	<b>Intervalo de tempo</b>	<b>Aplicações</b>
Curtíssimo prazo	Segundos a 30 minutos à frente	- <i>Electricity Market Clearing</i> - Ações regulatórias (Frequência e Tensão)
Curto prazo	30 minutos a seis horas à frente	- Planejamento do despacho econômico - Decisões de balanceamento de carga
Médio prazo	seis horas a um dia à frente	- Decisão de inclusão de geradores do despacho - Decisões de segurança do provisionamento para 1 dia à frente

Longo prazo	um dia a uma semana à frente	<ul style="list-style-type: none"> <li>- Decisão de despacho da geração</li> <li>- Requisitos de reserva de geração</li> <li>- Calendário de manutenção das unidades geradoras</li> </ul>
-------------	------------------------------	---

Fonte: SOMAN et al. (2010).

## 2.1 MÉTODOS DE PREVISÃO DE GERAÇÃO EÓLICA

Devido à grande intermitência da fonte eólica e seus impactos na operação de sistemas elétricos, tornou-se necessário o desenvolvimento de modelos capazes de prever a geração eólica em diversos horizontes de tempo e as incertezas associadas à estas previsões. Este capítulo possui o objetivo de introduzir o assunto a partir da experiência internacional.

### 2.1.1 Métodos de Previsão Determinísticos

O desenvolvimento e aperfeiçoamento de métodos determinísticos tem sido uma área de grande interesse e de intensa pesquisa (BAZIONIS; GEORGILAKIS, 2021a). Estes métodos obtêm como saída um valor determinístico de geração. Dentre os métodos determinísticos, os métodos estatísticos utilizam dados históricos de geração eólica e velocidade dos ventos para realizar suas previsões. Métodos mais específicos, como os denominados físicos, precisam de informações adicionais como topologia, temperatura, pressão atmosférica e obstáculos ao vento. Destacam-se também os métodos híbridos, que são modelos que utilizam inteligência computacional em conjunto com os dados fornecidos por outros métodos a fim de refinar o resultado. Tais métodos têm sido desenvolvidos para aumentar a acurácia das previsões. A Tabela A.1 (APÊNDICE A - Modelos determinísticos e probabilísticos de previsão de geração eólica) apresenta uma classificação de diferentes métodos de previsão determinísticos, contendo os horizontes de previsão em que são normalmente utilizados e suas respectivas referências.

### 2.1.2 Métodos de Previsão Probabilísticos

Os métodos probabilísticos levam em conta o fato de que gerar previsões exatas é uma tarefa árdua e praticamente impossível. Assim, possuem uma outra abordagem e obtêm como saída um intervalo de confiança de previsão (BAZIONIS; GEORGILAKIS, 2021a), i.e., um *range* de valores onde o valor observado pode realizar com uma probabilidade associada. Os métodos probabilísticos são divididos em duas classes: paramétricos e não paramétricos. Para os paramétricos é necessário predefinir uma distribuição de probabilidade. Ressalta-se que a

escolha da função de distribuição é realizada com base nas características da base de dados. De forma geral, as distribuições do tipo Gaussiana ou Beta são as distribuições mais amplamente utilizadas na literatura (BRUNINX; DELARUE, 2014).

No caso dos métodos não paramétricos, nenhum pressuposto a respeito da distribuição é necessário, visto que a distribuição será estimada diretamente a partir da base de dados fornecida. É importante ressaltar que devido à ausência de uma hipótese inicial a respeito da função de distribuição, a implementação destes métodos é mais complexa e necessita de uma base de dados mais ampla e detalhada em comparação aos métodos paramétricos (BAZIONIS; GEORGILAKIS, 2021a). A Tabela A.2 (APÊNDICE A - Modelos determinísticos e probabilísticos de previsão de geração eólica) apresenta a classificação dos métodos probabilísticos, assim como o horizonte de tempo comumente utilizado para sua exploração e referências associadas.

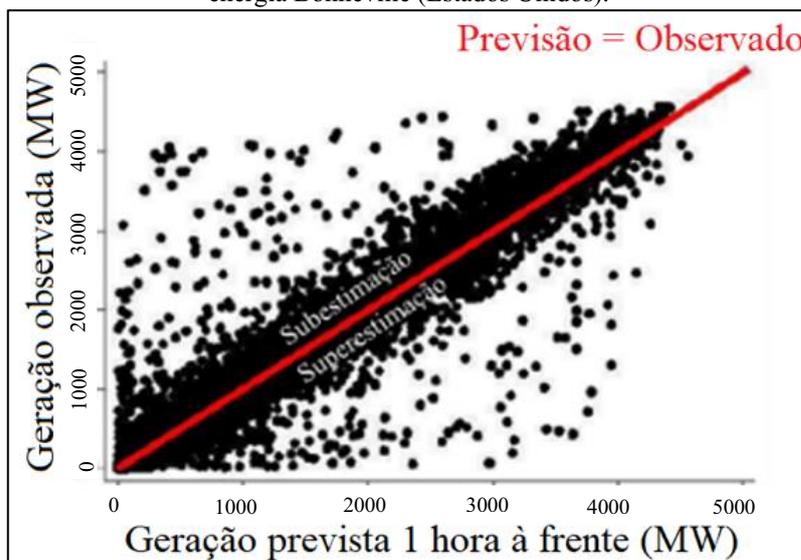
## 2.2 MÉTODOS DE DETERMINAÇÃO DOS DESVIOS DE PREVISÃO DE GERAÇÃO EÓLICA

Os métodos de determinação do erro de previsão de geração eólica, também conhecidos como *Wind Power Forecasting Error* (WPFE), são classificados como métodos probabilísticos visto que fornecem como saída uma função de densidade de probabilidade. Esta função descreve a variabilidade do erro de previsão a partir de dados históricos (GONZÁLEZ-APARICIO; ZUCKER, 2015). A necessidade de determinar o WPFE se deve ao fato de que nenhum método de previsão é capaz de gerar previsões de geração eólica (*Wind Power Forecast* - WPF) exatas, devido à incerteza inerente à esta fonte e aos métodos de previsão ainda em desenvolvimento. Desta forma, a WPFE é utilizada para caracterizar o erro das previsões, diminuindo o impacto da integração da fonte eólica nos sistemas elétricos.

### 2.2.1 Desvios de Previsão Eólica: Experiência Internacional

A experiência internacional em relação à previsão de energia eólica tem mostrado que, mesmo em um horizonte de tempo de apenas uma hora antes do despacho real, os métodos de previsão podem apresentar diferenças significativas com o valor observado, como apresentado na Figura 3.

Figura 3 – Comparação entre as previsões de geração uma hora à frente e a geração verificada do sistema de energia Bonneville (Estados Unidos).



Adaptado de FORBES; ZAMPELLI (2016).

A Tabela 2 apresenta a média e o desvio padrão dos desvios normalizados para diferentes sistemas elétricos e horizontes de previsão. Aproximando a distribuição dos erros com uma distribuição normal, obtêm-se os desvios de previsão de geração eólica apresentados para diferentes probabilidades de ocorrência.

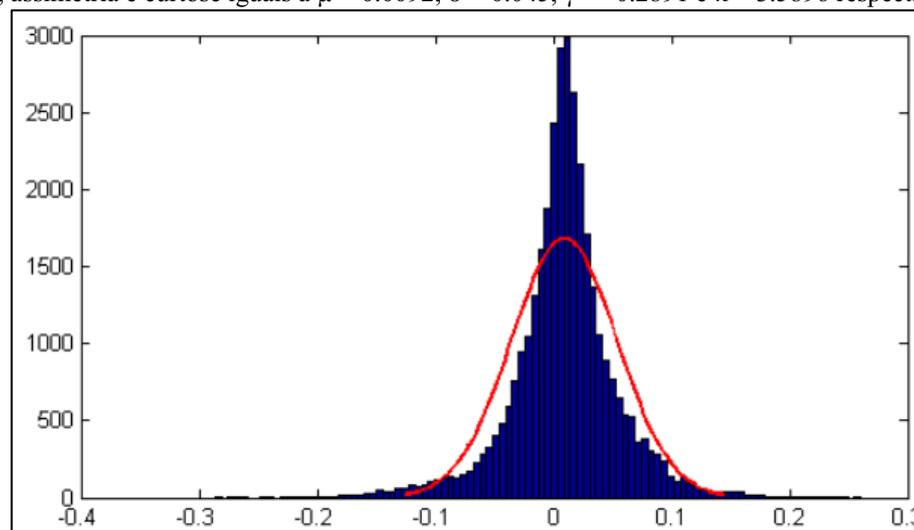
Tabela 2 – Comparação dos desvios de geração eólica em diferentes países e horizontes de tempo.

Sistema /Ano	Capacidade Eólica [MW]	Horizonte de Previsão	Média $\mu$	Desvio Padrão $\sigma$	Erro de Previsão de Geração Eólica [MW]		
					Probabilidade		
					5%	50%	95%
Texas/2010	9.000	um dia à frente	0,0117	0,1187	-1652	105	1862
Finlândia	102		-0,0155	0,0751	-14	-2	11
Espanha/2010	19.300		0,0162	0,0514	-1318	313	1945
Suécia/2011	2.899		-0,0052	0,0603	-303	-15	272
Dinamarca/2011	3.871		-0,0005	0,0534	-342	-2	338
Irlanda/2011	1.557		-0,0123	0,0827	-231	-19	193
Alemanha/2010	25.180		0,0092	0,045	-1632	232	2095
Espanha/2010	20.091	uma hora à frente	-0,0018	0,0133	-477	-36	404
Alemanha/2010	25.180		0,0004	0,0116	-470	10	491

Fonte: Adaptado de MILLIGAN et al. (2010).

A Figura 4 mostra o histograma e a distribuição gaussiana correspondente aos desvios de previsão um dia à frente do sistema da Alemanha 2010.

Figura 4 – Histograma dos desvios de previsão um dia à frente para o sistema alemão em 2010. Média, desvio padrão, assimetria e curtose iguais a  $\mu = 0.0092$ ,  $\sigma = 0.045$ ,  $\gamma = -0.2891$  e  $\kappa = 3.5896$  respectivamente.



Fonte : HODGE et al. (2012).

### 2.2.2 Métodos de Previsão de Erro Paramétricos

Os métodos de previsão de desvios paramétricos têm como característica principal o fato de explorarem distribuições de probabilidade já conhecidas, como as distribuições gaussiana, beta e gama. Para fins de contextualização, a

Tabela 3 mostra um panorama dos métodos paramétricos detalhados na literatura assim como os objetivos e a origem dos dados utilizados em cada estudo.

Tabela 3 – Métodos Paramétricos de Determinação da WPFE.

<b>Modelo de determinação da WPFE</b>	<b>Objetivo do Estudo</b>	<b>País Origem dos Dados</b>	<b>Referência</b>
Distribuição Gaussiana	Previsão da geração eólica para auxiliar geradores de energia eólica a propor ofertas de geração/preço ótimas no mercado no mercado <i>day-ahead</i> e <i>Intraday</i> .	Espanha Rede Elétrica da Espanha	GONZÁLEZ-APARICIO; ZUCKER (2015)
Distribuição Estável Lévy- $\alpha$	Impacto da integração da energia eólica na determinação dos requisitos de reserva de geração <sup>6</sup> .	Bélgica Operador do Sistema Elia	BRUNINX; DELARUE (2014)
Distribuição de Cauchy, Beta e Weibull	Análise de diferentes distribuições de desvios de geração em horizontes de tempo distintos.	Estados Unidos ERCOT	HODGE; MILLIGAN (2011)

<sup>6</sup> Controle de frequência, velocidade e tempo de rampa de subida/descida de geração.

Distribuição Gaussiana	Impacto da integração da energia eólica na determinação dos requisitos de reserva de geração levando em conta requisitos de transmissão e de rampa de geração.	Estados Unidos Cidade de Taylor Texas	SANG; ZHENG (2021)
Distribuição Gaussiana e Série de Taylor	Determinação da distribuição do erro de geração eólica usando série de Taylor a partir da previsão da velocidade do vento.	Alemanha Deutscher Wetterdienst	LANGE (2005)
Distribuição Gaussiana Truncada com média nula (TND)	Impacto da integração da energia eólica na regulação e determinação dos requisitos de reserva de geração.	Estados Unidos CAISO	MAKAROV et al. (2009)
Distribuição Beta	Dimensionamento de sistemas de estocagem de energia.	Não Informado	BLUDSZUWEIT; DOMINGUEZ-NAVARRO; LLOMBART (2008)

Fonte: do autor..

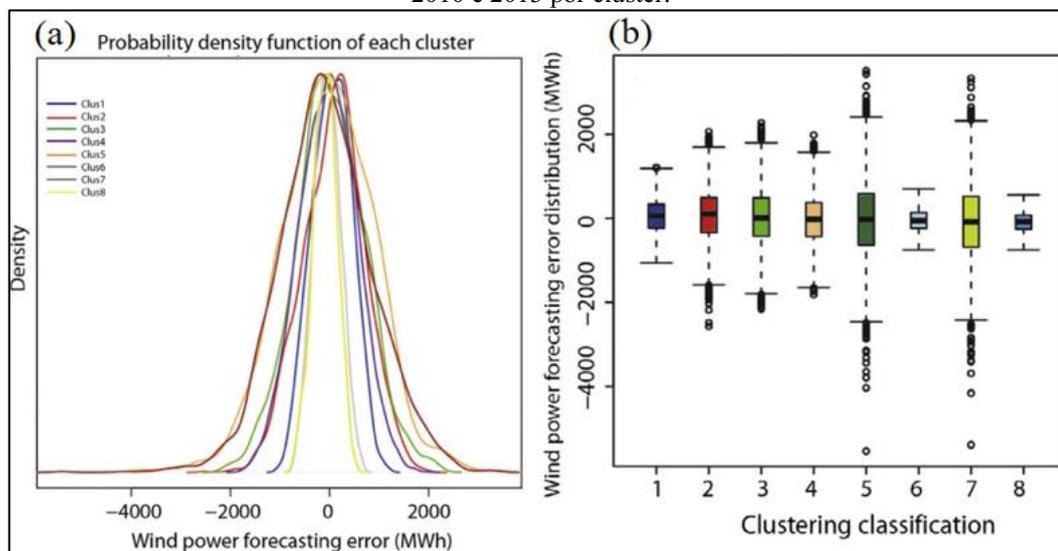
A simplicidade da implementação da distribuição normal para a obtenção de uma equação analítica dos desvios de previsão de geração eólica faz com que esta seja uma das distribuições mais exploradas na literatura. Alguns exemplos são os estudos (GONZÁLEZ-APARICIO; ZUCKER, 2015), (HODGE; MILLIGAN, 2011), (LANGE, 2005), (MAKAROV et al., 2009) e (SANG; ZHENG, 2021), os quais exploram a distribuição normal ou suas variantes para caracterizar o WPFE. Contudo, destaca-se que a distribuição normal pode ser uma aproximação grosseira dos desvios históricos de previsão, os quais usualmente podem apresentar níveis consideráveis de assimetria e curtose. Neste contexto, outras distribuições são analisadas na literatura tais como as distribuições Weibull, Beta, Cauchy e Lévy-  $\alpha$ .

A respeito do uso da distribuição normal para representar o WPFE, o estudo realizado em (GONZÁLEZ-APARICIO; ZUCKER, 2015) considera que os desvios variam de acordo com o período analisado e os principais *drivers* são o preço da energia (*Day-ahead* e *Intraday price*), o nível da carga, a velocidade típica do vento e a previsão de geração eólica de cada período. Além disso, este estudo utiliza um algoritmo de *clustering*<sup>7</sup> a fim de extrair estas informações e obter WPFE típicas. Ao final, oito distribuições normais são obtidas, cujas funções de distribuição de probabilidade e desvios de previsão são mostradas na Figura 5. Ressalta-se que cada WPFE é válida para níveis de geração eólica e para períodos de previsão distintos (horário,

<sup>7</sup> Denominado como *Expectations-Maximization method*.

diário, mensal).

Figura 5 – a) Função de densidade de Probabilidade, b) *Boxplots* dos desvios horário de geração eólica entre 2010 e 2013 por cluster.



Fonte: GONZÁLEZ-APARICIO; ZUCKER (2015).

Após a etapa de *clustering*, as distribuições de erro obtidas são empregadas em um modelo de previsão do preço *Intraday* baseado no método de Regressão Linear Múltipla<sup>8</sup>. Os resultados obtidos indicam que o modelo desenvolvido é adequado para a definição das estratégias de ofertas de energia das unidades geradoras eólicas no mercado *Intraday* visto que o erro de previsão de preço é em média 15%, um valor considerado aceitável neste estudo.

Similarmente à (GONZÁLEZ-APARICIO; ZUCKER, 2015), os artigos (LANGE, 2005), (MAKAROV et al., 2009) e (SANG; ZHENG, 2021) exploraram a distribuição gaussiana para capturar as características da distribuição dos desvios da geração eólica. Todavia, cada um destes artigos optou por uma metodologia distinta. Em (SANG; ZHENG, 2021), três métodos de previsão de geração eólica são utilizados (Gama, Rayleigh e Weibull). Para cada um destes métodos, uma distribuição gaussiana normalizada<sup>9</sup> é gerada a partir dos desvios obtidos em diversas simulações. A

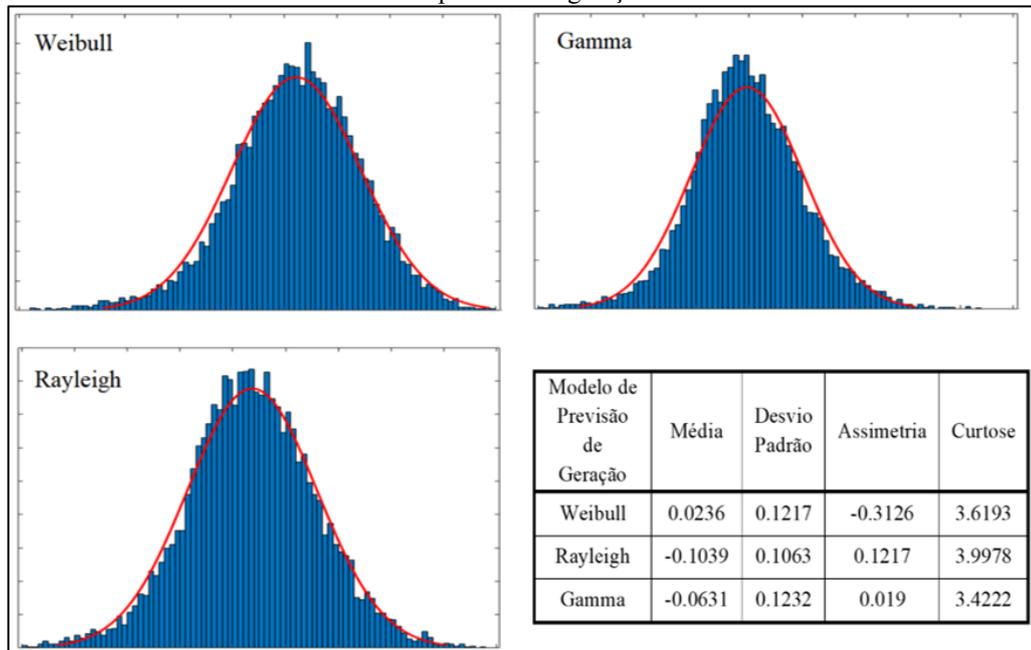
Figura 6 mostra que a distribuição gaussiana não é apropriada para capturar a forma da distribuição dos desvios visto que a curtose é maior que três, o que caracteriza uma distribuição

<sup>8</sup> *Multiple Linear Regression* – MLR

<sup>9</sup> Os erros são normalizados a partir da capacidade eólica instalada. Desta forma, os erros podem variar entre -1 e 1.

leptocúrtica ou com cauda pesada<sup>10</sup>.

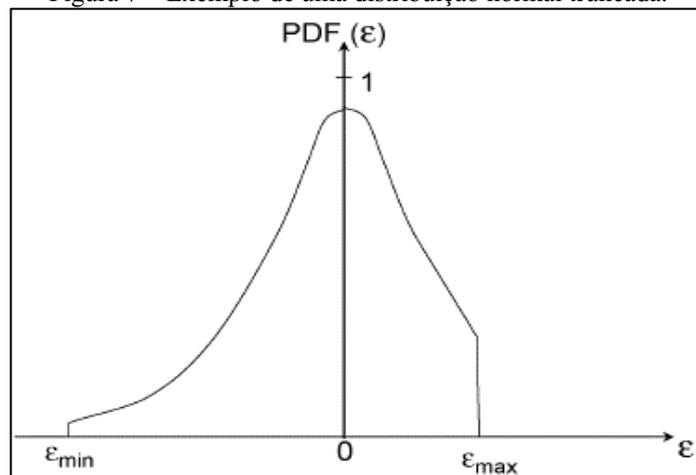
Figura 6 – Histogramas dos desvios normalizados e respectivos parâmetros das distribuições obtidas de cada modelo de previsão de geração eólica.



Adaptado de SANG; ZHENG (2021).

Ao invés de empregar uma distribuição normal com domínio infinito, o estudo mostrado por (MAKAROV et al., 2009) utiliza uma Distribuição Gaussiana Truncada (TND) com média nula como ilustrada na Figura 7. Neste trabalho, a distribuição normal foi limitada em  $\pm 3\sigma$ , cujos parâmetros são apresentados na Tabela 4. A distribuição TND é utilizada para gerar uma variável aleatória de erro que é empregada para prever a geração eólica uma hora à frente.

Figura 7 – Exemplo de uma distribuição normal truncada.



Fonte: MAKAROV et al. (2009)

<sup>10</sup> Distribuições com excesso positivo de curtoses (curtose > 3)

Tabela 4 – Erro de previsão de geração eólica uma hora à frente para o CAISO em 2006 expresso em fração da capacidade instalada por estação do ano.

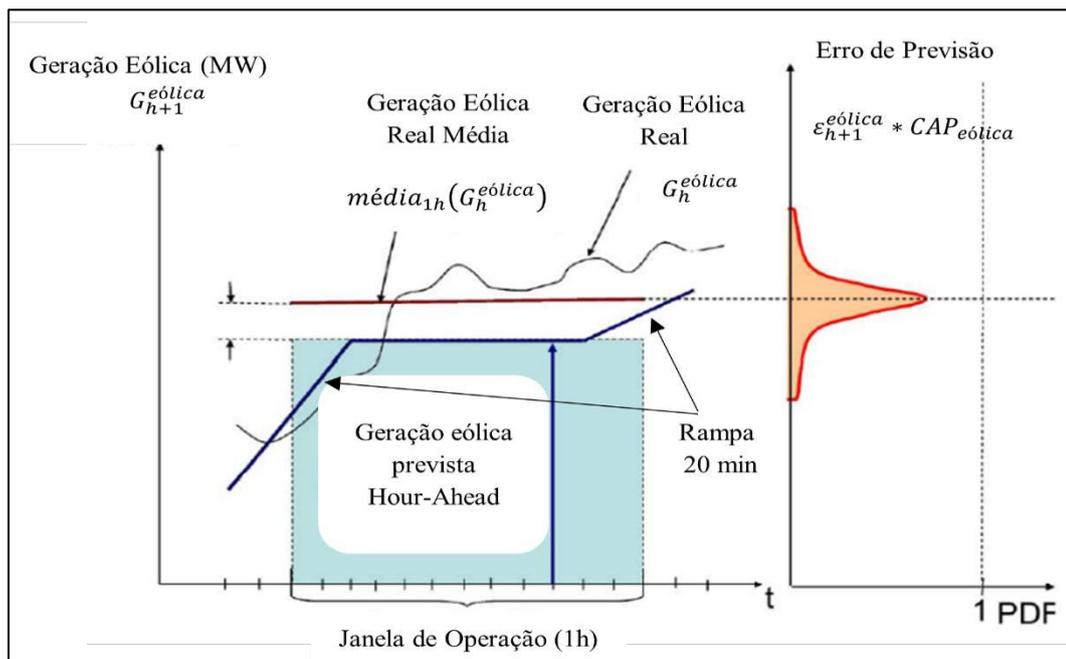
Estação	Inverno	Primavera	Verão	Outono
Média	0,00012	-0,0005	-0,00047	0,00058
Min	-0,35678	-0,43313	-0,32185	-0,31930
Max	0,30918	0,30840	0,30742	0,39663
Desvio Padrão	0,07233	0,08987	0,07955	0,07922
Autocorrelação	0,61062	0,70609	0,65185	0,59392

Fonte: MAKAROV et al. (2009).

A Figura 8 ilustra a metodologia aplicada em (MAKAROV et al., 2009) para prever a geração eólica uma hora à frente, a qual está descrita pela equação (1). Basicamente, a geração prevista para a hora  $h+1$  ( $G_{h+1}^{eólica}$ ) é obtida através da diferença entre a média da geração na hora  $h$  ( $G_h^{eólica}$ ) e o erro de geração. A primeira é obtida a partir do método de persistência (a geração real não é fornecida). A segunda parcela é obtida a partir da multiplicação entre o valor  $\varepsilon_{h+1}^{eólica}$  obtido da distribuição normal truncada e a capacidade eólica instalada ( $CAP_{eólica}$ ). Vale notar que uma rampa com duração de 20 minutos é adicionada no começo e no fim de cada hora operativa para manter a continuidade nas previsões.

$$G_{h+1}^{eólica} = Rampa_{20} \{ média_{1h}(G_h^{eólica}) - \varepsilon_{h+1}^{eólica} * CAP_{eólica} \} \quad (1)$$

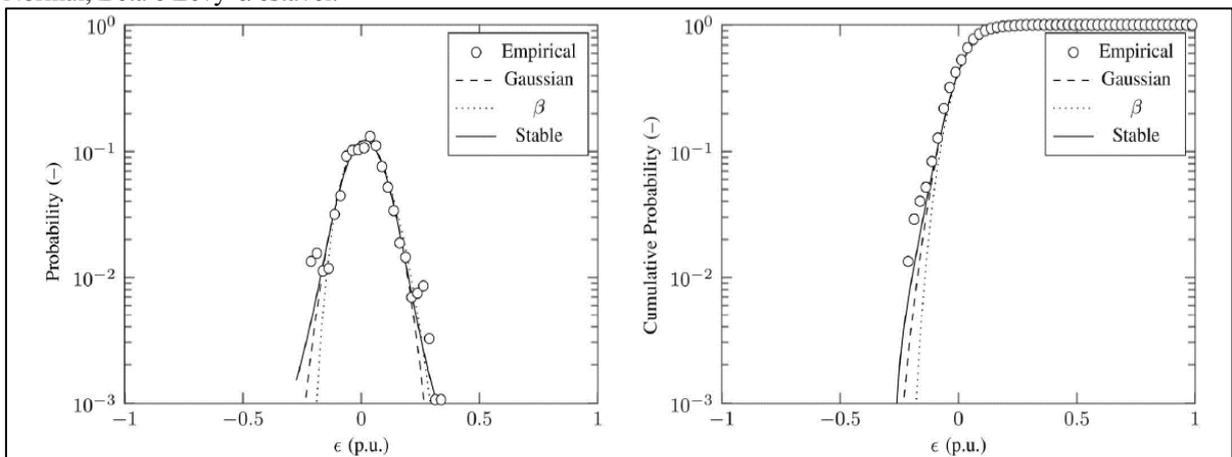
Figura 8 – Densidade de probabilidade do erro e simulação da geração eólica.



Adaptado de (MAKAROV et al., 2009).

Com o objetivo de capturar a característica de cauda pesada da distribuição do WPFE, diversos estudos analisaram diferentes distribuições probabilísticas. O trabalho desenvolvido por (BRUNINX; DELARUE, 2014) compara as distribuições de probabilidade Lévy- $\alpha$  estável, normal e beta para estimar a WPFE tendo como ponto de referência os dados observados. A partir da Figura 9, percebe-se que a distribuição Lévy- $\alpha$  estável se mostrou a mais apropriada para capturar a característica de cauda-pesada dos desvios.

Figura 9 – Comparação dos histogramas e função de densidade acumulada da WPFE estimada pela distribuição Normal, Beta e Lévy- $\alpha$  estável.



Adaptado de (BRUNINX; DELARUE, 2014).

Ressalta-se que o artigo (BRUNINX; DELARUE, 2014) classifica a distribuição do erro em cinco grupos para capturar a influência do nível de geração eólica na distribuição do erro. Assim, uma distribuição Lévy- $\alpha$  foi obtida para cada grupo. A Tabela 5 mostra os valores máximos e mínimos de cada um dos quatro momentos estatísticos da distribuição dos desvios normalizados.

Tabela 5 – Medidas normalizadas dos momentos estatísticos do erro.

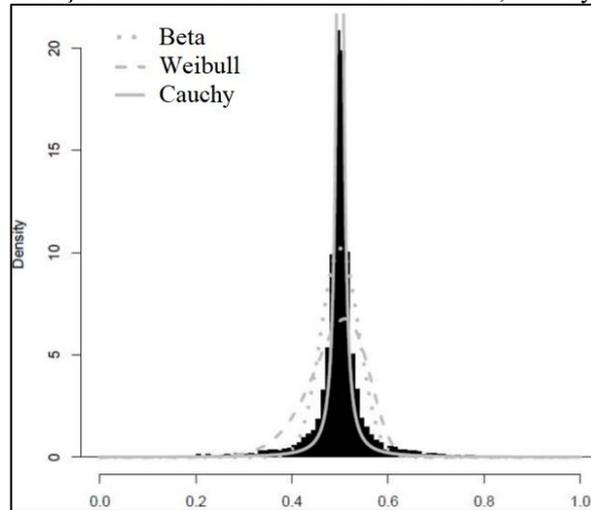
<b>Momento</b>	<b>Média</b>	<b>Desvio Padrão</b>	<b>Assimetria</b>	<b>Curtose</b>
Mínimo	-0,0388	0,0194	-3,5648	2,6257
Máximo	0,0448	0,1193	0,9068	16,5707
Média	0,0034	0,0833	-	-

Fonte: (BRUNINX; DELARUE, 2014).

Outro exemplo de uso de distribuições diferentes da gaussiana é o estudo descrito pelo artigo (HODGE; MILLIGAN, 2011) o qual compara três distribuições de probabilidade a fim de verificar qual captura melhor as características dos desvios de previsão da geração eólica. A principal conclusão retirada é que a distribuição de Cauchy possui um desempenho melhor do

que as distribuições Weibull e Beta, visto que a primeira é capaz de capturar de forma mais eficaz a característica de cauda pesada da distribuição empírica dos desvios. A Figura 10 mostra a distribuição empírica dos desvios, assim como as distribuições Beta, Weibull e Cauchy parametrizadas.

Figura 10 – Distribuição normalizada dos desvios observados, Cauchy, Weibull e Beta.



Fonte: (HODGE; MILLIGAN, 2011)

### 2.2.3 Métodos de Previsão de Desvios Não Paramétricos

Os métodos de previsão de erro não paramétricos possuem como característica principal o fato de explorarem distribuições de probabilidade empíricas ao invés de distribuições já conhecidas. Isso se deve ao fato de que, ao utilizar uma abordagem paramétrica, certas informações probabilísticas associadas às incertezas dos modelos de previsão são perdidas assim como a interdependência da estrutura dos desvios previstos (PINSON et al., 2009). Usualmente, os métodos não paramétricos buscam determinar a densidade de probabilidade do erro a partir da densidade empírica. Para fins de contextualização, a Tabela 6 mostra um panorama dos métodos não paramétricos detalhados na literatura assim como os objetivos e a origem dos dados utilizados em cada estudo.

Tabela 6 – Métodos Não-Paramétricos de Determinação da WPFE.

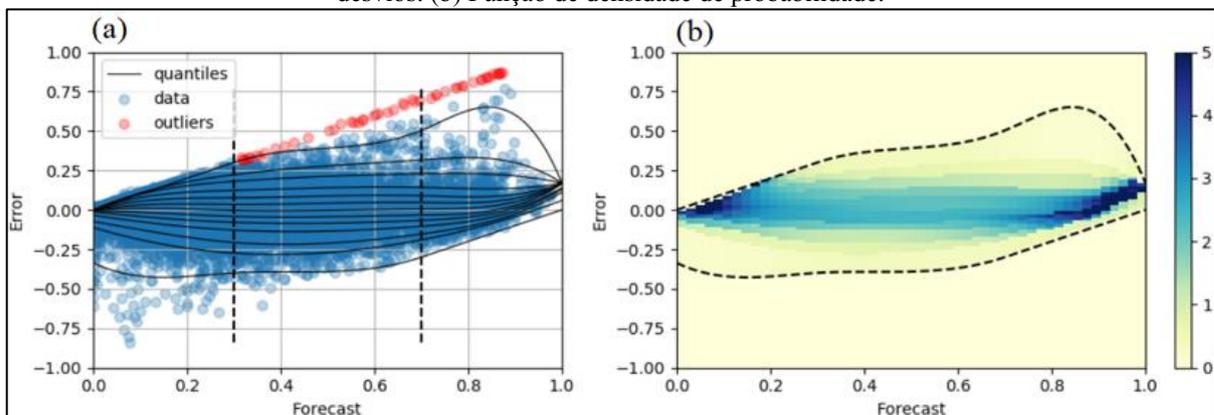
<b>Modelo de determinação da WPFE</b>	<b>Objetivo do Estudo</b>	<b>País/Origem dos Dados</b>	<b>Referência</b>
Regressão Quantílica e Distribuição Uniforme e Gaussiana	Dimensionamento dinâmico de estocagem de energia para compensação de desbalanços de energia de fazendas eólicas.	Dinamarca	PINSON et al. (2009)

Regressão Quantílica e Distribuição Uniforme e Gaussiana	Geração de cenários de geração eólica para decisões de despacho.	Austrália AEMO	NYCANDER; SÖDER (2021)
Quantis com intervalo de 5% a 95%	Impacto das incertezas da carga e geração eólica na determinação das reservas de balanço de energia no mercado <i>Intraday</i> e <i>day-ahead</i> .	Portugal Rede Elétrica Nacional	MATOS; BESSA (2011)
Estimativa de Densidade de Kernel	Determinação de intervalos de previsão de geração eólica com nível de confiança aceitáveis.	China Fazenda Eólica Específica	YANG; YANG (2016)
Função Cópula ( <i>Copula Theory</i> ) Cópula Gaussiana	Determinação de distribuições de desvios condicionais de erro de múltiplas fazendas eólicas.	Estados Unidos National Renewable Laboratory	ZHANG et al. (2014)

Fonte: do autor.

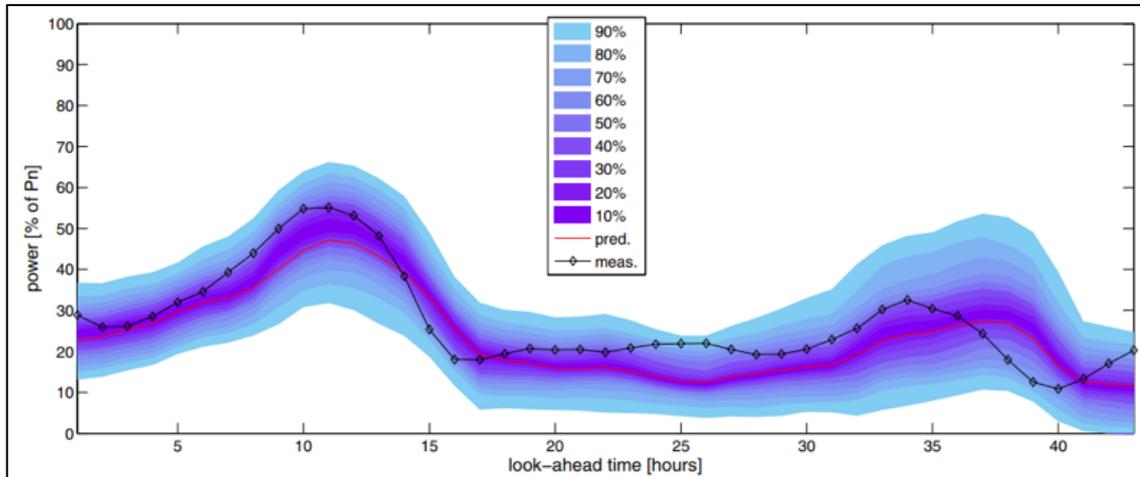
Dentre os artigos retratados neste capítulo, o uso de quantis para a representação do erro de geração eólica foi o método mais empregado. Os artigos (PINSON et al., 2009) e (NYCANDER; SÖDER, 2021) exploraram uma metodologia semelhante. Em ambos os artigos, o método de Regressão Quantílica foi utilizado para aproximar a distribuição empírica dos desvios (Figura 11-a) e construir uma função de densidade de probabilidade (Figura 11-b). Esta caracterização dos desvios de previsão é utilizada para a construção dos cenários de geração apresentados na Figura 12 e Figura 13.

Figura 11 – Distribuição dos desvios de previsão normalizados. (a) Curvas quantílicas sobre a distribuição dos desvios. (b) Função de densidade de probabilidade.



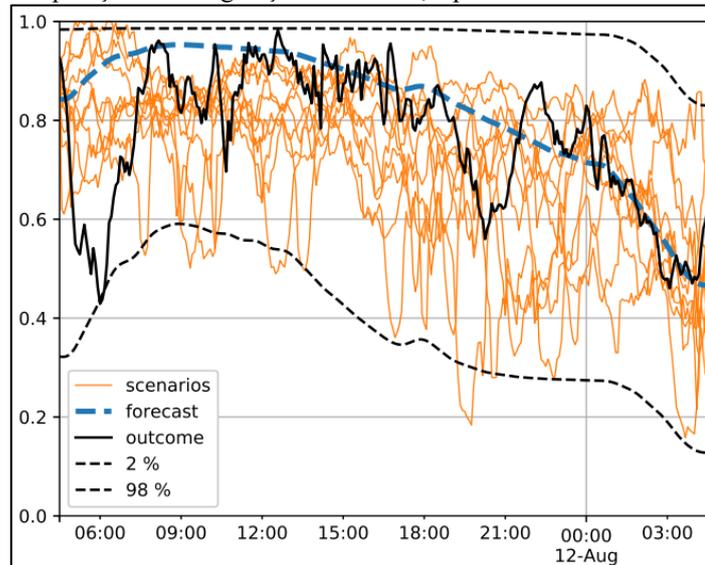
Fonte (NYCANDER; SÖDER, 2021).

Figura 12 – Conjunto de 25 Cenários de Previsão de Geração Eólica para múltiplas fazendas eólicas na Dinamarca para um horizonte de tempo de 43h.



Fonte: (PINSON et al., 2009)

Figura 13 – Comparação entre a geração eólica real, a prevista e os 10 cenários de previsão.



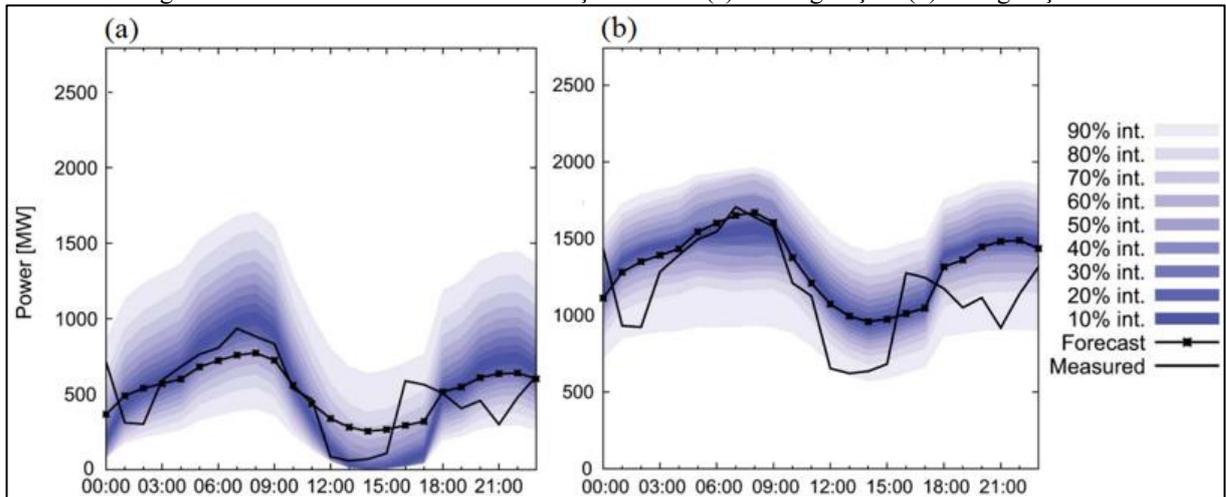
Fonte (NYCANDER; SÖDER, 2021).

A metodologia empregada em (MATOS; BESSA, 2011) é um pouco diferente daquela dos artigos (PINSON et al., 2009) e (NYCANDER; SÖDER, 2021). Isso se deve pelo fato de que as incertezas da velocidade dos ventos e das interrupções não planejadas das turbinas são integradas no modelo. A modelagem da primeira fonte de incerteza utiliza quantis com intervalo entre 5% à 95% e funções exponenciais para modelar os eventos extremos que se encontram na cauda das distribuições. Já para a segunda fonte de incerteza, o valor médio da COPT<sup>11</sup> é utilizado. Levando em conta ambas as fontes de incerteza, distribuições de previsão em base horária são geradas. A Figura 14 mostra um exemplo das previsões para cenários de baixa e

<sup>11</sup> Capacity Outage Probability Table – Tabela de Probabilidade de Interrupção da Capacidade de Geração

alta geração eólica contendo as duas fontes de incerteza.

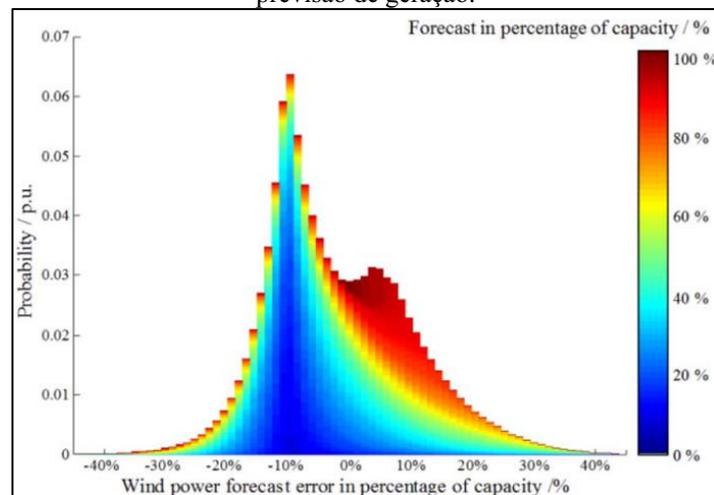
Figura 14– Cenários de Previsão de Geração Eólica. (a) Baixa geração. (b) Alta geração.



Adaptado de (MATOS; BESSA, 2011).

Destaca-se que o trabalho de (ZHANG et al., 2014) utiliza a cópula gaussiana para modelar a interdependência entre geradores e as variáveis de previsão e de geração observada. No final, uma função de distribuição acumulada do erro é construída como o intuito de gerar desvios de acordo a geração prevista. A Figura 15 apresenta a distribuição condicional acumulada do erro de previsão para diferentes níveis de previsão. Note que o contorno da distribuição forma a distribuição marginal do erro de previsão. O estudo concluiu que quando o valor previsto de geração é baixo, o erro de previsão possui uma média alta e variância baixa. Caso contrário, se o valor previsto é alto, a média é relativamente baixa e a variância da distribuição dos erros é alta.

Figura 15 – Gráfico acumulado da distribuição condicional dos desvios de previsão para diferentes níveis de previsão de geração.



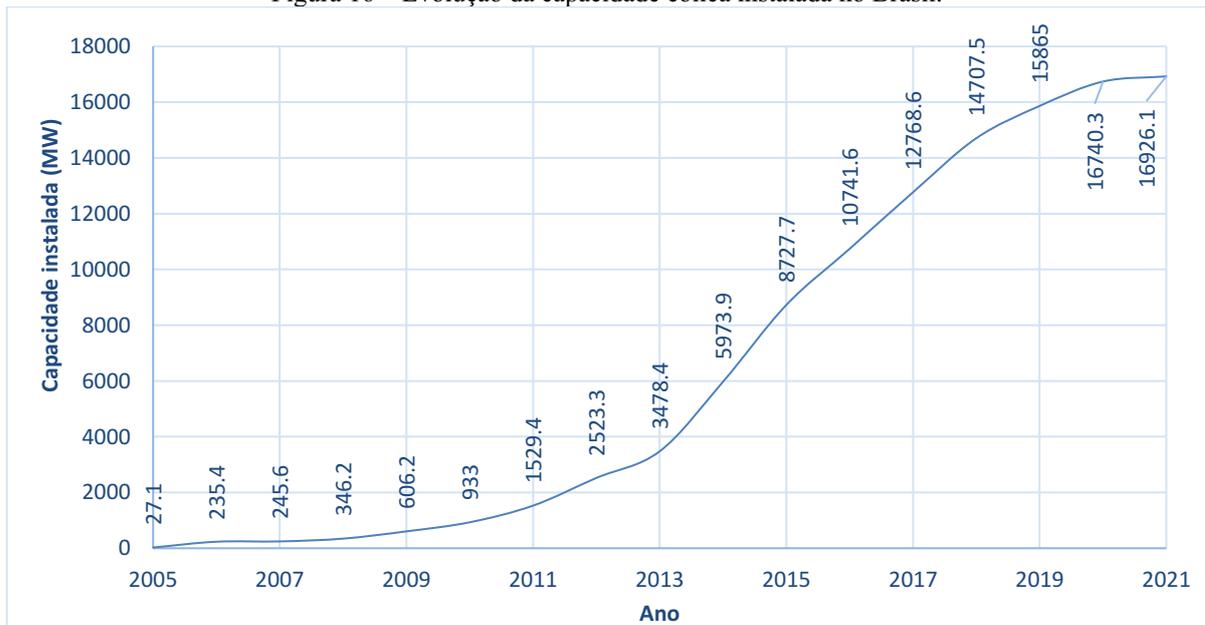
Fonte: (ZHANG et al., 2014).

### 2.3 CASO BRASILEIRO: MÉTODOS DE PREVISÃO DA GERAÇÃO EÓLICA PARA O SIN

Atualmente, a matriz elétrica brasileira apresenta uma tendência de maior participação das fontes de energia renováveis como a eólica, solar e biomassa. Conseqüentemente, o despacho ótimo do sistema torna-se mais complexo devido ao aumento das incertezas associadas às fontes de geração do sistema. Desta forma, torna-se necessário gerenciar tais incertezas em conjunto com as incertezas da carga de forma a manter os níveis de segurança elétrica e energética requeridos pelo sistema. Este capítulo tem como objetivo descrever como o ONS, entidade que realiza o despacho das unidades geradoras no Brasil, realiza as previsões de geração eólica com o intuito de garantir a segurança e a qualidade do suprimento elétrico do Sistema Elétrico Nacional (SIN).

Segundo a Abeeólica, a série de aumentos consistentes da capacidade instalada da fonte eólica (Figura 16) é devido à qualidade dos ventos no Brasil. Assim, com o aumento da penetração da fonte eólica no sistema, o ONS tem a necessidade de aprimorar os modelos de previsão de geração eólica para gerir as reservas de potência operativa e possibilitar o melhor uso dos recursos energéticos (ONS - SINTEGRE, 2018). A estratégia para prever a geração eólica no Brasil é baseada em quatro modelos de previsão de ventos. O objetivo de empregar diversos modelos é diminuir o efeito dos vieses na previsão de geração eólica final.

Figura 16 – Evolução da capacidade eólica instalada no Brasil.



Adaptado de Abeeólica ([s.d.]).

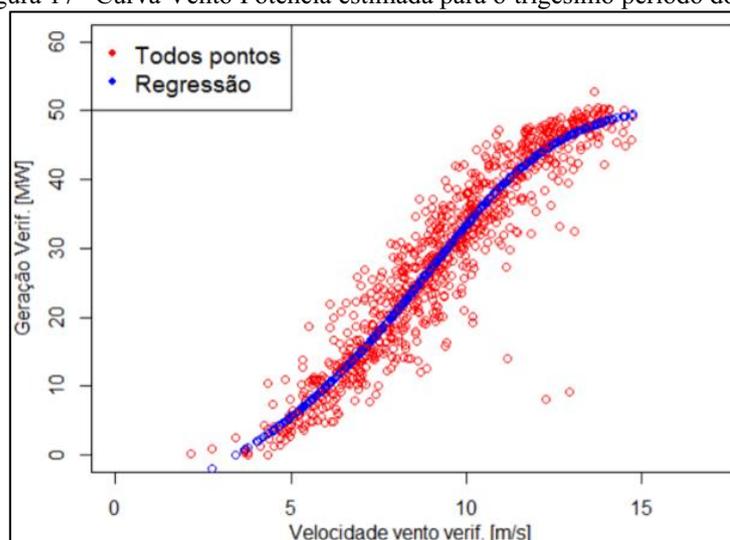
A metodologia de previsão de geração eólica é composta por três módulos (ONS - SINTEGRE, 2018): preparação dos dados, previsão e processos internos. A seguir são discutidos os principais conceitos das duas primeiras etapas.

### **2.3.1 Preparação dos Dados**

Este primeiro módulo tem como objetivo realizar um tratamento das informações coletadas e previstas que serão, futuramente, empregadas no módulo de previsão da geração eólica. De modo geral, os dados de entrada podem ser classificados em três tipos: Geração Verificada, Vento Verificado e Vento Previsto. Os dados referentes à geração verificada são oriundos dos dados coletados pelo CEPEL (Centro de Pesquisas de Energia Elétrica) e pelo ONS para cada complexo eólico. O tratamento destes dados consiste em efetuar uma filtragem inicial visto que os dados podem ter uma qualidade comprometida. Para isso, os valores espúrios são excluídos e as lacunas são preenchidas com aproximações. No final, uma série de geração verificada média composta de 48 intervalos de meia hora é obtida agrupando diversas usinas eólicas para simplificar o modelo. Ressalta-se que a mesma metodologia é empregada para os dados correspondentes ao vento verificado. Todavia, estes últimos são oriundos de medições feitas pela EPE (Empresa de Pesquisa Energética) e pelo ONS para cada usina eólica.

A partir dos dados de geração e vento verificado médio por grupo de usinas eólicas, curvas de vento por potência são ajustadas para melhorar a representação das séries temporais obtidas. Para isso, regressões logísticas são utilizadas para gerar curvas de vento e potência. Vale notar que uma curva de vento e potência é ajustada para cada intervalo de 30 minutos a fim de capturar as características particulares de cada período do dia típico. A Figura 17 mostra um exemplo da curva obtida pela regressão para o período 30, correspondente a um intervalo de tempo de 24 horas discretizado com um passo de meia hora.

Figura 17– Curva Vento Potência estimada para o trigésimo período do dia.



Fonte: (ONS - SINTEGRE, 2018).

Em relação ao vento previsto, ressalta-se que ele é o resultado da junção de quatro modelos:

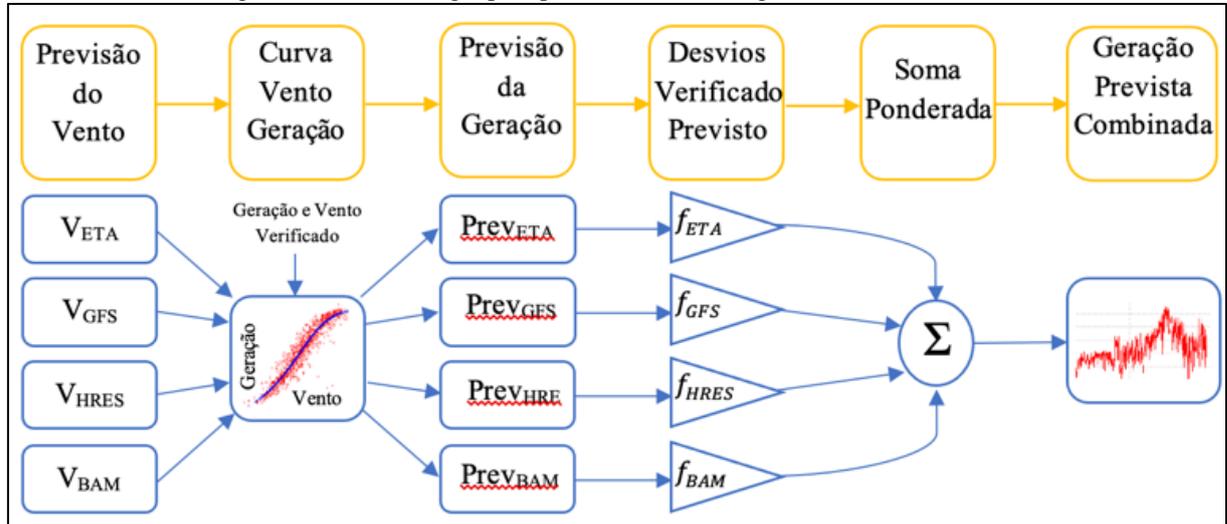
- ETA: desenvolvido pelo Centro de Previsão de Tempo e Estudos Climáticos (CPTEC); possui discretização temporal de uma hora e espacial de 5 km para um horizonte de tempo do dia corrente e dois dias à frente.
- GFS: desenvolvido pela Administração Oceânica e Atmosférica Nacional (NOAA); possui discretização temporal de uma hora e espacial de 25 km para um horizonte de tempo do dia corrente e quatro dias à frente.
- HRES: desenvolvido pelo Centro Europeu de Previsão de Tempo para médio Prazo (ECMWF); possui discretização temporal de três horas e espacial de 20 km para um horizonte de tempo do dia corrente e cinco dias à frente.
- BAM: desenvolvido pelo CPTEC e possui discretização temporal de 1 hora e espacial de 20 km para um horizonte de tempo do dia corrente e mais 10 dias à frente.

### 2.3.2 Previsão da Geração Eólica

Com a obtenção das curvas de vento e geração verificadas para cada intervalo de tempo, é possível prever a geração eólica a partir do vento previsto. De forma a se obter uma única previsão de geração, o ONS faz uma média ponderada cujos fatores de ponderação ( $f$ ) são inversamente proporcionais aos desvios ( $\epsilon$ ) associados a cada modelo de previsão baseado nos últimos três dias que antecedem o dia em que a geração é prevista. Desta forma, os modelos com menores desvios possuem uma maior participação na previsão final. Ressalta-se que esta ponderação é feita para cada intervalo de meia hora, ou seja, a cada meia hora, uma curva de

geração prevista é obtida. O princípio geral da metodologia é descrito pelo diagrama da Figura 18.

Figura 18 – Metodologia para previsão do vento e geração eólica do ONS.



Fonte: do autor..

O cálculo do erro  $\varepsilon$  associado a cada modelo e para cada meia hora é o resultado de duas ponderações, como mostrado pela Equação (2).

$$\varepsilon_{m,t}^{ponderado} = \frac{\sum_{t=-12}^{12} POND^{sinal}(\varepsilon_m^t) * POND^{linear}(t) * |\varepsilon_m^t|}{\sum_{t=-12}^{12} POND^{sinal}(\varepsilon_m^t) * POND^{linear}(t)} \quad (2)$$

$$POND^{sinal}(\varepsilon_m^t) = \begin{cases} 0.3 & \text{se } \varepsilon_m^t > 0 \\ 0.7 & \text{se } \varepsilon_m^t < 0 \end{cases}$$

onde:

$\varepsilon_{m,t}^{ponderado}$ : Erro ponderado para o modelo de previsão  $m$  e meia hora  $t$ .

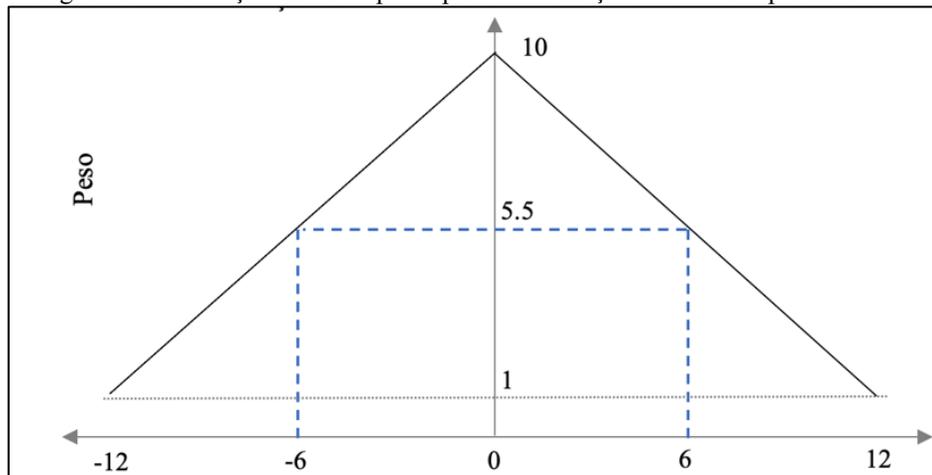
$\varepsilon_m^t$ : Erro do modelo de previsão  $m$  e meia hora  $t$ .

$POND^{sinal}$ : Ponderação por sinal.

$POND^{linear}$ : Ponderação linear.

A primeira é uma ponderação linear dos desvios das 12 meias horas à frente e atrás do instante em que se deseja efetuar a previsão. Para cada erro, um peso é atribuído de acordo com a distância temporal da hora central, como mostrado pela Figura 19.

Figura 19 – Gradação linear de pesos para a suavização dos desvios por intervalo.



Fonte: Adaptado de (ONS - SINTEGRE, 2018).

A segunda média ponderada é associada ao sinal do erro em cada intervalo de tempo: caso o erro for positivo, um peso igual à 0.3 é atribuído; caso o erro for negativo, o peso atribuído é igual à 0.7. A ponderação linear foi introduzida com o intuito de obter fatores de combinação cujos valores não variam abruptamente a cada meia hora. Já a ponderação por sinal é empregada com o objetivo de privilegiar os modelos que subestimem a geração visto que, caso a geração se verifique muito inferior à previsão, pode ser necessário o despacho de fontes de geração não programadas previamente, elevando os custos da operação.

A partir dos desvios ponderados de cada modelo, é possível calcular o fator de participação de cada modelo na previsão combinada. O fator de combinação e o cálculo da previsão combinada para cada intervalo de meia hora são mostrados pela Equação (3).

$$f_m^t = \frac{\frac{1}{\varepsilon_{m,t}^{\text{ponderado}}}}{\sum_{m'=1}^4 \left( \frac{1}{\varepsilon_{m',t}^{\text{ponderado}}} \right)} \quad (3)$$

$$Prev_{comb}^t = \sum_{m=1}^4 Prev_m^t \times f_m^t$$

sendo:

$\varepsilon_{m,t}^{\text{ponderado}}$ : Erro ponderado para o modelo de previsão  $m$  e meia hora  $t$ .

$f_m^t$ : Fator de participação do modelo de previsão  $m$  e meia hora  $t$ .

$Prev_{comb}^t$ : Previsão combinada para a meia hora  $t$ .

$Prev_m^t$ : Previsão do modelo  $m$  para a meia hora  $t$ .

Por fim, uma medida da qualidade da previsão combinada da geração é feita a partir de diversas métricas: erro médio (ME), erro médio absoluto (MAE), erro percentual absoluto médio (MAPE), erro médio quadrático médio (RMSE) e erro percentual absoluto médio normalizado (NMAPE). O cálculo dos desvios para a métrica é feito por meio das respectivas equações:

$$ME = \frac{1}{N} \sum_{i=1}^N (G_{verif}^i - G_{prev}^t) \quad (4)$$

$$MAE = \frac{1}{N} \sum_{i=1}^N |G_{verif}^i - G_{prev}^t| \quad (5)$$

$$MAPE = \frac{100}{N} \sum_{i=1}^N |G_{verif}^i - G_{prev}^t| / G_{verif}^i \quad (6)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (G_{verif}^i - G_{prev}^t)^2} \quad (7)$$

$$NMAPE = \frac{100}{N^{pnorm}} \sum_{i=1}^N |G_{verif}^i - G_{prev}^t| \quad (8)$$

sendo:

$N$ : Número de previsões no horizonte de tempo associado.

$G_{verif}^i$ : Geração verificada para o instante  $i$ .

$G_{prev}^i$ : Geração prevista para o instante  $i$ .

$P^{norm}$ : Fator de normalização.

Note que  $P^{norm}$  pode tomar diversas formas de acordo com a normalização desejada. A normalização pode ser feita pela potência instalada final do período de estudo (pf), potência instalada evolutiva ao longo do período de estudo (pe), pelo produto entre o fator de capacidade diário e a potência instalada evolutiva (fcd), pelo produto entre o fator de capacidade mensal e a potência instalada evolutiva (fcm) ou pelo produto entre o fator de capacidade anual e a potência instalada evolutiva (fca). Os desvios são calculados para cada modelo de previsão,

assim como para a previsão combinada. Os desvios de geração para as regiões Nordeste e Sul para o período entre outubro de 2017 e setembro de 2018 são mostrados, respectivamente, pela Tabela 7 e Tabela 8. Salienta-se que as métricas foram calculadas para as previsões efetuadas para um dia à frente com uma discretização horária. Ressalta-se igualmente que a combinação de diversos modelos de previsão melhora a performance da previsão final. Este efeito pode ser observado levando em consideração que as métricas do modelo combinado são inferiores aos valores das métricas dos modelos separados.

Tabela 7 – Desvios de previsão de fonte eólica no Nordeste entre outubro 2017 a setembro 2018.

<b>Métrica</b>	<b>Combinado</b>	<b>CPTEC</b>	<b>GFS</b>	<b>ECMWF1</b>	<b>ECMWF2</b>	<b>BAM</b>
ME [MW]	131,24	38,68	168,77	136,85	155,87	446,64
MAE [MW]	474,71	612,77	522,04	500,98	488,73	1338,89
MAPE [%]	13,73	18,20	15,26	14,34	13,81	46,36
RMSE [MW]	474,71	612,77	522,04	500,98	488,73	1338,89
NMAPE <sub>pf</sub> [%]	4,55	5,88	5,01	4,80	4,69	12,84
NMAPE <sub>pe</sub> [%]	4,92	6,35	5,42	5,20	5,08	13,58
NMAPE <sub>fcd</sub> [%]	13,07	17,04	14,58	13,69	13,32	42,46
NMAPE <sub>fcm</sub> [%]	12,35	15,86	13,77	13,01	12,73	37,86
NMAPE <sub>fca</sub> [%]	10,76	13,87	11,84	11,36	11,09	29,68

Fonte: (ONS - SINTEGRE, 2018).

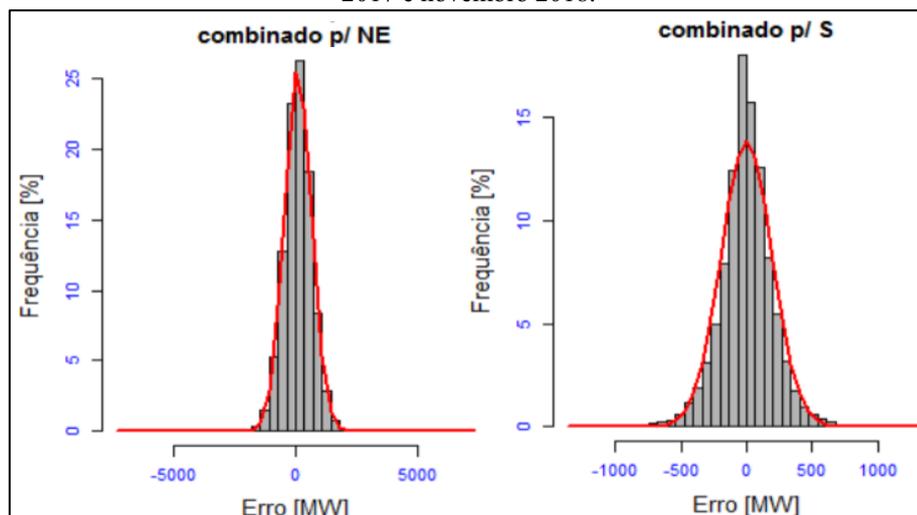
Tabela 8 – Desvios de previsão de fonte eólica no Sul entre outubro 2017 a setembro 2018.

<b>Métrica</b>	<b>Combinado</b>	<b>CPTEC</b>	<b>GFS</b>	<b>ECMWF1</b>	<b>ECMWF2</b>	<b>BAM</b>
ME [MW]	6,21	-19,70	30,19	1,57	9,88	42,11
MAE [MW]	144,30	181,11	159,38	163,08	165,93	346,93
MAPE [%]	39,06	53,55	39,05	43,57	41,29	189,28
RMSE [MW]	144,30	181,11	159,38	163,08	165,93	346,93
NMAPE <sub>pf</sub> [%]	7,69	9,65	8,49	8,69	8,84	18,49
NMAPE <sub>pe</sub> [%]	7,69	9,65	8,49	8,69	8,84	18,49
NMAPE <sub>fcd</sub> [%]	27,39	36,04	30,22	30,77	30,71	92,97
NMAPE <sub>fcm</sub> [%]	23,01	28,85	25,49	25,98	26,36	63,76
NMAPE <sub>fca</sub> [%]	21,94	27,54	24,23	24,80	25,23	52,75

Fonte: (ONS - SINTEGRE, 2018).

Como pode ser observado na Figura 20, tanto para o subsistema Nordeste quanto para o Sul, a maior frequência de desvios está centralizada em torno de zero. O pequeno viés que pode ser observado corresponde a uma pequena assimetria da distribuição dos desvios.

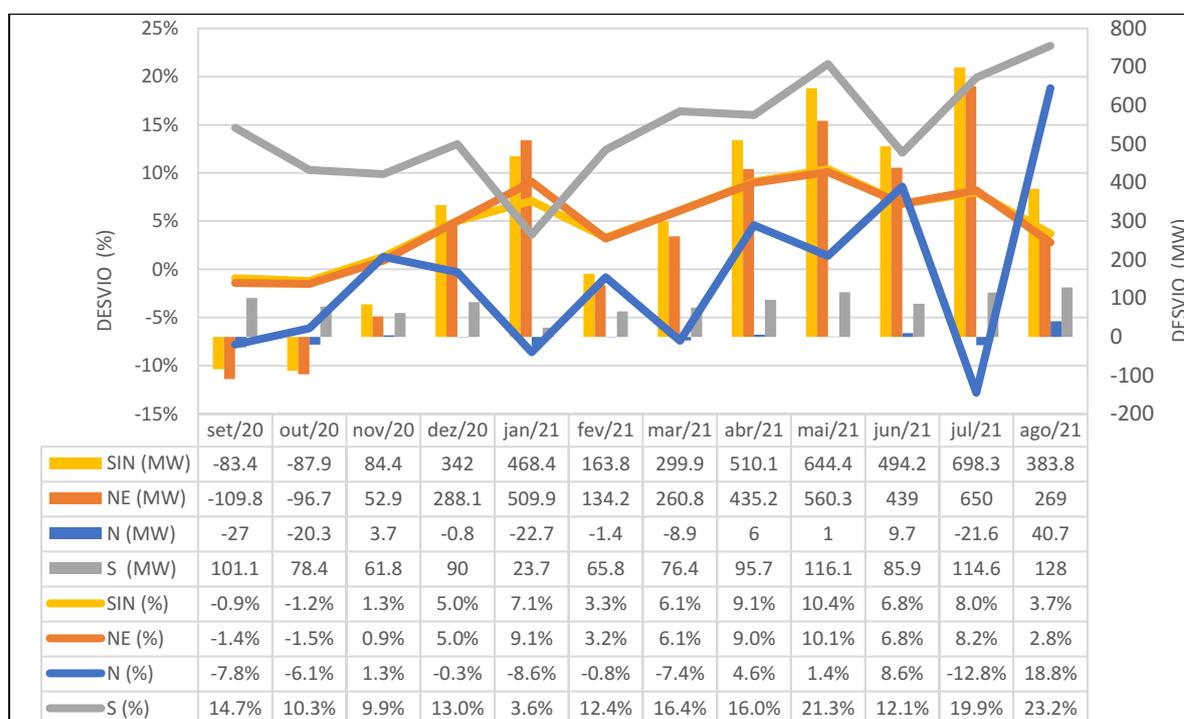
Figura 20 – Exemplo da distribuição dos desvios brutos do subsistema Nordeste (NE) e Sul (S) entre janeiro 2017 e novembro 2018.



Fonte: (ONS - SINTEGRE, 2018).

A Figura 21 mostra os desvios em MW (barras) e percentuais (linhas) entre a geração eólica verificada e a prevista para um dia à frente para o período entre setembro de 2020 e agosto de 2021. Os dados são oriundos da plataforma SINtegre-ONS. Nota-se que os subsistemas Sul e Norte possuem um erro relativo maior comparado às previsões de geração eólica no subsistema Nordeste.

Figura 21 – Desvio da Geração Eólica entre setembro de 2020 e agosto de 2021 para o SIN e subsistemas Nordeste, Norte e Sul.



Fonte: do autor..



### 3 METODOLOGIA PARA A PARAMETRIZAÇÃO DO MODELO DE GERAÇÃO DE CENÁRIOS

Conforme indicado no capítulo 1.2, este trabalho tem o objetivo de desenvolver um modelo capaz de integrar as incertezas nas previsões de geração eólica empregadas pelo ONS. Este modelo será capaz de gerar diversos cenários de desvios de geração para uma dada previsão. Ressalta-se que este modelo não possui o objetivo de fornecer um cenário de geração prevista corrigido, mas diversos cenários de desvios associados com suas respectivas probabilidades de ocorrência. Para atingir este objetivo, o modelo será baseado nos históricos de geração observada e de geração prevista do SIN para o horizonte de curto prazo. Este capítulo busca detalhar o processo metodológico empregado no desenvolvimento do modelo de geração de cenários para a fonte eólica. Primeiramente, são abordados os tópicos relativos à descrição e ao processamento dos dados geração eólica observados. Em seguida, são tratadas as questões relativas ao cálculo dos desvios de geração eólica. Por fim, o processo de parametrização e de validação dos modelos de geração de cenários é descrito.

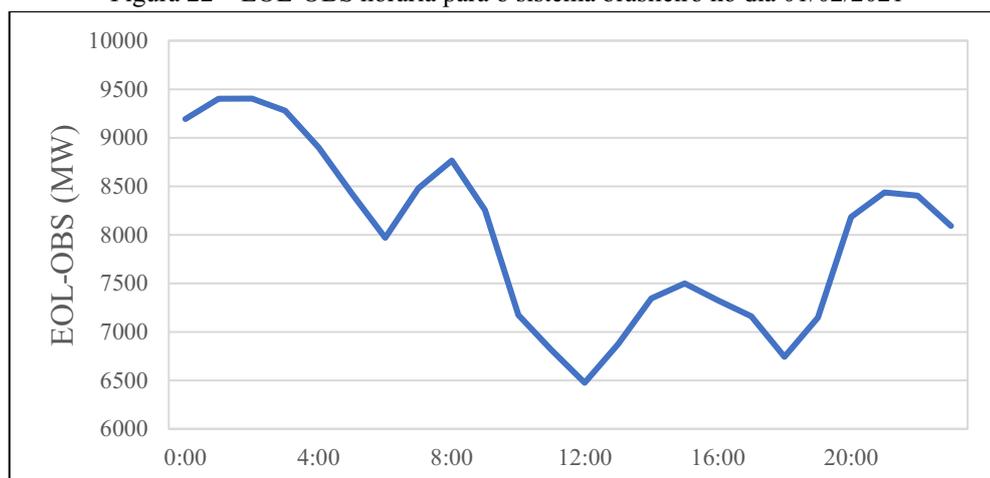
#### 3.1 DESCRIÇÃO E PROCESSAMENTO DOS DADOS

Neste capítulo são indicadas as principais características dos dados de geração eólica (observada e prevista) para um horizonte de curtíssimo prazo do SIN. Este capítulo apresenta, igualmente, o processamento necessário com o intuito de obter uma base dados consistente com as necessidades do modelo de geração de cenários. De forma geral, a base de dados pode ser classificada nas seguintes categorias:

- **Geração eólica observada (EOL-OBS):** geração eólica medida no sistema durante um determinado período de operação;
- **Geração eólica prevista (EOL-PREV):** geração eólica prevista para um dia à frente considerada no modelo da PDO.

A base de dados consolidada da EOL-OBS pelo ONS é de caráter público e está disponível em (ONS, 2022a). Esta informação pode ser filtrada para diferentes categorias: escala de tempo, subsistema do SIN, unidade federativa, período, entre outros. Salienta-se que a base de dados que compõe a EOL-OBS possui uma discretização horária. De forma ilustrativa, a Figura 22 apresenta a EOL-OBS do SIN para cada hora do dia 01/02/2021. Note que para um mesmo dia pode existir uma diferença de geração hora a hora consecutiva de  $\pm 1.8$  GW.

Figura 22 – EOL-OBS horária para o sistema brasileiro no dia 01/02/2021



Fonte: Adaptado de (ONS, 2022a).

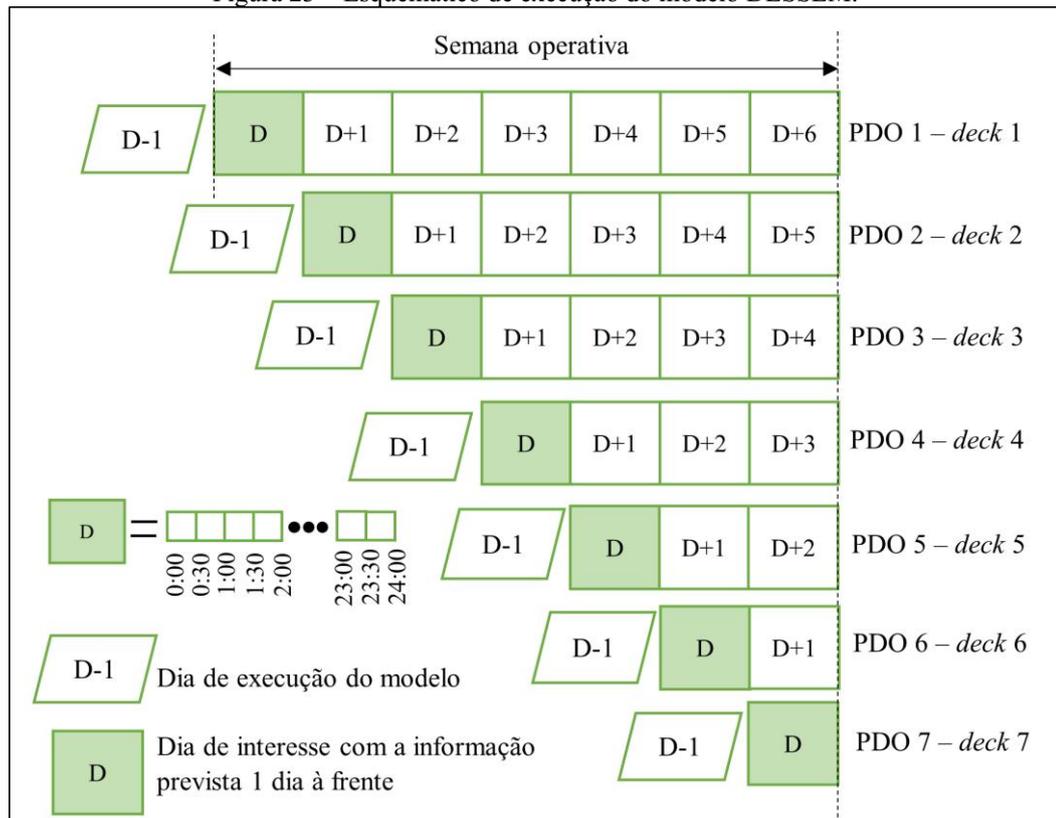
Por outro lado, a EOL-PREV para cada uma das usinas eólicas do SIN é encontrada nos *decks*<sup>12</sup> diários de entrada do modelo DESSEM<sup>13</sup> (SANTOS et al., 2020). Ressalta-se que o horizonte do modelo DESSEM pode ser de até uma semana, sendo o primeiro dia discretizado em períodos de 30 minutos e o restante dos dias em blocos de horas com patamares de carga sequenciais. Portanto, todas as informações previstas um dia à frente<sup>14</sup> têm uma granularidade semi-horária. Na Figura 23 é apresentado um esquemático da execução do modelo DESSEM para uma semana operativa. A execução do modelo acontece no dia D-1 e a informação do sistema um dia à frente, incluída a EOL-PREVS, encontra-se no dia D. Note que, durante uma semana operativa, sete problemas da PDO são executados. Destaca-se que as previsões são atualizadas e o número de dias do horizonte de planejamento é reduzido a cada dia da semana operativa. Neste contexto, para extrair a informação de um ano da EOL-PREV, é necessário ler 365 decks diários de entrada do modelo DESSEM.

<sup>12</sup> Conjunto de arquivos com os dados de entrada de um modelo.

<sup>13</sup> Modelo oficial utilizado pelo ONS para solucionar o problema da PDO do SIN.

<sup>14</sup> Por exemplo a carga, afluência nas hidrelétricas, geração eólica e geração solar

Figura 23 – Esquemático de execução do modelo DESSEM.



Fonte: do autor.

Dentre dos arquivos do *deck* de entrada do modelo DESSEM, a EOL-PREV encontra-se no arquivo “RENOVAVEIS.DAT” cuja formatação pode ser consultada em (CEPEL, 2020). Salienta-se que esse arquivo detalha as previsões de geração eólica, assim como as previsões para a geração das unidades fotovoltaicas, de pequenas hidrelétricas e térmicas à biomassa. Um exemplo da estrutura do arquivo de entrada do DESSEM é apresentada na Figura 24. É possível verificar que a EOL-PREV é informada para cada uma das usinas eólicas consideradas no problema da PDO e para cada período do horizonte de planejamento.

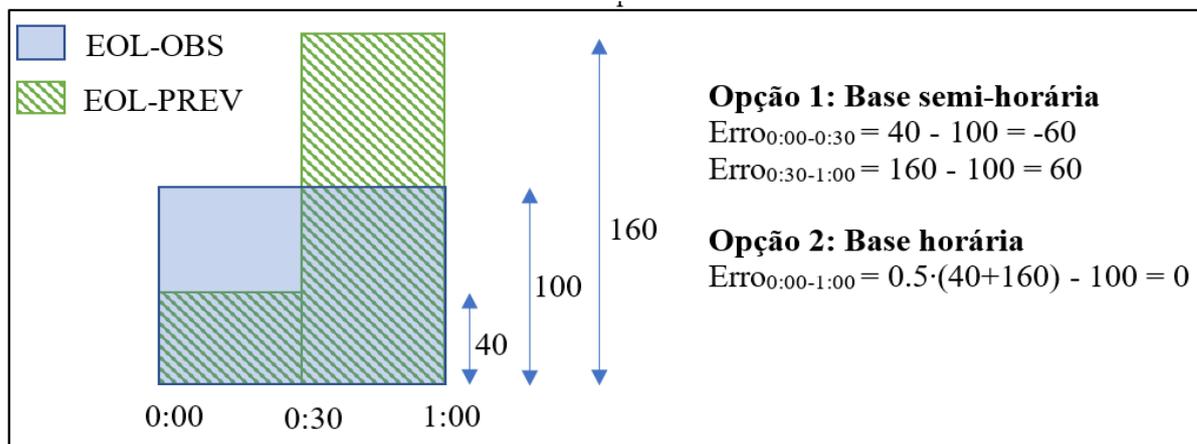
Figura 24 – Exemplo de Arquivo com a EOL-PREV no modelo DESSEM

Identificação da usina		Período	EOL-PREV
&XXXXXXXXXXXXXXXXXX	XXXXX	XX ;XX ;X ;XX ;XX ;X	XXXXXXXXXX ;
&	; CODIGO	DATA	GERACAO ;
&XXXXXXXXXXXXXXXXXX	XXXXX	XX ;XX ;X ;XX ;XX ;X	XXXXXXXXXX ;
EOLICA-GERACAO	; 1	24 ; 0 ; 0 ; 24 ; 19 ; 1	4 ;
EOLICA-GERACAO	; 1	24 ; 19 ; 1 ; 24 ; 22 ; 0	5 ;
EOLICA-GERACAO	; 1	24 ; 22 ; 0 ; 25 ; 19 ; 1	4 ;
EOLICA-GERACAO	; 1	25 ; 19 ; 1 ; 25 ; 22 ; 0	5 ;
EOLICA-GERACAO	; 1	25 ; 22 ; 0 ; 26 ; 19 ; 1	4 ;
EOLICA-GERACAO	; 1	26 ; 19 ; 1 ; 26 ; 22 ; 0	5 ;
EOLICA-GERACAO	; 1	26 ; 22 ; 0 ; 27 ; 19 ; 1	4 ;
EOLICA-GERACAO	; 1	27 ; 19 ; 1 ; 27 ; 22 ; 0	5 ;
EOLICA-GERACAO	; 1	27 ; 22 ; 0 ; 28 ; 19 ; 1	4 ;
EOLICA-GERACAO	; 1	28 ; 19 ; 1 ; 28 ; 22 ; 0	5 ;
EOLICA-GERACAO	; 1	28 ; 22 ; 0 ; 29 ; 0 ; 0	4 ;

Fonte: Adaptado de (ONS, 2022a).

Note que os dados no sistema brasileiro da EOL-OBS e da EOL-PREV têm uma granularidade diferente (horária e semi-horária, respectivamente). Uma primeira questão é a base temporal (horária ou semi-horária) que deve ser adotada para o cálculo do erro de previsão (diferença entre EOL-PREV e EOL-OBS). A Figura 25 apresenta um exemplo no qual a  $EOL-OBS_{0:00-1:00} = 100$  MW para um período de 1 hora e dois registros semi-horários da  $EOL-PREV_{0:00-0:30} = 40$  MW e  $EOL-PREV_{0:30-1:00} = 160$  MW para a mesma hora em questão. Considerando uma base semi-horária, são obtidos desvios de previsão iguais a -60 MW e 60 MW para cada período semi-horário. No entanto, estes desvios de previsão semi-horários podem não corresponder com a realidade já que a  $EOL-OBS_{0:00-1:00}$  corresponde à média horária dos valores medidos ao longo de uma hora. Portanto, a base temporal adequada para realizar o cálculo do erro de previsão é a granularidade temporal da informação observada. Neste contexto, é necessário obter a média horária da EOL-PREV para o período em análise a fim de manter a coerência entre as bases de dados.

Figura 25 – Exemplo cálculo do erro de previsão para EOL-OBS e da EOL-PREV com diferentes bases temporais.



Fonte: do autor..

Outro ponto que dever ser notado é o nível de detalhe da informação observada e prevista. A Tabela 9 apresenta as principais características associadas com a informação da EOL-OBS e EOL-PREV do sistema brasileiro. A principal discrepância se encontra no nível de agregação das usinas, i.e., a EOL-OBS está referenciada para um conjunto de usinas (as quais podem ser filtradas por unidade federativa, subsistemas do SIN, localização e ponto de conexão) e a informação da EOL-PREV pode ser reportada para cada uma das usinas ou por conjunto de usinas.

Tabela 9 – Características da estrutura dos dados da EOL-OBS e da EOL-PREV.

Características	EOL-OBS	EOL-PREV
Nível de agregação das usinas	Conjunto de usinas	Usinas individualizadas Conjunto de usinas
Informação adicional	Unidade Federativa Subsistema Localização (Interior/Litoral) Ponto de conexão <sup>15</sup>	Barra de conexão

Fonte: do autor.

A informação adicional de unidade federativa, subsistema e localização das usinas individualizadas da EOL-PREV pode ser complementada com informação georreferenciada no *Google Maps* e com mapas de redes de operação do ONS. Destaca-se que o máximo nível de granularidade comum entre a EOL-OBS e EOL-PREV é o ponto de conexão. Além disso, o processo de união das informações apresentadas em cada uma das usinas requer um

<sup>15</sup> Um ponto de conexão se refere ao local onde as usinas eólicas são conectadas eletricamente no SIN. Salienta-se que um ponto de conexão no sistema pode ser composto por uma ou mais barras de conexão.

mapeamento dos nomes das usinas visto que a nomenclatura das usinas na base da EOL-PREV difere da EOL-OBS. A partir da base de dados unificada, é possível agrupar as usinas a partir das informações adicionais. A Tabela 10

indica o número de agrupamentos para cada uma das possíveis combinações de menor à maior granularidade.

Tabela 10– Características da estrutura dos dados da EOL-OBS e da EOL-PREV.

Agrupamento	Número de agrupamentos	-
SIN	1	Granularidade
Subsistemas <sup>16</sup>	3	
Subsistemas + Localização <sup>17</sup>	5	
Unidades Federativas <sup>18</sup>	8	
Unidades Federativas + Localização <sup>19</sup>	11	
Ponto de conexão	58	+

Fonte: do autor.

### 3.2 PROCESSO DE CÁLCULO DOS DESVIOS DE PREVISÃO

Com o intuito de aproveitar o máximo das informações das bases de dados referentes à geração eólica, calculam-se os desvios de previsão para cada ponto de conexão do sistema com uma discretização temporal horária. Os desvios de previsão são obtidos a partir dos cálculos da diferença entre a EOL-PREV no D-1 para o dia D (Figura 23) e a EOL-OBS para o dia D. Desta forma, um desvio positivo (negativo) indica que o modelo de previsão empregado pelo ONS sobrestimou (subestimou) a geração observada. É importante ressaltar que a base de dados empregada é referente aos dados de previsão e de geração observada de janeiro de 2020 à junho de 2022. A Figura 26 mostra um exemplo dos desvios para cada uma das horas do mês janeiro de 2021 para o ponto de conexão ACARAU II-69 kV<sup>20</sup>.

<sup>16</sup> Nordeste, Sul e Norte.

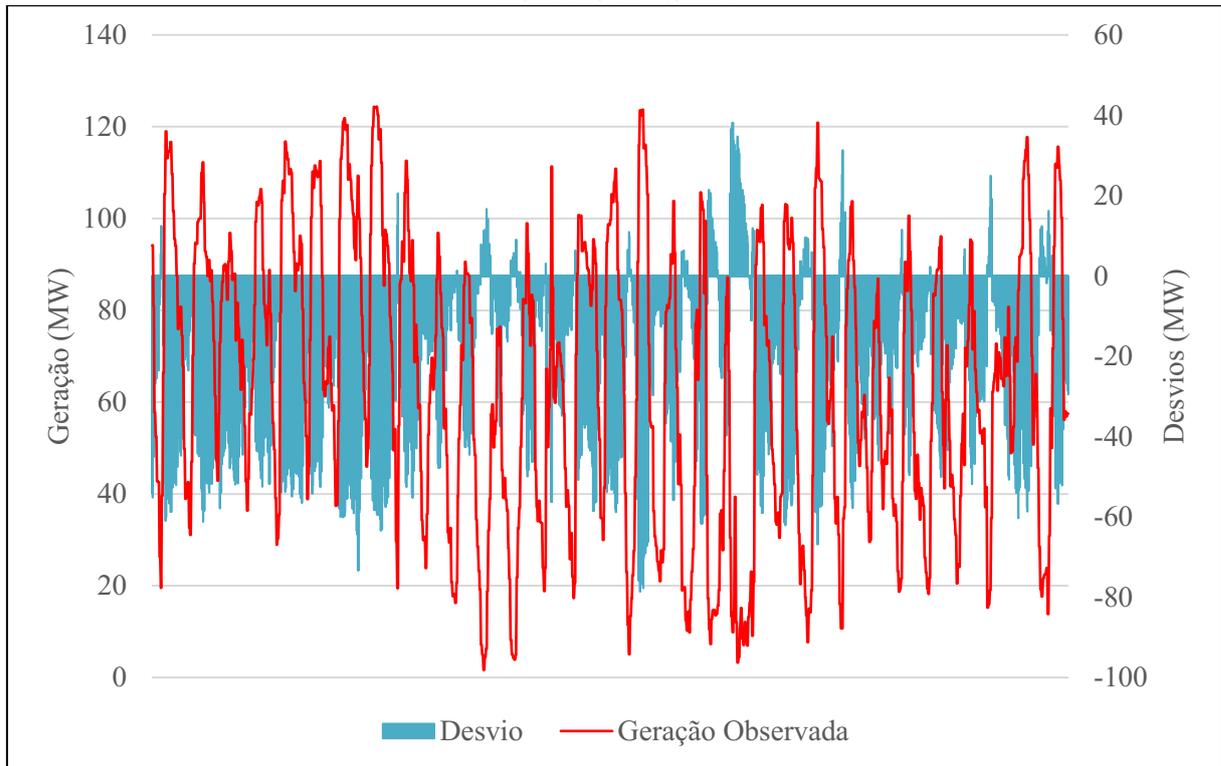
<sup>17</sup> Nordeste Litoral, Nordeste Interior, Sul Litoral, Sul Interior, Norte Interior.

<sup>18</sup> Bahia, Ceará, Maranhão, Pernambuco, Piauí, Rio Grande do Norte, Rio Grande do Sul, Santa Catarina.

<sup>19</sup> Bahia Interior, Ceará Interior, Ceará Litoral, Maranhão Interior, Pernambuco Interior, Piauí Interior, Rio Grande do Norte Interior, Rio Grande do Norte Litoral, Rio Grande do Sul Interior, Rio Grande do Sul Litoral, Santa Catarina Interior.

<sup>20</sup> Localização: litoral do Ceará.

Figura 26– Exemplo dos desvios observados referentes ao mês de janeiro de 2021 para o ponto de conexão ACARAU II – 69 kV.



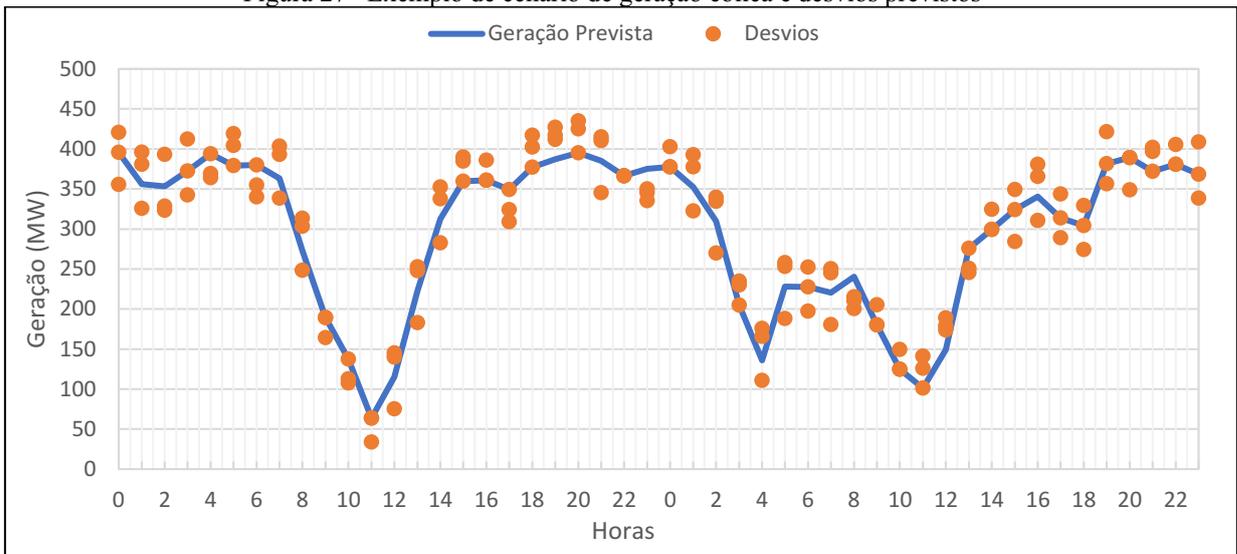
Fonte: do autor.

Após o cálculo dos desvios de geração, obtiveram-se duas bases de dados: geração observada e desvios de previsão. Ambas as bases possuem o mesmo número de amostras e uma discretização horária para cada um dos 58 pontos de conexão existentes no SIN.

### 3.3 DESCRIÇÃO DO MODELO DE GERAÇÃO DE CENÁRIOS DE DESVIOS

Conforme descrito na introdução do capítulo 3, o objetivo deste trabalho é gerar um modelo capaz de integrar as incertezas de geração eólica no modelo de previsão empregado pelo ONS. De uma forma geral, o modelo proposto gera uma nuvem de desvios para uma dada curva de geração prevista escolhida pelo usuário, como mostrado pela Figura 27.

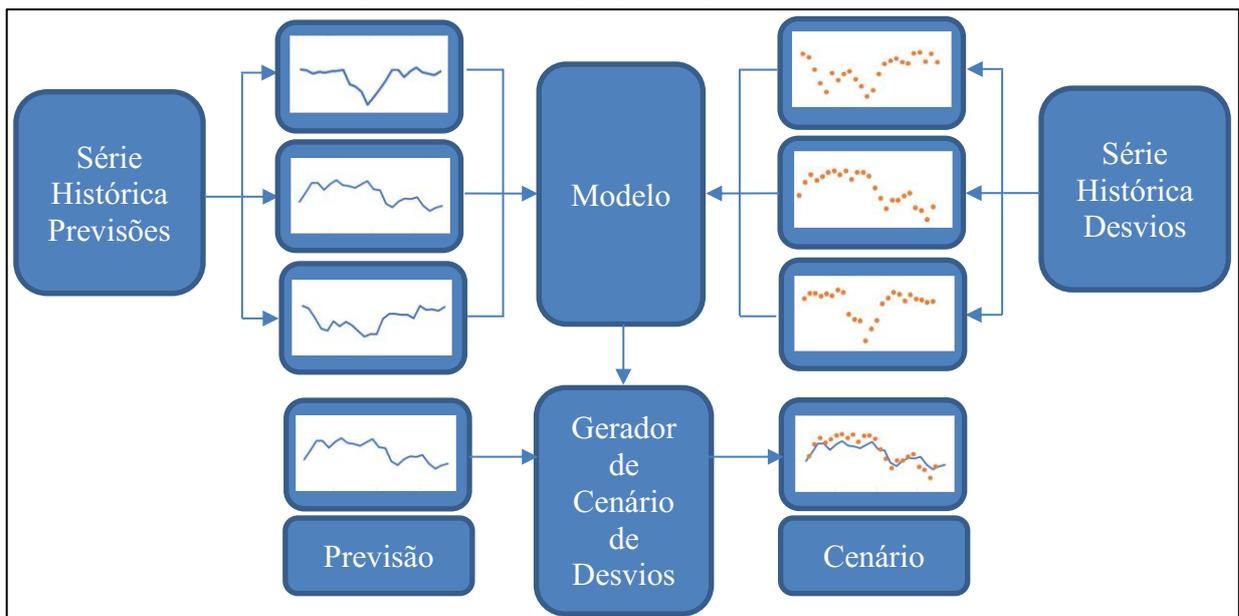
Figura 27– Exemplo de cenário de geração eólica e desvios previstos



Fonte: do autor.

De modo geral, o modelo será treinado a partir dos históricos de previsão e os desvios associados à cada série de previsão com o intuito de extrair as propriedades estatísticas dos dados para, posteriormente, replicá-las. Salienta-se que o modelo, após sua parametrização, toma como entrada uma série temporal de previsões de geração eólica. A partir desta série, o modelo replicará as propriedades estatísticas identificadas na base de dados de treinamento com o objetivo de obter uma série de desvios correlacionada com a curva de previsão. A Figura 28 mostra um esquemático do princípio de parametrização do modelo.

Figura 28 – Metodologia geral de geração de cenários



Fonte: do autor.

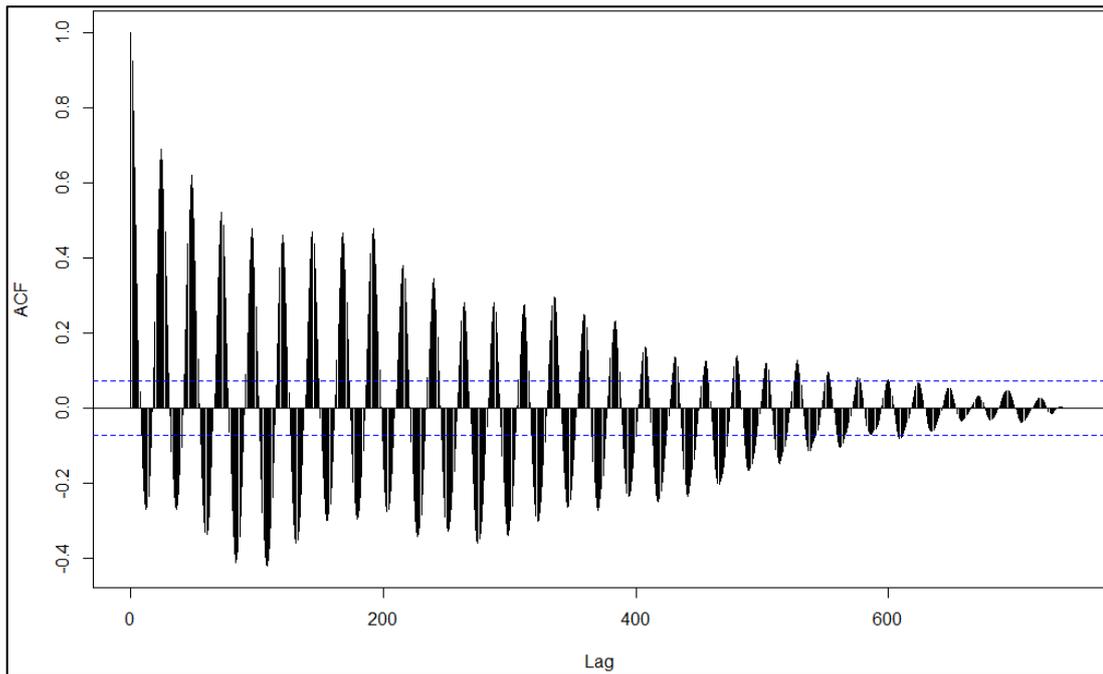
Dentre os possíveis modelos de geração de cenários, optou-se por um modelo de séries temporais denominado modelo Autorregressivo Periódico (PAR). A escolha de um modelo periódico é justificada pelo fato de que os dados de desvios, de geração eólica (verificada e prevista) indicam a existência de uma periodicidade horária. Esta propriedade pode ser verificada por meio do cálculo da função de autocorrelação (ACF) cuja definição é mostrada a seguir:

**Definição:** Seja  $x$  uma variável aleatória discreta,  $\mu_x$  a média e  $\sigma_x$  o desvio padrão das amostras, a função de autocorrelação para o um deslocamento no tempo de  $k$  ( $\rho_k$ ) é definida pela equação (9).

$$\rho_k = \frac{E[(x_t - \mu_x) \cdot (x_{t-k} - \mu_x)]}{\sigma_x^2} \quad (9)$$

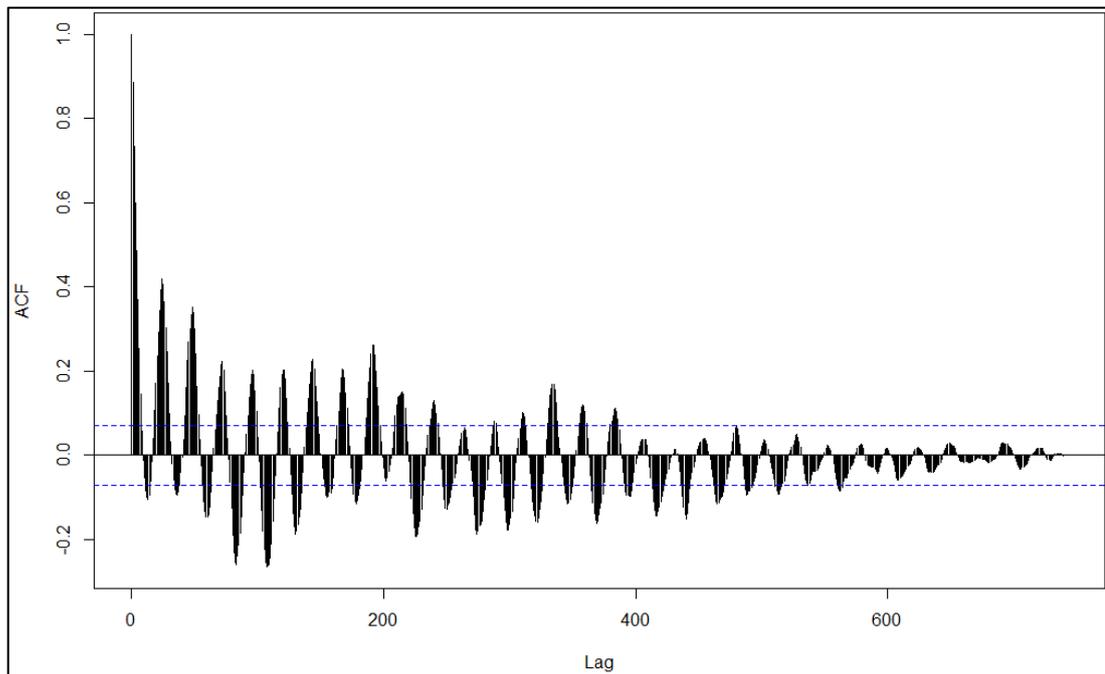
Ressalta-se que uma série temporal possui indícios de periodicidade quando o gráfico da ACF possui um caráter oscilatório/sinusoidal mal amortecido (K.W. HIPEL, 1994). De forma ilustrativa, a Figura 29 e a Figura 30 mostram, respectivamente, exemplos de ACF com indícios de periodicidade para as séries temporais de geração prevista e dos desvios observados para as amostras do mês de janeiro de 2021 para o ponto de conexão Acaraú-69kV. Vale notar que os valores das autocorrelações são significativos caso sejam maiores que um limiar de significância representado pelas linhas tracejadas nas figuras. Este limiar é escolhido de forma a obter um intervalo de confiança de 95% (K.W. HIPEL, 1994).

Figura 29 – Função de Autocorrelação obtido para o histórico de geração prevista de janeiro de 2021 para o ponto de conexão ACARAU II – 69 kV.



Fonte: do autor.

Figura 30 – Função de Autocorrelação obtido para os desvios de previsão de janeiro de 2021 para o ponto de conexão ACARAU II – 69 kV.



Fonte: do autor.

Na sequência é apresentado, em mais detalhes, a metodologia empregada para a parametrização do modelo PAR responsável pela geração de cenários de desvios de geração eólica.

### 3.3.1 Descrição do Algoritmo de Parametrização do modelo PAR

Com o objetivo de criar um modelo de geração de séries sintéticas por meio do modelo Autorregressivo Periódico, um algoritmo composto de 10 etapas foi desenvolvido. Este algoritmo integra as etapas de tratamento dos dados, assim como as etapas de parametrização do modelo PAR, a geração e validação do modelo. Salienta-se que a metodologia desenvolvida neste trabalho teve como base nas teorias de modelagem de séries sintéticas descritas em (K.W. HIPEL, 1994) e em (LARROYD, 2016).

O pseudocódigo referente à metodologia empregada para a parametrização do modelo PAR e a sua validação são mostrados a seguir:

- 1) Tratamento da base de dados;
- 2) Cálculo da matriz de correlação espacial dos desvios entre diferentes pontos de conexão;

Para cada Ponto de Conexão:

Para cada mês = 1,2, ... ,12:

- 3) Cálculo dos momentos estatísticos da base de dados de desvios e de geração prevista;
- 4) Parametrização do modelo PAR;
- 5) Geração das séries sintéticas de desvios;
- 6) Pós-tratamento das séries sintéticas de desvios;
- 7) Cálculo dos momentos estatísticos das séries sintéticas de desvios;
- 8) Cálculo da função de correlação periódica entre as séries sintéticas de desvios e a curva de tendência de geração prevista;
- 9) Cálculo da matriz de correlação espacial entre as séries sintéticas de entre diferentes pontos de conexão;

10) Validação do modelo para o mês em questão.
--

Antes de prosseguir com a descrição do processo de parametrização e de geração de séries sintéticas, é necessário ressaltar que a metodologia desenvolvida neste trabalho é aplicada de forma a obter 12 modelos PAR para cada um dos pontos de conexão. Isto é, um modelo PAR com periodicidade horária é parametrizado para cada um dos meses do ano para cada ponto de conexão da base de dados. A escolha pela segmentação das amostras de acordo com o mês se deve à necessidade de capturar as variações dos padrões de geração eólica e dos desvios de acordo com o período do ano.

1) Tratamento da base de dados

A primeira etapa da metodologia consiste em efetuar o tratamento da base de dados EOL-PREV e de desvios de geração com o intuito de obter uma base de dados coerente com as hipóteses do modelo proposto. De forma geral, o tratamento da base de dados consiste em duas etapas: (a) agrupamento dos 58 pontos de conexão; (b) aplicação de algumas transformações matemáticas.

(a) Agrupamento dos pontos de conexão

O agrupamento dos pontos de conexão é necessário devido ao fato de que o modelo proposto requer a inversão da matriz de correlação espacial. Assim, visto que a matriz de correlação espacial calculada com os 58 pontos de conexão é singular, torna-se necessário aplicar um agrupamento dos pontos de conexão com o intuito de eliminar as dependências.

O processo de agrupamento dos pontos de conexão considera a matriz de correlação espacial entre os pontos de conexão calculada a partir da base de geração observada. Casos em que a correlação entre dois pontos de conexão é elevada indicam um perfil de geração eólica semelhante. Neste contexto, o agrupamento não ocasiona uma perda de informação expressiva. Salienta-se que o uso da matriz de correlação espacial dos dados de geração eólica se deve ao fato de que os dados dos desvios herdam as hipóteses dos modelos de previsão empregados pelo ONS. Assim, os dados de geração são capazes de fornecer uma visão mais fiel das relações entre os dados de cada ponto de conexão.

Outro fator preponderante no agrupamento é o fato de que uma quantidade significativa de pontos de conexão compartilha nomes semelhantes. A diferença reside no nível de tensão da rede em que as usinas eólicas são conectadas. Assim, pontos de conexão com nomes similares e com proximidade geográfica são agrupadas.

#### (b) Transformação Matemática

Antes de iniciar o processo de parametrização, é imprescindível verificar se a estrutura e as propriedades estatísticas dos dados respeitam as hipóteses que regem o modelo a ser parametrizado. Caso os dados não respeitem essas hipóteses, torna-se necessário empregar algumas transformações de forma que os pressupostos sejam atendidos. No contexto do modelo PAR, os dados devem atender à 3 hipóteses principais: Sazonalidade; Estacionariedade e Normalidade. É importante ressaltar que essas hipóteses foram verificadas para os dados de cada um dos pontos de conexão agrupados na etapa anterior. A descrição de cada uma das hipóteses listadas acima, assim como as possíveis transformações dos dados podem ser encontradas em mais detalhes no APÊNDICE B - Requisitos para parametrização do modelo PAR.

#### 2) Cálculo da matriz de correlação espacial

A segunda etapa da metodologia consiste em calcular a matriz de correlação espacial entre os pontos de conexão referentes à base de dados desvios de geração. Ressalta-se que a correlação espacial dos dados é calculada com periodicidade horária para cada mês. Isto é, para cada mês, 24 correlações espaciais são calculadas correspondendo a cada uma das horas do dia. A definição da correlação espacial empregada neste trabalho é descrita a seguir:

**Definição:** Seja  $D_i^{m,h}$  os desvios de geração na hora  $h$  do mês  $m$  para o ponto de conexão  $i$ ,  $\mu D_i^{m,h}$  a média e  $\sigma D_i^{m,h}$  o desvio padrão das amostras, a correlação espacial entre o ponto de conexão  $i$  e  $j$  ( $j \neq i$ ) para a hora  $h$  é dado pela seguinte equação:

$$\rho_{espacial}^{m,h} = \frac{E[(D_i^{m,h} - \mu D_i^{m,h}) \cdot (D_j^{m,h} - \mu D_j^{m,h})]}{\sigma D_i^{m,h} \cdot \sigma D_j^{m,h}} \quad (10)$$

#### 3) Cálculo dos momentos estatísticos da base de dados de desvios e de geração prevista

A terceira etapa da metodologia consiste em calcular algumas métricas<sup>21</sup> com o intuito de identificar algumas propriedades estatísticas da base de dados de desvios e de geração prevista. É importante notar o cálculo é efetuado segmentando os dados de acordo com o ponto de conexão agrupado, mês e hora da amostra.

#### 4) Parametrização do modelo PAR

Na sequência, o modelo PAR com periodicidade horária é parametrizado para cada um dos pontos de conexão agrupados e para cada mês. O modelo PAR proposto neste trabalho é descrito pela equação (11). Salienta-se que o processo de parametrização do modelo PAR proposto é descrito no APÊNDICE C - Modelo Periódico Autorregressivo (PAR).

$$D_{d,h} - \mu D_h = \sum_{p=0}^{np^{(h)}} \phi_p^h \cdot (P_{d,h-p} - \mu P_{h-p}) + \eta_{d,h}(\omega) \quad (11)$$

$$\eta_{d,h}(\omega) = \sigma \eta_h^r \cdot \omega$$

É possível notar que o modelo PAR empregado neste trabalho difere em dois quesitos se comparado ao modelo PAR convencional (equação (C.1) do APÊNDICE C - Modelo Periódico Autorregressivo (PAR)). A primeira diferença é que, no caso deste trabalho, o modelo é multivariável. Isto é, busca-se gerar séries sintéticas de desvios de geração tendo como base a geração prevista.

A segunda diferença entre o modelo PAR convencional e o modelo PAR proposto é que o segundo integra, igualmente, as dependências da variável de geração prevista e desvios sem deslocamento no tempo, o que não é possível no modelo PAR convencional. Este fato é evidenciado pelo índice  $p$  da somatória da equação (11), o qual começa em zero no caso do modelo proposto.

#### 5) Geração das séries sintéticas de desvios

Após a parametrização do modelo PAR multivariável, é possível gerar as séries sintéticas de desvios de geração eólica para um dado ponto de conexão. Este processo resume-se na

---

<sup>21</sup> Média, Mediana, Desvio Padrão, Assimetria da distribuição perante uma distribuição Normal e os valores máximos e mínimos da série temporal.

utilização da equação (12) que corresponde à reestruturação da equação (11). Destaca-se que para gerar uma série sintética é necessário empregar uma série de previsões de geração como tendência. Isto é necessário devido ao fato de que o modelo gera desvios que são estatisticamente correlacionados à previsão de geração eólica.

$$D_{a,h} + \mu D_h + \sum_{p=0}^{np^{(h)}} \phi_p^h \cdot (P_{a,h-p} - \mu P_{h-p}) + \sigma \eta_h^r \cdot \omega \quad (12)$$

Uma das dificuldades para gerar cenários de desvios a partir do modelo PAR da equação (12) é que apenas os requisitos associados às dependências temporais entre a geração prevista e os desvios de um mesmo ponto de conexão são atendidos. No entanto, o modelo PAR, em sua forma atual, não integra as dependências espaciais entre desvios de diferentes pontos de conexão. Uma das estratégias de integração desta dependência espacial consiste em empregar uma matriz de carga nos ruídos brancos (variável  $\omega$ ) (LARROYD, 2016). Esta matriz de carga é obtida a partir da decomposição de Cholesky<sup>22</sup> da matriz de correlação espacial dos desvios estimada na segunda etapa. O APÊNDICE D - Integração da Correlação Espacial descreve em mais detalhes a metodologia de cálculo da matriz de carga a partir da matriz de correlação espacial dos desvios.

#### 6) Pós-tratamento das séries sintéticas de desvios

Após a geração dos cenários de desvios de geração, é necessário efetuar um tratamento das séries sintéticas. Isso se deve pelo fato de que os desvios gerados, somados com a curva de geração prevista empregada, pode desrespeitar os limites operativos superior e inferior das usinas eólicas que constituem o ponto de conexão. Neste contexto, é imprescindível limitar os valores de desvios gerados.

O processo de truncamento de desvios é efetuado a partir da identificação das amostras cujos valores da geração prevista somada com o desvio respectivo ultrapassam o limiar superior ou inferior para o dado ponto de conexão. Salienta-se que o limite inferior corresponde à geração

---

<sup>22</sup> Um dos requisitos para o cálculo da decomposição de Cholesky é que a matriz sendo decomposta seja definida positiva. Este é o motivo pelo qual é necessário efetuar o agrupamento dos pontos de conexão na primeira etapa.

nula, enquanto o limite superior corresponde ao valor máximo de geração encontrado no histórico de geração observada e prevista.

#### 7) Cálculo dos momentos estatísticos das séries sintéticas de desvios

Após o tratamento das séries sintéticas de desvios, as mesmas métricas calculadas na terceira etapa são calculadas nesta etapa. No entanto, ao invés de empregar os dados do histórico de desvios, as séries sintéticas de desvios são utilizadas. É importante ressaltar que as métricas calculadas nesta etapa são comparadas com as métricas obtidas na etapa 3 no processo de validação.

#### 8) Cálculo da função de correlação periódica entre as séries sintéticas de desvios e a curva de tendência de geração prevista

Outra métrica empregada no processo de validação é a correlação periódica entre a série sintética de desvios e a geração prevista empregada como curva de tendência. Note que a correlação temporal entre os desvios e a geração prevista do histórico é calculada durante o processo de parametrização do modelo PAR. Neste contexto, é necessário apenas efetuar o cálculo da correlação temporal entre a série sintética de desvios gerada pela curva de previsão de geração utilizada como tendência. A definição da função de correlação periódica é descrita pela equação (C.4) do APÊNDICE C - Modelo Periódico Autorregressivo (PAR).

#### 9) Cálculo da matriz de correlação espacial entre as séries sintéticas de entre diferentes pontos de conexão

A última métrica estatística necessária para a validação do modelo é a correlação espacial das séries sintéticas de desvios de diferentes pontos de conexão. O processo de cálculo da correlação espacial empregado é o mesmo da etapa 2.

#### 10) Validação do modelo

Por fim, após a parametrização do modelo PAR, a geração de séries sintéticas de desvios e o cálculo de algumas métricas estatísticas, torna-se possível validar o modelo. Para isso, as métricas calculadas a partir das séries sintéticas são comparadas com as métricas calculadas a partir das séries temporais históricas. Salienta-se que a comparação das métricas indica se o

modelo PAR parametrizado consegue capturar as propriedades estatísticas da base de dados de aprendizado e se conseguiu replicá-las nas séries sintéticas.

## 4 RESULTADOS

Este capítulo tem o objetivo de apresentar os resultados obtidos a partir da implementação da metodologia descrita no capítulo 3. Primeiramente, o tratamento da base de dados disponível é descrito. Em seguida, os resultados são apresentados para os dados de treinamento, assim como para os dados de validação final. Para concluir, as considerações finais são descritas com um foco nas limitações e possíveis aperfeiçoamentos para trabalhos futuros.

### 4.1 BASE DE DADOS

Neste trabalho é empregada uma base de dados composta pela geração prevista e verificada referente ao período entre janeiro e junho de 2022. Salienta-se que a base de dados referente ao período entre janeiro de 2020 e dezembro de 2021 foi utilizada para o treinamento do modelo. Já a base de dados de 2022 foi utilizada para a validação final do modelo. Isto é, o ano de 2022 é empregado com o intuito de verificar se o modelo é capaz de replicar os padrões dos desvios de geração para uma curva de previsão de geração não utilizada durante o treinamento. A validação final é utilizada para simular o uso do modelo pelo usuário.

#### 4.1.1 Tratamento da Base de Dados

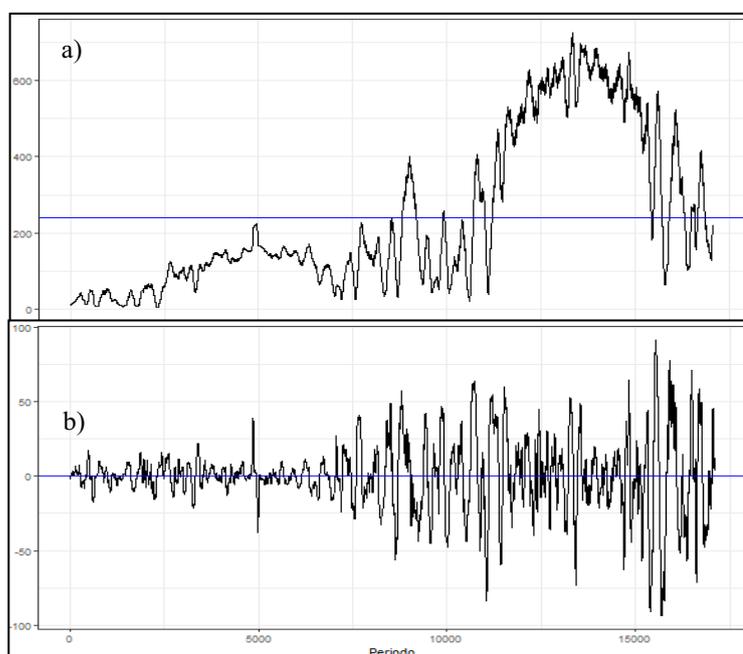
Como descrito no capítulo 3.3.1, é necessário realizar uma etapa de pré-processamento na base de dados. Esta etapa é essencial para a obtenção de uma base de dados robusta e condizente com as hipóteses do modelo PAR. Ressalta-se que os tratamentos da base de dados de treinamento são realizados separadamente da base de dados da validação final. A ideia é evitar que os resultados na etapa de validação final para que um viés não seja induzido no modelo. Isto é, caso ambas as bases de dados forem processadas em conjunto, o modelo teria a informação dos dados da base de dados de validação, e, por consequência induziria a um viés.

A primeira etapa do tratamento da base de dados consiste em agrupar os 58 pontos de conexão do SIN associados à geração eólica. Este agrupamento é necessário visto que a integração da correlação espacial dos desvios de geração requer que a matriz de correlação espacial de desvios seja definida positiva. No final, a partir dos 58 pontos de conexão iniciais, 28 pontos de conexão agrupados são obtidos. O APÊNDICE E - Lista de Pontos de Conexão mostra a relação dos pontos de conexão antes e depois do agrupamento. Ressalta-se que o agrupamento dos dados tem como base as correlações calculadas para os dados de geração observada de 2020 e de

2021. Após a identificação das regras de agrupamento para a base de treinamento, elas são replicadas na base de validação final.

A segunda etapa do tratamento de base de dados consiste em aplicar algumas transformações matemáticas nas séries temporais com o fim de garantir as hipóteses de estacionaridade e normalidade do modelo PAR. Este tratamento se faz necessário devido ao fato de que o parque eólico brasileiro está em forte expansão. Isto é, muitos parques eólicos tiveram sua capacidade instalada incrementada durante o período de 2020 e 2022. Assim, por construção, as séries de geração prevista e observada são crescentes (média crescente) e, por consequência, são não estacionárias. No final, verificou-se que a estacionaridade das séries de geração prevista e de desvios é suficiente para atender todas as hipóteses do modelo PAR. A Figura 31 e Figura 32 mostram, respectivamente, as médias móveis e os histogramas antes e depois da estacionaridade referentes às amostras de 2020 e 2021 do ponto de conexão Grupo João Piauí. Percebe-se que, após o tratamento, os dados são estacionários (média móvel estável) e possuem uma distribuição normal visto que a distribuição cumulativa dos dados estacionários se aproxima da distribuição cumulativa de uma curva normal<sup>23</sup>.

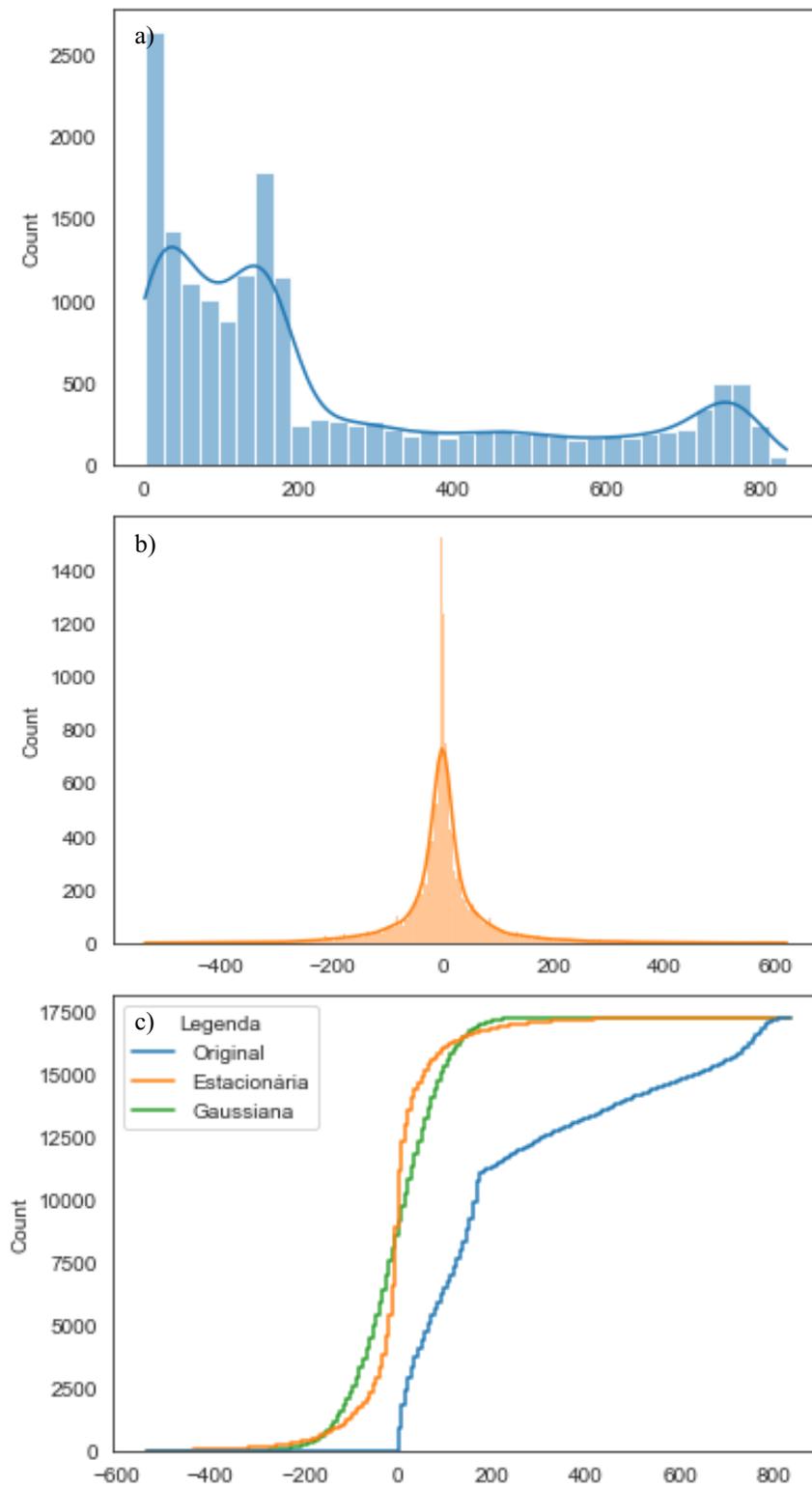
Figura 31 - Média móvel (curva em preto) e média global (curva em azul) antes (a) e depois (b) do processo de estacionarização para o ponto de conexão Grupo João Piauí



Fonte: do autor.

<sup>23</sup> A distribuição cumulativa normal foi obtida levando em consideração a média e o desvio padrão calculada a partir dos dados estacionarizados.

Figura 32- Distribuição das amostras antes (a) e depois (b) e a comparação da distribuição cumulativa (c) antes e depois da estacionarização para o ponto de conexão Grupo João Piauí



Fonte: do autor.

## 4.2 RESULTADOS DAS SIMULAÇÕES

Neste capítulo, os resultados obtidos para a base de dados de treinamento e de validação final são apresentados. Ressalta-se que os resultados apresentados neste capítulo utilizam as curvas de previsão de geração eólica de 2 a 31 de janeiro de 2021 para a base de dados de treinamento. Já para a base de dados de validação final, as curvas de previsão de geração correspondem ao período entre 2 e 31 de janeiro de 2022. Adicionalmente, 500 cenários independentes de desvios de geração eólica são gerados a partir de um modelo de ordem 1<sup>24</sup>.

Os resultados obtidos para cada um dos 28 pontos de conexão para o horizonte de estudo podem ser encontrados em sua integridade no repositório [GitHub](#) do trabalho.

### 4.2.1 Base de dados de Treinamento

Como mencionado na Etapa 9 da metodologia de parametrização do modelo PAR (capítulo 3.3.1), é necessário validar o modelo a partir das séries temporais provenientes da base de treinamento. A partir desta etapa, é possível verificar se o modelo parametrizado é capaz de replicar as propriedades estatísticas identificadas durante o treinamento. Assim, caso o modelo não tiver esta capacidade, será necessário revisitar a metodologia de parametrização, em especial a etapa de tratamento dos dados.

A validação do modelo é feita a partir da comparação de algumas propriedades estatísticas obtidas das séries sintéticas e dos desvios observados no histórico. As métricas empregadas são a média, o desvio padrão, a correlação espacial e a autocorrelação periódica. Note que apenas as 3 primeiras serão apresentadas neste capítulo para simplificar a interpretação dos resultados.

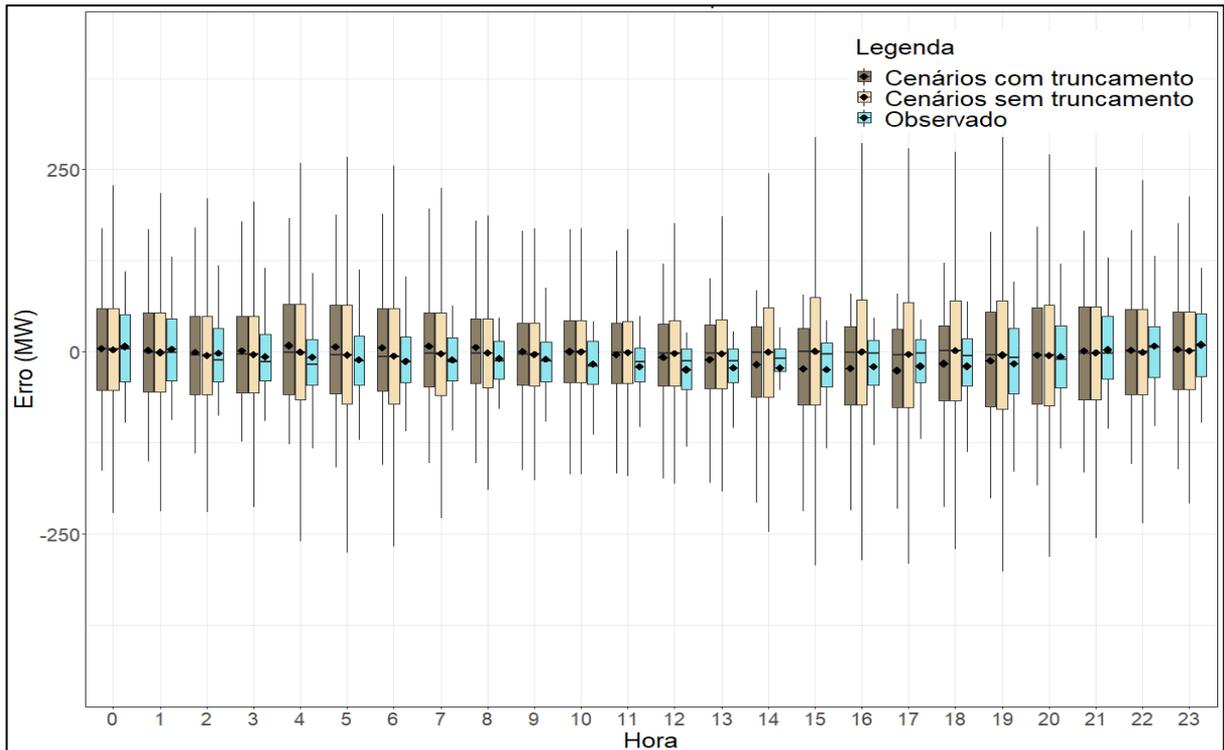
Antes de começar a analisar os resultados, é importante ressaltar que as séries sintéticas de desvios geradas na etapa 4 do capítulo 3.3.1 necessitam de um processamento. Isto se deve ao fato de que o modelo gera desvios sem levar em consideração os limites físicos do parque eólico, tais como a geração máxima e mínima. Neste sentido, é necessário efetuar um truncamento dos valores de desvios. Este procedimento é detalhado na etapa 5 (Pós-tratamento das séries sintéticas de desvios) do capítulo 3.3.1.

---

<sup>24</sup> A ordem do modelo se refere à variável  $np^{(h)}$ . Para um caso de ordem 1, o modelo utiliza a informação do presente e a informação da hora anterior. Mais detalhes podem ser encontrados no APÊNDICE C - Modelo Periódico Autorregressivo (PAR)

O efeito do truncamento pode ser observado comparando os boxplots da Figura 33, a qual mostra os desvios gerados pelo modelo PAR para o ponto de conexão Grupo Ibiapina. Analisando os gráficos, percebe-se que o processo de truncamento das séries possui um efeito significativo nos resultados. Note que após o truncamento, a média (pontos) das amostras de desvios é modificada visto que os outliers<sup>25</sup> presentes antes do truncamento são eliminados.

Figura 33 - Boxplots das amostras por hora para o ponto de conexão Grupo Ibiapina para o período entre 2 de janeiro a 31 de janeiro de 2021 com e sem truncamento.



Fonte: do autor..

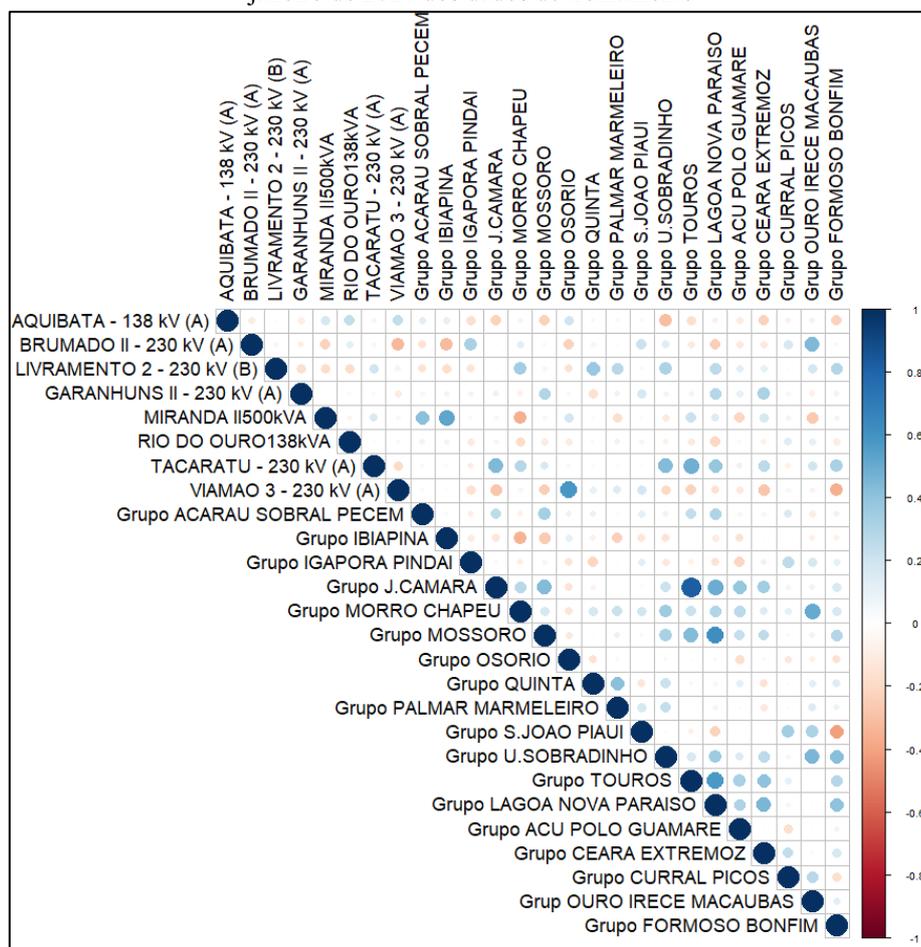
Outra análise que pode ser feita a partir da Figura 33 é que, em geral, o modelo é capaz de replicar as propriedades estatísticas da base de dados de treinamento (barras em azul) para o ponto de conexão em questão. Note que a média (pontos) e a mediana (traços) dos desvios “sintéticos” se aproximam da média e mediana dos desvios observados para o período entre 2 e 31 de janeiro de 2021. Além disso, percebe-se que o modelo PAR gera valores com uma maior dispersão, indicando uma capacidade de gerar casos de desvios mais extremos.

A segunda métrica analisada é a correlação espacial dos desvios entre os pontos de conexão agrupados. Essa métrica auxilia no processo de validação da metodologia de integração da correlação espacial no modelo PAR (APÊNDICE D - Integração da Correlação Espacial). A

<sup>25</sup> Os outliers são omitidos nos gráficos de boxplots para a facilitar a visualização dos resultados.

Figura 34 e Figura 35 mostram, respectivamente, a correlação espacial para os dados de treinamento e dos desvios gerados pelo modelo PAR. Os mapas de calor<sup>26</sup> apresentados se referem à correlação espacial na hora 0 (entre meia noite e 1 da manhã) para os dados entre 2 e 31 de janeiro de 2021. Note que, de forma geral, as séries sintéticas geradas reproduzem as correlações espaciais mais significativas<sup>27</sup> dos dados de treinamento. Contudo, a replicação da correlação espacial não é perfeita visto que as correlações espaciais entre alguns pontos de conexão podem ser ligeiramente diferentes. Para uma melhor representação, é necessário empregar a matriz de correlação espacial para cada um dos 58 pontos de conexão da base de dados. Todavia, isto não é possível visto que a matriz se torna semi-definida positiva ou definida negativa, o que impede o uso da técnica empregada neste trabalho.

Figura 34 – Mapa de calor correlação espacial entre os pontos de conexão agrupados referente à hora 0 de janeiro de 2021 dos dados de treinamento

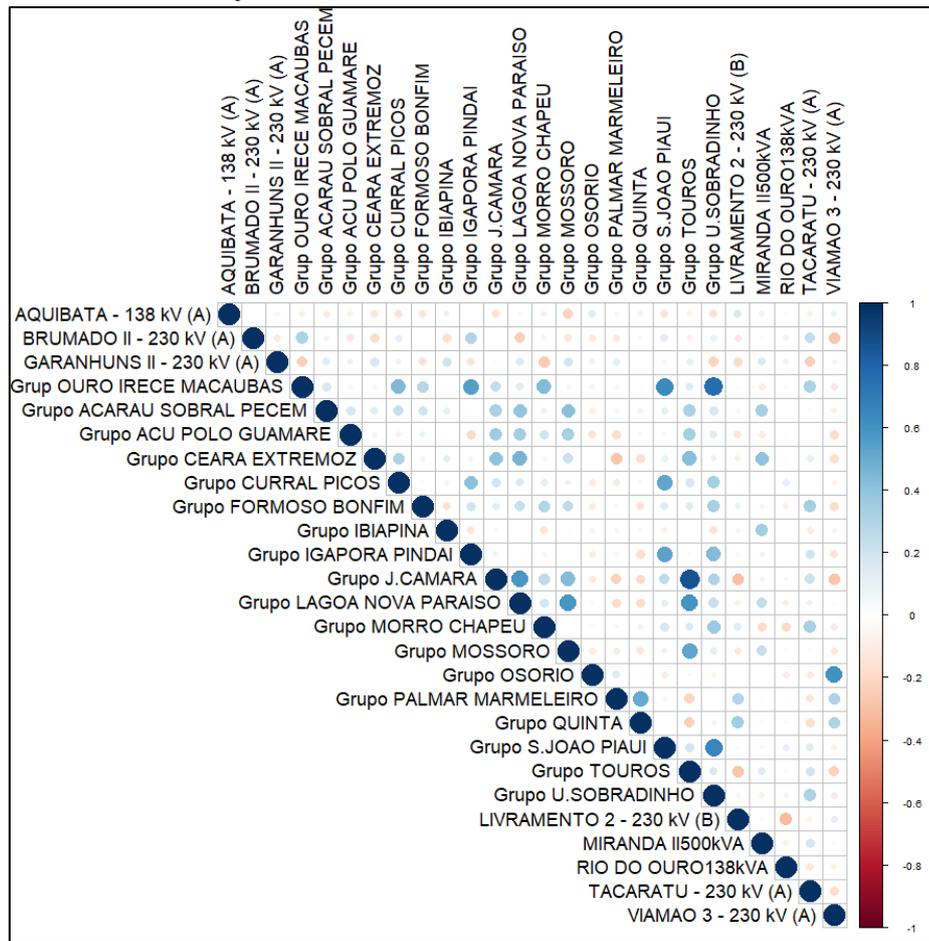


Fonte: do autor.

<sup>26</sup> Para a leitura do mapa de calor considere que o tamanho do círculo indica a intensidade da correlação e a cor indica o sentido (azul correlação positiva e laranja correlação negativa). A diagonal do mapa de calor se refere à autocorrelação dos desvios de cada ponto de conexão.

<sup>27</sup> Neste trabalho, considera-se que a correlação é significativa caso for maior em módulo que 0.4.

Figura 35 – Mapa de calor correlação espacial entre os pontos de conexão agrupados referente à hora 0 de janeiro de 2021 da série sintética de desvios

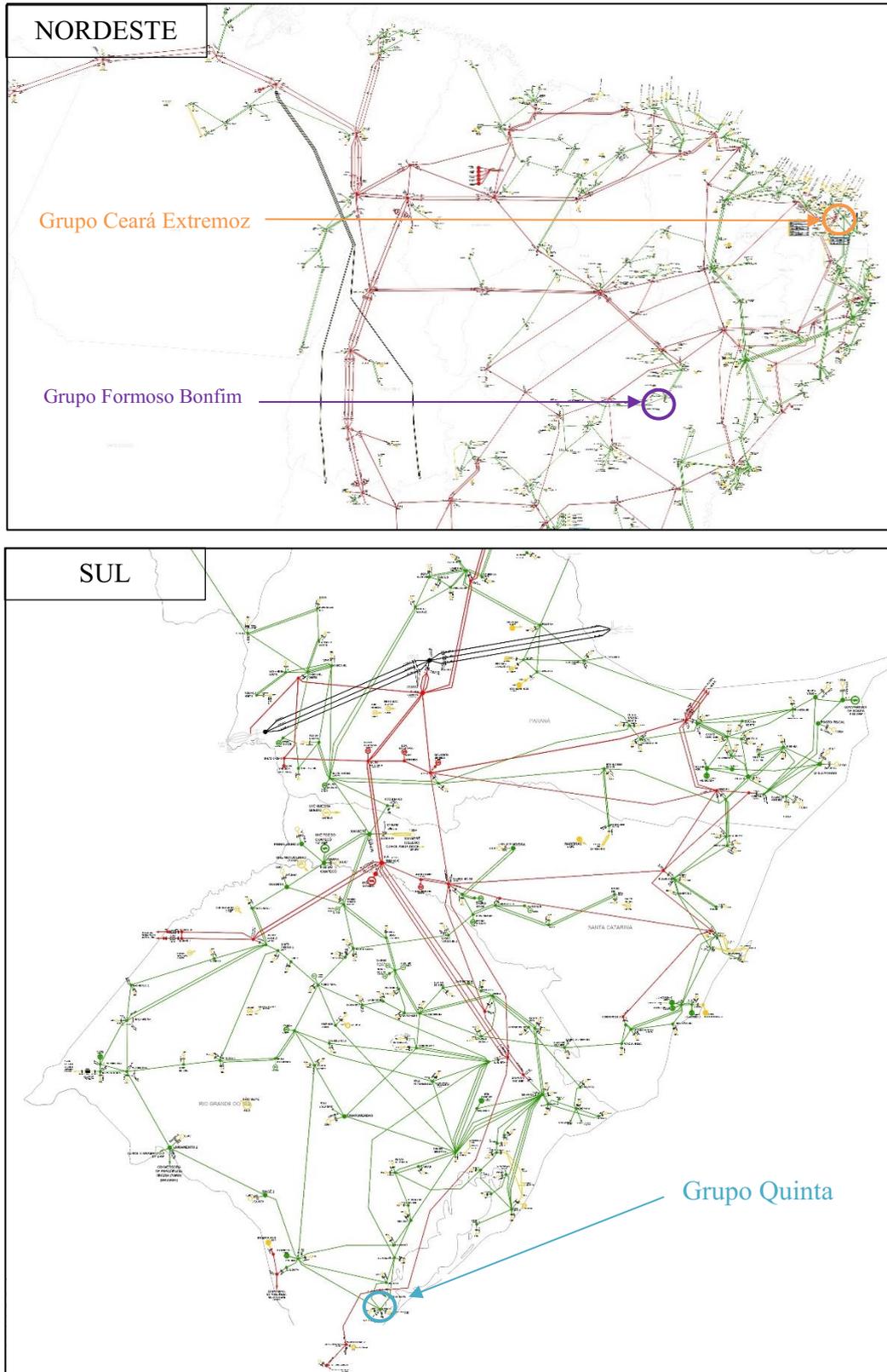


Fonte: do autor.

Para concluir a etapa de validação utilizando a base de dados de treinamento, são comparados os valores de geração prevista corrigidos<sup>28</sup> com os desvios gerados a partir da geração observada para cada ponto de conexão agrupado. A Figura 37, Figura 38 e a Figura 39 mostram, respectivamente, as comparações para os pontos de conexão Grupo Formos Bonfim, Grupo Quinta e Grupo Ceará Extremoz. O primeiro ponto de conexão encontra-se no interior da Bahia, enquanto o segundo e o terceiro, estão localizados no litoral do Rio Grande do Sul e litoral do Ceará. Estes pontos foram escolhidos devido à diversidade do padrão dos ventos observados entre o litoral/interior do Nordeste e dos ventos na região Sul. A Figura 36 mostra a localização no mapa de operação do SIN de cada um dos pontos de conexão listados acima.

<sup>28</sup> Geração prevista somada aos desvios dos cenários gerados

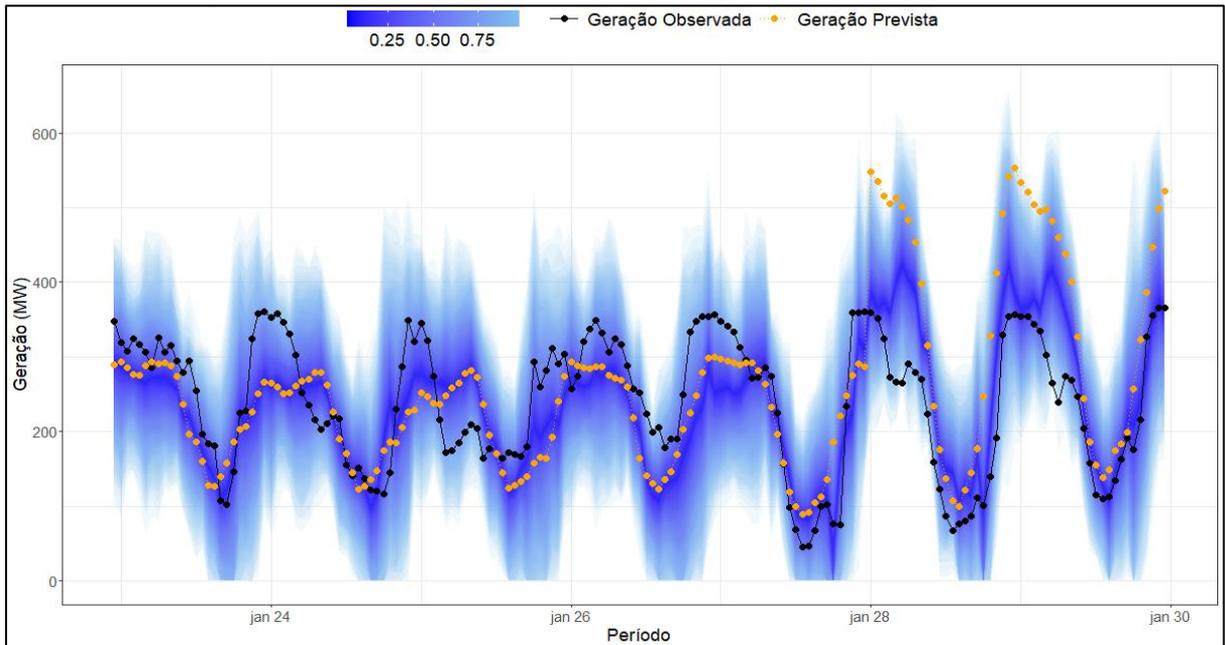
Figura 36 – Localização dos pontos de conexão agrupados no mapa do SIN



Fonte: adaptado de (ONS, 2022c)

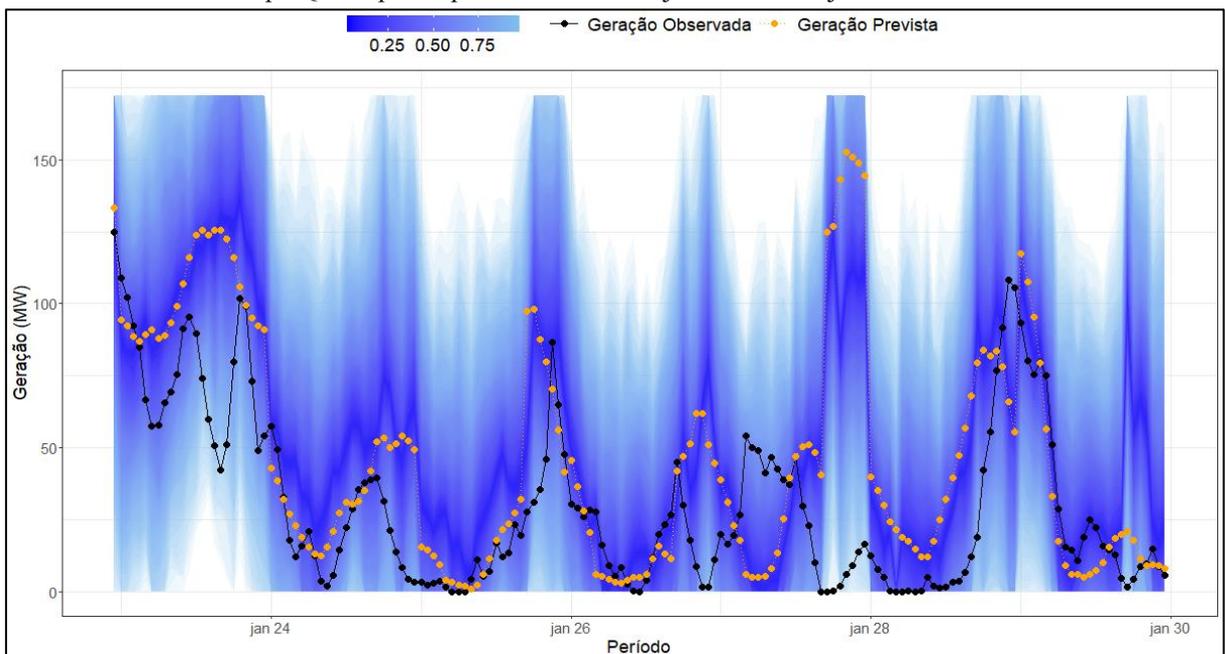
Ressalta-se que, para a leitura dos gráficos a seguir, deve se ter em mente que a área em tons de azul se refere aos desvios de geração somada com a geração prevista em laranja. A intensidade da cor indica as regiões com maior probabilidade. Isto é, quanto mais escuro, maior é a densidade dos desvios gerados pelo modelo PAR.

Figura 37- Comparação da distribuição dos desvios, geração prevista e observada para o ponto de conexão Grupo Formoso Bonfim para o período entre 23 de janeiro a 30 de janeiro de 2021



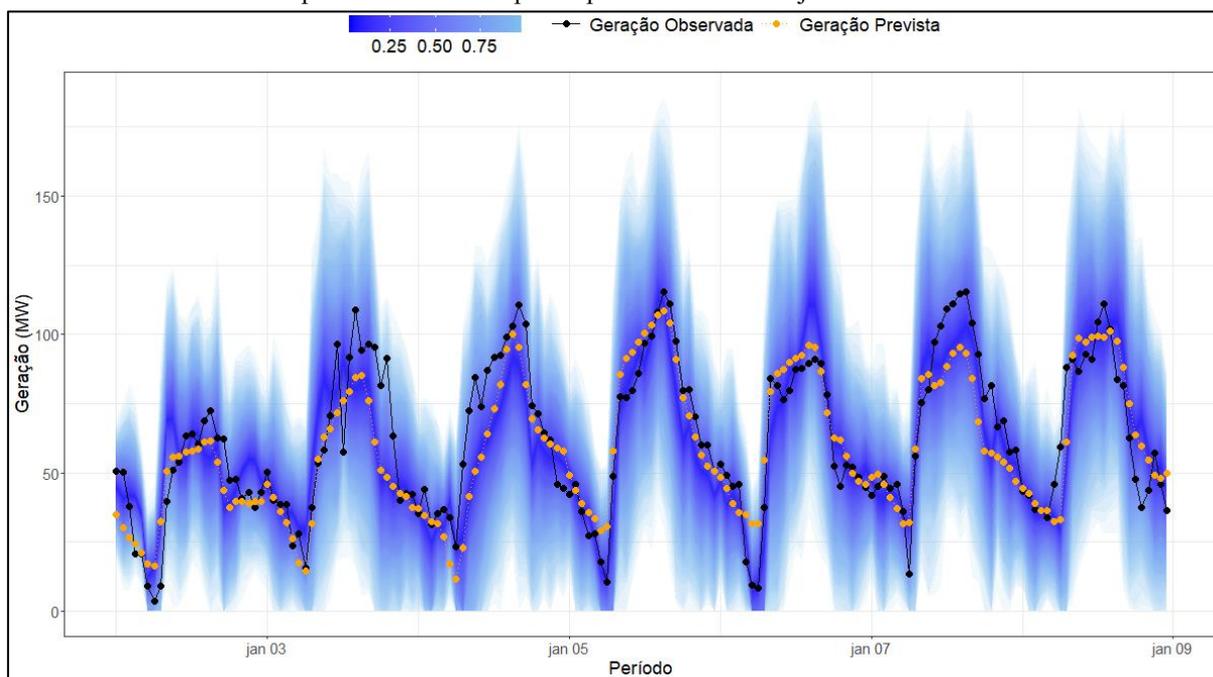
Fonte: do autor.

Figura 38- Comparação da distribuição dos desvios, geração prevista e observada para o ponto de conexão Grupo Quinta para o período entre 23 de janeiro a 30 de janeiro de 2021



Fonte: do autor.

Figura 39- Comparação da distribuição dos desvios, geração prevista e observada para o ponto de conexão Grupo Ceará Extremoz para o período de 2 a 9 de janeiro de 2021



Fonte: do autor.

Analisando os gráficos de geração prevista corrigida pelos desvios e comparando-os com os de geração prevista/observada, percebe-se que o modelo tende corrigir a tendência da geração prevista. Isto é, quando a geração prevista está em torno da geração máxima, o modelo interpreta que a geração observada está sobrestimada, e, conseqüentemente, os desvios gerados corrigem para baixo a geração prevista. Por outro lado, caso a geração prevista esteja em valores de geração mínima, o modelo tende a corrigir para cima a geração prevista. Note que o modelo pode gerar desvios que podem distanciar ainda mais dos pontos de geração observada. O modelo consegue gerar desvios mais condizentes com a geração verificada caso a previsão de geração não estiver sobrestimando ou superestimando em excesso a geração observada. Isto é, mesmo que a correção seja no sentido correto (aumento ou redução da geração prevista), o valor corrigido pode ainda estar distante do valor observado. Por exemplo, no caso da Figura 38 no final do dia 27 de janeiro percebe-se que a previsão está sobrestimando significativamente a geração observada. O modelo apenas consegue efetuar uma correção parcial visto que a diferença é considerável. Neste sentido, o desempenho do modelo ainda está atrelado à qualidade da previsão.

Para concluir, percebe-se que nos casos em que há um padrão de geração prevista, o modelo possui um desempenho melhor que nos casos em que a curva de geração prevista varia grandes amplitudes em um curto espaço de tempo.

## 4.2.2 Base de dados de Validação Final

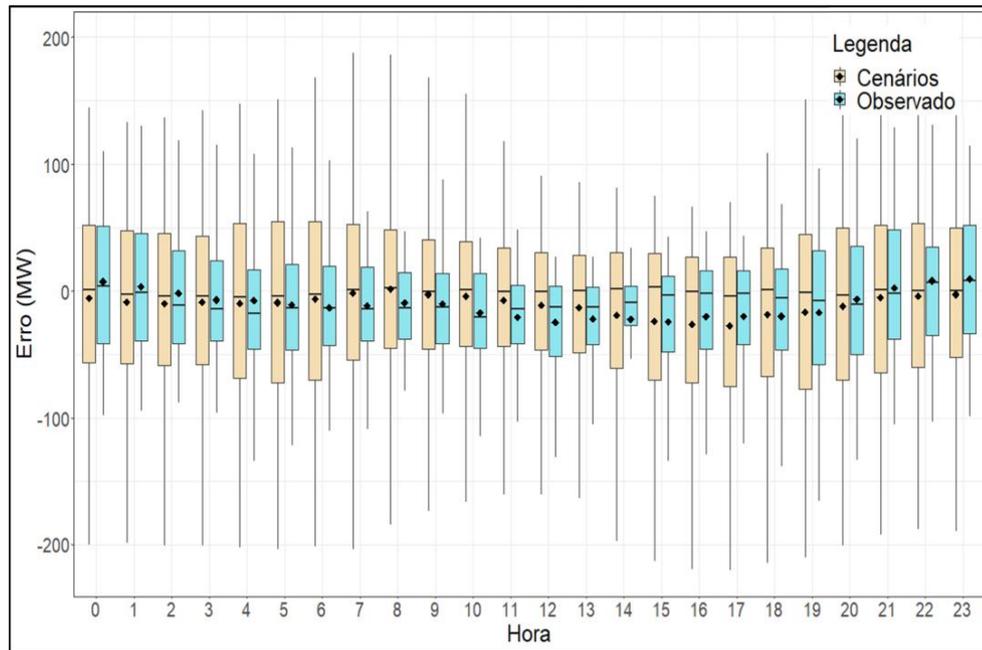
Esta etapa emula o uso do modelo: o usuário entra com uma curva de previsão de geração denominada de tendência que é desconhecida pelo modelo com o objetivo de obter cenários de desvios de geração correlacionados com os desvios observados no histórico.

De modo geral, para verificar se o modelo consegue gerar séries sintéticas de desvios de geração coerentes, é necessário verificar as mesmas métricas apresentados no capítulo 4.2.1. No entanto, neste capítulo, serão apenas apresentados os boxplots comparando a distribuição dos desvios observados e sintéticos após truncamento e os gráficos comparando a geração prevista corrigida e a geração observada. Note que os resultados mostrados a seguir utilizam os mesmos pontos de conexão agrupados da seção 4.2.1 para efeito de comparação. Os resultados para todos os pontos de conexão encontram-se no repositório [GitHub](#) do trabalho.

Analisando as Figura 40 à Figura 43, verifica-se que o modelo não é tão eficiente se comparado aos resultados do capítulo 4.2.1. Isto se deve pelo fato de que a geração observada tende a ficar fora da região associada à geração prevista corrigida pelos desvios gerados pelo modelo. No entanto, de forma geral, o modelo consegue capturar a média e a mediana dos desvios. Além disso, percebe-se que os desvios possuem uma dispersão maior neste caso se comparado com os resultados da seção 4.2.1.

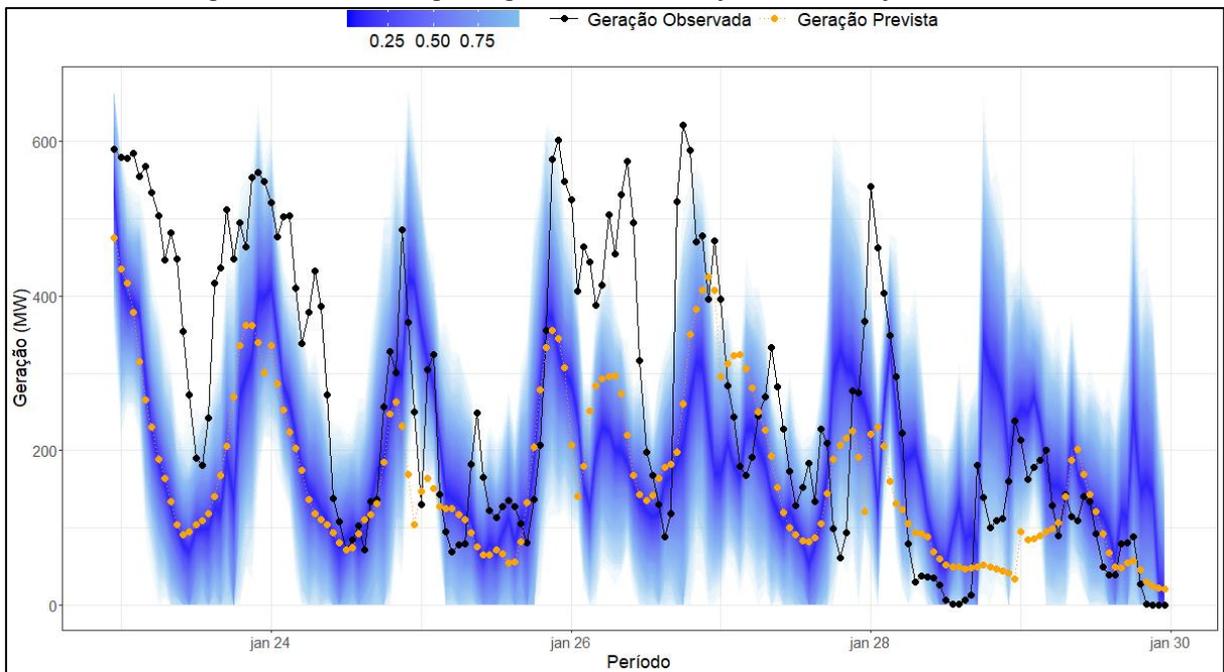
Um dos motivos da deterioração da performance é o uso de uma base de dados relativamente pequena. Isto é, como o modelo emprega apenas os dados de 2020 e 2021, ele não consegue lidar eficientemente com dados cujos padrões diferem muito dos padrões observados durante o treinamento. Isto pode ser verificado comparando a Figura 37 e a Figura 41. Note que o padrão de geração muda de forma considerável. Em 2021, a curva e geração é mais comportada se comparada com a curva de geração observada em 2022. Uma solução seria o uso de uma base de dados com o período maior, o que não foi possível neste trabalho devido à indisponibilidade dos dados necessários.

Figura 40 - Boxplots das amostras por hora para o ponto de conexão Grupo Ibiapina para o período entre 2 de janeiro a 31 de janeiro de 2022 após truncamento



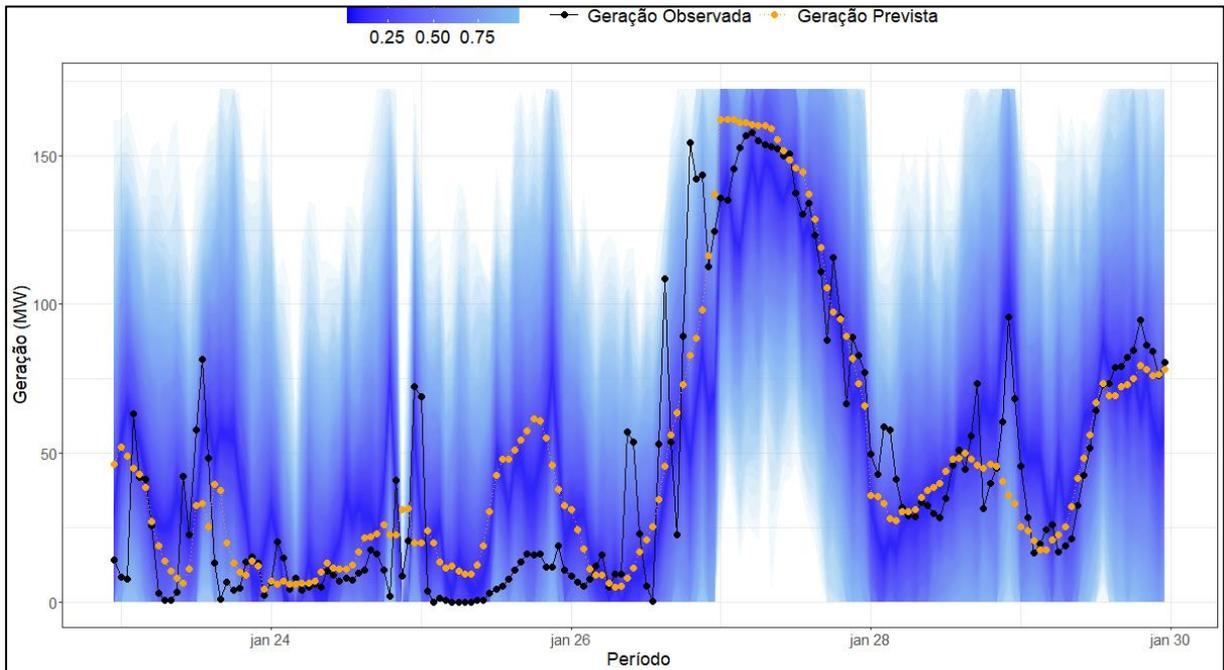
Fonte: do autor.

Figura 41- Comparação da distribuição dos desvios, geração prevista e observada para o ponto de conexão Grupo Formoso Bonfim para o período entre 23 de janeiro a 30 de janeiro de 2022



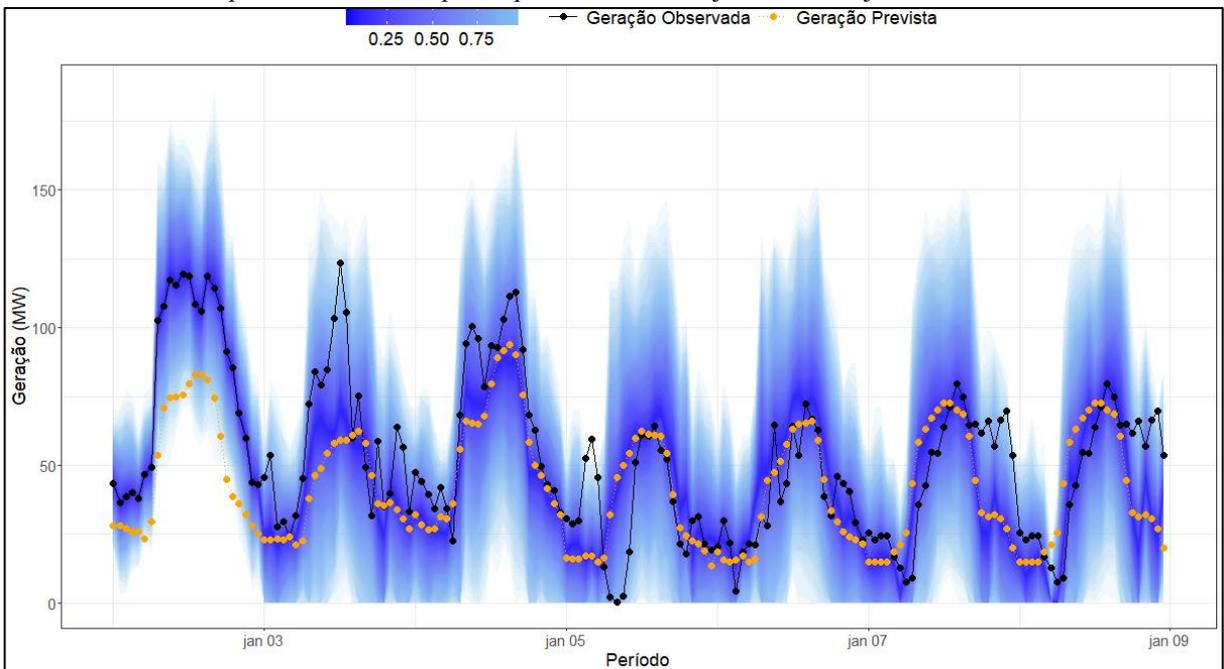
Fonte: do autor.

Figura 42- Comparação da distribuição dos desvios, geração prevista e observada para o ponto de conexão Grupo Quinta para o período entre 23 de janeiro a 30 de janeiro de 2022



Fonte: do autor.

Figura 43- Comparação da distribuição dos desvios, geração prevista e observada para o ponto de conexão Grupo Ceará Extremoz para o período entre 2 de janeiro a 9 de janeiro de 2021



Fonte: do autor.

### 4.3 CONSIDERAÇÕES FINAIS

Levando em consideração os resultados, percebe-se que o modelo é capaz de replicar as propriedades estatísticas da base de dados de treinamento. Como esperado, os resultados são

melhores quando uma curva de previsão de geração de 2021 é empregada para gerar séries sintéticas de desvios. Este fato mostra que seria necessário e desejável que a base de dados de treinamento incorporasse amostras de um horizonte de tempo maior. Isto não se mostrou possível no contexto brasileiro visto que os dados são obtidos por meio dos decks de entrada do modelo DESSEM cuja entrada na cadeia de modelos empregados pelo ONS e CCEE ocorreu apenas em janeiro de 2021. Assim, após uma análise dos dados disponíveis, julgou-se que apenas os dados a partir de 2020<sup>29</sup> eram confiáveis para utilização.

Outra limitação do modelo está associada ao fato de que as propriedades estatísticas entre geração prevista e de desvios pode mudar com o aumento da capacidade instalada de geração eólica em cada ponto de conexão. Assim, o modelo apresentado não será capaz de gerar séries sintéticas de desvios condizentes com a realidade caso a condições forem muito distintas daquelas observadas na base de dados de treinamento. Existem duas soluções para este problema:

1. Aumentar o horizonte de tempo da base de dados de treinamento;
2. Treinar o modelo a partir de valores de geração prevista e de desvios normalizados.

A primeira solução esbarra no problema já mencionado anteriormente. Isto é, não é possível aumentar a base de dados de forma imediata visto que seria necessário esperar por um longo período até possuir um número suficiente de dados, o que torna esta opção inviável.

A segunda solução se refere à normalização dos dados de geração prevista e de desvios. Para isto, seria necessário obter valores de geração prevista e de desvios relativos à capacidade instalada de geração eólica no ponto de conexão em questão. Ao invés de empregar dados em MW, seriam utilizados valores em p.u<sup>30</sup> tendo como base os valores capacidade instalada de cada ponto de conexão.

Para concluir, o modelo parametrizado neste trabalho é capaz de gerar séries sintéticas de desvios para o conjunto de pontos de conexão agrupados. Neste contexto, é necessário efetuar um desagrupamento para retornar ao número de pontos de conexão original. Esta etapa não foi desenvolvida visto que a regra de desagrupamento pode variar de acordo com a aplicação do modelo. É possível que, em algumas aplicações, não seja necessário desagrupar visto que a

---

<sup>29</sup> Os dados referentes ao ano de 2020 foram obtidos nos decks de operação sombra do DESSEM.

<sup>30</sup> p.u: por unidade

informação relevante é a geração eólica total do sistema e não a injeção por ponto de conexão. Por outro lado, o desagrupamento pode ser essencial para uma aplicação em que é necessário efetuar uma análise do sistema de um ponto de vista elétrico e não energético.

Uma sugestão de metodologia seria desagrupar de acordo com a contribuição relativa de cada ponto de conexão para o ponto agrupado. Para isso, seria necessário recorrer aos dados de treinamento. Note que esta regra teria que ser atualizada a cada vez que a capacidade instalada de um ponto de conexão aumenta.

## 5 CONCLUSÕES

Este trabalho teve o objetivo de desenvolver um modelo capaz de gerar cenários de desvios de geração eólica a partir dos históricos de previsão do ONS e de geração verificada do SIN. O modelo desenvolvido pode ser empregado como uma ferramenta para aprimorar a tomada de decisão em um contexto em que as incertezas de geração estão aumentando à medida que as fontes de energia renováveis variáveis, tais como a eólica, estão se expandindo.

O modelo proposto pode ser empregado em diversas aplicações. Por exemplo, para a geração de desvios que podem fazer parte dos decks de entrada do modelo DESSEM. A inclusão destas incertezas pode ser complementada com a adoção de algumas métricas de risco para inferir o risco associado a um desvio pré-determinado da geração prevista. Outra aplicação é a geração de cenários para o estudo da estabilidade do sistema elétrico em casos de alta penetração da fonte eólica na matriz elétrica.

Para atingir o objetivo deste trabalho, foram empregados os dados de geração prevista, de geração observada e de desvios de geração com uma discretização horária para o período entre janeiro de 2020 e junho de 2022. A primeira e a segunda base de dados foram obtidas, respectivamente, a partir dos decks de entrada do modelo DESSEM e do histórico de geração verificada pelo ONS. Já a última foi obtida a partir da diferença entre a geração prevista e observada. Salienta-se que os registros de geração e de desvios são agrupados em 58 pontos de conexão do SIN em que foi identificado a presença de geração eólica.

A primeira etapa da metodologia consiste em efetuar uma análise das propriedades estatísticas dos registros da base de dados. Esta análise mostra a existência de periodicidade temporal entre as amostras de um mesmo ponto de conexão e uma correlação espacial entre os registros de diferentes pontos de conexão. Com o intuito de integrar ambas as propriedades em um único modelo, é proposta uma metodologia baseada no uso do modelo PAR e no uso de uma matriz de carga nos resíduos. O primeiro é responsável pela integração das correlações temporais no modelo, enquanto, o segundo, pela integração das correlações espaciais.

Durante a implementação da metodologia proposta, notou-se que não era possível utilizar a base de dados discretizada em 58 pontos de conexão visto que as matrizes de correlação espacial eram semidefinidas positivas. Neste contexto, foi necessário efetuar um reagrupamento dos pontos de conexão para eliminar as linearidades existentes. No final, obteve-se uma base de

dados composta de 28 pontos de conexão agrupados a partir da qual o modelo foi parametrizado.

Após a parametrização, o modelo é validado em duas etapas. Na primeira, emprega-se uma curva de previsão de geração extraída da base de dados de treinamento. Na segunda etapa, uma previsão de geração desconhecida pelo modelo é utilizada. O objetivo de efetuar esta validação em duas etapas é de verificar se o modelo consegue replicar as propriedades estatísticas observadas ao usar uma curva empregada durante o treinamento e ao emular o uso do modelo pelo usuário. Em cada etapa de validação, 500 cenários de desvios foram gerados. Os resultados mostraram que o modelo consegue, de forma razoável, replicar as propriedades estatísticas observadas na base de dados de treinamento. Esta conclusão é feita a partir da análise das métricas de média e dos desvios padrões das amostras de desvios sintéticas e comparando-as com as métricas referentes aos dados de desvios observados.

Analisando os resultados, percebe-se que o modelo proposto neste trabalho possui diversas limitações oriundas da base de dados empregada no treinamento e de algumas premissas tomadas durante a parametrização. A primeira se refere ao fato de que a base de dados de treinamento é apenas composta pelos registros de geração prevista e de desvios de 2020 e de 2021. Esta limitação se deve pelo fato de que não existem registros de geração eólica consistentes antes de 2020, o que impede o aumento da base de dados. A segunda limitação se refere à premissa de que a capacidade instalada dos parques eólicos é constante. Isso se traduz na geração de cenário de desvios incoerentes caso o parque eólico em questão tenha sua capacidade instalada aumentada de forma significativa.

Sugere-se para trabalhos futuros, aumentar, na medida do possível, a base de dados de treinamento com o intuito de obter um modelo mais representativo. Sugere-se igualmente treinar o modelo com registros de geração prevista e de desvios normalizados com a capacidade instalada de geração eólica de cada ponto de conexão. Esta normalização pode ser empregada para remover a premissa de que a capacidade instalada não varia no tempo visto que os cenários finais podem ser escalonados pela capacidade instalada mais recente. Para concluir, sugere-se investigar a influência da ordem do modelo PAR com o intuito de verificar mudanças no desempenho do modelo.

## REFERÊNCIAS

- ABEEólica | Eólica já é a segunda fonte da matriz elétrica brasileira com 15 GW de capacidade instalada.** , 2021. Disponível em: <<http://abeeolica.org.br/noticias/eolica-ja-e-a-segunda-fonte-da-matriz-eletrica-brasileira-com-15-gw-de-capacidade-instalada/>>. Acesso em: 8 dez. 2021
- ALEXIADIS, M. C.; DOKOPOULOS, P. S.; SAHSAMANOGLU, H. S. Wind speed and power forecasting based on spatial correlation models. **IEEE Transactions on Energy Conversion**, v. 14, n. 3, p. 836–842, set. 1999.
- An Introduction to Non-Stationary Processes.** Disponível em: <<https://www.investopedia.com/articles/trading/07/stationary.asp>>. Acesso em: 16 abr. 2022.
- BAZIONIS, I. K.; GEORGILAKIS, P. S. Review of Deterministic and Probabilistic Wind Power Forecasting: Models, Methods, and Future Research. **Electricity**, v. 2, n. 1, p. 13–47, mar. 2021a.
- BAZIONIS, I. K.; GEORGILAKIS, P. S. Review of Deterministic and Probabilistic Wind Power Forecasting: Models, Methods, and Future Research. **Electricity**, v. 2, n. 1, p. 13–47, mar. 2021b.
- BESSA, R. J. et al. Time Adaptive Conditional Kernel Density Estimation for Wind Power Forecasting. **IEEE Transactions on Sustainable Energy**, v. 3, n. 4, p. 660–669, out. 2012.
- BESSA, R. J.; MIRANDA, V.; GAMA, J. Entropy and Correntropy Against Minimum Square Error in Offline and Online Three-Day Ahead Wind Power Forecasting. **IEEE Transactions on Power Systems**, v. 24, n. 4, p. 1657–1666, nov. 2009.
- BLUDSZUWEIT, H.; DOMINGUEZ-NAVARRO, J. A.; LLOMBART, A. Statistical Analysis of Wind Power Forecast Error. **IEEE Transactions on Power Systems**, v. 23, n. 3, p. 983–991, ago. 2008.
- BREMNES, J. B. Probabilistic wind power forecasts using local quantile regression. **Wind Energy**, v. 7, n. 1, p. 47–54, 2004.
- BRUNINX, K.; DELARUE, E. A Statistical Description of the Error on Wind Power Forecasts for Probabilistic Reserve Sizing. **IEEE Transactions on Sustainable Energy**, v. 5, n. 3, p. 995–1002, jul. 2014.
- CADENAS, E.; JARAMILLO, O. A.; RIVERA, W. Analysis and forecasting of wind velocity in chetumal, quintana roo, using the single exponential smoothing method. **Renewable Energy**, v. 35, n. 5, p. 925–930, 1 maio 2010.
- CEPEL. **Modelo DESSEM - Programação Diária da Operação de Sistemas Hidrotérmicos com Representação Detalhada das Unidades Geradoras, Consideração da Rede Elétrica e Restrições de Segurança - Manual do Usuário.** Brasil: [s.n.].
- EFRON, B. Bootstrap Methods: Another Look at the Jackknife. Em: KOTZ, S.; JOHNSON, N. L. (Eds.). . **Breakthroughs in Statistics: Methodology and Distribution.** Springer Series in

Statistics. New York, NY: Springer, 1992. p. 569–593.

FORBES, K. F.; ZAMPELLI, E. **The Accuracy of Wind and Solar Energy Forecasts and the Prospects for Improvement**. . Em: MANNHEIM ENERGY CONFERENCE. Germany: maio 2016.

GNEITING, T.; RAFTERY, A. E. Weather Forecasting with Ensemble Methods. **Science**, v. 310, n. 5746, p. 248–249, 14 out. 2005.

GONZÁLEZ-APARICIO, I.; ZUCKER, A. Impact of wind power uncertainty forecasting on the market integration of wind energy in Spain. **Applied Energy**, v. 159, p. 334–349, 1 dez. 2015.

HAQUE, A. U.; NEHRIR, M. H.; MANDAL, P. A Hybrid Intelligent Model for Deterministic and Quantile Regression Approach for Probabilistic Wind Power Forecasting. **IEEE Transactions on Power Systems**, v. 29, n. 4, p. 1663–1672, jul. 2014.

HODGE, B. M. et al. **Wind Power Forecasting Error Distributions: An International Comparison**. . Em: TRANSMISSION NETWORKS FOR OFFSHORE WIND POWER PLANTS. Portugal: National Renewable Energy Lab. (NREL), Golden, CO (United States), 1 set. 2012. Disponível em: <<https://www.osti.gov/biblio/1051129>>. Acesso em: 22 nov. 2021

HODGE, B.-M.; MILLIGAN, M. **Wind power forecasting error distributions over multiple timescales**. 2011 IEEE Power and Energy Society General Meeting. **Anais...** Em: 2011 IEEE POWER AND ENERGY SOCIETY GENERAL MEETING. jul. 2011.

HOLTTINEN, H. Impact of hourly wind power variations on the system operation in the Nordic countries. **Wind Energy**, v. 8, p. 197–218, 1 abr. 2005.

HU, J.; TANG, J.; LIN, Y. A novel wind power probabilistic forecasting approach based on joint quantile regression and multi-objective optimization. **Renewable Energy**, v. 149, p. 141–164, 1 abr. 2020.

HUANG, C.-Y. et al. **Short Term Wind Speed Predictions by Using the Grey Prediction Model Based Forecast Method**. 2011 IEEE Green Technologies Conference (IEEE-Green). **Anais...** Em: 2011 IEEE GREEN TECHNOLOGIES CONFERENCE (IEEE-GREEN). abr. 2011.

JUBAN, J.; FUGON, L.; KARINIOTAKIS, G. Probabilistic short-term wind power forecasting based on kernel density estimators. p. 12, 2007.

JUBAN, J.; FUGON, L.; KARINIOTAKIS, G. **Uncertainty estimation of wind power forecasts: Comparison of Probabilistic Modelling Approaches**. European Wind Energy Conference & Exhibition EWEC 2008. **Anais...** Brussels, Belgium: EWEC, mar. 2008. Disponível em: <<https://hal-mines-paristech.archives-ouvertes.fr/hal-00506297>>. Acesso em: 18 nov. 2021

KARINIOTAKIS, G. N.; STAVRAKAKIS, G. S.; NOGARET, E. F. Wind power forecasting using advanced neural networks models. **IEEE Transactions on Energy Conversion**, v. 11, n. 4, p. 762–767, dez. 1996.

KAVOUSI-FARD, A.; KHOSRAVI, A.; NAHAVANDI, S. A New Fuzzy-Based Combined Prediction Interval for Wind Power Forecasting. **IEEE Transactions on Power Systems**, v. 31, n. 1, p. 18–26, jan. 2016.

KHAN, A. A.; SHAHIDEHPOUR, M. **One day ahead wind speed forecasting using wavelets**. 2009 IEEE/PES Power Systems Conference and Exposition. **Anais...** Em: 2009 IEEE/PES POWER SYSTEMS CONFERENCE AND EXPOSITION. mar. 2009.

KHORRAMDEL, B. et al. A Fuzzy Adaptive Probabilistic Wind Power Prediction Framework Using Diffusion Kernel Density Estimators. **IEEE Transactions on Power Systems**, v. 33, n. 6, p. 7109–7121, nov. 2018.

KHOSRAVI, A.; NAHAVANDI, S. Combined Nonparametric Prediction Intervals for Wind Power Generation. **IEEE Transactions on Sustainable Energy**, v. 4, n. 4, p. 849–856, out. 2013.

K.W. HIPEL, A. I. M. **Time Series Modelling of Water Resources and Environmental Systems**. [s.l.] Elsevier Science B.V, 1994.

LANDBERG, L. Short-term prediction of the power production from wind farms. **Journal of Wind Engineering and Industrial Aerodynamics**, v. 80, n. 1, p. 207–220, 1 mar. 1999.

LANGE, M. On the Uncertainty of Wind Power Predictions—Analysis of the Forecast Accuracy and Statistical Distribution of Errors. **Journal of Solar Energy Engineering-transactions of The Asme - J SOL ENERGY ENG**, v. 127, 1 maio 2005.

LARROYD, P. V. **Um modelo de otimização estocástica com representação individualizada das usinas hidrelétricas no planejamento de médio prazo da operação hidrotérmica**, 2016. Disponível em: <<https://repositorio.ufsc.br/xmlui/handle/123456789/175312>>

LI, G.; SHI, J. On comparing three artificial neural networks for wind speed forecasting. **Applied Energy**, v. 87, n. 7, p. 2313–2320, 1 jul. 2010.

MAKAROV, Y. V. et al. Operational Impacts of Wind Generation on California Power Systems. **IEEE Transactions on Power Systems**, v. 24, n. 2, p. 1039–1050, maio 2009.

MATOS, M. A.; BESSA, R. J. Setting the Operating Reserve Using Probabilistic Wind Power Forecasts. **IEEE Transactions on Power Systems**, v. 26, n. 2, p. 594–603, maio 2011.

MILLIGAN, M. et al. **Operating reserves and wind power integration: An international comparison**. . Em: ELECTRIC POWER SYSTEMS RESEARCH - ELEC POWER SYST RES. 18 out. 2010.

MOHANDDES, M. A. et al. Support vector machines for wind speed prediction. **Renewable Energy**, v. 29, n. 6, p. 939–947, maio 2004.

MONFARED, M.; RASTEGAR, H.; KOJABADI, H. M. A new strategy for wind speed forecasting using artificial intelligent methods. **Renewable Energy**, v. 34, n. 3, p. 845–848, 1 mar. 2009.

MORI, H.; KURATA, E. **Application of Gaussian Process to wind speed forecasting for wind power generation**. 2008 IEEE International Conference on Sustainable Energy Technologies. *Anais...* Em: 2008 IEEE INTERNATIONAL CONFERENCE ON SUSTAINABLE ENERGY TECHNOLOGIES. nov. 2008.

NIELSEN, H. A.; MADSEN, H.; NIELSEN, T. S. Using quantile regression to extend an existing wind power forecasting system with probabilistic forecasts. *Wind Energy*, v. 9, n. 1–2, p. 95–108, 2006.

NOAA. **Climate Time Series: Seasonal Variability**, [s.d.].

NYCANDER, E.; SÖDER, L. **An Open Model for Generating High Resolution Wind Power Production Scenarios**. 2021 IEEE Madrid PowerTech. *Anais...* Em: 2021 IEEE MADRID POWERTECH. jun. 2021.

ONS. **Programação diária da operação - Decks de entrada e saída - Modelo DESSEM**. Disponível em: <[www.sintegre.ons.org.br](http://www.sintegre.ons.org.br)>. Acesso em: 16 fev. 2022a.

ONS. **Evolução da capacidade instalada do SIN - Abril 2022/ Dezembro 2026**. Disponível em: <<http://ons.org.br:80/paginas/sobre-o-sin/o-sistema-em-numeros>>. Acesso em: 8 abr. 2022b.

ONS. **Mapas da Rede ONS**. Disponível em: <<http://ons.org.br:80/paginas/sobre-o-sin/mapas>>. Acesso em: 5 jul. 2022c.

ONS - SINTEGRE. **NT 0151-2018 - Desenvolvimento Metodológico para Previsão de Geração de Fonte Eólica**. Disponível em: <[https://sintegre.ons.org.br/sites/6/27/48/\\_layouts/15/WopiFrame.aspx?sourcedoc={B6857DDF-CD23-43D4-9D27-E015608D6523}&file=NT%200151-2018%20-%20Desenvolvimento%20Metodol%C3%B3gico%20para%20Previs%C3%A3o%20de%20Gera%C3%A7%C3%A3o%20de%20Fonte%20E%C3%B3lica.pdf&action=default](https://sintegre.ons.org.br/sites/6/27/48/_layouts/15/WopiFrame.aspx?sourcedoc={B6857DDF-CD23-43D4-9D27-E015608D6523}&file=NT%200151-2018%20-%20Desenvolvimento%20Metodol%C3%B3gico%20para%20Previs%C3%A3o%20de%20Gera%C3%A7%C3%A3o%20de%20Fonte%20E%C3%B3lica.pdf&action=default)>. Acesso em: 7 dez. 2021.

OUR WORLD IN DATA. **Wind power generation**. Disponível em: <<https://ourworldindata.org/grapher/wind-generation>>. Acesso em: 23 jun. 2022a.

OUR WORLD IN DATA. **Share of electricity production from wind**. Disponível em: <<https://ourworldindata.org/grapher/share-electricity-wind>>. Acesso em: 23 jun. 2022b.

PINSON, P. et al. **Dynamic sizing of energy storage for hedging wind power forecast uncertainty**. 2009 IEEE Power Energy Society General Meeting. *Anais...* Em: 2009 IEEE POWER ENERGY SOCIETY GENERAL MEETING. jul. 2009.

PINSON, P. Very-short-term probabilistic forecasting of wind power with generalized logit-normal distributions. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, v. 61, n. 4, p. 555–576, 2012.

PINSON, P.; MADSEN, H. Ensemble-based probabilistic forecasting at Horns Rev. *Wind Energy*, v. 12, n. 2, p. 137–155, 2009.

QUAN, H.; SRINIVASAN, D.; KHOSRAVI, A. Short-Term Load and Wind Power Forecasting Using Neural Network-Based Prediction Intervals. **IEEE Transactions on Neural Networks and Learning Systems**, v. 25, n. 2, p. 303–315, fev. 2014.

RAMIREZ-ROSADO, I. J. et al. Comparison of two new short-term wind-power forecasting systems. **Renewable Energy**, v. 34, n. 7, p. 1848–1854, 1 jul. 2009.

REN, Y.; SUGANTHAN, P. N.; SRIKANTH, N. Ensemble methods for wind and solar power forecasting—A state-of-the-art review. **Renewable and Sustainable Energy Reviews**, v. 50, p. 82–91, 1 out. 2015.

RIAHY, G. H.; ABEDI, M. Short term wind speed forecasting for wind turbine applications using linear prediction method. **Renewable Energy**, v. 33, n. 1, p. 35–41, 1 jan. 2008.

SALCEDO-SANZ, S. et al. Hybridizing the fifth generation mesoscale model with artificial neural networks for short-term wind speed prediction. **Renewable Energy**, v. 34, n. 6, p. 1451–1457, 1 jun. 2009.

SANG, Y.; ZHENG, Y. Reserve Scheduling in the Congested Transmission Network Considering Wind Energy Forecast Errors. **2020 52nd North American Power Symposium (NAPS)**, 2021.

SANTOS, T. N. et al. Hourly pricing and day-ahead dispatch setting in Brazil: The dessem model. **Electric Power Systems Research**, v. 189, p. 106709, 1 dez. 2020.

SFETSOS, A. A novel approach for the forecasting of mean hourly wind speed time series. **Renewable Energy**, v. 27, n. 2, p. 163–174, 1 out. 2002.

SHI, J.; GUO, J.; ZHENG, S. Evaluation of hybrid forecasting approaches for wind speed and power generation time series. **Renewable and Sustainable Energy Reviews**, v. 16, n. 5, p. 3471–3480, 1 jun. 2012.

SLOUGHTER, J. M.; GNEITING, T.; RAFTERY, A. E. Probabilistic Wind Speed Forecasting Using Ensembles and Bayesian Model Averaging. **Journal of the American Statistical Association**, v. 105, n. 489, p. 25–35, 1 mar. 2010.

SOMAN, S. S. et al. A review of wind power and wind speed forecasting methods with different time horizons. **North American Power Symposium 2010**, 2010.

SUN, M.; FENG, C.; ZHANG, J. Multi-distribution ensemble probabilistic wind power forecasting. **Renewable Energy**, v. 148, p. 135–149, 1 abr. 2020.

WAN, C. et al. Direct Quantile Regression for Nonparametric Probabilistic Forecasting of Wind Power Generation. **IEEE Transactions on Power Systems**, v. 32, n. 4, p. 2767–2778, jul. 2017.

WU, Y.-K. et al. Probabilistic Wind-Power Forecasting Using Weather Ensemble Models. **IEEE Transactions on Industry Applications**, v. 54, n. 6, p. 5609–5620, nov. 2018.

XIA, J.; ZHAO, P.; DAI, Y. Neuro-fuzzy networks for short-term wind power forecasting. **In Proceedings of the International Conference on Power System Technology**, 24 out. 2010.

YANG, M.; YANG, C. **Uncertainty analysis of wind power prediction based on Granular Computing**. 2016 International Conference on Probabilistic Methods Applied to Power Systems (PMAPS). **Anais...** Em: 2016 INTERNATIONAL CONFERENCE ON PROBABILISTIC METHODS APPLIED TO POWER SYSTEMS (PMAPS). out. 2016.

ZHANG, N. et al. Modeling Conditional Forecast Error for Wind Power in Generation Scheduling. **IEEE Transactions on Power Systems**, v. 29, n. 3, p. 1316–1324, maio 2014.

ZHANG, Y.; WANG, J.; LUO, X. Probabilistic wind power forecasting based on logarithmic transformation and boundary kernel. **Energy Conversion and Management**, v. 96, p. 440–451, 15 maio 2015.

## APÊNDICE A - Modelos determinísticos e probabilísticos de previsão de geração eólica

Tabela A.1 - Comparação dos modelos determinísticos de previsão de ventos e geração eólica

<b>Métodos de Previsão</b>	<b>Classe</b>	<b>Exemplos</b>	<b>Horizonte de Previsão</b>	<b>Referencias</b>
Método de persistência ( <i>Naive Method</i> )	-	$P(t+\Delta t) = P(t)$	Curto e curtíssimo prazo	-
Métodos Físicos ( <i>Physical Approach</i> )	<i>Numeric Weather Predictors</i> (NWP)	<i>Global Forecasting System</i>	Longo Prazo	-
		MM5		(SALCEDO-SANZ et al., 2009)
		Prediktor		-
		HIRLAM		(LANDBERG, 1999)
Métodos Estatísticos ( <i>Statistical Approach</i> )	<i>Artificial Neural Networks</i> (ANN)	<i>Feed-Forward</i>	Curto e curtíssimo prazo	(LI; SHI, 2010)
		<i>Radial Basis Function</i>		
		<i>ADALINE</i>		
		<i>Recurrent</i>		(KARINIOTAKIS; STAVRAKAKIS; NOGARET, 1996)
		<i>Multilayer Perceptron</i>		(MORI; KURATA, 2008)
		SVM		(MOHANDES et al., 2004)
	<i>Time series Models</i>	ARMA	Curto e curtíssimo prazo	-
		ARMAX (ARX)		-
		ARIMA		(SFETSOS, 2002)
		<i>Grey Predictor</i>		(HUANG et al., 2011)
		<i>Linear Predictors</i>		(RIAHY; ABEDI, 2008)
		<i>Exponential Smoothing</i>		(CADENAS; JARAMILLO; RIVERA, 2010)
Outro Métodos	-	<i>Spatial Correlation</i>	Curto prazo	-
		<i>Fuzzy Logic</i>		(MONFARED; RASTEGAR; KOJABADI, 2009)
		<i>Wavelet Transform</i>	Curto à médio prazo	(KHAN; SHAHIDEHPOUR, 2009)
		<i>Entropy Based Training</i>	Médio à longo prazo	(BESSA; MIRANDA; GAMA, 2009)
Métodos Híbridos	-	NWP + ANN	Médio à longo prazo	(RAMIREZ-ROSADO et al., 2009)

<i>(Hybrid Methods)</i>		<i>ANN + Fuzzy Logic</i>	Curtíssimo prazo	(XIA; ZHAO; DAI, 2010)
		<i>Spatial Correlation + ANN</i>	Curto prazo	(ALEXIADIS; DOKOPOULOS; SAHSAMANOGLU, 1999)
		<i>NWP + Time-series</i>	-	(SHI; GUO; ZHENG, 2012)

Fonte: Adaptado de BAZIONIS; GEORGILAKIS (2021) e SOMAN et al. (2010).

Tabela A.2 - Comparação dos modelos probabilísticos de previsão de ventos e geração eólica

<b>Classe de métodos de previsão</b>	<b>Exemplos</b>	<b>Horizonte de previsão</b>	<b>Referências</b>
Métodos Paramétricos ( <i>Parametric Approach</i> )	Modelos com formas de distribuições de probabilidade determinadas previamente	Curtíssimo à Longo Prazo (depende da distribuição utilizada)	(HODGE et al., 2012; PINSON, 2012)
Métodos Não Paramétricos ( <i>Non-Parametric Approach</i> )	Regressão Quantílica - QR	Curto Prazo	(BREMNES, 2004; HAQUE; NEHRIR; MANDAL, 2014; HU; TANG; LIN, 2020; JUBAN; FUGON; KARINIOTAKIS, 2008; NIELSEN; MADSEN; NIELSEN, 2006; WAN et al., 2017)
	Estimação de Densidade de Kernel – KDE	Depende da função Kernel utilizada	(BESSA et al., 2012; JUBAN; FUGON; KARINIOTAKIS, 2007; KHORRAMDEL et al., 2018; ZHANG; WANG; LUO, 2015)
	<i>Ensemble Methods</i>	Depende do método	(GNEITING; RAFTERY, 2005; PINSON; MADSEN, 2009; REN; SUGANTHAN; SRIKANTH, 2015; SLOUGHTER; GNEITING; RAFTERY, 2010; SUN; FENG; ZHANG, 2020)
	<i>Lower Upper Bound Estimation - LUBE</i>	-	(KAVOUSI-FARD; KHOSRAVI; NAHAVANDI, 2016; KHOSRAVI; NAHAVANDI, 2013; QUAN; SRINIVASAN; KHOSRAVI, 2014; WU et al., 2018)
	<i>Bootstrap</i>	-	(EFRON, 1992; KHOSRAVI; NAHAVANDI, 2013)

Fonte: BAZIONIS; GEORGILAKIS (2021a)

## APÊNDICE B - Requisitos para parametrização do modelo PAR

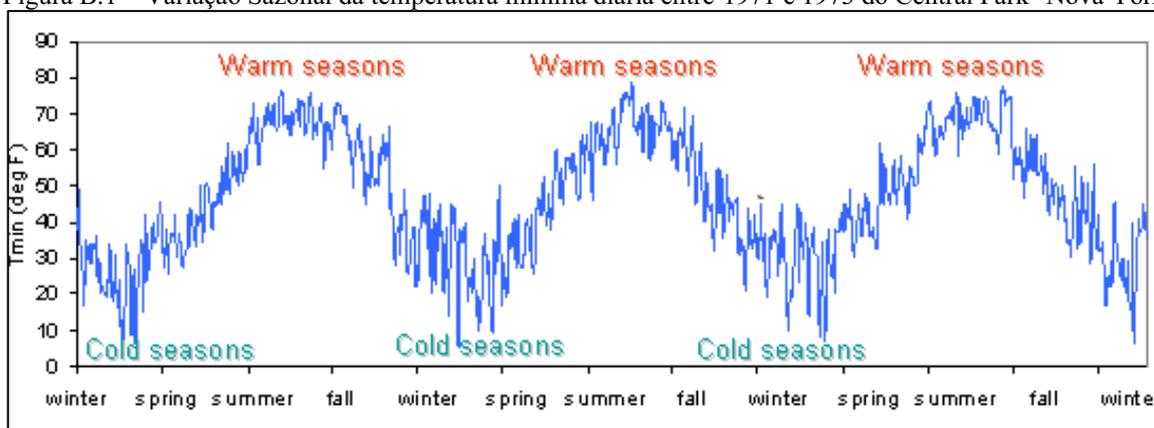
### B.1 INTRODUÇÃO

De forma geral, a qualidade da parametrização de modelos estatísticos é fortemente influenciada pela qualidade e pelas propriedades estatísticas dos dados. Por exemplo, o modelo de regressão linear possui como hipótese fundamental que os dados possuem uma distribuição normal. Caso essa hipótese seja violada, o modelo de regressão linear pode ser considerado inválido visto que os dados não são adequadamente representados. No contexto do modelo PAR, algumas hipóteses são feitas a respeito dos dados empregados no processo de aprendizagem do modelo. Estas hipóteses são: sazonalidade, estacionariedade e normalidade.

#### B.1.1 SAZONALIDADE

Uma série temporal é considerada sazonal caso as suas grandezas variam de acordo com a estação (ano, mês, estações do ano) que compõe os períodos da série temporal. Esta característica está de certa forma relacionada à periodicidade dos dados. Um exemplo de séries temporais com sazonalidade são os registros de temperatura (Figura B.1).

Figura B.1 – Variação Sazonal da temperatura mínima diária entre 1971 e 1973 do Central Park- Nova York



Fonte: (NOAA, [s.d.])

#### B.1.1 NORMALIDADE

A hipótese de normalidade das séries temporais se refere à distribuição das amostras. Caso a distribuição se aproxime de uma distribuição normal, a hipótese é respeitada. Esta hipótese é fundamental na modelagem do modelo PAR visto que a sua violação pode conduzir a um modelo incapaz de reproduzir as propriedades estatísticas da série temporal empregada em sua parametrização (K.W. HIPEL, 1994). Uma das maneiras de atender a hipótese de normalidade

é transformar os dados por meio da transformação de Box-Cox, cuja definição é descrita a seguir.

**Definição:** Seja  $x_t$  uma observação da série temporal e  $c$  e  $\lambda$  duas constantes, a transformação de Box-Cox de  $x_t$  é obtida por meio da equação (B.1):

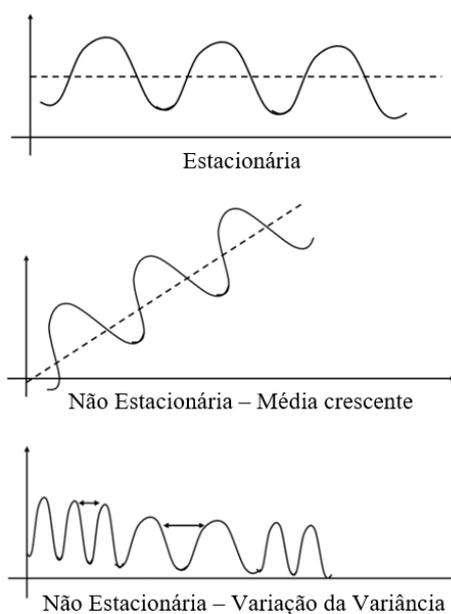
$$x_t = \begin{cases} \frac{(x_t + c)^\lambda - 1}{\lambda}, & \lambda \neq 0 \\ \ln(x_t + c), & \lambda = 0 \end{cases} \quad (\text{B.1})$$

Ressalta-se que a transformação é apenas válida para dados positivos. Conseqüentemente, caso os valores série temporal de aprendizado sejam positivos, a constante  $c$  é nula. No entanto, caso existam valores negativos na série temporal, a constante  $c$  é usualmente igual a um valor ligeiramente maior que o maior valor não positivo da série temporal (K.W. HIPEL, 1994).

### B.1.2 ESTACIONARIEDADE

Uma série temporal é dita estacionária caso as suas propriedades estatísticas (média e variância) não são uma função do tempo. Isto é, independentemente do período analisado, a média e a variância são invariáveis. A Figura B.2 mostram exemplos de série temporal estacionária e não estacionária. Salienta-se que para verificar a estacionariedade de uma série temporal é necessário calcular a média e a variância móvel.

Figura B.2 – Exemplo de séries temporais estacionárias e não estacionárias



Fonte: (“An Introduction to Non-Stationary Processes”, [s.d.]

Adicionalmente, a hipótese de estacionariedade pode ser violada dependendo do intervalo de tempo em análise. Neste sentido, a série temporal pode ser considerada estacionária para um dado período, mas pode ser não estacionária caso a janela de tempo for maior. Por exemplo, o consumo de água de uma cidade tende a ser estacionária para um período de análise de algumas semanas. No entanto, ela pode ser não estacionária se o período analisado for da ordem de anos visto que o consumo pode ser influenciado pelo aumento populacional e/ou industrialização da cidade.

Para estacionarizar uma série temporal periódica, é necessário estabilizar a variância e/ou a média. Esse resultado pode ser obtido por meio de duas transformações: (i) transformação de Box-Cox; (ii) diferenciação. A primeira transformação é descrita na seção B.1.1. Já a segunda é descrita a seguir.

**Definição:** Sejam  $x_t$  uma observação da série temporal no tempo  $t$ , e  $x_{t-s}$  no tempo  $t$  da estação  $s$ , e  $\eta$  o número de estações, a diferenciação por estação pode ser obtida por meio da equação (B.2).

$$x_{t+1} = x_t - x_{t-s} \quad \forall t = s + 1, s + 2, \dots, \eta \quad (\text{B.2})$$

Note que o número de observações da série temporal é decrescido de  $\eta$  unidades após a diferenciação visto que não é possível aplicá-la para as primeiras  $\eta$  observações devido à inexistência de uma observação em  $x_{s-1}$ . No contexto deste trabalho, as estações se referem às horas. Por exemplo, considerando dois dias consecutivos, as amostras estacionarizadas do segundo dia é dado pela seguinte equação (as primeiras 24 horas são perdidas) :

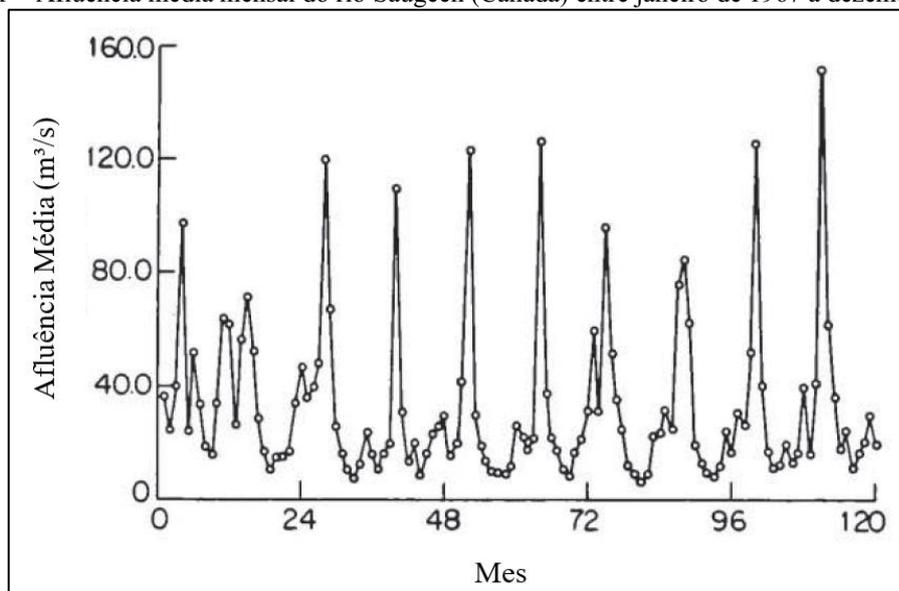
$$x_{d+1,h} = x_{d+1,h} - x_{d,h} \quad \forall h = 1, 2, \dots, 24$$

## APÊNDICE C - Modelo Periódico Autorregressivo (PAR)

### C.1 INTRODUÇÃO

Os modelos de geração de séries temporais considerados periódicos são empregados quando a série temporal de aprendizado do modelo exibe uma autocorrelação cuja estrutura mostra uma dependência não apenas entre as amostras consecutivas no tempo, mas também, entre amostras da mesma estação<sup>31</sup> (K.W. HIPEL, 1994). Um exemplo de série temporal com tal comportamento pode ser visualizado na Figura C.1, a qual mostra a afluência mensal média do rio Saugeen (Canadá) durante um período de 10 anos. Uma das características que pode ser verificada é o comportamento periódico evidenciado pela similaridade das afluências médias em função dos meses do ano.

Figura C.1 – Afluência média mensal do rio Saugeen (Canadá) entre janeiro de 1967 à dezembro de 1976



Fonte: Adaptado de (K.W. HIPEL, 1994).

Existem diversos modelos capazes de modelar séries temporais periódicas. No entanto, os modelos mais conhecidos são o modelo autorregressivo periódico (PAR) e o modelo autorregressivo de médias móveis (PARMA). No contexto deste trabalho, o modelo PAR é empregado devido à sua simplicidade, maior popularidade e pelo fato de ter uma base teórica mais desenvolvida (K.W. HIPEL, 1994).

<sup>31</sup> Estação pode ser um período na escala de ano, estação do ano, mês, dia, hora etc.

Na sequência deste apêndice será detalhado a teoria do modelo PAR, assim como o método de parametrização empregado. Vale mencionar que a teoria mostrada neste apêndice pode ser encontrada com mais detalhes em (K.W. HIPEL, 1994).

### C.1.1 MODELO AUTORREGRESSIVO PERIÓDICO

O modelo autorregressivo periódico busca capturar e replicar as características físicas de um fenômeno descrito a partir de uma série temporal. A sua modelagem parte do princípio de que um modelo autorregressivo (AR) deve ser parametrizado, de forma independente, para cada estação com periodicidade e agrega-los para, enfim, obter o modelo periódico completo (K.W. HIPEL, 1994).

Como evidenciado na seção 3.3, as séries temporais mensais de geração eólica e dos desvios de previsão indicam um comportamento periódico horário. Desta forma, para respeitar este comportamento, um modelo autorregressivo é parametrizado para cada uma das 24 horas do dia. De modo geral, a equação do modelo PAR pode ser escrita da seguinte forma:

$$x_{d,h} - \mu x_h = \sum_{p=1}^{np^{(h)}} \phi_p^h \cdot (x_{d,h-p} - \mu x_{h-p}) + \eta_{d,h}(\omega) \quad (C.1)$$

$$\eta_{d,h}(\omega) = \sigma \eta_h^r \cdot \omega$$

No contexto deste trabalho, um modelo PAR multivariável é proposto (equação (C.2)). Este modelo visa gerar séries sintéticas de desvios de geração tendo como base a geração prevista da base de dados. Para parametrizá-lo, é necessário extrair as informações relativas às dependências temporais entre os desvios observados e as previsões de geração. Esta dependência é expressada pela variável  $\phi_p^{(h)}$ , também conhecida como coeficiente autorregressivo.

$$D_{d,h} - \mu D_h = \sum_{p=0}^{np^{(h)}} \phi_p^h \cdot (P_{d,h-p} - \mu P_{h-p}) + \eta_{d,h}(\omega) \quad (C.2)$$

$$\eta_{d,h}(\omega) = \sigma \eta_h^r \cdot \omega$$

Salienta-se que as variáveis em verde são aquelas cujos valores são conhecidos e são obtidos diretamente da base de dados de aprendizado. Por outro lado, as variáveis em vermelho representam as variáveis cujos valores precisam ser determinados por meio cálculo da função

de correlação periódica (PECF). Outro ponto que deve ser ressaltado é que o método de parametrização do modelo PAR multivariável empregado neste trabalho se baseia na metodologia do modelo PAR convencional representado pela equação (C.1). O processo de parametrização do modelo PAR multivariável é detalhado a seguir.

### C.1.1.1 Função de Correlação Periódica

Antes de descrever o processo de parametrização do modelo PAR multivariável, é imprescindível definir a função de correlação periódica. Esta função é basicamente uma correlação cruzada de um sinal com um segundo sinal como função de um deslocamento no tempo denominado *lag*. Caso o segundo sinal for o mesmo do primeiro, a função se denomina função de autocorrelação periódica (PEACF).

Um dos objetivos do uso da PEACF e da PECF é a extração de propriedade estatísticas tais como a correlação temporal das amostras da base de dados sendo analisada. No caso deste, estas funções são utilizadas no processo de cálculo dos coeficientes autorregressivos ( $\phi^{(h)}$ ) e do desvio padrão dos resíduos  $\sigma\eta_h$  para cada hora  $h$ .

A PECF e a PEACF podem ser obtidas a partir da normalização da função de covariância periódica (PCVF) pela variância das amostras. No contexto deste trabalho, a PECF pode ser definida conforme a equação (C.4). Note que caso as variáveis  $x_{d,h}$  e  $y_{d,h}$  forem a mesma série temporal, a PECF se denomina PEACF. Outro ponto que deve ser ressaltado é que os valores  $\rho_k^{(h)}$  são, em módulo, limitados à 1.

**Definição:** Seja  $x_{d,h}$  e  $y_{d,h}$  duas variáveis aleatórias discretas, estacionárias e com distribuição normal, a função de correlação periódica para a hora  $h$  e para o *lag*  $k$  é definida pela equação a seguir:

$$\rho_k^h = \frac{\gamma_k^h}{\sqrt{\gamma_0^h \cdot \gamma_0^{h-k}}} \quad (C.3)$$

Com :  $\gamma_{h-k}^h = E[(x_{d,h} - \mu x_h) \cdot (y_{d,h-k} - \mu y_{h-k})]$

### C.1.1.1.1 Uso da PECF na parametrização do modelo PAR

Como dito anteriormente, o processo de parametrização do modelo PAR multivariável se baseia no processo do modelo PAR convencional. Primeiramente, multiplica-se a equação (C.2) pelo termo  $(D_{d,h-k} - \mu D_{h-k})$  e toma-se o valor esperado. Assim, obtêm-se a equação (C.4).

$$E[(D_{d,h} - \mu x_h) \cdot (D_{d,h-k} - \mu D_{h-k})] = \sum_{p=0}^{np^{(h)}} \phi_p^h \cdot E[(P_{d,h-p} - \mu P_{h-p}) \cdot (D_{d,h-k} - \mu D_{h-k})] + E[\eta_{d,h}(\omega) \cdot (D_{d,h-k} - \mu D_{h-k})] \quad (C.4)$$

Substituindo a definição da PCVF na equação (C.4), obtêm-se (C.5). Os termos  $\gamma D_{h-k}^h$  e  $\gamma PD_{h-k}^{h-p}$  representam, respectivamente, a autocorrelação dos desvios e a correlação entre os desvios e a geração prevista para os devidos deslocamentos no tempo.

$$\gamma D_{h-k}^h = \sum_{p=0}^{np^{(h)}} \phi_p^h \cdot \gamma PD_{h-k}^{h-p} + E[\eta_{d,h}(\omega) \cdot (D_{d,h-k} - \mu D_{h-k})] \quad (C.5)$$

Desenvolvendo a equação (C.5) para  $k = 0$ , é possível calcular o desvio padrão dos resíduos, conforme mostrado pela equação abaixo:

$$\gamma D_{h-0}^h = \sum_{p=0}^{np^{(h)}} \phi_p^h \cdot \gamma PD_{h-0}^{h-p} + E[\eta_{d,h}(\omega) \cdot (D_{d,h-0} - \mu D_{h-0})] \quad \sigma \eta_h^{r^2}$$

O termo  $E[\eta_{d,h}(\omega) \cdot (D_{d,h-0} - \mu D_{h-0})]$  é igual à variância dos resíduos do modelo ( $\sigma \eta_h^{r^2}$ ) (K.W. HIPEL, 1994). Assim, reescrevendo a equação acima, obtêm-se uma equação por meio da qual é possível obter o desvio padrão dos resíduos para uma dada hora  $h$ . Percebe-se a partir da equação (C.6) que ainda não é possível calcular o desvio padrão dos resíduos para a hora  $h$  visto que a ordem do modelo ( $np^{(h)}$ ) ainda é desconhecido.

$$\sigma \eta_h^r = \left( \gamma D_{h-0}^h - \sum_{p=0}^{np^{(h)}} \phi_p^h \gamma PD_{h-0}^{h-p} \right)^{1/2} \quad (C.6)$$

O segundo passo do processo de parametrização do modelo PAR multivariável é multiplicar a equação (C.2) pelo termo  $(P_{d,h-k} - \mu P_{h-k})$  e tomar, igualmente o valor esperado. Neste contexto, a equação (C.7) é obtida.

$$\begin{aligned}
& E[(D_{d,h} - \mu x_h) \cdot (P_{d,h-k} - \mu P_{h-k})] = \\
& \sum_{p=0}^{np^{(h)}} \phi_p^h \cdot E[(P_{d,h-p} - \mu P_{h-p}) \cdot (P_{d,h-k} - \mu P_{h-k})] + \\
& E[\eta_{d,h}(\omega) \cdot (P_{d,h-k} - \mu P_{h-k})]
\end{aligned} \tag{C.7}$$

Repetindo o mesmo processo para o caso anterior, ao substituir a definição da PCVF na equação (C.4), obtêm-se (C.8).

$$\gamma P D_{h-k}^{(h)} = \sum_{p=0}^{np^{(h)}} \phi_p^h \cdot \gamma P_{h-k}^{h-p} + E[\eta_{d,h}(\omega) \cdot (P_{d,h-k} - \mu P_{h-k})] \tag{C.8}$$

Desenvolvendo a equação (C.8) para  $k$  positivos, é possível obter uma formulação para calcular os coeficientes autorregressivos, como mostrado a seguir:

$$\gamma D_{h-k}^h = \sum_{p=0}^{np^{(h)}} \phi_p^h \cdot \gamma D_{h-k}^{h-p} + E[\eta_{d,h}(\omega) \cdot (P_{d,h-k} - \mu P_{h-k})] \rightarrow 0$$

O termo  $E[\eta_{d,h}(\omega) \cdot (P_{d,h-k} - \mu P_{h-k})]$  é nulo para qualquer valor de  $k$  visto que o termo  $\eta_{d,h}(\omega)$  é independente da série temporal de previsão de geração. Salienta-se que a independência dos ruídos (termo  $\eta_{d,h}(\omega)$ ) da série de geração prevista é uma hipótese do modelo (K.W. HIPEL, 1994). Neste contexto, obtêm-se a equação (C.9).

$$\gamma D P_{h-k}^h = \sum_{p=0}^{np^{(h)}} \phi_p^h \cdot \gamma P_{h-k}^{h-p} \tag{C.9}$$

A equação (C.9) pode ser reescrita em sua forma matricial, como mostrada abaixo. O sistema de equações evidenciado pela matriz é conhecido como equações de Yule-Walker do modelo. A obtenção dos valores dos coeficientes autorregressivos ( $\phi^h$ ) é feita após o conhecimento da ordem da matriz. Isto é, é necessário conhecer os valores da variável  $np^{(h)}$ .

$$\begin{bmatrix}
\gamma P_h^h & \gamma P_h^{h-1} & \cdots & \gamma P_h^{h-np^{(h)}} \\
\gamma P_{h-1}^h & \gamma P_{h-1}^{h-1} & \cdots & \gamma_{h-1}^{h-np^{(h)}} \\
\vdots & \vdots & \ddots & \vdots \\
\gamma P_{h-np^{(h)}}^h & \gamma_{h-np^{(h)}}^{h-1} & \cdots & \gamma P_{h-np^{(h)}}^{h-np^{(h)}}
\end{bmatrix}
\begin{bmatrix}
\phi_0^h \\
\phi_1^h \\
\vdots \\
\phi_{np^{(h)}}^h
\end{bmatrix}
=
\begin{bmatrix}
\gamma D P_{h-1}^h \\
\gamma D P_{h-2}^h \\
\vdots \\
\gamma D P_{h-np^{(h)}}^h
\end{bmatrix} \tag{C.10}$$

No contexto deste trabalho, a ordem do modelo  $e$ , conseqüentemente, a ordem da matriz da equação (C.10) será escolhida manualmente de forma a obter o melhor modelo possível. Ressalta-se que o processo de escolha da ordem do modelo é a principal diferença entre o modelo PAR convencional do modelo PAR multivariável apresentado neste trabalho. No caso do primeiro, a ordem do modelo pode ser identificada por meio do cálculo da função de correlação periódica parcial (K.W. HIPEL, 1994). No entanto, esse processo não é compatível com o modelo PAR multivariável.

### C.1.2 PSEUDOCÓDIGO PARA A PARAMETRIZAÇÃO DO MODELO PAR

De forma resumida, o pseudocódigo associado à parametrização do modelo PAR é mostrado pelo esquema a seguir:

Para cada hora  $h$ :

- 1) Cálculo da função de autocorrelação periódica dos desvios
- 2) Cálculo da função de autocorrelação periódica das previsões de geração
- 3) Cálculo da função de correlação periódica entre os desvios e as previsões de geração.
- 4) Cálculo dos coeficientes autorregressivos para uma dada ordem  $np^{(h)}$
- 5) Cálculo dos resíduos do modelo PAR multivariável

## APÊNDICE D - Integração da Correlação Espacial

### D.1 INTRODUÇÃO

A importância da integração da correlação espacial se deve ao fato de que o modelo de geração de cenários empregado neste trabalho, em sua forma original, pode se tornar susceptível a gerar cenários que não traduzem a realidade física.

Um dos métodos de integração da dependência espacial no modelo é por meio da ponderação dos ruídos brancos com a matriz de correlação espacial. Este apêndice busca descrever a metodologia empregada neste trabalho tendo como base a metodologia descrita em (LARROYD, 2016).

### D.2 CÁLCULO DOS RESÍDUOS CORRELACIONADOS

Seja  $\omega_h$  o vetor de ruídos brancos gerados de forma independente a partir de uma distribuição normal de média nula e variância unitária e  $RC_h$  uma matriz de carga triangular inferior, o vetor de ruídos brancos correlacionados  $\omega_h^c$  pode ser obtido por meio da equação (D. 1). Vale notar que os vetores  $\omega_h$  e  $\omega_h^c$  possuem um comprimento que corresponde ao número de pontos de conexão do problema ( $nr_{PC}$ ).

$$\begin{aligned}\omega_h^c &= RC_h \cdot \omega_h \rightarrow RC_h \cdot [\omega_h^1 \quad \omega_h^2 \quad \cdots \quad \omega_h^{nr_{PC}}] \\ \omega_h^c &= [\omega_{c,h}^1 \quad \omega_{c,h}^2 \quad \cdots \quad \omega_{c,h}^{nr_{PC}}]\end{aligned}\tag{D. 1}$$

Multiplicando o vetor  $\omega_h^c$  pelo seu transposto  $\omega_h^{cT}$  e tomando o valor esperado, obtêm-se a seguinte equação:

$$E[\omega_h^c \cdot \omega_h^{cT}] = E[(RC_h \cdot \omega_h) \cdot (RC_h \cdot \omega_h)^T] \rightarrow E[RC_h \cdot \omega_h \cdot \omega_h^T \cdot RC_h^T]\tag{D. 2}$$

A equação (D. 2) pode ser simplificada visto que o vetor  $\omega_h$  é obtido de forma independente a partir de uma distribuição normal de média nula e variância unitária. Assim, o termo  $E[\omega_h \cdot \omega_h^T]$  corresponde à matriz identidade quadrada de ordem  $nr_{PC}$  e, conseqüentemente, a equação (D. 2) pode ser reescrita da seguinte forma:

$$E[\omega_h^c \cdot \omega_h^{cT}] = RC_h \cdot RC_h^T\tag{D. 3}$$

Dado que  $\omega_h^c$  corresponde ao vetor de ruídos correlacionados, a operação  $E[\omega_h^c \cdot \omega_h^{cT}]$  corresponde à própria matriz de correlação para o período  $h$ ,  $\rho RC_h$ . Assim:

$$E[\omega_h^c \cdot \omega_h^{cT}] = RC_h \cdot RC_h^T = \rho RC_h \quad (\text{D. 4})$$

Analisando a equação (D. 4), percebe-se que a matriz de carga pode ser obtida a partir da matriz de correlação espacial da base de dados de aprendizado. Um dos métodos de obtenção da matriz de carga é a decomposição de Cholesky da matriz de correlações. Salienta-se que para a obtenção da matriz de carga via decomposição de Cholesky é necessário que a matriz de correlação espacial seja definida positiva.

### APÊNDICE E - Lista de Pontos de Conexão

Tabela E.1 Lista dos pontos de conexão da base de dados antes e depois do reagrupamento

Ponto de Conexão	Ponto de Conexão Agrupados
ACARAU II - 230 kV (A)	Grupo ACARAU SOBRAL PECEM
ACARAU II - 69 kV (A)	
PECEM II - 230 kV (A)	
SOBRAL III - 230 kV (A)	
ACU III500kVA	ACU III500kVA
ACU II - 230 kV (A)	Grupo ACU POLO GUAMARE
POLO GUAMARE - 138 kV (A)	
AQUIBATA - 138 kV (A)	AQUIBATA - 138 kV (A)
ATLANTIDA 2 - 69 kV (A)	Grupo OSORIO ATLANTIDA
OSORIO 2 - 230 kV (A)	
OSORIO 2 - 69 kV (A)	
BROT.MACAUBAS - 230 kV (A)	Grupo OURO IRECE MACAUBAS
IRECE - 230 kV (A)	
GENTIO OURO II230kVA	
OUROLANDIA II - 230 kV (A)	
BRUMADO II - 230 kV (A)	BRUMADO II - 230 kV (A)
CAMPO FORMOSO - 230 kV (A)	Grupo FORMOSO BONFIM
SR.BONFIM II - 230 kV (A)	
CURRAL PIAUI II - 500 kV (A)	Grupo CURRAL PICOS
PICOS - 230 kV (A)	
CEARA MIRIM 2500kVA	Grupo CEARA EXTREMOZ
EXTREMOZ II - 230 kV (A)	
GARANHUNS II - 230 kV (A)	GARANHUNS II - 230 kV (A)
IBIAPINA II - 230 kV (A)	Grupo IBIAPINA
IBIAPINA II - 69 kV (A)	
IGAPORA II - 230 kV (B)	Grupo IGAPORA PINDAI
IGAPORA II - 69 kV (A)	
PINDAI II - 230 kV (A)	
PINDAI II - 69 kV (A)	
JUAZEIRO II - 230 kV (A)	JUAZEIRO II - 230 kV (A)
J. CAMARA II - 230 kV (A)	Grupo J.CAMARA
J. CAMARA III138kVA	
J. CAMARA III230kVA	
JOAO CAMARA 269kVA	
LAGOA NOVA II - 69 kV (A)	Grupo LAGOA NOVA PARAISO
PARAISO138kVA	
LIVRAMENTO 2 - 230 kV (B)	LIVRAMENTO 2 - 230 kV (B)
MARMELEIRO 2 - 525 kV (A)	Grupo PALMAR MARMELEIRO

S V PALMAR 2 - 525 kV (A)	
MIRANDA II500kVA	MIRANDA II500kVA
MORRO CHAPEU II500kVA	Grupo MORRO CHAPEU
MORRO CHAPEU2 - 230 kV (A)	
MORRO CHAPEU2 - 69 kV (A)	
MOSSORO II - 230 kV (A)	Grupo MOSSORO RUSSAS
MOSSORO IV - 230 kV (A)	
RUSSAS II - 230 kV (A)	
QUINTA - 138 kV (A)	Grupo QUINTA
QUINTA - 69 kV (A)	
RIO DO OURO138kVA	RIO DO OURO138kVA
S.JOAO PIAUI - 500 kV (A)	Grupo S.JOAO PIAUI
S.JOAO PIAUI230kVA	
SANTA LUZIA II138kVA	SANTA LUZIA II138kVA
TACARATU - 230 kV (A)	TACARATU - 230 kV (A)
TOUROS - 230 kV (A)	Grupo TOUROS
TOUROS - 69 kV (A)	
U.SOBRADINHO - 230 kV (A)	Grupo U.SOBRADINHO
U.SOBRADINHO - 500 kV (A)	
VIAMAO 3 - 230 kV (A)	VIAMAO 3 - 230 kV (A)

Fonte: do autor.