



Aplicação da metodologia MCP a estimativas de produção anual de energia pós-operacional em parques eólicos

MARIANA MARQUES PASSOS ALVES DA COSTA

Setembro de 2023

APLICAÇÃO DA METODOLOGIA MCP A
ESTIMATIVAS DE PRODUÇÃO ANUAL DE ENERGIA
PÓS-OPERACIONAL EM PARQUES EÓLICOS

Mariana Marques Passos Alves da Costa
1180797

Setembro de 2023

Instituto Superior de Engenharia do Porto
Departamento de Engenharia Química



APLICAÇÃO DA METODOLOGIA MCP A ESTIMATIVAS DE PRODUÇÃO ANUAL DE ENERGIA PÓS-OPERACIONAL EM PARQUES EÓLICOS

Mariana Marques Passos Alves da Costa
1180797

Dissertação Submetida ao Instituto Superior de Engenharia do Porto como requisito parcial para a obtenção do grau de Mestre em Engenharia Química, área de especialização Energia e Biorrefinaria, realizada sob a orientação do Doutor Jorge Manuel Pires Mendonça

Setembro de 2023

Instituto Superior de Engenharia do Porto
Departamento de Engenharia Química

JÚRI

Presidente

<Grau Académico e Nome>

<Categoria, Instituição>

Orientador

Doutor Jorge Manuel Pires Mendonça

Professor Adjunto do Departamento de Matemática, ISEP

Coorientador

Doutora Rosa Maria Barbosa Rodrigues Pilão

Professor Adjunto do Departamento de Engenharia Química, ISEP

Orientador

Engenheiro Augusto Paulo Ferreira Pinto

CEO, MEGAJOULE S.A

Arguente

<Grau Académico e Nome>

<Categoria, Instituição>

AGRADECIMENTOS

Cada palavra escrita neste momento tenta expressar a gratidão que sinto. Esta tese seria apenas um sonho passageiro se não fossem os laços de amizade, os sorrisos de incentivo e a generosidade daqueles que estiveram ao meu lado, fortalecendo-me a cada passo desta jornada.

Agradeço ao Engenheiro Paulo Pinto, ao Doutor Jorge Mendonça e à Doutora Rosa Pilão pelo apoio ao longo destes meses. A vossa orientação e acompanhamento próximo foram essenciais para a conclusão desta tese. Agradeço por partilharem os vossos conhecimentos e experiências e por dedicarem tempo para me ajudar a enfrentar todos os desafios.

Um agradecimento à minha família e amigos: vocês foram a minha força nos momentos difíceis e a minha alegria nos momentos de comemoração. Obrigada por estarem ao meu lado, por me incentivarem e por compreenderem a importância deste projeto na minha vida.

A ti, Miguel, o meu sincero agradecimento pelo apoio inabalável em todos os momentos desta jornada. A tua presença constante, o teu incentivo e encorajamento foram essenciais para eu superar as adversidades e seguir em frente com determinação. És uma inspiração para mim, e sou grata por ter alguém tão especial como tu ao meu lado.

Tomás, não existem palavras suficientes para expressar a minha profunda gratidão. Tu foste muito além de um orientador e de um amigo; tornaste-te um pilar fundamental para a minha motivação e sucesso nesta tese. A tua presença constante, as tuas palavras de encorajamento e o teu compromisso em nunca me deixar desistir foram absolutamente inestimáveis. Foram inúmeros os momentos em que enfrentei desafios, dúvidas e momentos de desânimo, mas tu nunca deixaste que perdesse o foco ou a esperança. Acreditaste no meu potencial mesmo quando eu própria duvidava, e por isso sou-te eternamente grata. O teu conhecimento e disponibilidade para ajudar em todos os momentos foram determinantes para o sucesso desta tese. Esta conquista é tua também.

Com gratidão sincera,

Mariana

RESUMO

A implementação de um parque eólico envolve uma série de desafios, incluindo a avaliação da sua viabilidade económica e a obtenção de financiamento. Portanto, a análise do recurso eólico e da estimativa de energia associada desempenha um papel crucial. No entanto, a estimativa da produção anual de energia elétrica pré-construtiva enfrenta várias fontes de incerteza, que podem ser atribuídas a medições realizadas ou a pressupostos dos modelos utilizados.

O objetivo deste estudo consiste em identificar uma metodologia simples para estimar a produção anual de energia elétrica de Longo Termo pós construção. Isso foi realizado através da análise da correlação entre diversas fontes históricas de dados de vento e as produções mensais reais observadas em parques eólicos em operação. Além disso, este estudo investigou como o período de operação do parque eólico afeta a estimativa de produção de energia a Longo Termo em projetos com maior tempo de operação. Também avaliou o impacto das diferentes fontes históricas de dados de vento na estimativa de produção de energia de Longo Termo. Em parques eólicos que fazem parte de um complexo eólico, o método foi aplicado ao complexo e a estimativa de produção anual de energia foi comparada com a estimativa agregada da produção individual dos parques eólicos. O estudo desenvolveu-se em três fases distintas: a análise e validação da metodologia adotada, a análise da influência do tempo de correlação em parques com 20 e 4 anos de operação e o estudo da influência da agregação espacial em complexos eólicos.

A metodologia utilizada foi analisada através de um estudo da relação entre o coeficiente de correlação (R^2) e o desvio da produção de energia de Longo Termo em relação ao valor médio de produção real do parque, indicando que o R^2 é um indicador confiável para a seleção adequada da série de reanálise a ser utilizada no cálculo da estimativa de produção de energia. Para além disso, realizou-se uma avaliação do cumprimento dos pressupostos da regressão linear e os resultados revelaram que, de uma forma geral para todos os parques e complexos, o modelo de regressão linear empregado é estatisticamente significativo. Além disso, os resíduos gerados pelo modelo apresentam uma distribuição normal, e na maioria dos casos, não foi identificada auto correlação nos resíduos.

Ao analisar os parques com 20 anos de operação, observou-se uma tendência de redução gradual da dispersão dos dados à medida que o número de anos aumenta. Isso sugere que, para obter estimativas de Longo Termo mais precisas e próximas da realidade, é recomendável utilizar o maior número possível de anos de dados operacionais. Em relação à mediana, não foram identificadas diferenças significativas entre os grupos, pois as combinações resultaram em valores medianos semelhantes. As métricas de erro, como o RRMBE e RRMSE, seguiram o comportamento esperado, apresentando uma diminuição na dispersão à medida que o número de anos de operação aumenta. Isso sugere que as previsões estão, em média, mais próximas dos valores reais, com menos viés ao longo dos anos, indicando uma maior consistência nas estimativas. Quanto aos parques com 4 anos de operação, os resultados apontam para diferenças e nuances entre os grupos, principalmente em relação à dispersão dos resultados, no entanto, essas diferenças não se mostraram estatisticamente significativas. No contexto dos complexos eólicos, a análise revelou que não há uma diferença significativa entre a utilização da série de reanálise mais adequada para o complexo e a melhor para cada parque individualmente. Para além disso, parece não haver uma vantagem evidente na realização de estudos individuais para cada parque relativamente ao estudo para o complexo, visto que a diferença entre essas abordagens é insignificante em comparação com a incerteza típica do modelo e da própria estimativa de produção.

PALAVRAS CHAVE

Energia eólica; Estimativa de produção de energia de Longo Termo; Análise de dados; Testes estatísticos; Refinanciamento de projetos eólicos

ABSTRACT

Implementing a wind farm involves several challenges, including assessing its economic viability and obtaining financing. Therefore, the analysis of the wind resource and associated energy plays a crucial role. However, the estimation of pre-construction annual electricity production faces several sources of uncertainty, which can be attributed to measurements taken or to assumptions in the models used.

The aim of this study is to identify a simple methodology for estimating annual long-term electricity production. This was done by analyzing the correlation between various historical sources of wind data and the actual monthly productions observed in operating wind farms. In addition, this study investigated how the period of operation of the wind farm affects the estimate of Long-Term energy production in projects with a longer period of operation. The impact of different historical sources of wind data on the estimate of Long-Term Energy Production will also be assessed. In wind farms that are part of a wind complex, the method will be applied to the complex and the estimate of annual energy production will be compared with the aggregate estimate of individual wind farm production. The study was carried out in three distinct phases: validation of the methodology adopted, analysis of the influence of correlation time on wind farms with 20 and 4 years of operation, and study of the influence of spatial aggregation on wind complexes.

The methodology used was analyzed through a study of the relationship between the correlation coefficient (R^2) and the deviation of long-term energy production from the park's actual average production value, indicating that the R^2 is a reliable indicator for the appropriate selection of the reanalysis series to be used in calculating the energy production estimate. In addition, an assessment of compliance with the linear regression assumptions was carried out and the results revealed that, for all parks and complexes, the linear regression model employed is statistically significant. In addition, the residuals generated by the model have a normal distribution, and in most cases, no autocorrelation was identified in the residuals.

When analyzing parks with 20 years of operation, there was a tendency for data dispersion to gradually decrease as the number of years increased. This suggests that to obtain

long-term estimates that are more accurate and closer to reality, it is advisable to use as many years of operational data as possible. Regarding the median, no significant differences were identified between the groups, as the combinations resulted in similar median values. The error metrics, such as RRMBE and RRMSE, followed the expected behavior, showing a decrease in dispersion as the number of years of operation increased. This suggests that the forecasts are, on average, closer to the real values, with less bias over the years, indicating greater consistency in the estimates. As for the parks with 4 years of operation, the results point to differences and nuances between the groups, mainly in relation to the dispersion of the results. However, these differences were not statistically significant. In the context of wind complexes, the analysis revealed that there is no significant difference between using the most appropriate reanalysis series for the complex and the best one for each individual farm. Therefore, there does not seem to be a clear advantage in carrying out individual studies for each park, since the difference between these approaches is insignificant compared to the typical uncertainty of the model and the production estimate.

KEYWORDS

Wind energy; Long-term energy production estimation; Data analysis; Statistical tests; Refinancing of wind projects

ÍNDICE

Resumo.....	i
Palavras-chave.....	ii
Abstract.....	iii
Keywords.....	iv
Índice de Figuras.....	vii
Índice de Tabelas.....	viii
Lista de símbolos e abreviaturas.....	viii
1. INTRODUÇÃO.....	1
1.1. DESENVOLVIMENTO SUSTENTÁVEL E NEUTRALIDADE CARBÓNICA.....	1
1.2. ENERGIAS RENOVÁVEIS.....	3
1.3. ENERGIA EÓLICA.....	4
1.4. MOTIVAÇÃO E OBJETIVOS.....	5
1.5. A EMPRESA.....	6
1.6. DESCRIÇÃO GERAL DA TESE.....	7
2. REVISÃO BIBLIOGRÁFICA.....	8
2.1. NATUREZA DO RECURSO EÓLICO E VARIABILIDADE.....	8
2.2. CORRELAÇÃO DE LONGO TERMO.....	10
2.3. ESTIMATIVA DE PRODUÇÃO DE ENERGIA PÓS CONSTRUÇÃO.....	12
2.3.1. Método da curva de potência histórica.....	13
2.3.2. Método de índice.....	16
2.3.3. Método da regressão linear.....	19
3. METODOLOGIA.....	21
3.1. BASE DE DADOS.....	21
3.2. METODOLOGIA GERAL.....	23
3.2.1. Análise e alterações à Metodologia geral.....	25
3.2.2. Validação da Metodologia geral adotada.....	26
3.3. ANÁLISE ESTATÍSTICA DOS RESULTADOS.....	28
3.3.1. Definição das métricas de erro.....	28
3.3.2. Testes de hipótese e análises post-hoc.....	29
3.4. ANÁLISE DA INFLUÊNCIA DO TEMPO DE CORRELAÇÃO.....	32
3.5. ANÁLISE DA INFLUÊNCIA DA AGREGAÇÃO ESPACIAL EM COMPLEXOS EÓLICOS... 33	33
4. RESULTADOS.....	35
4.1. VALIDAÇÃO DA METODOLOGIA UTILIZADA.....	35
4.2. ANÁLISE DA INFLUÊNCIA DO TEMPO DE CORRELAÇÃO.....	38
4.3. ANÁLISE DA INFLUÊNCIA DA AGREGAÇÃO ESPACIAL EM COMPLEXOS EÓLICOS... 47	47
5. CONCLUSÕES E SUGESTÕES DE TRABALHOS FUTUROS.....	50
5.1. CONCLUSÕES.....	50

5.2. SUGESTÃO DE TRABALHOS FUTUROS	51
BIBLIOGRAFIA.....	52
ANEXO A – VALIDAÇÃO DA METODOLOGIA.....	56
ANEXO B – ANÁLISE DA INFLUÊNCIA DO TEMPO DE CORRELAÇÃO.....	61
B.1. PARQUES COM 20 ANOS DE DADOS OPERACIONAIS.....	61
B.2. PARQUES COM 4 ANOS DE DADOS OPERACIONAIS.....	77
ANEXO C – COMPLEXOS EÓLICOS.....	88

ÍNDICE DE FIGURAS

Figura 1.1. Distribuição das emissões de GEE, por setor, em Portugal [10].....	2
Figura 1.2. Evolução da potência instalada em Portugal do ano 2002 a 2022 [13].....	3
Figura 1.3. Perspetiva de evolução da produção de energia por tecnologia - hídrica, eólica e solar [14].....	4
Figura 2.1. Representação esquemática da circulação geral da atmosfera terrestre [23]	8
Figura 2.2. Efeitos térmicos que influenciam o recurso eólico: brisas marítimas (esquerda) e brisas terrestres (direita) [24].....	9
Figura 2.3. Aumento da velocidade do vento devido a efeitos orográficos [26].....	9
Figura 2.4. Demonstração esquemática da correlação realizada entre os dados de produção do parque eólico e os dados da fonte de dados históricos, no método MCP.....	11
Figura 2.5. Representação da curva de potência média para uma turbina, em três períodos distintos, para a primeira abordagem [31].....	14
Figura 2.6. Representação da curva de potência para um aerogerador, para oito setores, para a segunda abordagem [31].....	16
Figura 2.7. Relação entre a produção semanal de energia corrigida observada e o índice de produção semanal [30].....	18
Figura 2.8. Curva de potência típica de um aerogerador que relaciona a velocidade do vento e a potência elétrica [34].....	20
Figura 3.1. Localização geográfica dos parques e complexos eólicos selecionados.....	21
Figura 3.2. Representação esquemática da metodologia geral.....	23
Figura 3.3. Representação esquemática da metodologia adotada a nível de análise estatística.....	31
Figura 4.1. Relação entre o coeficiente de correlação (R^2) e o desvio percentual da estimativa de produção de Longo Termo em relação à média, para o parque P_P5, P_P6 e P_P7.....	37
Figura 4.2. Desvio da estimativa de produção de Longo Termo em relação à média real do parque, para diferentes períodos de correlação, de 1 a 20 anos, para o parque P_P5, P_P6 e P_P7.....	38
Figura 4.3. Desvios máximos para cada um dos grupos, de 1 a 20 anos, para todos os parques com 20 anos de dados operacionais em estudo (P_P5, P_P6 e P_P7).....	39
Figura 4.4. Representação gráfica do erro RRMBE, para todas as combinações realizadas de 1 a 20 anos, para os parques P_P5, P_P6 e P_P7.....	41

Figura 4.5. Representação gráfica do erro RRRMSE, para todas as combinações realizadas de 1 a 20 anos, para os parques P_P5, P_P6 e P_P7.....	41
Figura 4.6. Representação gráfica do erro RMBE (à esquerda) e RRMSE (à direita), para todas as combinações realizadas de 1 a 4 anos, para o parque P_E8.....	44
Figura 4.7. Representação gráfica do erro RMBE (à esquerda) e RRMSE (à direita), para todas as combinações realizadas de 1 a 4 anos, para o parque P_P2.....	45
Figura 4.8. Desvio entre o valor estimado de Longo Termo para o complexo obtido através dos dados dos parques individuais com a melhor série para o complexo ou com a melhor série para cada parque e o valor estimado com os dados do próprio complexo.....	47

ÍNDICE DE TABELAS

Tabela 3.1. Informação genérica relativa a cada parque eólico, o que inclui a designação utilizada, a potência instalada em MW e os anos de dados de operação.....	22
Tabela 3.2. Informação genérica relativa a cada complexo eólico, o que inclui a designação utilizada, a potência instalada em MW e os anos de dados de operação.....	23

LISTA DE SIMBOLOS E ABREVIATURAS

GEE - Gases de efeito de estufa
MCP - Método estatístico *Measure-Correlate-Predict*
PNEC 2030 - Plano Nacional Energia de Clima 2030
RRMSE - Raíz do Erro Quadrático Médio Relativo
RMBE - Erro Absoluto Médio Relativo
UE – União Europeia
WRF - *Weather Research and Forecast*

1. INTRODUÇÃO

1.1. Desenvolvimento Sustentável e Neutralidade Carbónica

Em 1987, foi publicado o Relatório Brundtland, o qual refere que a pobreza dos países menos desenvolvidos e o consumismo exagerado dos países de primeiro mundo são causas que impedem um desenvolvimento igualitário no Mundo, e que, conseqüentemente produzem graves crises ambientais. O conceito de Desenvolvimento Sustentável foi introduzido neste documento, sendo apresentado como “o desenvolvimento que procura satisfazer as necessidades da geração atual, sem comprometer a capacidade das gerações futuras de satisfazerem as suas próprias necessidades” [1].

A crescente preocupação com o impacto ambiental das atividades humanas, bem como a finitude dos recursos naturais, tem impulsionado o desenvolvimento sustentável como um objetivo essencial da sociedade. A partir da Conferência das Nações Unidas sobre Meio Ambiente e Desenvolvimento realizada no Rio de Janeiro em 1992, o desenvolvimento sustentável tornou-se um objetivo central para a maioria dos países do mundo [2].

O *Acordo de Paris*, um tratado internacional assinado por 195 países em 2015, apresentou uma estratégia para diminuir as emissões de gases de efeito de estufa (GEE), tendo como objetivo fundamental a neutralidade carbónica até 2050 [3] [4]. Isto significa que, a quantidade de emissões de GEE deve ser igual à quantidade de carbono retirado da atmosfera [5].

A nível europeu, foi criada uma diretiva denominada *European Green Deal*, na qual a Comissão Europeia adota um conjunto de propostas para tornar as políticas de clima, energia, transporte e tributação da União Europeia (UE) adequadas para reduzir as emissões de GEE em pelo menos 55% até 2030, em comparação com os níveis de 1990 [6]. Para além disso, foi criado também um pacote legislativo designado *Fit For 55*, que se trata de um conjunto de propostas destinadas a rever e a atualizar a legislação da UE, referindo-se à meta defendida pelo *European Green Deal* [7].

O *Plano Nacional Energia de Clima 2030* (PNEC 2030) é um instrumento de política nacional importante para a definição das estratégias a seguir na área da energia. O PNEC 2030

estabelece metas nacionais de redução de emissões de GEE (45 a 55% em relação a 2005), incorporação de energias renováveis e eficiência energética [8, 9].

O *Inventário Nacional de Emissões de Gases com Efeito de Estufa* contabiliza todas as emissões de origem humana, constituindo um instrumento que permite monitorizar e verificar o cumprimento nacional face às metas assumidas, sendo por isso um elemento-chave da política climática. O inventário é elaborado anualmente, sendo que o último inventário realizado em 2022 é relativo às emissões de 2020. A Figura 1.1 ilustra a distribuição das emissões de GEE, por setor, em Portugal. O setor energético, que inclui os transportes e a produção e transformação de energia, é o que mais contribui para as emissões de GEE em Portugal, representando 67,1% das emissões nacionais [10].

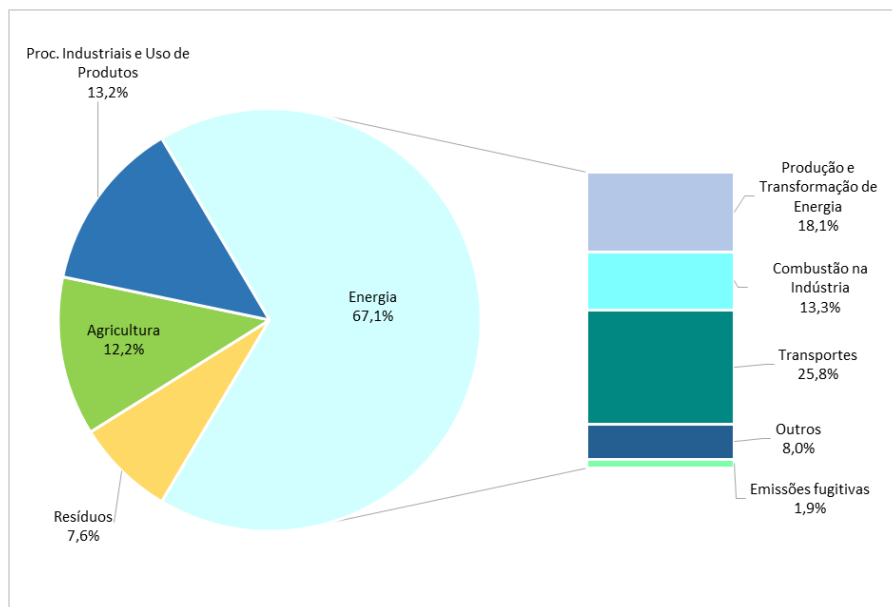


Figura 1.1. Distribuição das emissões de GEE, por setor, em Portugal [10].

As enormes quantidades de GEE na atmosfera, decorrentes da atividade humana, reforçam o efeito de estufa e o aquecimento global [11]. Para diminuir o aquecimento do planeta, é urgente alcançar a neutralidade carbónica, adotando comportamentos e padrões de produção e consumo de energia a favor da sustentabilidade. Analisando o impacto significativo que o setor da energia tem na quantidade de emissões nacionais de GEE, é possível perceber a urgência em adotar um modelo de produção e consumo de energia mais eficiente, que priorize o uso de fontes de energia renováveis [2, 10].

1.2. Energias renováveis

As fontes renováveis de energia são recursos naturais capazes de se regenerar num curto espaço de tempo e de um modo sustentável, ou seja, impactando o mínimo possível a Natureza. Estas fontes de energia contribuem de forma positiva a nível da sustentabilidade económica e ambiental dos países, uma vez que evitam a utilização de combustíveis fósseis, como o gás natural e o carvão, para gerar eletricidade e, conseqüentemente, evitam a emissão de GEE [12].

Tendo em vista o cumprimento do objetivo de atingir a neutralidade carbónica até 2050, a utilização de energias renováveis é um contributo muito importante e nesse sentido Portugal tem vindo a aumentar a utilização de recursos endógenos e renováveis. Na Figura 1.2 encontra-se esquematizada a evolução da potência instalada em Portugal, em MW, de 2002 a 2022. Através da análise da Figura 1.2 é possível aferir que desde 2011 a potência instalada de fontes fósseis tem vindo a diminuir, apresentando uma descida drástica em 2021 devido à desativação das centrais a carvão de Sines e do Pego [13]. Destaca-se também o aumento da potência instalada de fontes renováveis ao longo dos anos, nomeadamente hídrica, eólica, bioenergia e solar.

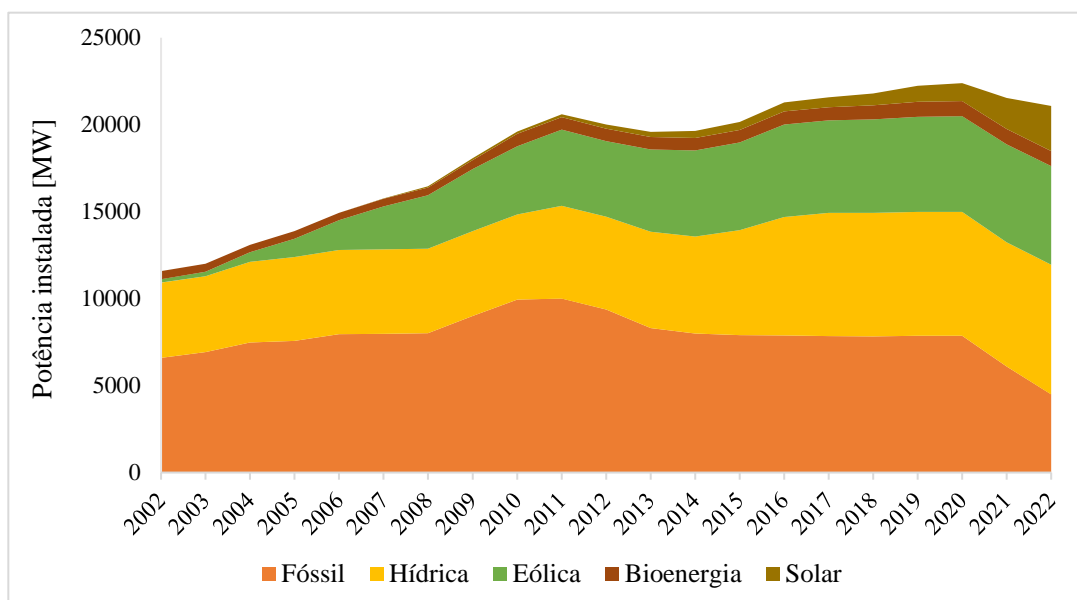


Figura 1.2. Evolução da potência instalada em Portugal, de 2002 a 2022 [13].

O PNEC 2030 demonstra um forte impulso à descarbonização da produção de energia, reforçando a utilização de fontes de energia renováveis, com ênfase nas tecnologias eólica e solar. Neste sentido, apresenta-se na Figura 1.3. a perspectiva de evolução da produção de energia elétrica por tecnologia, nomeadamente hídrica, eólica e solar [14]. O crescimento da potência hídrica instalada será devido à conclusão da construção do complexo hidroelétrico do Alto Tâmega que representa um aumento de 1,16 GW de potência instalada [14]. Relativamente à energia eólica, perspectiva-se que em 2030 estejam disponíveis 9,3 GW de potência instalada. A tecnologia solar fotovoltaica terá um aumento significativo até 2030, sendo a tecnologia para a qual se perspectiva um maior crescimento.

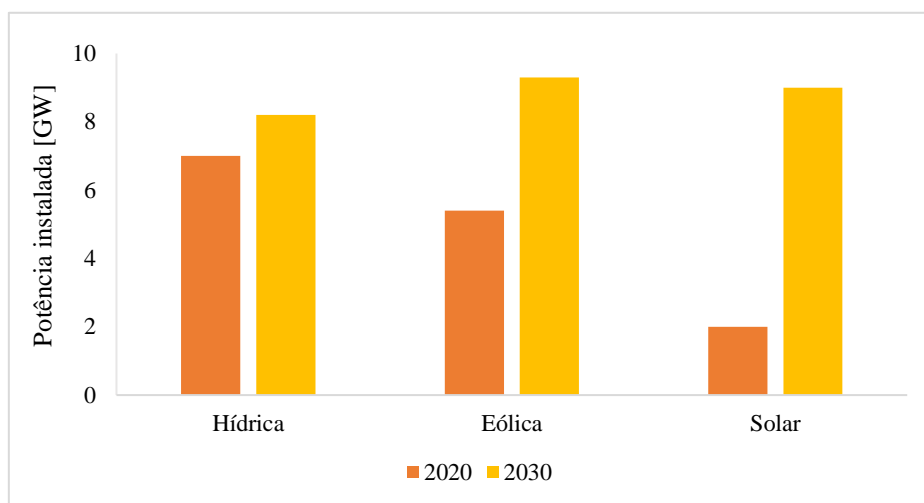


Figura 1.3. Perspetiva de evolução da produção de energia por tecnologia - hídrica, eólica e solar [14].

1.3. Energia eólica

A energia eólica é obtida a partir da conversão da energia cinética do vento em energia elétrica. A produção de energia elétrica através da energia do vento é realizada através de aerogeradores, que funcionam com o movimento das pás, capturando a energia do vento e a convertendo em energia mecânica. Essa energia é transferida para um gerador, que a transforma em energia elétrica e a envia para a rede elétrica [15].

A energia eólica tem demonstrado um crescimento acelerado durante a última década, sendo amplamente adotada por vários países como uma opção para diversificar recursos energéticos, reduzir emissões de dióxido de carbono, estimular novas indústrias e criar oportunidades de emprego [16].

A energia eólica é uma fonte de energia renovável tecnologicamente muito desenvolvida e com elevado potencial. Muitas das vantagens da energia eólica são comuns a outras renováveis, nomeadamente o combate às mudanças climáticas, com a redução da utilização de combustíveis fósseis e a emissão de GEE para além do facto de serem um auxílio para a autossuficiência energética, diminuindo as importações de recursos para produção de energia elétrica [17]. No caso particular da energia eólica, é possível ressaltar o facto de o vento ser uma fonte inesgotável de energia e que se encontra disponível em grande parte da superfície terrestre; a geração de energia através da energia eólica é descentralizada, levando à criação de emprego fora dos grandes centros e geração de investimento em zonas mais desfavorecidas, para além de que esta geração de energia descentralizada é menos vulnerável em termos geoestratégicos (em caso de guerra, por exemplo) [17] [18].

A fonte de energia eólica também apresenta desvantagens entre as quais a sua dependência das condições climáticas, nomeadamente a variabilidade e imprevisibilidade do recurso eólico e o facto de a construção, funcionamento e desmantelamento dos parques eólicos poderem causar impactos ambientais, até porque os aerogeradores apresentam peças com dificuldades ao nível da sua reciclagem, especialmente as pás (que são feitas de resinas reforçadas com fibra de vidro e, por isso, é particularmente difícil fazer a reciclagem) [19] [20]. É possível considerar como desvantagem a interferência na migração de aves, a poluição visual em paisagens naturais [16] [17] [18], além do facto de a geração distribuída de eletricidade causar maiores problemas de gestão da rede elétrica, dado que aumenta a sua complexidade e o número de pontos de injeção de energia.

1.4. Motivação e Objetivos

A aposta no desenvolvimento das tecnologias associadas às energias renováveis é cada vez mais importante e torna-se imprescindível para garantir os objetivos traçados para a neutralidade carbónica sendo que a energia eólica se reveste da maior importância neste caminho devido a todas as vantagens enumeradas anteriormente, mas principalmente pelo facto de ser uma tecnologia muito desenvolvida e com muito potencial [17] [18].

O desenvolvimento de um parque eólico inclui diversos desafios, entre os quais a avaliação da sua viabilidade económica e a obtenção de financiamento, pelo que o estudo de

caracterização do recurso eólico e de energia é uma tarefa fundamental. No entanto, a estimativa de produção anual de energia elétrica pré-construção apresenta muitas fontes de incerteza, que podem ser atribuídas às medições realizadas ou aos pressupostos dos modelos utilizados, por exemplo, fazendo com que o desvio padrão da produção anual de energia seja elevado.

Após a construção e quando o parque entra em operação normal, habitualmente após um ano de funcionamento, a reavaliação da estimativa futura da produção anual de energia elétrica, com dados operacionais, é um objetivo do proprietário do projeto uma vez que pode levar à eliminação de parte das incertezas de um estudo energético pré-construção, como as referidas anteriormente, e reveste-se da maior importância no caso de o proprietário pretender refinar o projeto ou caso haja a intenção de compra ou venda de um parque eólico já em operação.

A existência de um grande número de parques eólicos em funcionamento há vários anos permite a análise de dados operacionais, uma ferramenta fundamental para diminuir a incerteza global da estimativa de produção de energia.

Este trabalho teve como objetivo identificar uma metodologia simples e eficaz para estimar a produção anual de energia elétrica de Longo Termo, através da correlação entre diversas fontes históricas de dados de vento e produções mensais reais de parques eólicos em operação. Nos projetos com maior período de operação, foi avaliado em que medida o período de operação do parque eólico influencia a estimativa de produção de energia de Longo Termo. Foi também estudado de que forma as diferentes fontes históricas de dados de vento afetam a estimativa de produção de energia de Longo Termo. Em parques eólicos que integram um complexo eólico, o método foi aplicado ao complexo e comparada a estimativa de produção anual de energia com a estimativa agregada da produção individual dos parques eólicos.

1.5. A Empresa

O presente trabalho de tese foi conduzido sob a supervisão e colaboração da empresa *Megajoule*. Fundada em Portugal em 2004 por três engenheiros especializados em energia eólica e avaliação de recursos de vento, a *Megajoule* é uma empresa independente de consultoria e engenharia que oferece uma ampla gama de serviços. Com 19 anos de atividade, a empresa já operou mais de 400 mastros meteorológicos, realizou estudos de avaliação de recursos eólicos para mais de 50 GW de capacidade eólica, e participou em vários processos de

auditoria para financiamento ou negociação de projetos de parques eólicos. Além disso, a *Megajoule* também se dedica à avaliação de recursos solares, tendo realizado estudos e auditorias em cerca de 4 GW de projetos. Entre os serviços de engenharia oferecidos pela empresa, destacam-se o processo de desenvolvimento de centrais eólicas, análise de contratos, serviços de supervisão e gestão de projeto. Os clientes da *Megajoule* incluem importantes promotores de centrais eólicas, investidores e instituições de financiamento ativos nos principais mercados mundiais [21].

1.6. Descrição geral da tese

A presente tese encontra-se organizada da seguinte forma: inicia com um capítulo de Introdução no qual são apresentados o contexto geral e os principais objetivos do estudo, destacando a relevância do problema estudado. O segundo capítulo é dedicado à Revisão Bibliográfica no qual é realizada uma análise da literatura relevante para o tópico de pesquisa, sendo identificados os principais temas, teorias e modelos que fundamentam o estudo, servindo como base teórica sólida para a compreensão dos resultados apresentados posteriormente. O terceiro capítulo descreve a Metodologia adotada no decorrer do estudo, sendo detalhados os procedimentos de análise e as ferramentas utilizadas para estudar os principais objetivos assim como a justificação das escolhas metodológicas. No quarto capítulo, são apresentados os Resultados do estudo, sendo que a relação entre os resultados e os objetivos da pesquisa é discutida em detalhes. No quinto capítulo, as Conclusões e sugestões de trabalhos futuros são apresentadas e este capítulo é fundamental para consolidar as principais contribuições da pesquisa e responder aos objetivos propostos, bem como a apresentação das limitações do estudo e sugestões para trabalhos futuros que possam surgir a partir desta tese. Na Bibliografia são listadas todas as fontes de pesquisa citadas ao longo da tese e por fim, os Anexos, contêm os dados que complementam os resultados e conclusões presentes no corpo do relatório.

2. REVISÃO BIBLIOGRÁFICA

2.1. Natureza do recurso eólico e variabilidade

A disponibilidade do recurso eólico num local depende da circulação geral, do clima regional e dos efeitos locais, incluindo fatores térmicos e mecânicos [22].

De uma forma simplificada, a circulação geral da atmosfera terrestre pode ser explicada como um sistema de ventos que transportam o calor necessário entre latitudes tropicais e polares, sendo que sem este transporte os trópicos ficam cada vez mais quentes e os polos cada vez mais frios [23]. O vento circula em cada hemisfério em três células distintas que ajudam a transportar energia e calor do equador para os polos [23]. Na Figura 2.1 está uma representação esquemática da circulação geral da atmosfera terrestre.

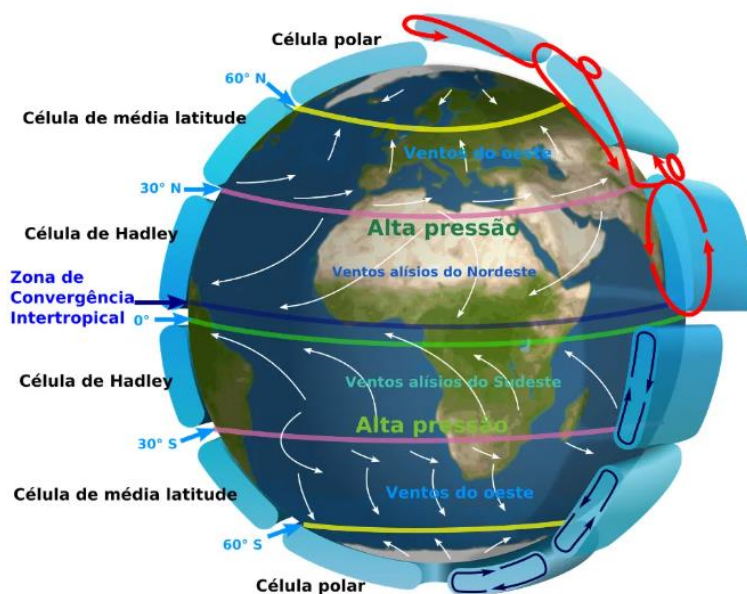


Figura 2.1. Representação esquemática da circulação geral da atmosfera terrestre [23].

Os efeitos térmicos estão relacionados com o aquecimento desigual da superfície da Terra. Por exemplo, nas regiões costeiras, como mostra a Figura 2.2, as superfícies terrestres têm uma menor capacidade de absorção de calor, aquecendo e arrefecendo mais rapidamente do que as superfícies aquáticas. Conseqüentemente, durante o dia, a superfície terrestre mais quente força o ar a subir e deixa uma zona de baixa pressão por baixo, criando ventos que se deslocam da água para a terra (brisas marítimas). À noite, ocorre o inverso: a superfície terrestre

está agora mais fria do que a água, forçando o ar a descer; as brisas terrestres afastam-se da superfície terrestre em direção à superfície da água [24].



Figura 2.2. Efeitos térmicos que influenciam o recurso eólico: brisas marítimas (esquerda) e brisas terrestres (direita) [24].

O recurso eólico é igualmente influenciado pela interação mecânica com a superfície da Terra, afetando a velocidade e a direção do vento. A Figura 2.3 mostra o efeito de um elemento orográfico moderado - uma colina - no perfil vertical do vento. A colina reduz a área da secção transversal do fluxo de ar, forçando o ar próximo do solo a acelerar à medida que sobe a colina. Depois de ultrapassar a colina, o ar desacelera novamente e o perfil vertical do vento a jusante é semelhante ao de montante [25].

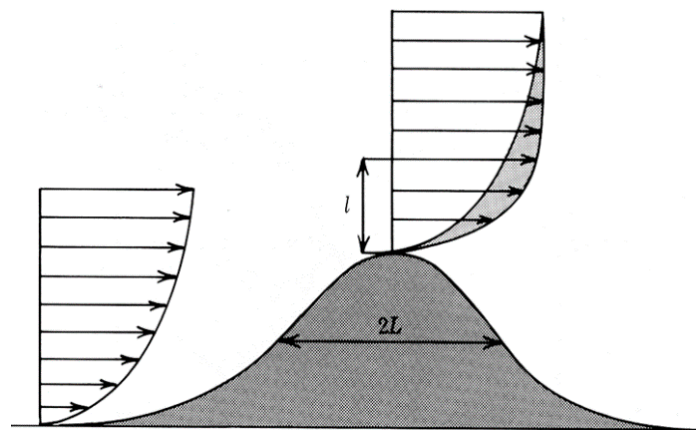


Figura 2.3. Aumento da velocidade do vento devido a efeitos orográficos [25].

A rugosidade da superfície terrestre também afeta o clima de vento local, uma vez que introduz fricção superficial e influencia o perfil vertical do vento. As alterações nas características da ocupação do solo distorcem o perfil vertical do vento, acelerando ou desacelerando o ar que se desloca junto ao solo. Além disso, os obstáculos também têm um impacto significativo no clima do vento, promovendo a recirculação e introduzindo turbulência no fluxo de vento [25].

Outro fator importante que afeta o aproveitamento da energia do vento é a densidade do ar, que varia com a pressão e temperatura e por conseguinte difere dos locais mais altos para o nível do mar, uma vez que existe relação linear entre a densidade do ar e a quantidade de energia que pode ser gerada a partir do vento. No entanto, as variações sazonais da densidade do ar e da velocidade do vento não coincidem necessariamente, o que pode afetar a precisão do cálculo da potência disponível em um determinado local [22].

A variabilidade temporal do recurso eólico é manifestada em diferentes escalas de tempo, sendo uma das mais relevantes, a variabilidade inter-anual. A medição do regime de vento ao longo de um curto período requer uma interpretação cautelosa: por um lado, quando as medidas sobrestimam a estimativa de produção de energia de Longo Termo, o regime de vento medido é otimista e pode levar a uma sobrestimação da produtividade e viabilidade económica de um projeto de parque eólico e por outro lado, uma subestimação do regime de vento de Longo Termo fornece uma visão mais conservadora, embora possa levar à rotulagem incorreta de um projeto de parque eólico como inviável [26].

Numa escala de tempo menor, o recurso eólico pode mostrar um forte padrão sazonal. Isso significa que, além das variações inter-anuais, é possível identificar um padrão de velocidade do vento ao longo do ano. A sazonalidade do vento é altamente dependente da localização e pode ser influenciada por fatores como a orografia e a proximidade do mar ou outras superfícies de água [26, 27].

2.2. Correlação de Longo Termo

Devido à variabilidade inter-anual da velocidade do vento, a existência de dados de produção de apenas alguns anos é insuficiente para espelhar o regime de ventos de um determinado local e daí a necessidade de realizar uma correlação de Longo Termo. Assim, para parques eólicos que tenham entrado em operação há pouco tempo (1 ou 2 anos), os dados de produção podem não ser suficientes para estimar com precisão a produção média de energia elétrica do parque eólico que operará por um longo período, tipicamente 20 ou 25 anos. Portanto, é possível subestimar ou sobrestimar a geração de energia com um grande erro ao basear as estimativas em um período curto de dados, devido à elevada variabilidade sazonal e inter-anual do vento, na sua velocidade e direção.

O método estatístico *Measure-Correlate-Predict (MCP)*, representado esquematicamente na Figura 2.4, pode ser utilizado para reduzir as incertezas de um projeto eólico, relacionando os dados de produção do parque existentes com uma base de dados de vento de Longo Termo próximo dos dados locais, chamados de fontes de dados históricos. Esta abordagem deriva uma correlação entre os dados de produção de energia do parque e a base de dados históricos para um período simultâneo e utiliza esta correlação para estimar a produção de energia de Longo Termo.

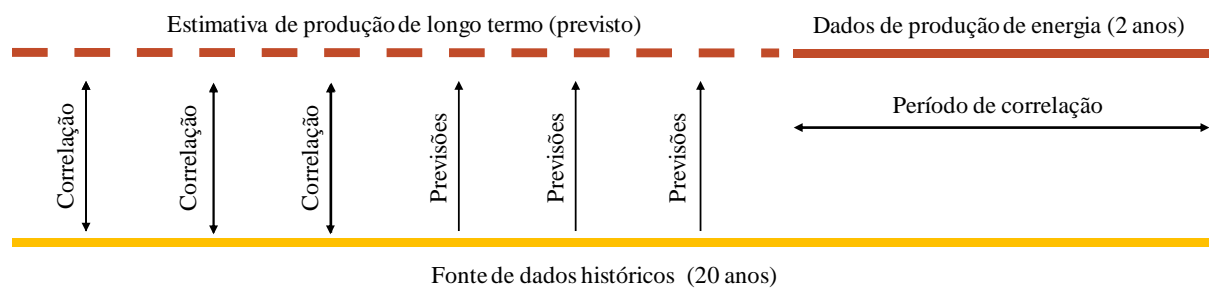


Figura 2.4. Demonstração esquemática da correlação realizada entre os dados de produção do parque eólico e os dados da fonte de dados históricos, no método MCP.

As fontes de dados históricas usadas no método MCP podem ser de vários tipos, sendo o mais comum a reanálise. As séries de reanálise combinam observações passadas de diferentes sistemas de medições (satélites, estações meteorológicas, radiossondas, entre outras) com modelos numéricos do comportamento da atmosfera para gerar séries temporais consistentes de várias variáveis climáticas, fornecendo uma descrição abrangente do clima observado e em que medida evoluiu nas últimas décadas [28]. As séries de reanálise mais utilizadas são as MERRA-2 e a ERA5, sendo que a principal distinção entre as mesmas reside na resolução espacial, em termos de graus, e na altura de medição das características do vento: a série ERA5 fornece uma grelha horizontal com resolução de $0,25^\circ$ por $0,25^\circ$ que na latitude da Europa corresponde a, aproximadamente, 30 quilómetros, medida a 100 metros de altura, enquanto a série MERRA-2 facultava a grelha com resolução aproximada de 50 quilómetros (correspondente a uma resolução de $0,5^\circ$ por $0,67^\circ$), avaliada a 50 metros, em relação ao nível do solo [29]. Para além das alturas de medição referidas anteriormente para a série ERA5, às quais é medida a velocidade do vento, existem outras alturas, nomeadamente 10 e 50 metros, que medem outras variáveis como a temperatura, pressão e humidade relativa, por exemplo.

2.3. Estimativa de produção de energia pós construção

A velocidade do vento é uma das principais variáveis que afeta a geração de energia eólica. Essa velocidade aumenta com a altura ao solo e é amplamente influenciada pela topografia e cobertura do solo local além de poder variar significativamente ao longo do tempo, tornando a sua medição essencial para determinar o potencial eólico de uma determinada localização. Devido à relação cúbica entre a velocidade do vento e a energia gerada, pequenas flutuações na velocidade do vento podem resultar em grandes variações na potência disponível. Como resultado, calcular a velocidade média anual do vento para estimar a energia elétrica gerada por via eólica pode ser um processo complexo [22].

A estimativa de produção de energia de um parque eólico pode ser baseada em dados de vento recolhidos no local, modelagem numérica do comportamento do vento e do efeito de esteira que as turbinas podem provocar e sofrer. Esta abordagem pode ser designada de estudo pré-construtivo, uma vez que é realizado anteriormente à construção do parque eólico e não são utilizados dados do parque em operação [30].

Por outro lado, numa fase posterior ao comissionamento de um parque eólico, a estimativa de produção de energia pode ser realizada tendo por base os dados reais do parque eólico, nomeadamente a produção de energia e disponibilidade dos aerogeradores, realizando uma projeção de Longo Termo, normalmente para 20 ou 25 anos, tendo em consideração pressupostos fundamentados das condições de vento e performance energética, que serão abordados adiante. Esta abordagem designa-se por estudo de produção de energia pós-construtivo [30].

Existem diversos métodos para estimar a produção de energia de Longo Termo de um parque em operação, entre os quais o método da curva de potência histórica, o método de índice e o método da regressão linear tendo em comum o facto de utilizarem um conjunto de dados de uma fonte de dados históricos de Longo Termo e a aplicação do método MCP para correlacionar os dados de medição reais com os dados históricos, nomeadamente as séries de reanálise [30].

Nos métodos descritos em seguida, é imperativo realizar uma filtragem aos dados medidos nos aerogeradores, de modo a excluir dados erróneos que possam afetar os cálculos da estimativa de produção de energia. Para períodos em que exista limitação da capacidade de

injeção na rede elétrica da energia produzida, situações em que determinadas grandezas medidas possam estar afetadas por condições climáticas mais extremas (com formação de gelo, por exemplo) e, de uma forma geral, todas as incidências que provoquem códigos de alarme nos aerogeradores, devem ser excluídas da análise [30].

2.3.1. Método da curva de potência histórica

O método da curva de potência histórica prende-se com a construção de uma curva de potência para o aerogerador ou para o parque, baseado em dados de períodos concorrentes de velocidade do vento e produção observada.

Uma possível formulação do método usa dados de velocidade de vento medidos em cada aerogerador, recolhidos no anemómetro instalado no próprio aerogerador. Para o cálculo de uma curva de potência histórica são necessários alguns meses de dados operacionais do aerogerador a funcionar em pleno desempenho, contudo, os períodos para os quais esta curva é calculada não devem ser muito longos de forma a levar em consideração possíveis mudanças no aerogerador. A curva de potência média é obtida através da determinação dos valores de potência média para diferentes valores de velocidade do vento, a cada 0,5 m/s, por exemplo. Este intervalo não deve ser muito alargado, uma vez que originaria uma caracterização muito grosseira da curva de potência, nem muito curto dado que qualquer desvio face ao comportamento normal do aerogerador teria muito peso na definição da curva de potência [30].

Um exemplo da variação da curva de potência média em diferentes períodos é apresentado na Figura 2.5. Diferenças observadas nas curvas de potência calculadas podem ser indicativas de alterações no aerogerador, devido à calibração do anemómetro ou a substituição do mesmo durante o período de dados analisado, ou mesmo variações nas características do aerogerador, decorrentes do desgaste [30].

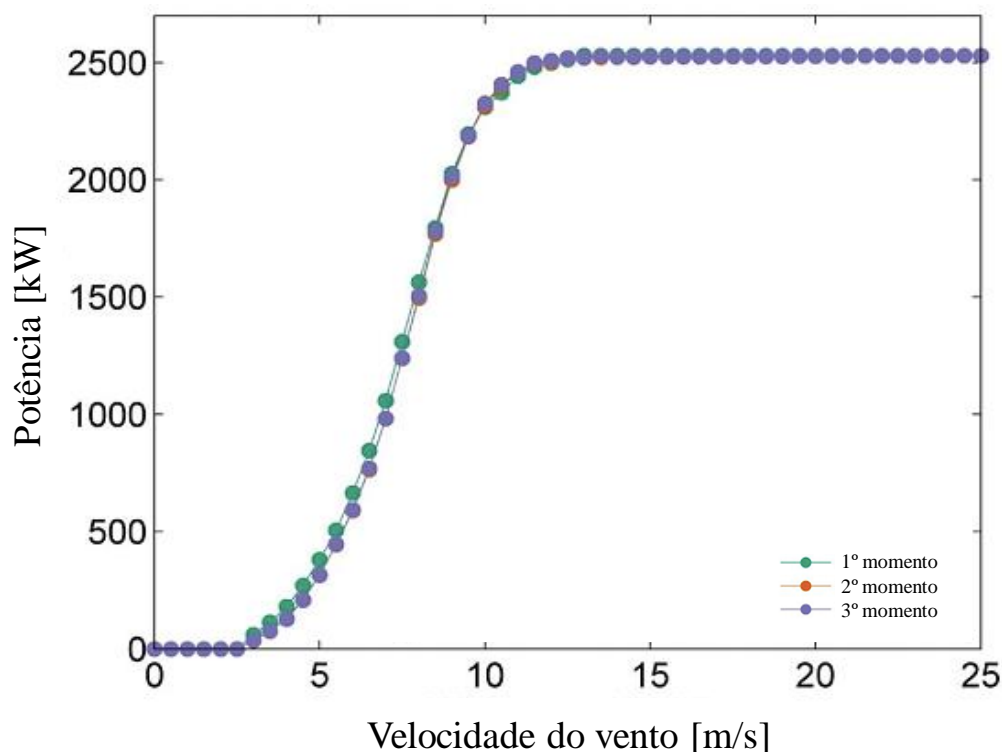


Figura 2.5. Representação da curva de potência média para uma turbina, em três períodos distintos, para a primeira abordagem [30].

Para cada um dos períodos para os quais as curvas de potência foram derivadas, são calculadas funções de transferência¹ que relacionam a velocidade do vento simulada por um modelo de circulação atmosférica (*WRF*²) com a velocidade do vento no aerogerador, setor a setor, quando o mesmo está a funcionar em pleno [30]. Aplicam-se estas funções de transferência à série de longo prazo da *WRF* e seguidamente relaciona-se a curva de potência histórica calculada com velocidade do vento *WRF* ajustada, obtendo-se uma série de produção ajustada de Longo Termo [30]. A produção anual de energia bruta após esteira, ou seja, afetada das influências dos restantes aerogeradores que constituem o parque eólico ou de parques vizinhos, é calculada como o valor médio anual da série de produção de energia de Longo Termo [30].

¹ Uma função de transferência é a representação matemática da relação entre a entrada e a saída de um sistema físico. Neste contexto, a função de transferência representa a relação entre a velocidade de vento simulada por um modelo de circulação atmosférica e a velocidade do vento medida no aerogerador, tendo como função obter valores de velocidade de vento ajustados, tendo em consideração as duas fontes de dados.

² O modelo de circulação atmosférica *WRF* (*Weather Research and Forecast*) é um sistema de modelagem numérica da atmosfera que tem como objetivo principal a investigação de fenómenos atmosféricos de micro e meso escala, bem como a realização de previsões meteorológicas para essas escalas [23].

O anemómetro instalado no aerogerador mede a velocidade do vento a montante do mesmo, fazendo com que a medição seja influenciada pelo efeito de esteira provocado pelo próprio aerogerador e por outros aerogeradores que se encontrem localizados nas proximidades, sendo uma das desvantagens desta abordagem. Outra desvantagem prende-se no facto de a curva de potência ser construída relacionando os dados de potência de cada um dos aerogeradores com uma série de dados de vento do *WRF*, fazendo com que a modelação do clima de ventos local seja realizada com menor detalhe [30].

Em alternativa, também é possível usar dados de velocidade de vento simulados por um modelo de circulação atmosférica (*WRF*) e os dados de produção filtrados do aerogerador ou do parque. Com base nos dados de produção e na velocidade e direção do vento, é calculada uma curva de potência para cada uma das turbinas que constituem o parque eólico. Neste método, as curvas de potência derivadas têm valores de potência para um determinado número de setores direcionais (oito, por exemplo) e para intervalos de velocidade de 1 m/s, por exemplo. Os valores de potência são calculados como o valor médio de todas as ocasiões para cada bin de velocidade-direção do vento. Na Figura 2.6, encontram-se representadas oito curvas de potência para um determinado aerogerador de um parque eólico em análise. Facilmente se identifica que as curvas de potência têm um comportamento com uma tendência menos definida quando as velocidades de vento são mais elevadas, devido à escassez dos dados que lhes deram origem [30].

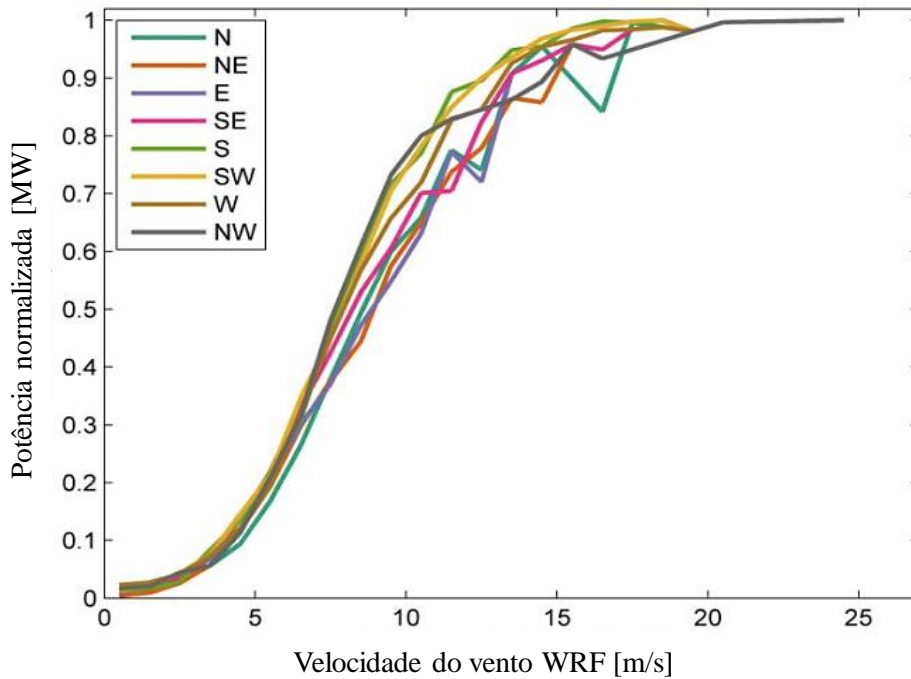


Figura 2.6. Representação da curva de potência para um aerogerador, para oito setores, para a segunda abordagem [30].

Em comparação com as curvas de potência comerciais reais, as curvas de potência derivadas sobrestimam a potência a baixas velocidades e subestimam a potência a velocidades mais altas. Esta é uma característica inerente ao método e é resultado, por um lado, do facto de a potência de saída do aerogerador variar entre zero e a potência nominal e, por outro, da imprecisão dos dados do *WRF*, resultando numa das desvantagens desta abordagem. Outras desvantagens desta abordagem estão, em primeiro lugar, no facto de a velocidade do vento modelada pelo *WRF* não ser afetada por efeitos de esteira e em segundo, a curva de potência histórica é construída tendo em consideração as diferentes direções do vento, pelo que é necessária uma maior quantidade de dados operacionais válidos para obter uma curva de potência bem definida [30].

2.3.2. Método de índice

Existem outras formas de obter a produção de energia bruta após esteira, nomeadamente os denominados métodos de índice, nos quais é determinada uma função que correlaciona a produção de energia observada, corrigida para os períodos de indisponibilidade das turbinas ou desempenho parcial, com um índice baseado numa série temporal de Longo Termo da produção ou da velocidade do vento. Este método correlaciona, numa qualquer base temporal, desvios da

produção face a um valor médio com desvios na velocidade do vento também face a um valor médio, sendo calculada a quantidade de energia que seria produzida se os aerogeradores estivessem em pleno funcionamento [30].

O índice de produção pode ser calculado numa base semanal, através da equação seguinte, mas também pode ser calculado em outra base temporal:

$$\text{Índice (semanal)} = \frac{\text{Produção (semanal)}}{\frac{1}{n} \sum_{i=1}^n \text{Produção (semanal)}} \quad (1)$$

De uma forma simplificada, a produção de energia corrigida é calculada através do quociente entre a produção real observada e a disponibilidade dos aerogeradores. Tendo esta premissa como mote, é possível dividir este método em duas abordagens, sendo que a diferença entre elas é o índice utilizado. É importante referir que, independentemente da abordagem em questão, devem ser consideradas semanas ou meses com elevada cobertura de dados para garantir boa representatividade do local (superior a 95%, por exemplo) [30].

Na primeira abordagem, o índice de produção é derivado de uma curva de potência do parque, calculada no *WindPro*³, que relaciona as séries temporais de velocidade e direção do vento do conjunto de dados de Longo Termo com a produção de energia do parque. Antes de calcular a curva de potência do parque, o modelo de fluxo *WASP*⁴ é utilizado para extrapolar horizontalmente as velocidades do vento do modelo para as localizações das turbinas, tendo em conta a topografia local e as características da superfície. Assim, para cada direção e velocidade do vento, a curva de potência do parque fornece a potência total produzida pelo parque eólico [30].

A Figura 2.7 demonstra de que forma é deduzida a produção normal esperada, para uma determinada base temporal, neste caso semanal. A produção total corrigida produzida pelo parque é correlacionada com o índice semanal de produção, e através de uma correlação linear, é determinada a produção semanal esperada [30].

³ O *windPRO* é uma ferramenta utilizada para diversos fins, entre os quais a análise de dados de vento, cálculo de produção de energia, quantificação de incertezas, avaliação da adequação do local e até ao cálculo e visualização do impacto ambiental [55].

⁴ O modelo computacional *WASP*® (*Wind Atlas Analysis Application Program*) é usado para gerar dados para a elaboração de atlas eólicos, estimar o potencial de uma região e calcular a produção de energia [54].

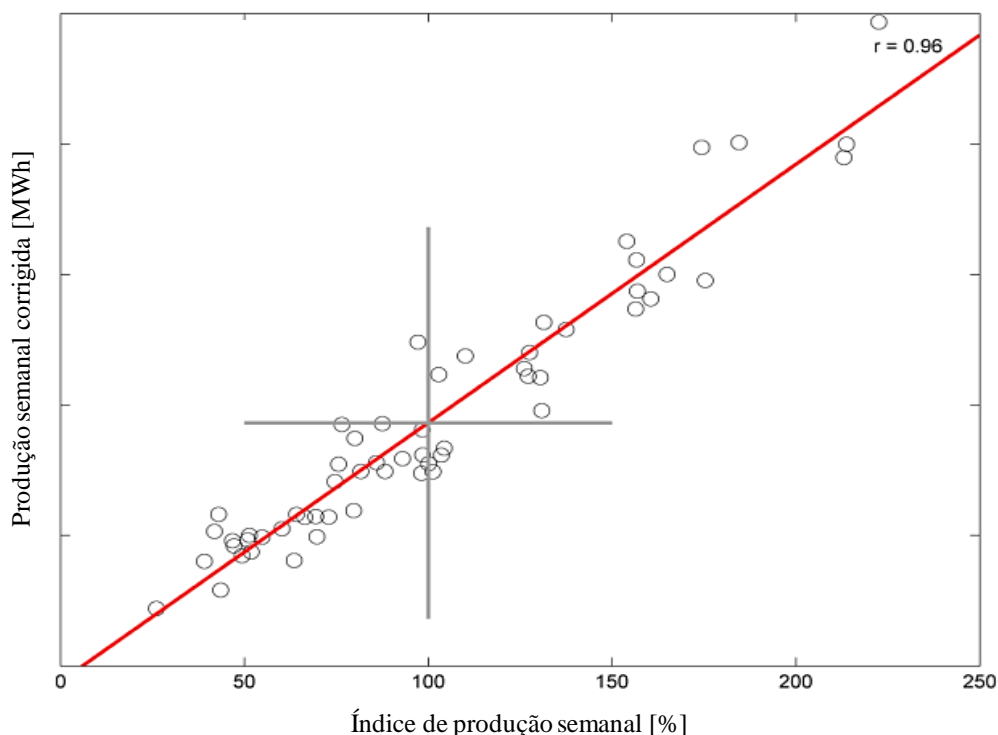


Figura 2.7. Relação entre a produção semanal de energia corrigida observada e o índice de produção semanal [30].

Desta forma, utilizando um índice produção semanal, a estimativa de produção anual de energia é calculada através da equação (2). O termo 52,18 que é apresentado na Equação 2, refere-se ao número médio de semanas anual [30].

$$\text{Produção anual de energia} = 52,18 \times \text{Produção semanal de energia esperada} \quad (2)$$

Usando um índice de produção mensal, a estimativa de produção de energia anual de energia é calculada através da equação (3). O termo 12 que é apresentado na Equação 3, refere-se ao número de meses anual [30].

$$\text{Produção anual de energia} = 12 \times \text{Produção mensal de energia esperada} \quad (3)$$

Relativamente à segunda abordagem, a única diferença reside no facto de o índice de produção modelado ser calculado a partir da série de dados de vento de Longo Termo, enquanto na primeira abordagem é obtida através da série de produção do parque eólico. Em seguida, utiliza-se o mesmo método que se encontra descrito anteriormente e esquematizado na Figura 2.4 [30].

2.3.3. Método da regressão linear

A regressão linear é uma técnica estatística que estabelece uma relação entre duas ou mais variáveis. No contexto mais comum, a regressão linear envolve duas variáveis: a variável que se pretende prever (dependente) e a variável utilizada para fazer a previsão (independente). A equação 4 descreve o método da regressão linear [31, 32].

$$y = \alpha x + \beta \quad (4)$$

Na qual,

y é a variável dependente

x é a variável independente

α é o coeficiente da regressão que representa a inclinação da reta

β é o ponto de interseção da reta com o eixo vertical

Para aplicar o método da regressão linear de maneira adequada, é importante verificar se alguns pressupostos são cumpridos, nomeadamente a linearidade entre as variáveis, a independência, homocedasticidade e normalidade dos resíduos e a ausência de multicolinearidade.

O primeiro pressuposto que é necessário considerar é a linearidade uma vez que a relação entre as variáveis dependente e independente deve ser linear e isso significa que o efeito de uma mudança da variável independente sobre a variável dependente é constante em todas as faixas de valores das variáveis [31, 32]. A relação não linear entre a velocidade do vento e a potência, representada na Figura 2.8, pode ser aproximada a várias expressões polinomiais, sendo o modelo mais simplificado baseado numa curva linear, que descreve a região 2 da curva de potência através de uma reta [33]. Assim, através de uma aproximação, é possível considerar que existe uma relação linear entre a potência elétrica e a velocidade do vento.

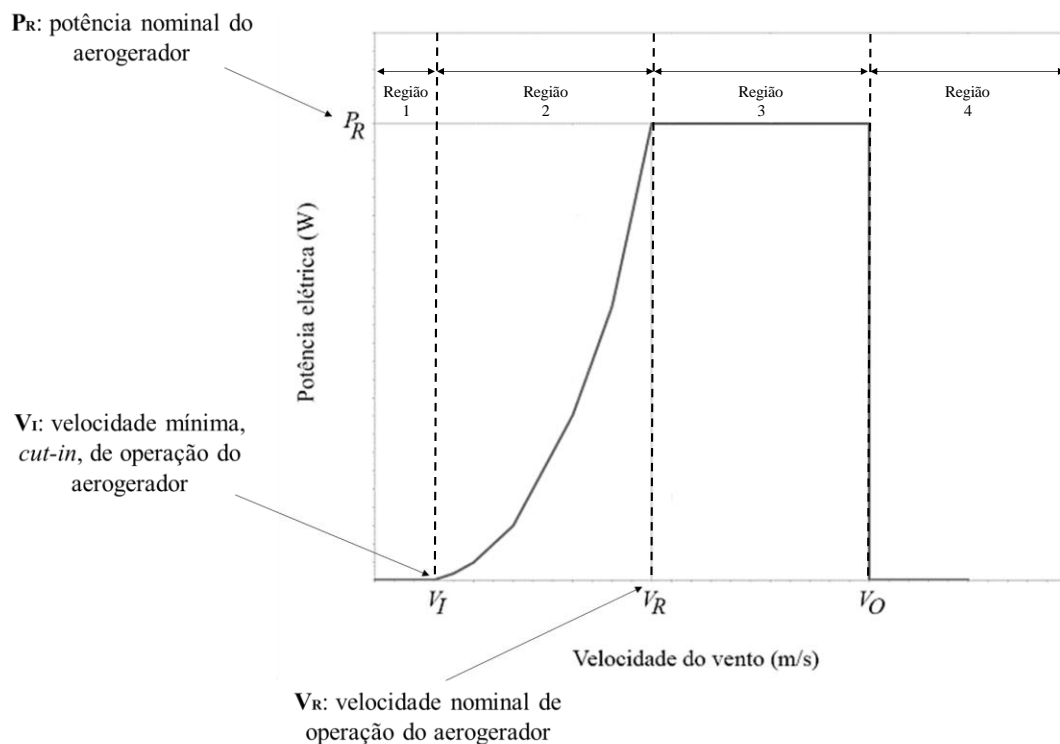


Figura 2.8. Curva de potência típica de um aerogerador que relaciona a velocidade do vento e a potência elétrica [33].

Os resíduos, definidos como a diferença entre os valores observados e os valores previstos pela regressão, devem ser analisados com precaução uma vez que existem vários pressupostos da regressão linear relacionados com as características dos resíduos. Os resíduos devem ser independentes, ou seja, a ocorrência de um resíduo não deve influenciar a ocorrência de outro. A variabilidade dos resíduos deve ser constante em todas as faixas de valores das variáveis independentes pelo que não deve haver um padrão claro de aumento ou diminuição da dispersão dos resíduos à medida que os valores das variáveis independentes mudam. Para além disso, os resíduos devem seguir uma distribuição normal [33].

Por fim, não deve existir multicolinearidade, ou seja, as variáveis independentes não devem estar altamente correlacionadas entre si [33].

Este método apresenta algumas desvantagens nomeadamente a sua simplicidade, fazendo com que o método possa não conseguir capturar algumas nuances dos dados e pela sensibilidade a *outliers*, uma vez que estes pontos podem distorcer a linha de ajuste e afetar significativamente o coeficiente de correlação. Para além disso, ainda existe o problema da extrapolação dos resultados dado que fazer previsões fora da gama de valores inicial pode ser arriscado [33].

3. METODOLOGIA

Utilizando um conjunto de parques eólicos em operação em diferentes partes do globo, foram estudados os objetivos propostos na secção 1.4. pela aplicação da metodologia de cálculo utilizada pela *Megajoule*. Neste capítulo é apresentada a base de dados em estudo, assim como a metodologia utilizada para o cumprimento dos objetivos propostos.

3.1. Base de dados

A base de dados do presente estudo é constituída por 21 parques eólicos e 6 complexos eólicos, espalhados por vários países do globo, nomeadamente Portugal, Espanha, Roménia, Canadá e Brasil. Na Figura 3.1 está identificada a localização geográfica dos parques e complexos eólicos selecionados. A base de dados foi selecionada tendo em conta os estudos pós construção realizados pela *Megajoule* no passado com um número mínimo de quatro anos de dados operacionais. A seleção de parques e complexos eólicos em diferentes pontos do mundo teve como objetivo captar diferentes regimes de vento e o efeito que possam ter na qualidade da própria série de reanálise e estimativa de produção de energia de Longo Termo.

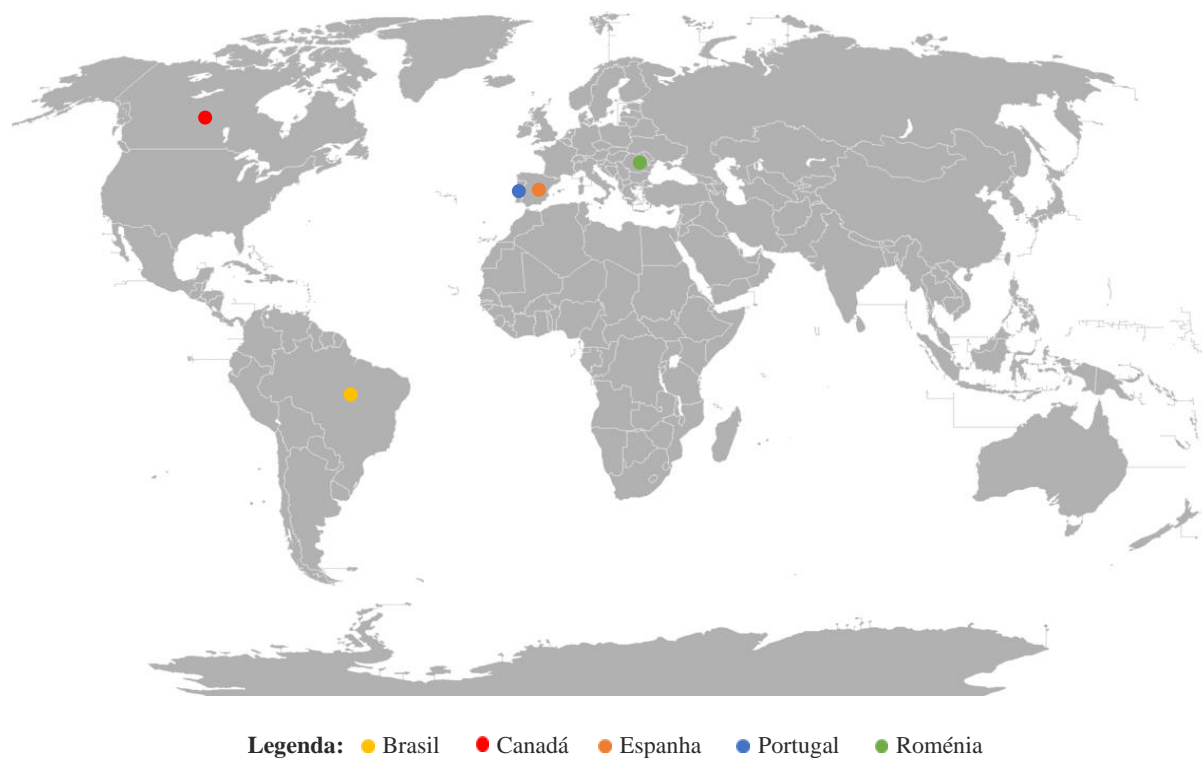


Figura 3.1. Localização geográfica dos parques e complexos eólicos selecionados.

A informação genérica relativa a cada parque e complexo eólico utilizados no presente estudo encontra-se descritas nas Tabelas 3.1 e 3.2. Por questões de confidencialidade, os parques e complexos eólicos estudados são apenas identificados com a distinção entre parque (P) ou complexo eólico (C), seguido da inicial do país e de um número, e é apenas apresentada uma informação muito genérica em relação aos mesmos, nomeadamente a potência instalada e os anos de operação estudados, para os parques eólicos, e o número de parques eólicos constituintes, para os complexos.

*Tabela 3.1. Informação genérica relativa a cada **parque eólico**, o que inclui a designação utilizada, a potência instalada em MW e os anos de dados de operação*

País	Designação	Potência instalada (MW)	Anos de operação
Portugal	P_P1	26	6
	P_P2	15	12
	P_P3	40	6
	P_P4	102	5
	P_P5	12	20
	P_P6	21,6	20
	P_P7	18	20
Espanha	P_E1	64,7	4
	P_E2	100,7	4
	P_E3	54,4	4
	P_E4	128	4
	P_E5	45	4
	P_E6	20	4
	P_E7	49,5	4
	P_E8	49,5	4
	P_E9	30,6	4
	P_E10	25,2	4
	P_E11	30	4
Roménia	P_R1	33,6	4
	P_R2	8,4	4
Canadá	P_C1	100	9

Tabela 3.2. Informação genérica relativa a cada **complexo eólico**, o que inclui a designação utilizada, a potência instalada em MW e os anos de dados de operação

País	Designação	Nº de parques	Potência instalada (MW)	Anos de operação
Brasil	C_B1	5	160	8
	C_B2	8	267	5
	C_B3	7	223	3
	C_B4	8	205,1	6
	C_B5	2	40	2
	C_B6	3	70	2
Espanha	C_E12	2	101,4	4
	C_E13	3	94,5	4

3.2. Metodologia geral

A metodologia de cálculo de estimativas de produção de energia de parques eólicos em operação seguida pela *Megajoule*, resulta de uma combinação dos métodos publicados em bibliografia recente, e que foram descritas na secção 2.3. A Figura 3.2 mostra uma representação esquemática desta metodologia.

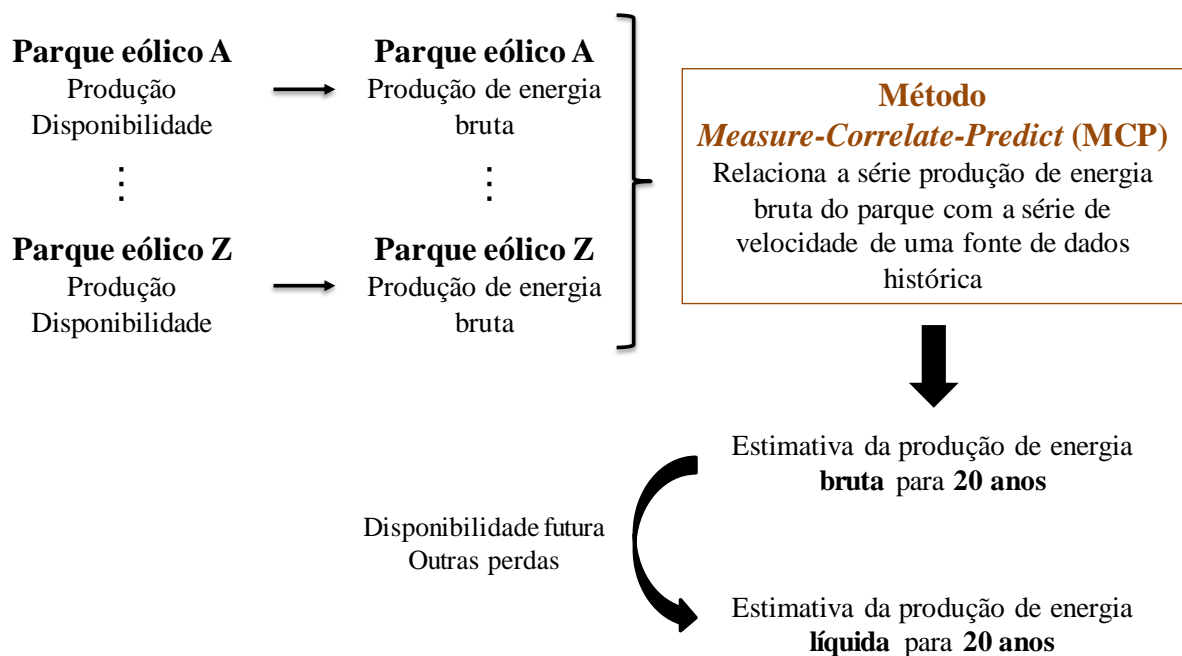


Figura 3.2. Representação esquemática da metodologia geral.

Os dados de entrada utilizados são a produção de energia real dos parques eólicos selecionados, em base mensal, e a disponibilidade dos aerogeradores que o constituem, para o período de operação. Para além disso, para alguns dos parques existe também um terceiro ponto, designado de perdas, que inclui todas as perdas alocadas ao parque eólico como, por exemplo, perdas por limitação da rede, que podem ser calculadas ou fornecidas diretamente como dados de entrada. Os dados de produção real do parque eólico são corrigidos para uma disponibilidade do parque de 100%, ou seja, calcula-se o quociente entre o valor de produção mensal e disponibilidade do parque nesse mês. Assim, obtém-se uma série de dados de produção de energia, em base mensal, equivalente a 100% de disponibilidade do parque eólico.

Para uma determinada localização central de um projeto eólico representativa do parque ou complexo em estudo extraíram-se as séries de reanálise para os pontos da grelha de reanálise ERA5 e MERRA-2 mais próximas desse ponto central. Assim, obtiveram-se 4 séries de reanálise ERA5 e 4 séries de reanálise MERRA-2, totalizando 8 séries para cada parque ou complexo em estudo.

Para cada série de reanálise, os dados reais de produção mensal de cada parque eólico foram relacionados, através de uma correlação linear, com as séries de dados de vento históricos de Longo Termo para o período simultâneo de dados, obtendo-se 8 correlações para cada parque/complexo. Para melhorar a correlação entre os dados da série de produção real e da série de reanálise, os pontos muito afastados da linha de tendência ou que apresentem um valor de disponibilidade mensal inferior a 85% são filtrados. A eliminação de pontos erróneos é feita, tendo em vista um equilíbrio entre o maior número de registos válidos e a melhor correlação possível, para que os dados com a mesma base temporal possam permanecer comparáveis, ou seja, é importante que ao comparar a estimativa de produção de Longo Termo de dois parques realizada com a mesma base temporal (4 anos, por exemplo), o número de registos mensais válidos seja o mais próximo possível. As correlações obtidas foram analisadas com o objetivo de selecionar o ponto de reanálise mais adequado à realização do estudo. O critério de seleção baseou-se na escolha da correlação com maior coeficiente de correlação R^2 .

Após a seleção da fonte de dados de reanálise mais adequada, a equação da correlação estabelecida entre a velocidade da série de reanálise e a série de dados de produção real do parque é aplicada à série de dados de reanálise, para o período de 20 anos de dados, gerando uma série de produção mensal reconstruída para os 20 anos de operação futura do parque eólico.

Posteriormente, a série mensal de produção é convertida numa série anual de produção, através do somatório das produções mensais de cada ano. O valor da estimativa de produção de energia de Longo Termo, para os 20 anos, é calculado através da média das produções anuais da série de 20 anos.

Numa última etapa, é necessário voltar a afetar o valor final da estimativa de produção de energia de Longo Termo de todas as perdas que sejam necessárias considerar, como a disponibilidade e limitações da rede, por exemplo.

3.2.1. Análise e alterações à Metodologia geral

Na metodologia seguida pela Megajoule, o critério de seleção do ponto de reanálise a usar no estudo baseia-se na escolha da correlação com maior coeficiente de correlação R^2 . Neste estudo foram feitas algumas modificações na metodologia geral, tendo em vista o cumprimento dos objetivos propostos. Foi, também, efetuada uma análise ao critério de escolha da série de reanálise baseado no valor do R^2 .

Análise do coeficiente de correlação R^2 como indicador de seleção da fonte histórica de dados

Utilizando como base de dados os parques com 20 anos de dados operacionais, procedeu-se à análise de todas as fontes de dados de reanálise extraídas para esses três parques. O objetivo primordial desta análise consistiu em investigar a relação entre o valor do coeficiente de correlação R^2 e o valor do desvio percentual entre a estimativa de produção de Longo Termo obtida para cada uma das séries de reanálise e a média real do parque. Para cada uma das séries de reanálise extraídas, foi estimada a série de produção reconstruída, calculando o valor da estimativa de produção de energia de Longo Termo para essa série de reanálise, sendo posteriormente calculado um desvio percentual entre esse valor e o valor médio real dos 20 anos de operação. Assim, é possível relacionar o valor do coeficiente de correlação de cada uma das séries de reanálise com o desvio entre a estimativa de produção de energia de Longo Termo para essa série de reanálise e a média real do parque.

Alterações à metodologia geral

A primeira diferença está na escolha da série de reanálise mais adequada para o estudo: para além de analisar o valor do coeficiente de correlação, foi estudada também a aleatoriedade da distribuição dos resíduos entre a série de produção real do parque e a reconstruída com a correlação associada a cada ponto de reanálise, para o período simultâneo de dados. Assim, a série de reanálise mais adequada ao estudo considerou-se ser a associada a combinação de resíduos aleatórios e o maior valor de coeficiente de correlação. Esta análise adicional dos resíduos reveste-se da maior importância uma vez que uma correlação com coeficiente de correlação elevado e que apresente tendências nos resíduos pode levar a previsões enviesadas e imprecisas.

Outra distinção é o facto de não se realizar a etapa final de alocação de perdas uma vez que é possível obter as conclusões necessárias, numa base de estimativa de produção de energia bruta, sem a afetação das perdas.

3.2.2. Validação da Metodologia geral adotada

A validação da metodologia adotada é de extrema importância e representa um processo crucial para garantir a confiabilidade dos resultados obtidos neste estudo. Com esse objetivo, a verificação do cumprimento dos pressupostos da regressão linear da correlação obtida utilizando a série de reanálise selecionada é de clara importância.

Verificação dos pressupostos do modelo de regressão linear

Na primeira fase de validação, os dados dos parques foram analisados de modo a perceber se todos os pressupostos da regressão linear são cumpridos, através de análises estatísticas que avaliam cada um dos parâmetros mais relevantes. Estas análises foram aplicadas a todos os parques com 4 e 20 anos de dados operacionais e aos complexos (tendo sido escolhido um parque representativo do complexo para realizar este estudo). Os testes estatísticos utilizados foram o *teste de variâncias (ANOVA) tradicional* de modo a verificar se o modelo é estatisticamente significativo, o *teste de Kolmogorov-Smirnov* para avaliar a normalidade dos

dados, o teste de *Durbin-Watson* para verificar a presença de auto-correlação nos resíduos do modelo de regressão e o teste de *Breusch-Pagan* para avaliar a constância das variâncias.

A *análise de variância (ANOVA)* serve para avaliar a significância estatística do modelo de regressão, ajudando a determinar se a regressão linear ajustada aos dados é estatisticamente significativa e se as variáveis independentes estão a contribuir de maneira significativa para explicar o comportamento da variável dependente (hipótese nula). O valor de p é a medida estatística que ajuda a determinar se existem evidências para rejeitar a hipótese nula, pelo que se o valor de p for inferior ao nível de significância (nomeadamente 0,05) é possível afirmar que o modelo é estatisticamente significativo [34].

O teste de *Kolmogorov-Smirnov* compara a distribuição empírica dos dados com a distribuição normal teórica, auxiliando na identificação de possíveis desvios da normalidade. Se o resultado do teste indicar discrepâncias significativas, pode sugerir que a normalidade dos dados não está plenamente satisfeita. Se o valor- p for menor do que o nível de significância (em geral 0,05), é possível rejeitar a hipótese nula e concluir que os dados não seguem a distribuição teórica [35]. No âmbito deste trabalho, este teste foi aplicado para avaliar a normalidade dos resíduos.

O teste de *Durbin-Watson* é um procedimento estatístico utilizado para verificar a presença de auto-correlação nos resíduos de um modelo de regressão, que ocorre quando os resíduos de um modelo não são independentes entre si, ou seja, há uma relação entre o erro num ponto e o erro em pontos anteriores. Este teste calcula uma estatística que varia entre 0 e 4, onde um valor próximo a 2 indica que não há auto correlação significativa nos resíduos e um valor menor ou maior que 2 sugere que os resíduos adjacentes estão correlacionados [36].

O teste de *Breusch-Pagan* é utilizado na análise de regressão para avaliar a variabilidade dos resíduos do modelo. O teste supõe que a variabilidade dos resíduos em modelos de regressão é constante, calculando uma estatística sob a hipótese nula de variabilidade constante e se o valor- p associado à estatística D for menor do que um certo nível de significância (geralmente 0,05), pode-se rejeitar a hipótese nula e concluir que há evidências de variabilidade não constante [37].

No caso de todos os pressupostos do modelo da regressão linear serem cumpridos, a escolha da utilização deste modelo é validada. Assim, e após a validação do método, o fator mais relevante que passa a ser tido em consideração para a escolha da série de reanálise mais adequada é o valor do coeficiente de correlação R^2 .

3.3. Análise estatística dos resultados

A análise estatística desempenha um papel essencial no presente estudo uma vez que proporciona uma estrutura mais sólida para a interpretação de resultados e obtenção de conclusões significativas. Assim, e após a validação da metodologia geral adotada, o estudo avançou para as análises que permitem dar resposta aos objetivos propostos, nomeadamente o estudo da influência do tempo de correlação e a análise da influência da agregação espacial em complexos eólicos na estimativa de produção de energia de Longo Termo, que são explicadas com maior pormenor nas secções 3.4 e 3.5. Nesta secção são explorados dois aspetos fundamentais de análise estatística, que são aplicados nos estudos abordados nas secções 3.4 e 3.5: a avaliação de desempenho e precisão por meio de definição de métricas de erro e a investigação de diferenças significativas entre grupos com o uso de testes de hipótese e as respetivas análises *post hoc*⁵.

3.3.1. Definição das métricas de erro

A *Raíz do Erro Quadrático Médio Relativo* (RRMSE) e o *Erro Absoluto Médio Relativo* (RMBE) são métricas frequentemente utilizadas na avaliação do desempenho de modelos de previsão ou regressão. Ambos são utilizados para medir a qualidade das previsões em relação aos valores reais, mas focam aspetos diferentes da precisão e exatidão do modelo. A principal diferença entre o RRMSE e o RMBE está nas suas abordagens para medir a precisão do modelo. O RRMSE concentra-se na magnitude dos erros, enfatizando mais os erros maiores, enquanto o RMBE foca na direção dos erros e determina se há um viés no modelo [38].

⁵ As comparações *post hoc* são análises estatísticas realizadas após a conclusão de um teste de hipóteses como a *ANOVA tradicional*, a *ANOVA de Welch* ou o *teste de Kruskal-Wallis*, quando o teste original indica diferenças significativas entre os grupos. Esses testes adicionais são realizados para identificar quais grupos específicos diferem uns dos outros.

O RRMSE é uma medida de dispersão dos erros entre as previsões e os valores, fornecendo uma noção da precisão média do modelo e penalizando erros maiores devido ao fator quadrático [38]. O RRMSE é calculado através da Equação 5, na qual n corresponde ao número de amostras, y_i ao valor observado para a amostra i e p_i o valor previsto para a amostra i :

$$RRMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - p_i)^2} \quad (5)$$

Por outro lado, o RMBE mede o viés ou tendência sistemática das previsões em relação aos valores reais. Um valor de RMBE próximo de zero indica que, em média, as previsões estão próximas dos valores reais, não apresentando viés sistemático. Valores positivos do RRMBE indicam que o modelo tende a sobrestimar as previsões, enquanto valores negativos indicam uma tendência de subestimação [38]. O RMBE é calculado através da Equação 6:

$$RMBE = \frac{1}{n} \sum_{i=1}^n (y_i - p_i) \quad (6)$$

3.3.2. Testes de hipótese e análises *post-hoc*

A análise de variância (ANOVA) tradicional é uma técnica estatística essencial para comparar as médias de dois ou mais grupos independentes de um estudo, permitindo inferências sobre as relações e efeitos entre as variáveis em análise. A hipótese nula (H_0) afirma que não há diferenças significativas entre as médias dos grupos. O objetivo é testar essa hipótese, avaliando a significância estatística dos resultados obtidos e para isso, calcula-se o valor de p . O valor de p é a probabilidade de observar os resultados da ANOVA, assumindo que a hipótese nula (H_0) é verdadeira. Quando o valor de p é inferior a um limiar previamente determinado (tipicamente 5%), é comum rejeitar a hipótese nula em favor da hipótese alternativa, a qual postula que existe, de facto, uma diferença ou efeito real entre os grupos sob investigação [39, 40].

No entanto, antes de aplicar a ANOVA, é necessário verificar se os pressupostos de independência das observações, normalidade e homogeneidade de variância são atendidos, pelo que no estágio inicial da análise é necessário utilizar métodos que avaliem estes parâmetros, nomeadamente o *teste de Levene* e o teste de *Kolmogorov-Smirnov* [41, 35].

O *teste de Levene* é uma ferramenta estatística que verifica se as variâncias das amostras são estatisticamente semelhantes. Isso é essencial, pois a homogeneidade das variâncias é uma suposição fundamental para a aplicação adequada da análise de variância (ANOVA) tradicional. Caso o *teste de Levene* indique a violação da homogeneidade das variâncias, pode ser necessário recorrer a alternativas que sejam mais adequadas para tal cenário, ao invés da análise ANOVA tradicional [41].

Quando os pressupostos da homogeneidade das variâncias e da normalidade não são verificados, é necessário adotar estratégias alternativas à ANOVA tradicional, nomeadamente a *ANOVA Welch* ou o *teste de Kruskal-Wallis* (teste não paramétrico) [42, 43].

A *ANOVA Welch* (Análise de Variância com Diferenças de Variâncias) é uma extensão do teste de ANOVA tradicional, utilizada quando as variâncias entre os grupos não são homogêneas. Este teste corrige a estimativa dos graus de liberdade e usa um cálculo ponderado das variâncias dos grupos para lidar com as disparidades. A estatística F é calculada como a razão entre a variabilidade entre os grupos e a variabilidade dentro dos grupos, tal como na ANOVA tradicional. No entanto, na *ANOVA Welch*, essa variabilidade é ajustada tendo em consideração as diferentes variâncias dos grupos [42].

O *teste de Kruskal-Wallis* é uma alternativa não paramétrica à ANOVA, usada quando os pressupostos da ANOVA tradicional não são atendidos e os dados não seguem uma distribuição normal. Este teste é aplicado quando se deseja comparar grupos independentes para determinar se eles têm medianas diferentes e é baseado na classificação dos dados combinados de todos os grupos, calculando uma estatística de teste que leva em consideração as diferenças nas classificações médias entre os grupos e a variação dentro dos grupos. Essa estatística é comparada a uma distribuição conhecida para determinar se há diferenças significativas entre os grupos. Se o *teste de Kruskal-Wallis* indicar diferenças significativas entre os grupos, isso sugere que pelo menos um grupo difere em termos de mediana [43].

Uma vez concluída a fase dos testes de hipóteses, e independentemente do método empregado, é fundamental explorar as diferenças entre os grupos. Para tal, recorre-se às comparações *post hoc*. No âmbito da ANOVA tradicional ou da *ANOVA de Welch*, testes como o *teste de Tukey*, podem ser utilizados para identificar quais grupos apresentam médias significativamente distintas. No caso do *teste de Kruskal-Wallis*, as comparações *post hoc* podem ser realizadas por meio do teste de *Dunn* [44, 45]. O *teste de Tukey* é o método estatístico usado para identificar quais grupos específicos têm diferenças significativas entre as médias dos grupos quando a ANOVA tradicional indica uma diferença global significativa [44]. O *teste de Dunn* é um teste de comparações múltiplas utilizado para determinar quais grupos específicos apresentam diferenças significativas, após estabelecido que pelo menos um grupo tem uma diferença significativa em relação aos outros no teste de hipóteses realizado anteriormente [45].

Na Figura 3.3 está ilustrado um diagrama que esquematiza a metodologia adotada a nível de análise estatística, representando a sequência de procedimentos estatísticos empregados e explicados anteriormente.

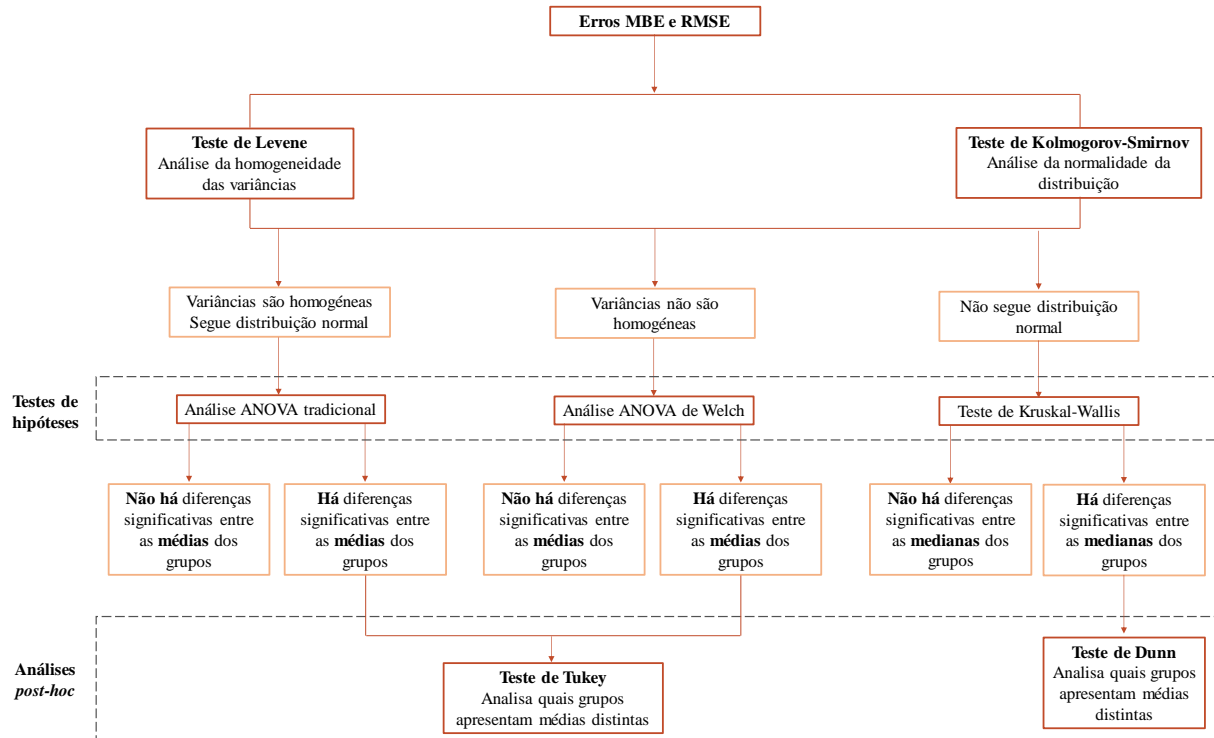


Figura 3.3. Representação esquemática da metodologia adotada a nível de análise estatística.

3.4. Análise da influência do tempo de correlação

A presente tese teve como um dos objetivos analisar a influência do tempo de correlação na estimativa de produção de energia de Longo Termo de um parque eólico e também na previsão da produção real do parque eólico.

Numa primeira fase, foram usados dados de produção dos 3 parques eólicos com 20 anos de operação, proporcionando uma compreensão abrangente dos padrões de produção de energia e sua variabilidade temporal. Para aferir o impacto do tempo de correlação nas estimativas de Longo Termo, os dados de vento dos parques foram divididos em diferentes períodos, considerando combinações possíveis de anos consecutivos, variando de 1 a 20 anos. Para períodos de 1 ano foi usado separadamente cada um dos anos de operação (ano 1, ano 2, ..., ano 20), para períodos de 2 anos foram usados períodos de 2 anos consecutivos (anos 1 e 2, anos 2 e 3, ..., anos 19 e 20) até completar o período total de 20 anos de dados operacionais (anos 1 a 20). Para cada período de anos consecutivos, foi aplicada a metodologia definida na secção 3.2 resultando numa série de produção reconstruída para 20 anos, com a qual foi calculada a estimativa de produção de Longo Termo dessa série. Em seguida, foi determinado o desvio percentual desse valor em relação à média da série real dos 20 anos de operação do parque. As diferentes séries de dados reconstruídas foram também comparadas com a série de produção real do parque para os 20 anos de operação do mesmo, utilizando as métricas de erro, RMBE e o RRMSE (referidas na secção 3.3.1). Foram ainda realizados alguns testes estatísticos às métricas de erro (abordados na secção 3.3.2) para perceber se os grupos de anos (1 a 20 anos) apresentam diferenças significativas, agregando os resultados obtidos para os três parques com 20 anos de operação.

O número de parques utilizado nesta fase do estudo, apresenta-se como uma limitação e é manifestamente insuficiente para efetuar uma análise sustentada sobre a influência do período de dados de correlação na previsão da produção real de um parque eólico. Por outro lado, é importante avaliar com maior pormenor o comportamento dos dados operacionais para um período mais curto de dados, uma vez que usualmente não se encontra disponível um período de dados operacionais de 20 anos, mas sim períodos de tempo mais curtos (como 3 ou 4 anos, por exemplo) que apresentam muitas oscilações, e detetar se há melhorias consistentes ao longo do tempo ou se há momentos específicos em que ocorrem variações significativas nos erros, para além de reiterar determinadas nuances que possam existir nos dados e que podem não ser

refletidas em análises com períodos mais longos de dados. Assim, numa segunda fase do estudo foram usados os dados de 18 parques eólicos distintos cada um com um histórico de pelo menos 4 anos de medições de produção de energia, sendo que a seleção do tamanho do período de análise (4 anos) resultou de um compromisso entre um número suficiente de anos e um número satisfatório de parques (dado que uma parte significativa dos parques disponíveis tem 4 anos de operação).

Esses 4 anos de dados operacionais foram subdivididos em todas as combinações possíveis de 1 a 4 anos, resultando em 15 combinações para cada parque (para períodos de 2 anos foram usadas as combinações: anos 1 e 2, anos 1 e 3, anos 1 e 4, anos 2 e 3, anos 2 e 4 e anos 3 e 4; para 3 anos: anos 1,2 e 3, anos 2, 3 e 4, anos 1, 3 e 4 e anos 1, 2 e 4). A metodologia descrita na secção 3.2 foi aplicada para cada combinação de anos a todos os parques, resultando numa série de produção de energia reconstruída para cada combinação, para cada parque. Foram também calculados o RMBE e do RRMSE (referidos na secção 3.3.2) para avaliar a precisão e exatidão das projeções em comparação com os dados reais de 4 anos de produção de energia. Foi também, realizado um conjunto de análises estatísticas, garantindo interpretações mais precisas, abordadas na secção 3.3.2.

3.5. Análise da influência da agregação espacial em complexos eólicos

Tendo como objetivo analisar qual a influência do estudo do complexo eólico como um todo, relativamente ao estudo individual dos parques que o constituem, para a estimativa de produção de energia desse complexo, foram realizados dois estudos, alterando os dados de entrada: no primeiro, foram utilizados os dados do complexo eólico como um todo e no segundo os dados dos parques individuais que o constituem. A base de dados utilizada neste estudo encontra-se caracterizada na tabela 3.2.

A metodologia descrita na secção 3.2 foi aplicada para cada um dos complexos em estudo, considerando os dados operacionais do complexo como um todo e os dados dos parques que o constituem, resultando numa série de produção de energia reconstruída para o complexo e para cada um dos parques. Depois disto, as séries individuais dos parques foram agregadas resultando na série de produção do complexo. Assim, são comparadas as séries de produção do

complexo provenientes de diferentes dados de entrada com a série de produção real do complexo.

Para o segundo estudo, no qual se utilizou os dados dos parques individualmente, foi ainda estudada a influência da fonte de reanálise escolhida – realizou-se a estimativa de produção de energia com a fonte de reanálise mais adequada (a que apresentou erros aleatórios e maior valor de coeficiente de correlação) para o parque e com a fonte de reanálise mais adequada para o complexo eólico. Foram calculados o RMBE e o RRMSE (abordados na secção 3.3.2) com o objetivo de avaliar a exatidão e a precisão das projeções, respetivamente, em comparação com os dados reais de produção de energia, quer para o complexo, quer para os parques individuais que o constituem. Foram ainda realizados um conjunto de análises estatísticas, que incluem a análise da homogeneidade das variâncias e da normalidade da distribuição, os testes de hipóteses e as respetivas análises *post hoc*, para avaliar se existem diferenças significativas em utilizar os dados operacionais dos complexos ou dos parques individuais, explicadas na secção 3.3.2.

4. RESULTADOS

Neste capítulo são expostas a validação da metodologia aplicada e a análise pormenorizada dos resultados, seguindo a abordagem metodológica descrita no capítulo 3. O objetivo deste capítulo é proporcionar uma compreensão aprofundada dos resultados dos estudos realizados.

4.1. Validação da metodologia utilizada

Para simplificar a linguagem e facilitar o entendimento da discussão, o termo “*coeficiente de correlação da série de reanálise*” deve ser entendido como o coeficiente de correlação R^2 da regressão linear obtida, para o período de dados simultâneo, entre os dados reais de produção de energia do parque e os dados de vento da série de reanálise.

Na primeira fase de análise e validação da metodologia, foi investigada a relação entre o valor do coeficiente de correlação R^2 de cada uma das séries de reanálise extraídas e o valor do desvio percentual entre a estimativa de produção de Longo Termo obtida para cada uma das séries de reanálise e a média real do parque. Os resultados obtidos para esta análise estão representados graficamente na Figura 4.1, sendo que os dados que permitiram a construção da mesma se encontram descritos na Tabela A.2 no Anexo A. A análise realizada revelou que existe uma correlação elevada, acima de 80% para os parques em estudo, entre o coeficiente de correlação das séries de reanálise e o desvio percentual da estimativa de produção de Longo Termo em relação à média real do parque. Essa correlação indica uma relação forte entre essas duas variáveis, ou seja, à medida que o valor de R^2 aumenta para uma série de reanálise, o desvio percentual em relação à média de Longo Termo tende a diminuir, e vice-versa. No entanto, para o parque P_P7 verifica-se que os pontos interseam o eixo das ordenadas pelo que, ao selecionar a série de reanálise com 59 % de coeficiente de correlação (primeiro ponto apresentado na série do parque P7) leva a um desvio menor da estimativa de Longo Termo em relação ao valor real do parque, em relação a escolher a série de reanálise com 90 % de coeficiente de correlação (sendo o desvio -0,6% e 1,3%, respetivamente). Embora o coeficiente de correlação meça a relação linear entre duas variáveis, os fenómenos reais nem sempre são estritamente lineares, pelo que pode haver fatores não lineares que estejam a afetar a estimativa de produção de energia do parque P7. Apesar desta exceção, a informação contida na

representação gráfica da Figura 4.1 sugere que o coeficiente de correlação R^2 pode ser um bom indicador para a escolha da série de reanálise mais adequada para o estudo do parque eólico em questão.

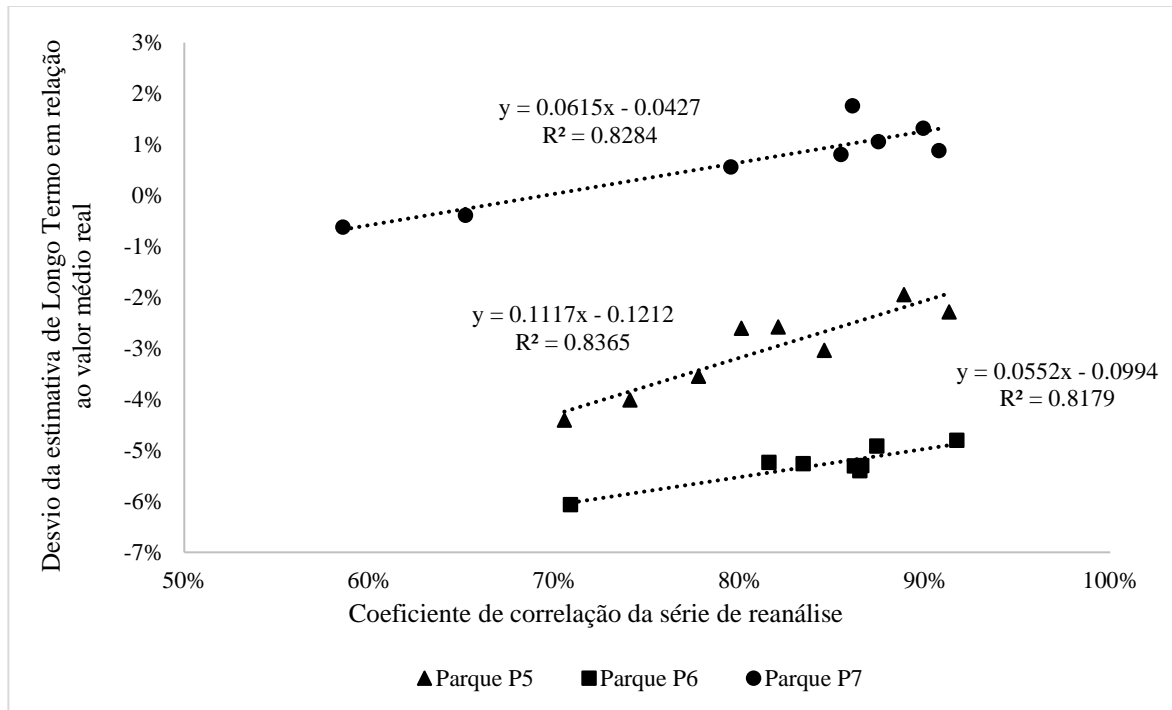


Figura 4.1. Relação entre o coeficiente de correlação (R^2) e o desvio percentual da estimativa de produção de Longo Termo em relação à média, para o parque P_P5, P_P6 e P_P7.

Tendo como objetivo verificar o cumprimento dos pressupostos do modelo de regressão linear, todos os parques e complexos foram submetidos a um conjunto de testes estatísticos que permitiram tirar conclusões relevantes para a validação da metodologia utilizada. Na Tabela A.1 do Anexo A estão apresentados os resultados das análises estatísticas realizadas a todos os parques e complexos em estudo. É de notar que para os complexos foi selecionado um dos parques que o constituem como sendo representativo de todo o complexo, tendo em consideração a sua localização (aproximadamente central) e características (nomeadamente apresentar a maior semelhança possível com o complexo).

A análise de variância (ANOVA) realizada mostrou que, para todos os parques e complexos, o modelo da regressão linear utilizado é estatisticamente significativo, uma vez que os valores de p são inferiores ao nível de significância para todos os grupos em estudo. O teste de Kolmogorov-Smirnov revelou que os resíduos de todos os parques e complexos seguem uma distribuição normal, dado que todos os valores tabelados são superiores aos valores calculados.

Segundo o *teste de Durbin-Watson*, não existe auto correlação nos resíduos na maior parte dos parques, dado que os valores são próximos do valor 2 (considerando como próximo a gama de valores entre 1,5 e 2,5), à exceção dos parques P_P2 e P_E7. Por fim, o *teste de Breusch-Pagan* demonstrou que existem cinco parques e um complexo (P_P4, P_P5, P_P6, P_E4, P_R1 e C_E12) que apresentam evidências de variâncias não constantes. Apesar de existirem alguns parques que violam os pressupostos da auto-correlação nos resíduos e de variâncias constantes, é possível afirmar que, de uma forma geral, a metodologia pode ser validada.

Para resolver o problema da auto correlação dos resíduos, seria necessário avaliar se existem outros parâmetros a influenciar a variável dependente, além da própria variável independente, e no caso de existirem, incluí-los no modelo de regressão linear e realizar novamente o *teste de Durbin-Watson* para analisar se o problema foi resolvido. No caso das variâncias não constantes, seria necessário transformar a equação que representa o modelo da regressão linear para o tipo logarítmico e repetir o teste de o *teste de Breusch-Pagan*, com o objetivo de perceber se o problema das variâncias não constantes foi eliminado [32]. Estas soluções não foram testadas na presente tese, devido ao período de trabalho disponível, apresentando-se como uma limitação.

As conclusões dos dois momentos de validação da metodologia utilizada sugerem que, de uma forma geral para todos os parques e complexos, os pressupostos da regressão linear são cumpridos e que o coeficiente de correlação R^2 mostra-se como uma métrica confiável para a seleção adequada da série de reanálise a utilizar para o cálculo da estimativa de produção de energia de Longo Termo.

4.2. Análise da influência do tempo de correlação

Esta análise permitiu uma avaliação mais detalhada da precisão das estimativas de Longo Termo obtidas através de correlações obtidas com diferentes bases temporais de dados. Com esta análise pretendeu-se compreender a confiabilidade das estimativas e identificar padrões de comportamento à medida que o número de anos de dados operacionais aumenta.

Na Figura 4.2 é mostrada a representação gráfica do desvio da estimativa de produção de Longo Termo em relação à média real, para as 630 combinações realizadas, de 1 a 20 anos, para os parques P_P5, P_P6 e P_P7, respetivamente, sendo que os dados que lhe deram origem estão descritos na Tabela B.1 a B.3 do Anexo B.1. À medida que o número de anos considerados na correlação aumenta, a amplitude interquartil tende a diminuir de forma geral, indicando uma redução na dispersão dos dados. Essa tendência era esperada, uma vez que o aumento do período de análise pode resultar numa menor variabilidade dos dados, o que é refletido pela redução da amplitude interquartil, uma vez que, por um lado, à medida que o número de anos aumenta, a média dos dados tende a representar mais precisamente o comportamento médio do parque, o que suaviza as flutuações de curto prazo e também contribui para a redução da variabilidade; por outro lado, um maior número de dados operacionais tende a demonstrar melhor o regime de ventos do local, contribuindo igualmente para a menor variabilidade dos resultados e consequente menor dispersão. À medida que o número de anos aumenta, a presença de *outliers* tende a diminuir, indicando uma maior estabilidade nas estimativas, o que sugere que a estimativa de produção de energia de Longo Termo pode ser mais suscetível a valores atípicos quando baseada em períodos mais curtos. De uma forma geral, a análise revela uma tendência de redução gradual da dispersão à medida que o número de anos aumenta, sugerindo que, para obter estimativas de Longo Termo mais precisas e próximas da realidade, é recomendável utilizar o maior número possível de anos de dados operacionais. Relativamente à mediana é possível aferir que não existem diferenças significativas entre os grupos uma vez que, devido às combinações que foram realizadas, os efeitos negativos e positivos anulam-se, resultando em valores de medianas semelhantes. Apesar da tendência geral de diminuição de dispersão com o aumento do número de anos operacionais, existem alguns dados erróneos entre os quais o grupo dos 13 anos no parque P_P7 ou o grupo dos 8 anos no parque P_P6, que apresentam dispersão inferior aos grupos seguintes, por exemplo. Este comportamento atípico pode acontecer uma vez que apenas foram realizadas combinações de anos consecutivos e não

todas as combinações possíveis, apresentando-se como uma limitação do estudo, e que pode levar a resultados enviesados.

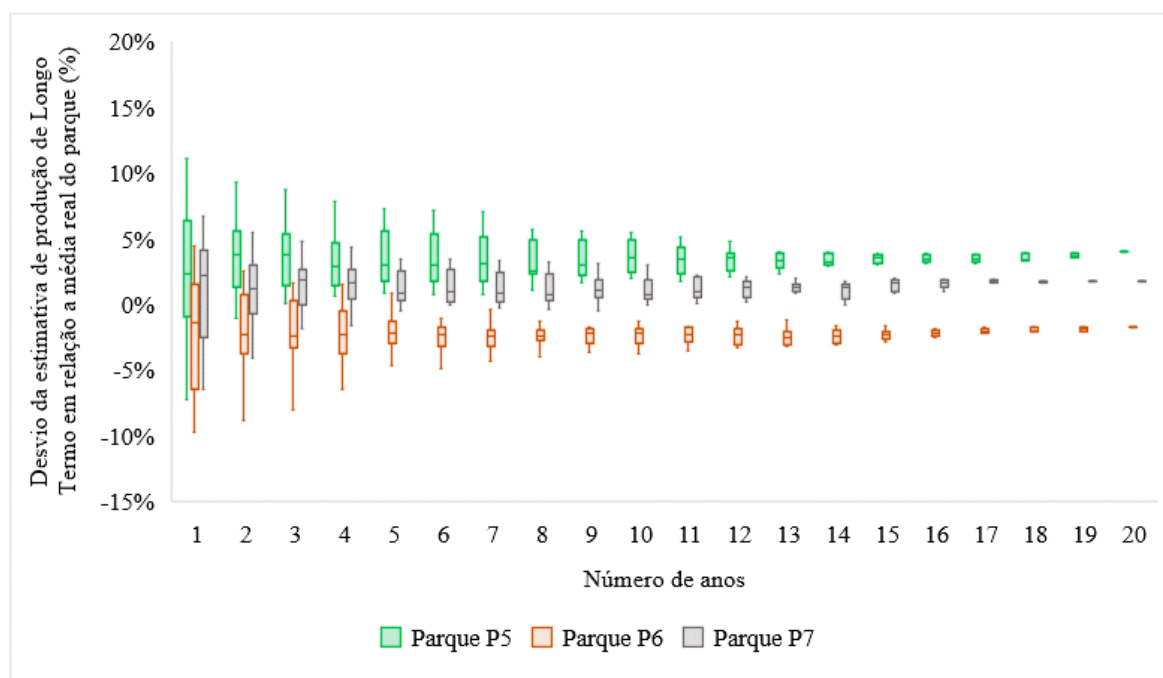


Figura 4.2. Desvio da estimativa de produção de Longo Termo em relação à média real do parque, para diferentes períodos de correlação, de 1 a 20 anos, para o parque P_P5, P_P6 e P_P7.

Uma alternativa interessante aos gráficos de blocos passa pela utilização de gráficos de dispersão sendo possível observar mais detalhadamente as flutuações extremas dos dados ao longo dos diferentes períodos de análise, realçando as variações mais acentuadas nos resultados e permitindo uma compreensão mais nítida das discrepâncias entre os grupos. A Figura 4.3 apresenta o gráfico de dispersão dos desvios máximos para cada um dos grupos, de 1 a 20 anos, para todos os parques com 20 anos de dados operacionais em estudo, nomeadamente os parques P_P5, P_P6 e P_P7, e os dados que lhe deram origem estão contemplados na Tabela B.4 do Anexo B.1. É observável que, de maneira congruente com as representações gráficas em blocos previamente analisadas, ocorre uma tendência de redução dos desvios em relação à média de Longo Termo. Este padrão reafirma a tendência geral de maior estabilidade e coesão nos resultados à medida que o período de análise se estende. Especificamente, é importante atentar no parque P_P5, no qual essa diminuição dos desvios se manifesta de maneira mais acentuada nas combinações de 1 a 4 anos. De facto, a observação de uma diminuição mais acentuada nos primeiros 4 anos dos desvios em relação à média de Longo Termo, particularmente no parque P_P5, sugere uma dinâmica temporal específica que merece uma análise mais detalhada. Diante dessa constatação, é plausível considerar a extensão desta análise para outros parques,

nomeadamente com um período de dados de operação mais pequeno, com o intuito de verificar se essa tendência é recorrente ou se está limitada a circunstâncias particulares do parque P_P5.

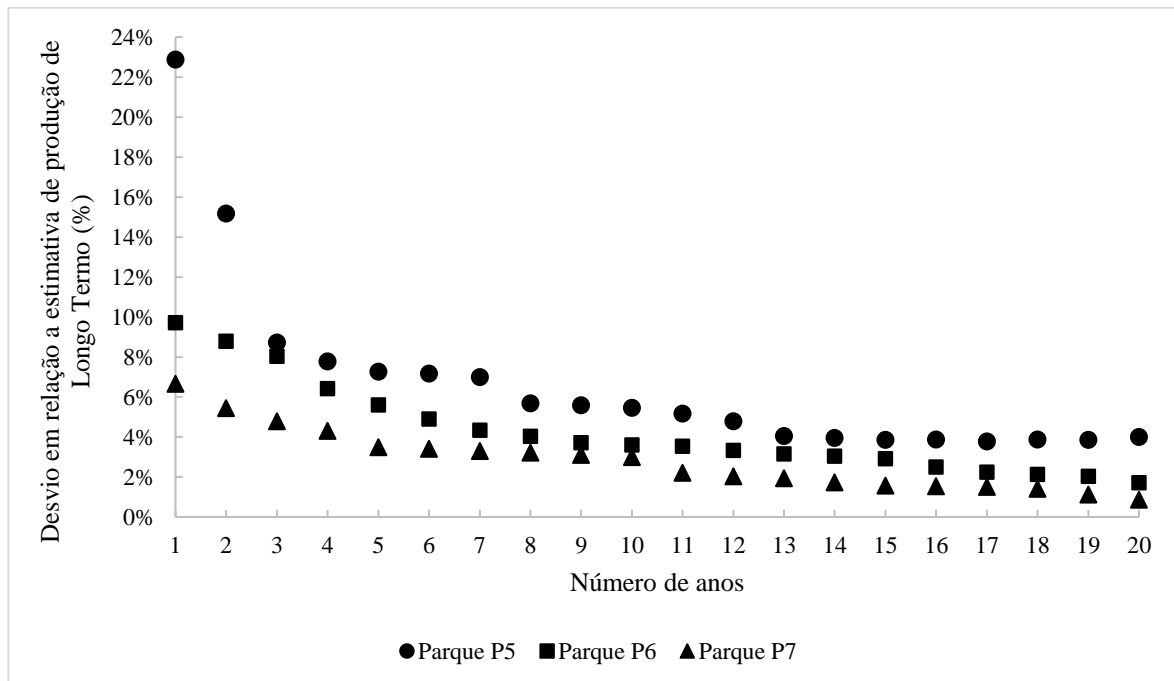


Figura 4.3. Desvios máximos para cada um dos grupos, de 1 a 20 anos, para todos os parques com 20 anos de dados operacionais em estudo (P_P5, P_P6 e P_P7).

Para além do estudo das diferenças nas estimativas de Longo Termo para cada um dos grupos (1 a 20 anos), foram analisados os erros RMBE e RRMSE com o objetivo de perceber qual é o comportamento que estas métricas de erro seguem à medida que o número de anos de dados operacionais aumenta. O que seria de esperar seria, tal como no desvio em relação ao Longo Termo, que as métricas de erro sigam uma tendência de diminuição da dispersão e haja uma aproximação do valor médio ao valor médio de 20 anos à medida que o número de anos aumenta uma vez que a inclusão de mais anos de dados ajuda a capturar melhor os padrões e tendências subjacentes dos dados, reduzindo assim o efeito de flutuações aleatórias e sazonais. Nas Figuras 4.4 e 4.5 estão representados os erros RMBE e RRMSE para todas as combinações de 1 a 20 anos realizadas para os parques P_P5, P_P6 e P_P7, respetivamente, sendo que os dados que lhe deram origem estão descritos nas Tabelas B.5 a B.10 do Anexo B.1. De uma forma geral, as métricas de erro apresentam o comportamento esperado, diminuindo a dispersão dos erros à medida que o número de anos de operação aumenta, havendo algumas oscilações provenientes de possíveis *outliers* nos dados. No caso do RRMBE, a diminuição da dispersão indica que as previsões estão, em média, mais próximas dos valores reais, com menos viés ao

longo dos anos, sugerindo que as estimativas estão a aproximar-se do valor médio real de forma mais consistente. Para o RRMSE, a diminuição da dispersão significa que as variações entre as previsões e os valores reais vão-se tornando cada vez mais estáveis e consistentes à medida que mais anos de dados são incorporados no estudo, indicando que as estimativas de produção de energia de Longo Termo se tornam mais precisas e confiáveis em relação à média real. Relativamente ao valor da mediana, para ambas métricas de erro, verificou-se que existem pequenas diferenças entre os grupos, mas que sugerem não ser muito significativas. A tendência geral é a de diminuição de dispersão com o aumento do número de anos operacionais, no entanto, existem alguns dados fora dessa tendência entre os quais o grupo dos 11 anos no parque P_P7 ou o grupo dos 8 anos no parque P_P6, que apresentam dispersão inferior aos grupos seguintes, por exemplo. Este comportamento anómalo pode acontecer (tal como no desvio da produção de Longo Termo em relação à média real do parque) já que não foram realizadas todas as combinações possíveis, apenas as combinações de anos consecutivos, apresentando-se como uma limitação do estudo, e que pode levar, tal como anteriormente, a resultados erróneos.

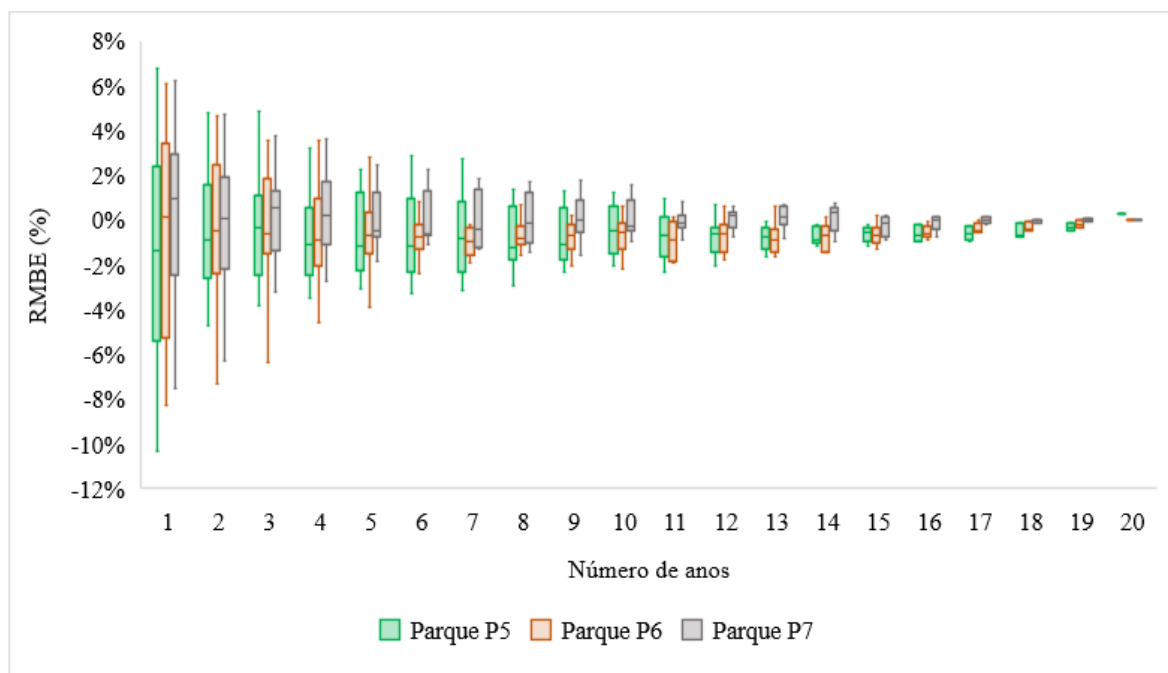


Figura 4.4. Representação gráfica do erro **RMBE**, para todas as combinações realizadas de 1 a 20 anos, para os **parques P_P5, P_P6 e P_P7**.

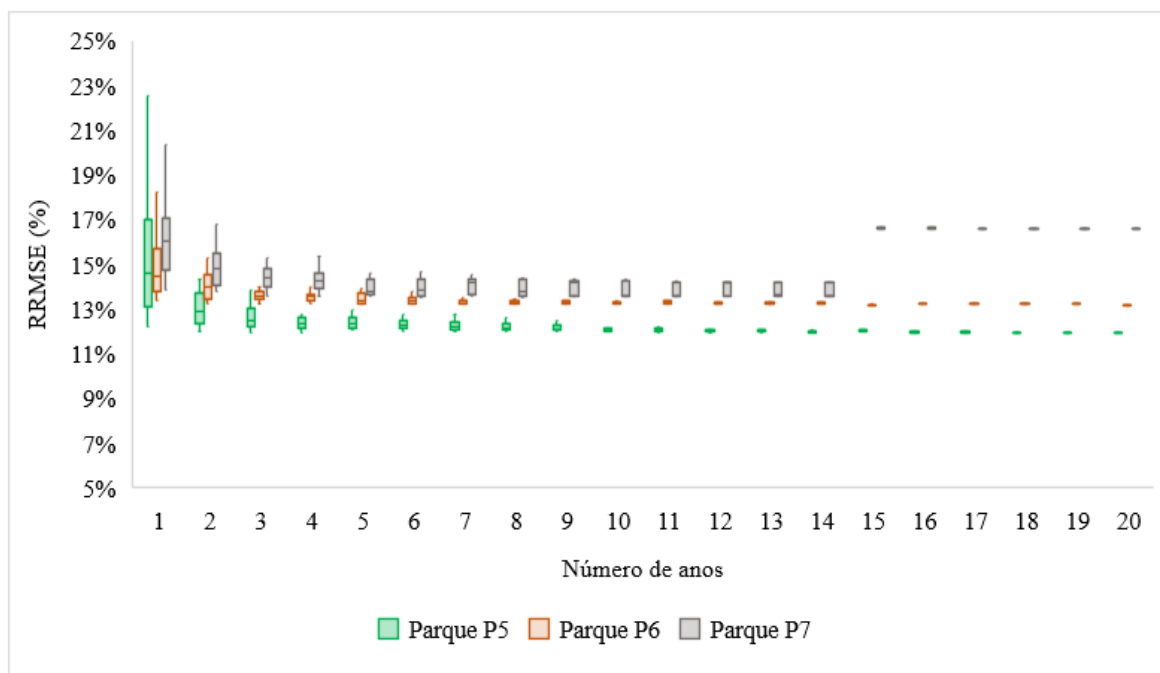


Figura 4.5. Representação gráfica do erro **RRMSE**, para todas as combinações realizadas de 1 a 20 anos, para os **parques P_P5, P_P6 e P_P7**.

Para fundamentar as conclusões relativas às análises gráficas apresentadas anteriormente, foram realizados alguns testes estatísticos aos erros RMBE e RRMSE para perceber se os grupos (1 a 20 anos) apresentam diferenças significativas, agregando os resultados obtidos para os três parques com 20 anos de operação. Foi realizado o *teste de Levene* para avaliar a homogeneidade das variâncias dos erros e com o *teste de Kolmogorov-Smirnov* foi analisada a normalidade da distribuição dos mesmos (realizado a cada um dos grupos estudados, de 1 a 20 anos). Dependendo do resultado destes testes, realiza-se o teste de hipótese adequado e a análise *post-hoc* respectiva. Na Tabela B.11 do Anexo B.1. estão descritos os resultados dos *testes de Levene e de Kolmogorov Smirnov*, assim como a escolha de teste de hipóteses realizada e respectiva análise *post-hoc*. Ambas as métricas de erro (RMBE e RRMSE) apresentam variâncias homogêneas (o valor de p é inferior ao nível de significância de 0,05). Relativamente à distribuição dos erros, todos os grupos da métrica RMBE seguem uma distribuição normal uma vez que o *teste de Kolmogorov-Smirnov* mostrou que todos os valores calculados são inferiores aos tabelados; para a métrica RRRMSE, o teste demonstrou que dois grupos (10 e 15 anos) não seguem uma distribuição normal. Assim, o teste mais adequado para o RRMBE é a *ANOVA tradicional*, seguida do *teste de Tukey*, e para o RRMSE, o *teste de Kruskal-Wallis* e o

teste de Dunn como análise *post-hoc*, já que existem dois grupos que não seguem uma distribuição normal.

Na Tabela B.12 do Anexo B.1 estão contemplados os resultados obtidos para a análise de variância (*ANOVA*) tradicional realizada ao grupo da métrica de erro RMBE e do *teste de Kruskal-Wallis* ao grupo da métrica RRMSE. Ambos os testes demonstraram que existe pelo menos um grupo que apresenta valor de mediana distinta dos restantes grupos, já que os valores de *p* são inferiores ao nível de significância estipulado. Assim, foram realizadas as análises *post-hoc* para ambas métricas de erro. Para a métrica de erro RMBE foi realizado o *teste de Tukey* com o objetivo de perceber quais grupos apresentam medianas diferentes, fazendo comparações entre todos os grupos, totalizando 190 comparações enquanto que, para o RRMSE foi efetuado o teste de Dunn com o mesmo objetivo e completando o mesmo número de comparações. Os resultados obtidos para os *testes de Tukey e Dunn* estão descritos nas Tabelas B.13 e B.14 do Anexo B.1, respetivamente. Para o RMBE, 21 comparações resultaram num valor de *p* inferior ao nível de significância estipulado, não se rejeitando a hipótese nula, pelo que existem diferenças significativas nas medianas dos grupos comparados. O que é possível observar é que à medida que o número de anos de operação aumenta, o número de comparações onde não se rejeita a hipótese nula são cada vez menores, tal como seria de esperar. Concretizando, com o grupo “1 ano” são realizadas 19 comparações, de entre as quais 18 não rejeitam a hipótese nula; para o grupo “2 anos” são realizadas 18 comparações, de entre as quais 2 não rejeitam a hipótese nula e com o grupo “3 anos” são realizadas 17 comparações, e neste caso apenas 1 não rejeita a hipótese nula. As observações que não rejeitam a hipótese nula sucedem até ao grupo dos 3 anos (inclusive), pelo que a partir dos 4 anos não se observa diferenças significativas entre as medianas deste grupo e os grupos de 5 a 19 anos. Relativamente ao RRMSE, não foi rejeitada a hipótese nula (existindo por isso diferenças significativas nas medianas dos grupos) para 38 das 190 comparações realizadas e neste caso as observações que não rejeitam a hipótese nula acontecem até ao grupo dos 4 anos (inclusive), pelo que a partir dos 5 anos não se observa diferenças significativas entre as medianas deste grupo e os grupos de 6 a 19 anos. As conclusões obtidas com ambos testes sugerem que as métricas de erro RMBE e RRMSE são mais sensíveis à menor quantidade de dados operacionais pelo que à medida que o número de dados operacionais aumenta, as flutuações nos erros tendem a estabilizar, pelo aumento do número de anos e possivelmente devido a uma maior consistência operacional e otimização da operação do parque. Assim, e tendo em consideração os resultados obtidos para a base de dados reduzida que foi utilizada, o número

mínimo de anos de dados operacionais para que não existam diferenças significativas na estimativa de produção de energia de Longo Termo são 4 anos. No entanto, do ponto de vista comercial, a escolha do período mínimo pode ser direcionada para o grupo dos 3 anos, uma vez que, em termos de refinanciamento de um parque eólico, a espera adicional de 1 ano para atingir os 4 anos pode resultar em perdas financeiras significativas.

O estudo progrediu em direção a uma análise mais detalhada, focando num número menor de dados operacionais, nomeadamente 4 anos, devido aos resultados obtidos na investigação feita para os parques com 20 anos de dados operacionais.

Nas Figuras 4.6 e 4.7 estão representados graficamente a distribuição das métricas de erro RMBE e RRMSE, para os parques P_E8 e P_P2, sendo que as representações gráficas dos restantes parques estão apresentadas nas Figuras B.1 a B.16 do Anexo B.2, assim como os dados que deram origem a todas as representações gráficas (na Tabela B.13). O objetivo deste estudo foi verificar com maior pormenor se existe uma diferença significativa entre os grupos de 1 a 4 anos de dados operacionais, sendo que o comportamento esperado é a diminuição da dispersão dos erros RMBE e RRMSE à medida que o número de anos aumenta e a mediana do grupo se vá aproximando do valor da mediana do grupo com 4 anos de dados operacionais. O que se verifica é que na maior parte dos parques, a dispersão dos dados para as duas métricas de erro segue o comportamento esperado, à exceção do parque P_P2 para o RMBE. Relativamente ao valor da mediana: no erro RMBE, apenas alguns parques seguem o comportamento esperado, nomeadamente os parques P_E6, P_E7 e P_E9; no RRMSE a generalidade dos parques segue o comportamento esperado. De uma forma geral para ambas métricas de erro, o que se verifica é que há uma melhoria no comportamento dos erros à medida que o número de anos aumenta, o que é um indicador de que a melhor escolha passará por ter como dados de entrada o maior número de anos de operação do parque possível. O que distingue este estudo com 4 anos de dados operacionais do anterior com 20 anos é facto de capturar nuances como é o caso do parque P_P2, que demonstra uma diferença significativa entre o período de 2 e 3 anos, pelo que é preferível utilizar os 3 anos como dados de entrada ao invés de 1 ou 2 anos, enquanto nos restantes parques não há uma diferença tão significativa entre grupos. Um ponto que é muito importante estudar é o grupo dos 4 anos: em primeiro lugar, reiterar o facto de o RMBE (embora seja apenas um valor) se situar muito próximo do 0 % em todos os parques, à exceção do P_P2 que ronda os 2 %, o que sugere que os valores estimados pelo modelo estão muito próximos dos valores reais dos parques, não apresentando viés

sistemático; relativamente ao RRMSE, este apresenta valores um pouco mais dispersos, situados na gama entre 7 e 15 %, indicando que a estimativa do modelo apresenta uma certa dispersão em relação aos valores reais. A combinação dos resultados obtidos para ambas métricas de erros pode indicar que o modelo de regressão está a conseguir, em média, capturar bem os valores reais, mas ainda há uma certa variabilidade nas estimativas individuais.

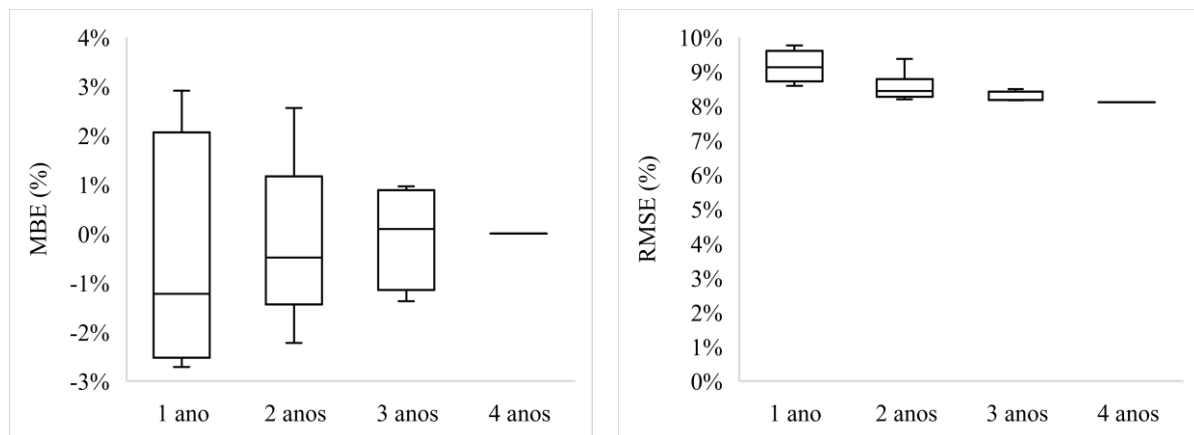


Figura 4.6. Representação gráfica do erro **RMBE** (à esquerda) e **RRMSE** (à direita), para todas as combinações realizadas de 1 a 4 anos, para o **parque P_E8**.

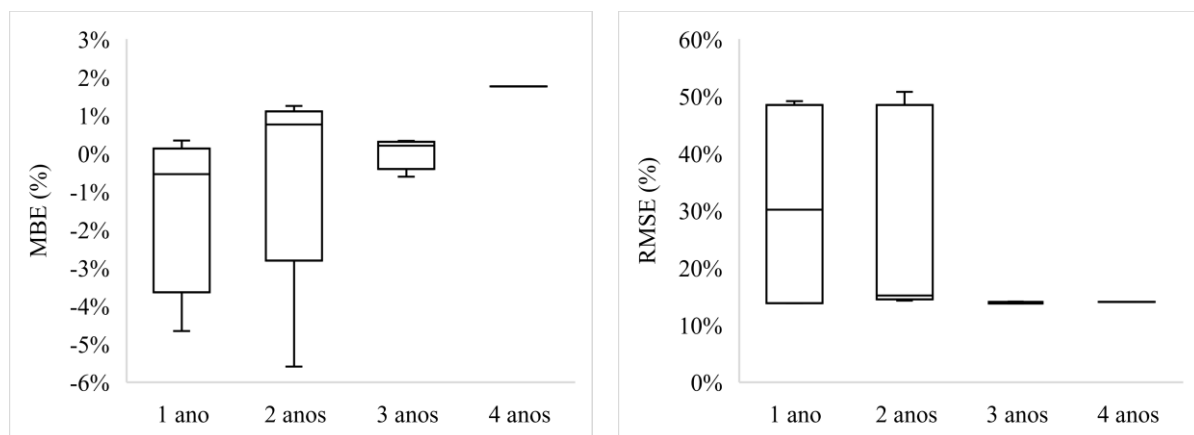


Figura 4.7. Representação gráfica do erro **RMBE** (à esquerda) e **RRMSE** (à direita), para todas as combinações realizadas de 1 a 4 anos, para o **parque P_P2**.

Na Tabela B.14 do anexo B.2 estão apresentados em detalhe os resultados dos *testes de Levene e Kolmogorov-Smirnov*, juntamente com a seleção do teste de hipótese apropriado e respetiva análise *post-hoc*. Verificou-se que o RMBE apresenta homogeneidade de variâncias (com valor de p inferior ao nível de significância de 0,05) e o RRMSE mostra variâncias não homogêneas. O teste de *Kolmogorov-Smirnov* revelou que o RMBE segue uma distribuição normal (onde o valor da estatística de teste calculada é superior ao valor tabelado), enquanto o

RRMSE apresenta uma distribuição não normal (com a estatística de teste calculada superior ao valor tabelado). Como resultado, optou-se por utilizar o *teste de Kruskal-Wallis* para as duas métricas de erro, sendo o *teste de Dunn* selecionado como método de análise *post-hoc* apropriado. É de notar que os testes *post-hoc* apenas são realizados no caso de os testes de hipóteses rejeitarem a hipótese nula.

Na Tabela B.15 do Anexo B.2 estão apresentados os resultados obtidos na *ANOVA tradicional* para o erro RMBE e no *teste de Kruskal-Wallis* para o RRMSE. O resultado esperado para estes testes seria haver uma diferença estatisticamente significativa entre os grupos, sendo mais considerável entre os grupos de 1 e 4 anos, tal como a análise visual realizada aos gráficos individuais anteriores demonstra. No entanto, os testes de hipóteses revelaram que, para o RMBE, as médias dos grupos são estatisticamente diferentes, rejeitando a hipótese nula; para o RRMSE, a hipótese nula do teste também foi rejeitada, levando à conclusão de que as medianas dos grupos são estatisticamente semelhantes. Esta é uma conclusão diferente da que seria de esperar, mas que pode ser justificada pela gama de valores de trabalho uma vez que, apesar de os gráficos individuais dos parques que representam RMBE e RRMSE apresentarem diferenças, essas não são significativas a nível estatístico devido à magnitude das variações e do tamanho da amostra, que é muito mais reduzido quando comparada com a amostra dos parques com 20 anos de dados operacionais.

Com esta análise aos parques com 4 anos de dados operacionais é possível concluir que existem diferenças e nuances entre os grupos, principalmente a nível da dispersão dos resultados, mas a nível estatístico essa diferença não se mostra significativa. Assim, apesar de as diferenças entre grupos não serem estatisticamente significativas, o grupo de 4 anos mostrou-se como um candidato muito interessante no que diz respeito aos valores de RMBE e RRMSE, pelo que a melhor decisão passará sempre por utilizar o maior número de anos possível para realizar uma estimativa de Longo Termo de um parque em operação.

4.3. Análise da influência da agregação espacial em complexos eólicos

A estimativa de produção de energia de um complexo eólico foi realizada de duas formas, alterando os dados de entrada, como foi referido no capítulo 3.5. O objetivo principal deste estudo não foi avaliar a influência da utilização dos dados do complexo ou dos parques individuais na estimativa de produção de Longo Termo, uma vez que não é possível comparar com a média real do complexo (já que nenhum dos complexos tem 20 anos de dados operacionais), mas sim perceber se as diferenças que existem entre a utilização de diferentes bases de dados têm impacto significativo na estimativa de Longo Termo.

A importância deste estudo prende-se com o facto de haver a possibilidade de não ser necessário estudar os parques eólicos que constituem um complexo de forma individual, e realizar apenas uma estimativa de produção de Longo Termo ao próprio complexo eólico, tornando mais simples e expedita a metodologia utilizada para analisar um complexo eólico, revestindo-se da maior importância em contexto empresarial.

Na Figura 4.8 está apresentado um gráfico de barras com duas séries de dados:

- a primeira, traduz o desvio entre a estimativa de produção de Longo Termo para o complexo, obtido através dos dados dos parques individuais com a série de reanálise mais adequada para o complexo, e o valor estimado de Longo Termo com dados do próprio complexo;
- a segunda, descreve o desvio entre o valor estimado de Longo Termo para o complexo, obtido com a melhor série para cada parque, e o valor estimado de Longo Termo com os dados do próprio complexo.

Os dados que originaram o gráfico da Figura 4.8 estão descritos nas Tabelas C.1 a C.8 do Anexo C. Para os parques C_B3, C_B6 e C_E12, a estimativa de produção de Longo Termo aproxima-se mais da estimativa realizada com os dados do complexo eólico quando é utilizada a mesma fonte de reanálise que a escolhida para o complexo eólico; para os complexos C_B4 e C_E11, a melhor reanálise para cada parque é a mesma que a utilizada para o complexo; para os restantes complexos, existe um desvio em relação à estimativa de Longo Termo superior quando é utilizada a melhor reanálise para cada parque. Para além disso, os desvios obtidos entre a estimativa de produção de energia de Longo Termo com os dados dos parques

individuais e com os do complexo eólico são baixos (o valor máximo observado de entre todos os complexos é inferior a 3%). Assim, é possível concluir que não existe uma diferença significativa entre utilizar a série de reanálise mais adequada para o complexo ou a melhor para cada parque. Para além disso, é exequível afirmar que não existe nenhuma vantagem evidente na realização de estudos aos parques individuais, pelo que os estudos poderão ser conduzidos com os dados do complexo, diminuindo assim o tempo despendido, já que a diferença para o estudo dos parques individuais é diminuta, apresentando uma ordem de grandeza menor que a incerteza típica do modelo e da estimativa de produção.

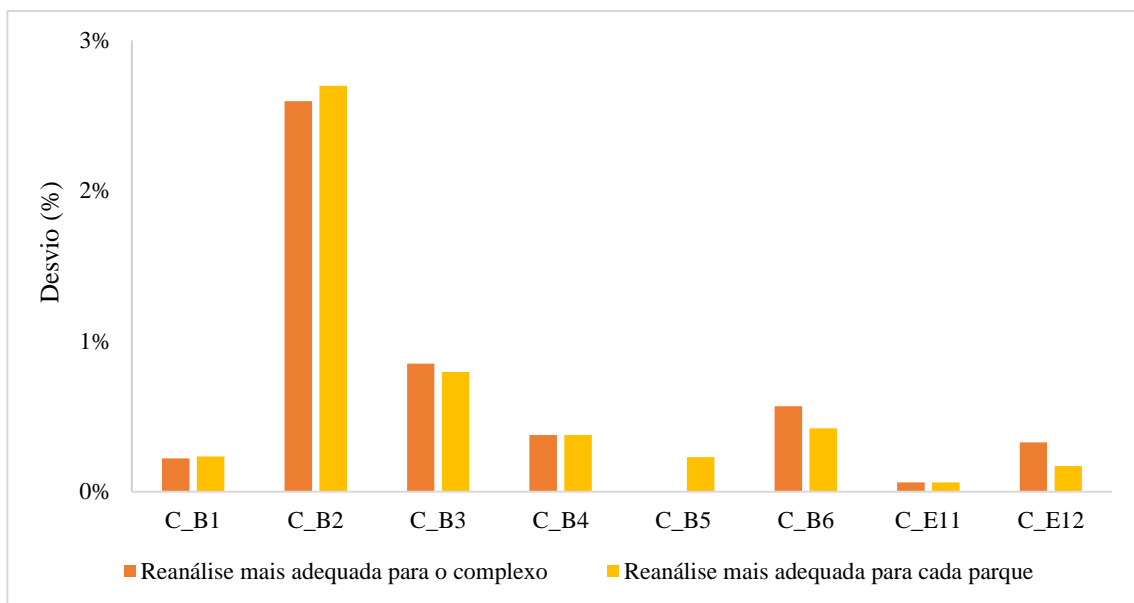


Figura 4.8. Desvio entre o valor estimado de Longo Termo para o complexo obtido através dos dados dos parques individuais com a melhor série para o complexo ou com a melhor série para cada parque e o valor estimado com os dados do próprio complexo.

Para fundamentar a análise gráfica apresentada anteriormente, foram conduzidos diversos testes estatísticos sobre as métricas de erro RMBE e RRMSE, visando compreender a possível divergência nas médias entre os grupos. A Tabela C.9 e C.10 do Anexo C lista os resultados dos testes de Levene e de Kolmogorov-Smirnov, juntamente com a seleção dos testes de hipóteses e a análise *post-hoc* correspondente, para a análise feita ao complexo com os dados do próprio complexo e os dados da série de reanálise mais adequada para o complexo, assim como a análise feita ao complexo com os dados do próprio complexo e os dados da série de reanálise mais adequada para cada parque, respetivamente. Ambas as métricas de erro (RRMBE e RRRMSE) revelam variações heterogêneas nas suas variâncias (o valor de p ultrapassa o nível de significância de 0,05) e uma distribuição normal para as duas fontes de dados (a estatística

de teste calculada é superior ao valor tabelado). Como tal, a *ANOVA de Welch* emerge como o teste de hipóteses mais conveniente para ambas as métricas de erro, com o *teste de Tukey* assumindo o papel de análise *post-hoc* caso a hipótese nula seja rejeitada no teste de hipótese.

Na Tabela C.11 do Anexo C estão apresentados os resultados obtidos na *ANOVA de Welch* para as duas métricas de erro, para os complexos em estudo. Conclui-se que, para ambos os erros, as medianas dos grupos são estatisticamente semelhantes uma vez que o valor de p é superior ao nível de significância. Este resultado reitera a conclusão de que não existem diferenças significativas entre a utilização dos dados do complexo ou dos dados dos parques individuais, quer a série de reanálise seja a melhor para o complexo quer seja a melhor para cada parque que constitui o complexo.

5. CONCLUSÕES E SUGESTÕES DE TRABALHOS FUTUROS

5.1. Conclusões

O objetivo deste estudo consistiu em identificar uma solução simples e eficaz para estimar a produção anual de energia elétrica de Longo Termo. O estudo efetuado resultou em diversas conclusões importantes no contexto da indústria da energia eólica, contribuindo para uma estimativa de produção confiável e próxima da realidade.

Em primeiro lugar, confirmou-se que o coeficiente de correlação R^2 é uma métrica eficaz para a seleção da série de reanálise apropriada, desempenhando um papel crucial na estimativa precisa de produção de energia de Longo Termo. Para além disso, os resultados demonstraram que o aumento do período de operação de um parque eólico está diretamente relacionado com uma redução na dispersão dos erros, indicando que a utilização de um conjunto de dados operacionais mais extenso resulta em estimativas mais precisas e próximas da realidade. Tendo em consideração os resultados obtidos para a base de dados reduzida que foi utilizada, o número mínimo de anos de dados operacionais para que não existam diferenças significativas na estimativa de produção de energia de Longo Termo são 4 anos. Contudo, sob uma perspetiva financeira, a opção pelo período mínimo pode ser orientada para o grupo de 3 anos, uma vez que aguardar mais 1 ano para atingir os 4 anos pode acarretar em prejuízos financeiros consideráveis no contexto de refinanciamento de um parque eólico.

No contexto de complexos eólicos, concluiu-se que a realização de estudos individuais para cada parque dentro do complexo não proporciona benefícios substanciais em relação ao uso de dados agregados do complexo, resultando em economia de tempo e recursos sem comprometer significativamente a precisão das estimativas.

O impacto deste estudo vai além do âmbito académico, tendo implicações práticas no setor de energia eólica. As conclusões obtidas são importantes para otimizar as estratégias para o cálculo da estimativa de produção de Longo Termo, contribuindo para decisões informadas na operação de parques eólicos e no planeamento de novos empreendimentos. O trabalho também destaca a importância de abordagens integradas que combinam análise gráfica,

métodos estatísticos e análises comparativas para uma compreensão abrangente e confiável das tendências em estudos de energia eólica.

5.2. Sugestão de trabalhos futuros

O presente estudo oferece conclusões relevantes para a melhoria das estimativas de produção de energia de Longo Termo pós-construção, no entanto, existem sugestões que podem levar a outras conclusões relevantes para o tópico.

No estudo dos parques com 20 anos de dados operacionais foram feitas combinações de anos consecutivos, resultando em 210 combinações para cada parque estudado. O facto de apenas terem sido realizadas combinações de anos consecutivos e não todas as combinações possíveis com 20 anos de dados é uma limitação do estudo, pelo que uma sugestão para trabalho futuro é a realização de todas as combinações possíveis e o estudo do impacto que essas combinações podem ter na estimativa de produção de energia de Longo Termo. Outra limitação do estudo é o facto de não terem sido estudadas as soluções no caso de existir auto correlação nos resíduos ou variâncias não constantes, pelo que a análise dessas soluções é mais uma sugestão de trabalho futuro interessante.

Este estudo concentrou-se na correlação entre dados de vento de Longo Termo e a produção de energia do parque. Um próximo passo relevante seria considerar se os dados climáticos regionais, como pressão atmosférica e a temperatura, influenciam a estimativa de produção de energia de Longo Termo. Outra recomendação seria realizar um estudo ainda mais abrangente, com uma base de dados com mais parques e complexos eólicos, principalmente parques que tenham 20 ou mais anos de dados operacionais.

BIBLIOGRAFIA

- [1] Comissão Mundial sobre Meio Ambiente e Desenvolvimento, 17 Maio 2023. [Online]. Available: <https://sustainabledevelopment.un.org/content/documents/5987our-common-future.pdf>.
- [2] Conferência das Nações Unidas, “Declaração do Rio sobre Ambiente e Desenvolvimento,” Junho 1992. [Online]. Available: https://apambiente.pt/sites/default/files/_A_APA/Cidadania_ambiental/AssuntosInternacionais/1992_Declaracao_Rio.pdf. [Acedido em 17 Maio 2023].
- [3] Conselho Europeu, “Acordo de Paris sobre alterações climáticas,” 2023. [Online]. Available: <https://www.consilium.europa.eu/pt/policies/climate-change/paris-agreement/>. [Acedido em 17 Maio 2023].
- [4] Parlamento Europeu, “O que é a neutralidade das emissões de carbono e como pode ser atingida até 2050?,” 2023 Abril 13. [Online]. Available: <https://www.europarl.europa.eu/news/pt/headlines/society/20190926STO62270/como-a-ue-podera-atingir-a-neutralidade-carbonica-ate-2050>. [Acedido em 17 Maio 2023].
- [5] FLORESTAS.PT, “O que significa carbono zero e neutralidade carbónica?,” 27 Abril 2021. [Online]. Available: <https://florestas.pt/saiba-mais/o-que-significa-carbono-zero-e-neutralidade-carbonica/>. [Acedido em 17 Maio 2023].
- [6] European Commission, “A European Green Deal,” [Online]. Available: https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/european-green-deal_en. [Acedido em 12 Fevereiro 2023].
- [7] Conselho Europeu, “Objetivo 55,” [Online]. Available: <https://www.consilium.europa.eu/pt/policies/green-deal/fit-for-55-the-eu-plan-for-a-green-transition/>. [Acedido em 12 Fevereiro 2023].
- [8] Agência portuguesa do ambiente, “Plano Nacional de Energia e Clima (PNEC),” 2021. [Online]. Available: <https://apambiente.pt/clima/plano-nacional-de-energia-e-clima-pnec>. [Acedido em 18 Maio 2023].
- [9] “Plano Nacional Energia e Clima 2021-2030,” [Online]. Available: <https://eurocid.mne.gov.pt/artigos/plano-nacional-energia-e-clima-2021-2030>. [Acedido em 12 Fevereiro 2023].
- [10] APA - Agência Portuguesa do Ambiente, “Emissões GEE,” [Online]. Available: <https://apambiente.pt/clima/emissoes-gee>. [Acedido em 12 Fevereiro 2023].
- [11] Comissão Europeia, “Causas das alterações climáticas,” [Online]. Available: https://climate.ec.europa.eu/climate-change/causes-climate-change_pt. [Acedido em 12 Fevereiro 2023].
- [12] APREN - Associação de Energias Renováveis, “Energias Renováveis,” [Online]. Available: <https://www.apren.pt/pt/energias-renovaveis/destaques>. [Acedido em 17 Maio 2023].

- [13] INEGI, “Evolução da Potência instalada em Portugal,” [Online]. Available: <https://www.apren.pt/pt/energias-renovaveis/potencia>. [Acedido em 15 Fevereiro 2023].
- [14] Observatório da Energia, “Metas PNEC 2030: Potência instalada de energias renováveis,” 9 Outubro 2021. [Online]. Available: <https://www.observatoriodaenergia.pt/pt/comunicar-energia/post/8617/metas-pnec-2030-potencia-instalada-de-energias-renovaveis>. [Acedido em 15 Fevereiro 2023].
- [15] Cabeólica, “Energia Eólica: O que é?,” [Online]. Available: <https://cabeolica.com/energia-eolica/>. [Acedido em 15 Fevereiro 2023].
- [16] R. S. d. N. I. B. S. João Paulo Minardi de Azevedo, “Energia eólica e os impactos ambientais: um estudo de revisão,” *Revista UNINGÁ*, vol. 51, pp. 101-106, 2017.
- [17] Enel - Green Power, “Todas as vantagens da energia eólica,” [Online]. Available: <https://www.enelgreenpower.com/pt/learning-hub/energias-renoveveis/energia-eolica/vantagens-energia-eolica>. [Acedido em 18 Maio 2023].
- [18] Portal Energia, “Vantagens e desvantagens da energia eólica,” [Online]. Available: <https://www.portal-energia.com/vantagens-desvantagens-da-energia-eolica/>. [Acedido em 18 Maio 2023].
- [19] IBERDROLA, “O que é uma turbina eólica e como funciona,” [Online]. Available: <https://www.iberdrola.com/sustentabilidade/pas-aerogeradores>. [Acedido em 13 Fevereiro 2023].
- [20] ENEL - Green Power, “Nova vida e novos materiais: o desafio da inovação e sustentabilidade do vento,” [Online]. Available: <https://www.enelgreenpower.com/pt/midias/news/2021/03/reciclagem-turbinas-eolicas>. [Acedido em 13 Fevereiro 2023].
- [21] MEGAJOULE - Rethinking Energy Worldwide, “Quem somos,” [Online]. Available: <https://megajoule.pt/pt/aboutmegajoule/>. [Acedido em 5 Fevereiro 2023].
- [22] J. Carneiro, “Energia Eólica - 3º Ano da Licenciatura em Ciências do Ambiente,” 9 Janeiro 2013. [Online]. Available: <https://hdl.handle.net/1822/22408>. [Acedido em 10 Fevereiro 2023].
- [23] METEORED, “A circulação geral da atmosfera terrestre,” [Online]. Available: <https://www.tempo.com/noticias/ciencia/circulacao-geral-da-atmosfera-terrestre.html>. [Acedido em 13 Fevereiro 2023].
- [24] Science Sparks, “What is a sea breeze?,” [Online]. Available: <https://www.science-sparks.com/what-is-a-sea-breeze/>. [Acedido em 10 fevereiro 2023].
- [25] Global Wind Atlas, “Methodology - Orography,” [Online]. Available: <https://globalwindatlas.info/en/about/method>.
- [26] D. S. N. J. E. B. Tony Burton, “The Nature of the Wind,” em *Wind Energy Handbook*, JOHN WILEY & SONS, LTD, 2001, pp. 11-12.
- [27] D. S. N. J. E. B. Tony Burton, “GEographical Variation in the Wind Resource,” em *Wind Energy Handbook*, JOHN WILEY & SONS, LTD, 2001, pp. 12-13.

- [28] Climate Change Service, “Climate reanalysis,” [Online]. Available: <https://climate.copernicus.eu/climate-reanalysis>. [Acedido em 10 Fevereiro 2023].
- [29] S. P. H. H. J. S. I. S. C. B. K. G. A. N. H. M. J. M. K. N. L. X. G. L. J. L. B. P. Russell McKenna, “High-resolution large-scale onshore wind energy assessments: A review of potential definitions, methodologies and future research needs,” em *Renewable Energy*, ELSEVIER, 2022, pp. 659-684.
- [30] J. H. O. U. K. V. Johannes Lindvall, “Post-construction production assessment of wind farms - Assessment and optimization of the energy production of operational wind farms: Part 1,” ENERGIFORSK, 2016.
- [31] J. Fox, *Applied Regression Analysis and Generalized Linear Models*, Sage Publications, 2015.
- [32] U. A. Vivek Krishnamoorthy, “Linear Regression: Assumptions and Limitations,” Abril 2022. [Online]. Available: <https://blog.quantinsti.com/linear-regression-assumptions-limitations/>. [Acedido em 2023 julho 25].
- [33] L. B. Damas, “Aplicação da metodologia MCP para análise da geração em dois parques eólicos em operação no Brasil,” Florianópolis, 2013.
- [34] Department of Statistics and Data Science - Yale University, “ANOVA for Regression,” [Online]. Available: <http://www.stat.yale.edu/Courses/1997-98/101/anovareg.htm>. [Acedido em 15 julho 2023].
- [35] H. W. Lilliefors, “On the Kolmogorov-Smirnov Test for Normality with Mean and Variance Unknown,” *Journal of the American Statistical Association*, vol. 62, pp. 339-402, 1967.
- [36] W. Kenton, “Durbin Watson Test: What It Is in Statistics, With Examples,” maio 2023. [Online]. Available: <https://www.investopedia.com/terms/d/durbin-watson-statistic.asp>. [Acedido em julho 2023].
- [37] STATOLOGY, “The Breusch-Pagan Test: Definition & Example,” dezembro 2020. [Online]. Available: <https://www.statology.org/breusch-pagan-test/>. [Acedido em julho 2023].
- [38] C. J. & M. K. Willmott, “Advantages of the mean absolute error (MAE) over the root mean square error (RRMSE) in assessing average model performance,” *Climate research*, pp. 79-82, 2005.
- [39] J. I. Hoffman, “Hypothesis Testing: The Null Hypothesis, Significance, and Type I Error,” em *Biostatistics for Medical and Biomedical Practitioners*, Academic Press, 2015, pp. 153-158.
- [40] C. Andrade, “The P Value and Statistical Significance: Misunderstandings, Explanations, Challenges, and Alternatives,” em *Indian J Psychol Med.*, 2019, pp. 210-215.
- [41] W. M. Gastwirth Yulia R. Gel, “The Impact of Levene's Test of Equality of Variances on Statistical Theory and Practice,” *Institute of Mathematical Statistics*, vol. 24, pp. 343-360, 2009.

- [42] D. B. S. J.F. Reed III, “Robust alternatives to traditional analysis of variance: Welch W*, James II*, James III*, Brown-Forsythe BF*,” *Computer Methods and Programs in Biomedicine*, vol. 26, pp. 233-237, 1988.
- [43] O. O. a. J. K. Eva Ostertagová, “Methodology and Application of the Kruskal-Wallis Test,” *Applied Mechanics and Materials*, vol. 611, pp. 115-120, 2014.
- [44] L. J. W. Herv'e Abdi, “Tukey’s Honestly Significant Difference (HSD) Test,” *Encyclopedia of Research Design*, 2010.
- [45] A. Dinno, “Nonparametric Pairwise Multiple Comparisons in Independent Groups using Dunn's Test,” *The Stata Journal, University Portland*, pp. 292-300, 2015.
- [46] APAmbiente, “Agência Portuguesa do Ambiente,” 2021. [Online]. Available: <https://apambiente.pt/clima/roteiro-para-neutralidade-carbonica-2050>. [Acedido em 12 Fevereiro 2023].
- [47] H. L. S. d. Reis, “Estudo sobre a aplicação do método MCP em dados de longo período para a estimativa do recurso eólico e solar,” Brasília, 2013.
- [48] A. P. G. F. Ferreira, “Sensibilidade às parameterizações físicas do WRF nas previsões à superfície em Portugal Continental,” Universidade de Aveiro, Departamento de Física, 2007.
- [49] Global Wind Energy Council (GWEC), “Global Wind Report 2022,” 2023. [Online]. Available: <https://gwec.net/global-wind-report-2022/>. [Acedido em 10 Fevereiro 2023].
- [50] D. S. N. J. E. B. Tony Burton, “Long-term Wind speed Variations,” em *Wind Energy Handbook*, JOHN WILEY & SON , LTD, 2001, pp. 13-14.
- [51] A. J. Nunes, “Implementação e comparação de processos de medição da curva de potência de um aerogerador,” 2018.
- [52] GoldEnergy, “Energia eólica,” [Online]. Available: <https://goldenergy.pt/glossario/energia-eolica/>. [Acedido em 10 Fevereiro 2023].
- [53] V. Sohoni, “A Critical Review on Wind Turbine Power Curve Modelling Techniques and Their Applications in Wind Based Energy Systems,” *Journal of Energy Hindawi*, 2016.
- [54] R. V. Silva, “Estudo da Sensibilidade da Estimativa de vento no modelo WAsP®,” Recife, 2013.
- [55] EMD International, “WindPro,” [Online]. Available: <https://www.emd-international.com/windpro/>. [Acedido em 20 agosto 2023].

ANEXO A – VALIDAÇÃO DA METODOLOGIA

Neste anexo são apresentados os dados utilizados para a realização do segundo momento de validação da metodologia adotada.

Na Tabela A.1 estão apresentados os resultados das análises estatísticas realizadas a todos os parques e complexos em estudo. A coluna denominada Resultados descreve o resultado do teste, ou seja: para a análise de variância (*ANOVA*), utilizada para verificar a significância do modelo, o resultado pode ser “Estatisticamente significativo” quando o valor de p é inferior a 0,05 ou “Não estatisticamente significativo” no caso contrário; no caso do *teste de Kolmogorov-Smirnov*, empregado para avaliar a normalidade da distribuição dos dados, o resultado pode ser “Distribuição normal” quando o valor calculado é inferior ao tabelado ou “Distribuição não normal” no caso contrário; o *teste de Durbin-Watson* é usado para estudar a existência de autocorrelação nos resíduos e o resultado pode ser “Não existe auto correlação” quando o valor da estatística D é próximo de 2 ou “Existe auto correlação” no caso contrário e o *teste de Breusch-Pagan*, utilizado para verificar a variabilidade dos resíduos (o resultado pode ser “Variabilidade constante” quando o valor de p é inferior a 0,05 ou “Variabilidade não constante” no caso contrário).

Tabela A.1. Resultados das análises estatísticas realizadas para todos os parques e complexos em estudo, que incluíram a análise de variância (ANOVA) para verificar a significância do modelo (o resultado pode ser “Estatisticamente significativo” quando o valor de p é inferior a 0,05 ou “Não estatisticamente significativo” no caso contrário); o teste de Kolmogorov-Smirnov para avaliar a normalidade da distribuição dos dados (o resultado pode ser “Distribuição normal” quando o valor calculado é inferior ao tabelado ou “Distribuição não normal” no caso contrário); o teste de Durbin-Watson para estudar a existência de auto-correlação nos resíduos (o resultado pode ser “Não existe auto correlação” quando o valor da estatística D é próximo de 2 ou “Existe auto correlação” no caso contrário) e o teste de Breusch-Pagan para verificar a heteroscedasticidade dos resíduos (o resultado pode ser “Variabilidade constante” quando o valor de p é inferior a 0,05 ou “Variabilidade não constante” no caso contrário)

Parque/ Complexo	Análise de variância (ANOVA)		Teste de Kolmogorov-Smirnov			Teste de Durbin-Watson		Teste de Breusch-Pagan	
	Valor de p	Resultado	Tabelado	Calculado	Resultado	Estatística D	Resultado	Valor de p	Resultado
P_P1	<0,01	Estatisticamente significativo	0,207	0,074	Distribuição normal	1,70	Não existe auto correlação	0,716	Variabilidade constante
P_P2	<0,01	Estatisticamente significativo	0,196	0,069	Distribuição normal	2,53	Existe auto correlação	0,054	Variabilidade constante
P_P3	<0,01	Estatisticamente significativo	0,224	0,116	Distribuição normal	1,78	Não existe auto correlação	0,158	Variabilidade constante
P_P4	<0,01	Estatisticamente significativo	0,225	0,116	Distribuição normal	2,48	Não existe auto correlação	0,009	Variabilidade não constante
P_P5	<0,01	Estatisticamente significativo	0,097	0,037	Distribuição normal	1,68	Não existe auto correlação	0,0002	Variabilidade não constante
P_P6	<0,01	Estatisticamente significativo	0,094	0,079	Distribuição normal	1,84	Não existe auto correlação	0,001	Variabilidade não constante
P_P7	<0,01	Estatisticamente significativo	0,094	0,060	Distribuição normal	1,64	Não existe auto correlação	0,065	Variabilidade constante
P_E1	<0,01	Estatisticamente significativo	0,201	0,113	Distribuição normal	1,75	Não existe auto correlação	0,421	Variabilidade constante
P_E2	<0,01	Estatisticamente significativo	0,196	0,107	Distribuição normal	2,41	Não existe auto correlação	0,168	Variabilidade constante
P_E3	<0,01	Estatisticamente significativo	0,231	0,107	Distribuição normal	1,55	Não existe auto correlação	0,814	Variabilidade constante
P_E4	<0,01	Estatisticamente significativo	0,207	0,117	Distribuição normal	1,95	Não existe auto correlação	0,008	Variabilidade não constante

P_E5	<0,01	Estatisticamente significativo	0,201	0,081	Distribuição normal	1,67	Não existe auto correlação	0,522	Variabilidade constante
P_E6	<0,01	Estatisticamente significativo	0,210	0,080	Distribuição normal	2,38	Não existe auto correlação	0,278	Variabilidade constante
P_E7	<0,01	Estatisticamente significativo	0,198	0,094	Distribuição normal	1,02	Existe auto correlação	0,092	Variabilidade constante
P_E8	<0,01	Estatisticamente significativo	0,203	0,078	Distribuição normal	1,87	Não existe auto correlação	0,068	Variabilidade constante
P_E9	<0,01	Estatisticamente significativo	0,205	0,082	Distribuição normal	1,96	Não existe auto correlação	0,156	Variabilidade constante
P_E10	<0,01	Estatisticamente significativo	0,203	0,122	Distribuição normal	1,94	Não existe auto correlação	0,646	Variabilidade constante
P_E11	<0,01	Estatisticamente significativo	0,210	0,137	Distribuição normal	2,35	Não existe auto correlação	0,502	Variabilidade constante
P_R1	<0,01	Estatisticamente significativo	0,213	0,122	Distribuição normal	1,70	Não existe auto correlação	0,004	Variabilidade não constante
P_R2	<0,01	Estatisticamente significativo	0,215	0,123	Distribuição normal	1,58	Não existe auto correlação	0,191	Variabilidade constante
P_C1	<0,01	Estatisticamente significativo	0,207	0,077	Distribuição normal	2,12	Não existe auto correlação	0,312	Variabilidade constante
C_B1	<0,01	Estatisticamente significativo	0,194	0,075	Distribuição normal	1,19	Não existe auto correlação	0,444	Variabilidade constante
C_B2	<0,01	Estatisticamente significativo	0,189	0,079	Distribuição normal	1,43	Não existe auto correlação	0,351	Variabilidade constante
C_B3	<0,01	Estatisticamente significativo	0,250	0,107	Distribuição normal	1,19	Não existe auto correlação	0,144	Variabilidade constante
C_B4	<0,01	Estatisticamente significativo	0,198	0,091	Distribuição normal	1,59	Não existe auto correlação	0,152	Variabilidade constante
C_E12	<0,01	Estatisticamente significativo	0,203	0,077	Distribuição normal	1,92	Não existe auto correlação	0,040	Variabilidade não constante

C_E13	<0,01	Estatisticamente significativo	0,198	0,094	Distribuição normal	1,05	Não existe auto correlação	0,309	Variabilidade constante
-------	-------	--------------------------------	-------	-------	---------------------	------	----------------------------	-------	-------------------------

Na Tabela A.2 estão apresentados os dados relativos ao coeficiente de correlação de todas as séries de reanálise e o desvio entre a produção de Longo Termo estimada com cada uma das séries e a média real do parque, para os parques P_P5, P_P6 e P_P7, respetivamente, que deram origem à representação gráfica da Figura 4.1.

Tabela A.2. Dados relativos a o coeficiente de correlação de todas as séries de reanálise em estudo para cada parque, R², e o desvio entre a produção de Longo Termo estimada com cada uma das séries e a média real do parque, desvio (%), para os parque P_P5, P_P6 e P_P7.

Parque	R²	Desvio (%)
P5	80%	-2,6%
	78%	-3,5%
	82%	-2,6%
	74%	-4,0%
	91%	-2,3%
	89%	-1,9%
	71%	-4,4%
	85%	-3,0%
P6	87%	-5,3%
	71%	-6,1%
	83%	-5,3%
	86%	-5,3%
	87%	-4,9%
	82%	-5,2%
	92%	-4,8%
	86%	-5,4%
P7	87%	-5,3%
	71%	-6,1%
	83%	-5,3%
	86%	-5,3%
	87%	-4,9%
	82%	-5,2%
	92%	-4,8%
	91%	1,0%

ANEXO B – ANÁLISE DA INFLUÊNCIA DO TEMPO DE CORRELAÇÃO

Neste anexo são apresentados os dados relativos à análise da influência do tempo de correlação para os parques com 20 anos de operação na secção B.1 e para os parques com 4 anos de operação na secção B.2.

B.1. Parques com 20 anos de dados operacionais

Para o caso dos parques com 20 anos de operação, começou-se por analisar o desvio da estimativa de Longo Termo em relação à média real do parque, para todas as combinações estudadas. Nas Tabelas B.1 a B.3 estão descritos os dados que deram origem às representações gráficas da Figura 4.2 do capítulo 3. Em seguida foram analisados os desvios máximos entre a estimativa de produção de Longo Termo e a média real do parque, apresentados na Tabela B.4 e representados graficamente na Figura 4.3 do capítulo 3. Depois, foram analisados os erros RRMBE e RRRMSE para cada um dos parques e os dados que originaram as representações gráficas das Figuras 4.4 e 4.5 estão apresentados nas tabelas B.5, B.6 e B.7 e B.8, B.9 e B.10, respetivamente. Para fundamentar as conclusões, foram analisadas as métricas de erro recorrendo ao *teste de Levene* e *teste de Kolmogorov-Smirnov*, sendo que os resultados obtidos nestes testes estão descritos nas Tabelas B.11 e B.12. Por fim, o *teste de Dunn* foi utilizado como análise *post-hoc* e os resultados obtidos estão descritos na Tabela B.13 e B.14 para os erros RRMBE e RRRMSE, respetivamente.

Tabela B.1. Desvio da estimativa de Longo Termo em relação à média real do parque, para cada uma das combinações realizadas, de 1 a 20 anos, para o **parque P_P5**.

1 ano	2 anos	3 anos	4 anos	5 anos	6 anos	7 anos	8 anos	9 anos	10 anos	11 anos	12 anos	13 anos	14 anos	15 anos	16 anos	17 anos	18 anos	19 anos	20 anos
-6,3%	6,3%	3,7%	0,8%	7,3%	7,2%	7,0%	5,7%	5,6%	5,5%	5,2%	4,8%	4,1%	4,0%	3,9%	3,9%	3,6%	3,4%	3,5%	4,0%
4,5%	5,0%	4,5%	4,8%	6,5%	6,9%	5,3%	4,8%	5,4%	5,2%	4,8%	3,9%	3,7%	3,9%	3,7%	3,4%	3,2%	3,3%	3,9%	-
6,4%	2,2%	5,0%	6,8%	6,4%	5,4%	5,1%	5,7%	5,0%	4,9%	4,2%	3,7%	3,9%	3,7%	3,4%	3,1%	3,3%	3,9%	-	-
3,1%	4,5%	6,9%	7,8%	4,0%	4,9%	5,6%	4,9%	4,5%	4,0%	3,5%	3,7%	3,5%	3,2%	3,0%	3,2%	3,8%	-	-	-
5,9%	8,8%	7,3%	5,2%	4,4%	5,9%	5,1%	4,9%	3,8%	3,7%	3,8%	3,5%	3,2%	2,9%	3,1%	3,8%	-	-	-	-
11,1%	9,3%	6,5%	4,6%	5,9%	5,0%	4,5%	3,6%	3,3%	3,6%	3,3%	3,0%	2,7%	3,0%	3,6%	-	-	-	-	-
6,3%	3,4%	2,0%	4,1%	3,6%	2,4%	2,5%	2,3%	2,7%	2,5%	2,3%	2,0%	2,3%	3,1%	-	-	-	-	-	-
-1,6%	1,3%	3,9%	2,7%	1,8%	1,7%	1,9%	2,4%	2,3%	2,0%	1,8%	2,1%	2,9%	-	-	-	-	-	-	-
1,6%	5,6%	3,9%	4,0%	2,1%	2,4%	2,8%	2,6%	2,2%	2,0%	2,3%	3,2%	-	-	-	-	-	-	-	-
9,7%	4,8%	4,8%	2,9%	2,4%	2,9%	2,7%	2,3%	1,9%	2,4%	2,9%	-	-	-	-	-	-	-	-	-
-1,2%	3,8%	0,7%	0,6%	1,6%	1,6%	0,8%	1,1%	1,6%	2,7%	-	-	-	-	-	-	-	-	-	-
0,8%	1,3%	1,0%	2,1%	1,9%	1,5%	1,2%	1,8%	2,4%	-	-	-	-	-	-	-	-	-	-	-
-1,8%	-0,5%	1,6%	1,2%	0,9%	0,7%	1,3%	2,6%	-	-	-	-	-	-	-	-	-	-	-	-
1,1%	3,9%	2,3%	1,6%	1,2%	1,9%	3,4%	-	-	-	-	-	-	-	-	-	-	-	-	-
7,8%	2,9%	1,9%	1,3%	1,5%	3,1%	-	-	-	-	-	-	-	-	-	-	-	-	-	-
1,1%	0,4%	0,1%	1,4%	3,5%	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
-7,3%	-1,1%	0,9%	3,6%	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
-0,2%	2,6%	8,7%	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
6,4%	15,2%	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
22,9%	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Tabela B.2. Desvio da estimativa de Longo Termo em relação à média real do parque, para cada uma das combinações realizadas, de 1 a 20 anos, para o **parque P_P6**.

1 ano	2 anos	3 anos	4 anos	5 anos	6 anos	7 anos	8 anos	9 anos	10 anos	11 anos	12 anos	13 anos	14 anos	15 anos	16 anos	17 anos	18 anos	19 anos	20 anos
3,6%	1,7%	1,3%	1,5%	1,9%	1,4%	-0,3%	-1,3%	-1,8%	-1,3%	-1,8%	-1,2%	-1,2%	-1,7%	-1,6%	-1,8%	-1,7%	-1,7%	-1,7%	-1,7%
-0,5%	0,1%	0,7%	1,4%	0,9%	-1,1%	-2,1%	-2,5%	-1,9%	-2,4%	-1,8%	-1,7%	-2,1%	-2,0%	-2,3%	-2,1%	-2,1%	-2,0%	-2,0%	-
-2,7%	0,8%	1,7%	1,1%	-1,2%	-2,3%	-2,8%	-2,1%	-2,6%	-1,9%	-1,8%	-2,3%	-2,1%	-2,4%	-2,2%	-2,2%	-2,2%	-2,1%	-	-
1,9%	2,6%	1,5%	-1,4%	-2,7%	-3,2%	-2,4%	-2,9%	-2,1%	-2,0%	-2,5%	-2,3%	-2,5%	-2,4%	-2,4%	-2,3%	-2,2%	-	-	-
3,3%	1,2%	-2,7%	-4,0%	-4,3%	-3,2%	-4,3%	-2,6%	-2,4%	-3,0%	-2,7%	-3,0%	-2,7%	-2,7%	-2,6%	-2,5%	-	-	-	-
-1,6%	-5,8%	-6,6%	-6,4%	-4,7%	-4,9%	-3,6%	-3,2%	-3,7%	-3,4%	-3,6%	-3,3%	-3,2%	-3,0%	-2,9%	-	-	-	-	-
-9,7%	-8,8%	-8,0%	-5,5%	-5,6%	-4,0%	-3,5%	-4,0%	-3,6%	-3,8%	-3,5%	-3,3%	-3,2%	-3,0%	-	-	-	-	-	-
-8,7%	-7,1%	-3,8%	-4,4%	-2,6%	-2,4%	-3,1%	-2,7%	-3,1%	-2,8%	-2,7%	-2,5%	-2,5%	-	-	-	-	-	-	-
-6,4%	-1,8%	-3,2%	-1,2%	-1,1%	-2,2%	-1,9%	-2,4%	-2,2%	-2,1%	-2,0%	-2,0%	-	-	-	-	-	-	-	-
4,4%	-2,0%	0,1%	-0,1%	-1,6%	-1,3%	-2,0%	-1,8%	-1,8%	-1,7%	-1,8%	-	-	-	-	-	-	-	-	-
-7,0%	-1,6%	-1,3%	-2,8%	-2,2%	-2,8%	-2,5%	-2,4%	-2,3%	-2,2%	-	-	-	-	-	-	-	-	-	-
4,4%	1,7%	-1,3%	-0,9%	-2,0%	-1,7%	-1,7%	-1,7%	-1,7%	-	-	-	-	-	-	-	-	-	-	-
-0,9%	-4,1%	-2,8%	-3,6%	-2,9%	-2,7%	-2,5%	-2,4%	-	-	-	-	-	-	-	-	-	-	-	-
-7,6%	-3,8%	-4,4%	-3,3%	-3,0%	-2,7%	-2,6%	-	-	-	-	-	-	-	-	-	-	-	-	-
0,3%	-3,0%	-2,0%	-2,0%	-1,9%	-1,9%	-	-	-	-	-	-	-	-	-	-	-	-	-	-
-5,0%	-2,9%	-2,7%	-2,3%	-2,2%	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
-1,2%	-2,3%	-2,3%	-3,1%	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
-1,8%	-3,5%	-2,7%	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
-6,4%	-3,6%	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
-0,8%	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Tabela B.3. Desvio da estimativa de Longo Termo em relação à média real do parque, para cada uma das combinações realizadas, de 1 a 20 anos, para o **parque P_P7**.

1 ano	2 anos	3 anos	4 anos	5 anos	6 anos	7 anos	8 anos	9 anos	10 anos	11 anos	12 anos	13 anos	14 anos	15 anos	16 anos	17 anos	18 anos	19 anos	20 anos
-3,7%	-0,7%	-1,9%	-1,6%	-0,4%	0,2%	-0,3%	-0,4%	-0,5%	0,0%	0,0%	0,2%	0,1%	-0,1%	0,9%	1,0%	1,6%	1,8%	1,7%	1,8%
2,2%	0,7%	-0,8%	1,4%	1,0%	0,3%	0,1%	-0,1%	0,4%	0,4%	0,5%	0,4%	1,2%	0,4%	1,0%	1,6%	1,8%	1,8%	1,8%	-
-3,8%	-1,2%	2,1%	3,4%	0,2%	0,0%	-0,2%	0,4%	1,2%	0,6%	1,4%	1,3%	1,1%	1,4%	1,5%	1,7%	1,6%	1,6%	-	-
2,1%	2,1%	2,3%	0,9%	0,5%	0,1%	0,8%	0,8%	0,9%	0,7%	0,4%	0,7%	0,9%	1,2%	1,9%	1,9%	1,9%	-	-	-
4,6%	3,9%	1,5%	0,8%	0,3%	1,0%	1,0%	1,1%	0,9%	0,6%	0,8%	1,0%	1,4%	1,5%	1,9%	1,9%	-	-	-	-
2,2%	-0,6%	3,2%	-0,8%	0,3%	0,3%	0,7%	0,3%	1,0%	0,8%	0,6%	1,6%	1,3%	1,7%	1,8%	-	-	-	-	-
3,5%	-2,5%	-1,9%	0,1%	-0,5%	0,2%	0,2%	0,3%	0,1%	0,4%	1,1%	1,6%	1,9%	1,2%	-	-	-	-	-	-
-3,8%	-3,3%	-1,3%	0,2%	0,5%	0,8%	0,8%	0,6%	1,1%	1,6%	2,1%	2,0%	1,5%	-	-	-	-	-	-	-
-6,1%	-0,4%	0,2%	0,6%	0,9%	1,0%	0,8%	1,2%	1,8%	2,3%	2,2%	1,8%	-	-	-	-	-	-	-	-
4,7%	3,2%	2,5%	2,5%	2,3%	1,8%	2,3%	2,8%	3,1%	3,0%	2,0%	-	-	-	-	-	-	-	-	-
4,2%	-4,1%	2,0%	1,8%	1,3%	2,0%	2,6%	3,2%	2,9%	1,8%	-	-	-	-	-	-	-	-	-	-
-6,5%	2,5%	2,4%	1,8%	0,7%	2,7%	3,3%	3,0%	1,9%	-	-	-	-	-	-	-	-	-	-	-
2,1%	1,9%	1,2%	1,6%	2,7%	3,4%	3,1%	1,9%	-	-	-	-	-	-	-	-	-	-	-	-
2,2%	1,1%	1,1%	2,8%	2,8%	3,3%	2,4%	-	-	-	-	-	-	-	-	-	-	-	-	-
3,8%	3,0%	4,8%	4,0%	3,5%	3,1%	-	-	-	-	-	-	-	-	-	-	-	-	-	-
6,7%	5,2%	4,3%	4,3%	3,0%	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
6,4%	5,4%	3,7%	2,1%	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
1,8%	1,5%	1,8%	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
1,8%	1,2%	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
1,2%	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Tabela B.4. Dados relativos aos desvios máximos entre a estimativa de produção de Longo Termo e a média real do parque, para os parques P_P5 a P_P7

Nº de anos	Parque P_P5	Parque P_P6	Parque P_P7
1	22,9%	9,7%	6,7%
2	15,2%	8,8%	5,4%
3	8,7%	8,0%	4,8%
4	7,8%	6,4%	4,3%
5	7,3%	5,6%	3,5%
6	7,2%	4,9%	3,4%
7	7,0%	4,3%	3,3%
8	5,7%	4,0%	3,2%
9	5,6%	3,7%	3,1%
10	5,5%	3,6%	3,0%
11	5,2%	3,5%	2,2%
12	4,8%	3,3%	2,0%
13	4,1%	3,2%	1,9%
14	4,0%	3,0%	1,7%
15	3,9%	2,9%	1,6%
16	3,9%	2,5%	1,5%
17	3,8%	2,2%	1,5%
18	3,9%	2,1%	1,4%
19	3,9%	2,0%	1,1%

Tabela B.5. Dados relativos ao erro **RMBE** calculado para cada uma das combinações realizadas, de 1 a 20 anos, para o **parque P_P5**.

1 ano	2 anos	3 anos	4 anos	5 anos	6 anos	7 anos	8 anos	9 anos	10 anos	11 anos	12 anos	13 anos	14 anos	15 anos	16 anos	17 anos	18 anos	19 anos	20 anos
-9,4%	2,1%	-0,3%	-2,6%	2,2%	2,9%	2,8%	1,4%	1,3%	1,2%	0,9%	0,7%	-0,1%	-0,2%	-1,3%	-1,3%	-1,6%	-1,8%	-1,6%	-0,8%
1,2%	1,3%	0,1%	0,4%	2,1%	2,5%	0,9%	0,4%	1,0%	0,8%	0,6%	-0,2%	-0,5%	-0,3%	-1,6%	-1,8%	-2,1%	-1,8%	-1,2%	-
2,0%	-2,2%	0,5%	2,4%	2,0%	0,9%	0,7%	1,4%	0,7%	0,6%	0,0%	-0,5%	-0,3%	-0,4%	-1,8%	-2,1%	-1,8%	-1,2%	-	-
-0,4%	0,0%	2,5%	3,2%	-0,5%	0,5%	1,3%	0,5%	0,3%	-0,2%	-0,7%	-0,5%	-0,6%	-0,9%	-2,3%	-2,0%	-1,3%	-	-	-
1,4%	4,4%	3,0%	0,7%	0,0%	1,6%	0,8%	0,7%	-0,3%	-0,5%	-0,4%	-0,6%	-0,9%	-1,2%	-2,0%	-1,3%	-	-	-	-
6,8%	4,8%	1,9%	0,2%	1,7%	0,7%	0,3%	-0,5%	-0,9%	-0,6%	-0,7%	-1,0%	-1,3%	-1,0%	-1,4%	-	-	-	-	-
1,9%	-1,2%	-2,4%	-0,1%	-0,7%	-1,6%	-1,6%	-1,8%	-1,4%	-1,5%	-0,7%	-2,0%	-1,7%	-0,9%	-	-	-	-	-	-
-6,7%	-3,1%	-0,3%	-1,5%	-2,1%	-2,3%	-2,2%	-1,7%	-1,8%	-2,0%	-1,8%	-1,9%	-1,0%	-	-	-	-	-	-	-
-2,3%	1,6%	-0,2%	0,0%	-1,8%	-1,7%	-1,2%	-1,4%	-1,7%	0,0%	-2,3%	-0,7%	-	-	-	-	-	-	-	-
5,5%	0,7%	0,8%	-1,1%	-1,7%	-1,1%	-1,3%	-1,7%	-2,1%	-1,6%	-1,6%	-	-	-	-	-	-	-	-	-
-5,9%	-0,9%	-3,3%	-3,5%	-2,4%	-2,4%	-3,2%	-3,0%	-2,3%	-1,2%	-	-	-	-	-	-	-	-	-	-
-2,4%	-2,6%	-3,0%	-1,9%	-2,0%	-2,4%	-2,7%	-2,1%	-1,4%	-	-	-	-	-	-	-	-	-	-	-
-5,7%	-4,7%	-2,6%	-2,8%	-3,1%	-3,3%	-2,6%	-1,2%	-	-	-	-	-	-	-	-	-	-	-	-
-3,5%	-0,4%	-1,6%	-2,3%	-2,8%	-2,0%	-0,4%	-	-	-	-	-	-	-	-	-	-	-	-	-
2,5%	-0,9%	-2,0%	-2,6%	-2,3%	-0,5%	-	-	-	-	-	-	-	-	-	-	-	-	-	-
-2,6%	-3,5%	-3,8%	-2,3%	-0,2%	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
-10,3%	-4,5%	-2,2%	0,6%	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
-4,5%	-1,1%	4,8%	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
3,5%	11,5%	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
18,3%	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Tabela B.6. Dados relativos ao erro **RRMSE** calculado para cada uma das combinações realizadas, de 1 a 20 anos, para o **parque P_P5**.

1 ano	2 anos	3 anos	4 anos	5 anos	6 anos	7 anos	8 anos	9 anos	10 anos	11 anos	12 anos	13 anos	14 anos	15 anos	16 anos	17 anos	18 anos	19 anos	20 anos
21,1%	13,7%	13,8%	17,7%	15,0%	12,6%	12,4%	12,3%	12,3%	12,2%	12,2%	12,0%	12,0%	12,0%	12,3%	12,3%	12,3%	12,3%	12,3%	12,2%
15,9%	13,9%	13,5%	12,5%	12,9%	12,8%	12,8%	12,6%	12,5%	12,4%	12,1%	12,0%	12,1%	12,1%	12,3%	12,3%	12,4%	12,3%	12,2%	-
12,7%	13,6%	12,7%	12,7%	12,6%	12,7%	12,5%	12,5%	12,4%	12,1%	12,1%	12,1%	12,1%	12,0%	12,3%	12,4%	12,3%	12,2%	-	-
17,2%	12,7%	12,7%	14,0%	12,7%	12,5%	12,4%	12,4%	12,1%	12,1%	12,1%	12,1%	12,0%	12,0%	12,4%	12,3%	12,2%	-	-	-
13,0%	13,3%	12,7%	12,7%	12,5%	12,4%	12,4%	12,1%	12,0%	12,1%	12,1%	12,0%	12,0%	12,0%	12,3%	12,2%	-	-	-	-
14,1%	14,3%	13,7%	12,4%	12,3%	12,3%	12,0%	12,0%	12,0%	12,0%	12,0%	12,0%	12,0%	12,0%	12,3%	-	-	-	-	-
13,6%	13,7%	13,7%	12,2%	12,3%	12,0%	12,1%	12,1%	12,1%	12,0%	12,0%	12,1%	12,0%	11,9%	-	-	-	-	-	-
17,3%	12,9%	12,3%	12,3%	12,1%	12,1%	12,2%	12,1%	12,0%	12,1%	12,1%	12,1%	12,0%	-	-	-	-	-	-	-
12,2%	12,0%	11,9%	11,9%	12,1%	12,0%	12,0%	12,0%	12,0%	12,1%	12,1%	12,0%	-	-	-	-	-	-	-	-
13,2%	12,0%	11,9%	12,0%	12,0%	12,0%	12,0%	12,0%	12,1%	12,0%	12,0%	-	-	-	-	-	-	-	-	-
15,1%	13,8%	12,4%	12,5%	12,2%	12,1%	12,3%	12,3%	12,1%	12,0%	-	-	-	-	-	-	-	-	-	-
16,0%	12,2%	12,3%	12,1%	12,1%	12,2%	12,2%	12,1%	12,2%	-	-	-	-	-	-	-	-	-	-	-
13,2%	12,9%	12,3%	12,2%	12,3%	12,4%	12,2%	12,1%	-	-	-	-	-	-	-	-	-	-	-	-
13,5%	12,5%	12,0%	12,1%	12,2%	12,1%	12,1%	-	-	-	-	-	-	-	-	-	-	-	-	-
16,2%	12,1%	12,1%	12,2%	12,4%	12,3%	-	-	-	-	-	-	-	-	-	-	-	-	-	-
12,6%	12,5%	12,5%	12,4%	12,3%	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
22,4%	12,8%	12,3%	12,4%	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
13,1%	12,4%	12,9%	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
16,5%	16,9%	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
22,5%	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Tabela B.7. Dados relativos ao erro **RMBE** calculado para cada uma das combinações realizadas, de 1 a 20 anos, para o **parque P_P6**.

1 ano	2 anos	3 anos	4 anos	5 anos	6 anos	7 anos	8 anos	9 anos	10 anos	11 anos	12 anos	13 anos	14 anos	15 anos	16 anos	17 anos	18 anos	19 anos	20 anos
6,1%	4,0%	3,4%	3,6%	3,9%	3,5%	1,7%	0,7%	0,2%	0,6%	0,1%	0,6%	0,6%	0,1%	-3,3%	-3,5%	-3,5%	-3,6%	-3,6%	-3,5%
1,6%	2,1%	2,4%	3,2%	2,8%	0,8%	-0,2%	-0,6%	-0,1%	-0,5%	0,0%	0,0%	-0,4%	-0,3%	-4,0%	-4,0%	-4,0%	-4,0%	-3,9%	-
-1,2%	2,5%	3,4%	3,0%	0,6%	-0,5%	-0,9%	-0,3%	-0,8%	-0,2%	-0,1%	-0,6%	-0,5%	-0,7%	-4,1%	-4,1%	-4,1%	-4,0%	-	-
4,0%	4,6%	3,5%	0,5%	-0,8%	-1,3%	-0,5%	-1,0%	-0,3%	-0,2%	-0,8%	-0,6%	-0,8%	-0,7%	-4,2%	-4,2%	-4,1%	-	-	-
5,2%	3,3%	-0,7%	-2,1%	-2,4%	-1,3%	-1,8%	-0,9%	-0,7%	-1,3%	-1,0%	-1,2%	-1,1%	-1,1%	-4,5%	-4,4%	-	-	-	-
0,7%	-4,0%	-4,9%	-4,6%	-2,8%	-3,1%	-1,9%	-1,6%	-2,1%	-1,7%	-1,9%	-1,7%	-1,6%	-1,4%	-4,8%	-	-	-	-	-
-8,3%	-7,3%	-6,4%	-3,7%	-3,9%	-2,4%	-1,9%	-2,5%	-2,0%	-2,2%	-1,9%	-1,8%	-1,6%	-1,5%	-	-	-	-	-	-
-7,4%	-5,3%	-2,0%	-2,7%	-1,0%	-0,7%	-1,5%	-1,2%	-1,4%	-1,2%	-1,9%	-1,0%	-0,9%	-	-	-	-	-	-	-
-4,2%	0,2%	-1,3%	0,5%	0,5%	-0,6%	-0,3%	-0,7%	-0,6%	0,0%	-1,2%	-0,4%	-	-	-	-	-	-	-	-
5,7%	-0,3%	1,6%	1,4%	-0,1%	0,2%	-0,4%	-0,3%	-0,3%	-0,3%	-0,5%	-	-	-	-	-	-	-	-	-
-5,6%	-0,2%	0,1%	-1,4%	-0,8%	-1,2%	-1,0%	-1,0%	-0,8%	-0,7%	-	-	-	-	-	-	-	-	-	-
5,6%	3,1%	0,1%	0,5%	-0,4%	-0,2%	-0,3%	-0,3%	-0,2%	-	-	-	-	-	-	-	-	-	-	-
0,6%	-2,7%	-1,3%	-1,9%	-1,3%	-1,2%	-1,0%	-0,8%	-	-	-	-	-	-	-	-	-	-	-	-
-6,3%	-2,3%	-2,7%	-1,8%	-1,6%	-1,3%	-1,0%	-	-	-	-	-	-	-	-	-	-	-	-	-
1,8%	-1,1%	-0,5%	-0,6%	-0,4%	-0,3%	-	-	-	-	-	-	-	-	-	-	-	-	-	-
-3,2%	-1,4%	-1,2%	-0,9%	-0,6%	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
-0,3%	-0,5%	-0,3%	-1,0%	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
-0,7%	-2,3%	-1,4%	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
-5,7%	-2,4%	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
1,0%	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Tabela B.8. Dados relativos ao erro **RRMSE** calculado para cada uma das combinações realizadas, de 1 a 20 anos, para o **parque P_P6**.

1 ano	2 anos	3 anos	4 anos	5 anos	6 anos	7 anos	8 anos	9 anos	10 anos	11 anos	12 anos	13 anos	14 anos	15 anos	16 anos	17 anos	18 anos	19 anos	20 anos
15,0%	14,1%	13,7%	13,7%	13,8%	13,7%	13,4%	13,3%	13,4%	13,3%	13,3%	13,2%	13,2%	13,2%	13,8%	13,9%	13,9%	13,9%	13,9%	13,9%
13,5%	14,3%	13,4%	13,6%	13,5%	13,3%	13,3%	13,3%	13,3%	13,3%	13,2%	13,2%	13,2%	13,2%	14,0%	14,0%	14,1%	14,1%	14,0%	-
17,9%	13,6%	13,7%	13,5%	13,2%	13,2%	13,3%	13,3%	13,3%	13,2%	13,2%	13,2%	13,2%	13,2%	14,1%	14,1%	14,1%	14,1%	-	-
13,8%	14,0%	13,7%	13,3%	13,3%	13,5%	13,4%	13,4%	13,2%	13,2%	13,2%	13,2%	13,2%	13,2%	14,1%	14,1%	14,1%	-	-	-
14,3%	13,7%	13,4%	13,5%	13,7%	13,5%	13,5%	13,3%	13,2%	13,3%	13,2%	13,3%	13,2%	13,2%	14,2%	14,2%	-	-	-	-
13,8%	14,2%	14,4%	14,5%	13,8%	13,8%	13,4%	13,3%	13,4%	13,3%	13,4%	13,3%	13,3%	13,3%	14,3%	-	-	-	-	-
15,8%	15,3%	15,0%	14,0%	13,9%	13,4%	13,3%	13,4%	13,4%	13,4%	13,4%	13,3%	13,3%	13,3%	-	-	-	-	-	-
15,2%	14,8%	13,6%	13,6%	13,2%	13,2%	13,3%	13,3%	13,3%	13,3%	13,3%	13,3%	13,2%	-	-	-	-	-	-	-
15,4%	13,4%	13,5%	13,2%	13,3%	13,2%	13,2%	13,2%	13,2%	0,0%	13,3%	13,2%	-	-	-	-	-	-	-	-
1,7%	13,2%	13,6%	13,5%	13,4%	13,4%	13,2%	13,3%	13,3%	13,4%	13,3%	-	-	-	-	-	-	-	-	-
14,4%	13,5%	13,4%	13,4%	13,4%	13,3%	13,3%	13,4%	13,4%	13,3%	-	-	-	-	-	-	-	-	-	-
16,6%	14,5%	13,6%	13,6%	13,2%	13,4%	13,4%	13,4%	13,3%	-	-	-	-	-	-	-	-	-	-	-
13,4%	13,5%	13,3%	13,4%	13,3%	13,3%	13,4%	13,3%	-	-	-	-	-	-	-	-	-	-	-	-
14,6%	13,4%	13,6%	13,3%	13,4%	13,4%	13,3%	-	-	-	-	-	-	-	-	-	-	-	-	-
13,8%	13,3%	13,3%	13,3%	13,4%	13,2%	-	-	-	-	-	-	-	-	-	-	-	-	-	-
13,9%	13,3%	13,4%	13,4%	13,2%	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
14,7%	14,5%	14,0%	16,6%	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
14,1%	17,1%	16,8%	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
18,2%	16,9%	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
20,2%	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Tabela B.9. Dados relativos ao erro **RMBE** calculado para cada uma das combinações realizadas, de 1 a 20 anos, para o **parque P_P7**.

1 ano	2 anos	3 anos	4 anos	5 anos	6 anos	7 anos	8 anos	9 anos	10 anos	11 anos	12 anos	13 anos	14 anos	15 anos	16 anos	17 anos	18 anos	19 anos	20 anos
-4,9%	-2,8%	-3,0%	-2,8%	-1,3%	-0,8%	-1,3%	-1,5%	-1,6%	-1,0%	-0,9%	-0,8%	-0,9%	-1,0%	-0,2%	0,0%	0,6%	0,7%	0,6%	0,7%
1,5%	-0,5%	-1,8%	0,6%	0,1%	-0,7%	-1,0%	-1,1%	-0,5%	-0,5%	-0,4%	-0,5%	-0,1%	-0,5%	0,2%	0,8%	0,8%	0,7%	0,8%	-
-6,8%	-2,2%	1,3%	2,4%	-0,8%	-1,1%	-1,2%	-0,5%	0,0%	-0,3%	0,1%	0,0%	-0,3%	0,1%	0,5%	0,5%	0,5%	0,6%	-	-
0,2%	1,5%	1,5%	-0,1%	-0,6%	-0,8%	-0,1%	-0,1%	0,0%	-0,2%	-0,4%	-0,2%	0,1%	0,5%	0,9%	0,8%	0,9%	-	-	-
4,5%	3,1%	0,6%	-0,2%	-0,6%	0,1%	0,2%	0,3%	0,0%	-0,3%	0,0%	0,2%	0,7%	0,8%	0,8%	0,9%	-	-	-	-
1,0%	-1,8%	1,1%	-2,0%	-0,7%	-0,6%	-1,2%	-0,6%	-0,5%	-0,8%	-0,2%	0,2%	0,4%	0,5%	0,7%	-	-	-	-	-
1,5%	-3,8%	-3,0%	-1,9%	-1,8%	-0,7%	-1,3%	-1,2%	-1,4%	-0,3%	-0,2%	0,2%	0,2%	0,3%	-	-	-	-	-	-
-3,4%	-6,3%	-3,2%	-1,1%	-1,0%	-0,7%	-0,7%	-0,9%	-0,3%	0,2%	-0,2%	0,6%	0,7%	-	-	-	-	-	-	-
-7,5%	-2,2%	-1,2%	-1,0%	-0,6%	-0,6%	-0,9%	-0,2%	0,4%	0,9%	0,7%	0,4%	-	-	-	-	-	-	-	-
2,9%	1,9%	0,9%	1,0%	0,7%	0,1%	0,9%	1,4%	1,8%	1,6%	0,8%	-	-	-	-	-	-	-	-	-
3,1%	-6,3%	0,5%	0,2%	-0,4%	0,7%	1,3%	1,7%	1,5%	1,0%	-	-	-	-	-	-	-	-	-	-
-7,5%	0,7%	0,5%	-0,1%	0,1%	1,3%	1,9%	1,6%	1,1%	-	-	-	-	-	-	-	-	-	-	-
0,8%	0,3%	-0,5%	0,4%	1,4%	2,1%	1,7%	1,1%	-	-	-	-	-	-	-	-	-	-	-	-
0,2%	-0,9%	0,5%	1,6%	2,0%	1,6%	1,5%	-	-	-	-	-	-	-	-	-	-	-	-	-
3,0%	2,4%	3,8%	3,3%	2,5%	2,3%	-	-	-	-	-	-	-	-	-	-	-	-	-	-
6,2%	4,7%	3,6%	3,6%	2,2%	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
4,9%	4,7%	3,0%	1,8%	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
1,3%	0,1%	0,5%	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
0,7%	0,1%	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
0,8%	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Tabela B.10. Dados relativos ao erro **RRMSE** calculado para cada uma das combinações realizadas, de 1 a 20 anos, para o **parque P_P7**.

1 ano	2 anos	3 anos	4 anos	5 anos	6 anos	7 anos	8 anos	9 anos	10 anos	11 anos	12 anos	13 anos	14 anos	15 anos	16 anos	17 anos	18 anos	19 anos	20 anos
14,5%	13,8%	14,0%	14,0%	13,6%	13,6%	13,7%	13,8%	13,7%	13,6%	13,6%	13,6%	13,6%	13,6%	16,2%	16,3%	16,2%	16,2%	16,2%	16,2%
18,8%	14,3%	14,7%	15,0%	13,6%	13,6%	13,7%	13,6%	13,6%	13,5%	13,5%	13,6%	14,2%	13,6%	16,3%	16,3%	16,2%	16,2%	16,2%	-
20,4%	14,8%	15,1%	14,6%	13,7%	13,8%	13,7%	13,6%	14,2%	13,5%	14,2%	14,2%	14,2%	14,2%	16,2%	16,2%	16,2%	16,2%	-	-
16,7%	13,8%	13,6%	13,6%	13,7%	13,6%	13,5%	13,5%	13,5%	13,5%	13,6%	13,6%	13,6%	13,6%	16,2%	16,2%	16,2%	-	-	-
16,1%	13,9%	13,6%	13,7%	13,6%	13,5%	13,6%	13,6%	13,6%	13,6%	13,6%	13,6%	13,6%	13,6%	16,2%	16,2%	-	-	-	-
14,1%	14,0%	17,1%	13,9%	13,6%	13,6%	16,7%	13,6%	14,3%	14,3%	13,6%	14,2%	13,6%	14,2%	16,2%	-	-	-	-	-
16,9%	14,9%	14,0%	16,8%	14,3%	13,6%	14,2%	14,3%	14,3%	13,6%	13,6%	14,2%	14,2%	13,6%	-	-	-	-	-	-
16,0%	17,8%	16,9%	14,2%	14,3%	14,3%	14,3%	14,4%	14,2%	14,2%	14,2%	14,2%	13,6%	-	-	-	-	-	-	-
16,1%	16,8%	14,2%	14,5%	14,4%	14,5%	14,6%	14,2%	14,2%	14,3%	14,2%	14,2%	-	-	-	-	-	-	-	-
16,9%	14,3%	14,6%	14,4%	14,6%	14,7%	14,2%	14,3%	14,4%	14,3%	14,2%	-	-	-	-	-	-	-	-	-
14,6%	21,7%	14,3%	14,5%	14,6%	14,2%	14,2%	14,3%	14,3%	13,6%	-	-	-	-	-	-	-	-	-	-
21,7%	15,1%	15,3%	15,3%	13,8%	14,3%	14,4%	14,3%	13,6%	-	-	-	-	-	-	-	-	-	-	-
14,2%	14,5%	14,7%	14,2%	14,2%	14,3%	14,3%	13,6%	-	-	-	-	-	-	-	-	-	-	-	-
15,8%	15,7%	13,9%	14,2%	13,7%	16,7%	13,6%	-	-	-	-	-	-	-	-	-	-	-	-	-
13,9%	14,2%	14,4%	14,0%	13,8%	13,7%	-	-	-	-	-	-	-	-	-	-	-	-	-	-
17,6%	14,9%	14,1%	14,6%	13,7%	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
15,2%	15,5%	14,7%	13,9%	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
16,0%	15,1%	14,4%	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
15,8%	14,0%	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
17,1%	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Tabela B.11. Resultados das análises estatísticas realizadas para todos os parques com 20 anos de dados operacionais, que incluíram o teste de Levene para analisar a homogeneidade das variâncias (o resultado pode ser “Variâncias homogêneas” quando o valor de p é inferior a 0,05 ou “Variâncias não homogêneas” no caso contrário) e o teste de Kolmogorov-Smirnov para avaliar a normalidade da distribuição dos dados (o resultado pode ser “Distribuição normal” quando o valor calculado é inferior ao tabelado ou “Distribuição não normal” no caso contrário); e o teste de hipótese e respetiva análise post-hoc selecionados

Métrica de erro	Teste de Levene		Teste de Kolmogorov-Smirnov				Teste de hipótese e análise <i>post-hoc</i>
	Valor de p	Resultado	Período	Tabelado	Calculado	Resultado	
RRMBE	0	Variâncias não homogêneas	1 ano	0,176	0,173	Distribuição normal	ANOVA de Welch e teste de Tukey
			2 anos	0,180	0,133	Distribuição normal	
			3 anos	0,185	0,125	Distribuição normal	
			4 anos	0,190	0,073	Distribuição normal	
			5 anos	0,196	0,100	Distribuição normal	
			6 anos	0,203	0,091	Distribuição normal	
			7 anos	0,210	0,091	Distribuição normal	
			8 anos	0,213	0,105	Distribuição normal	
			9 anos	0,221	0,097	Distribuição normal	
			10 anos	0,231	0,094	Distribuição normal	
			11 anos	0,242	0,119	Distribuição normal	
			12 anos	0,254	0,168	Distribuição normal	
			13 anos	0,269	0,091	Distribuição normal	
			14 anos	0,287	0,133	Distribuição normal	
			15 anos	0,309	0,108	Distribuição normal	
			16 anos	0,338	0,158	Distribuição normal	
			17 anos	0,375	0,134	Distribuição normal	
			18 anos	0,430	0,248	Distribuição normal	
			19 anos	0,519	0,224	Distribuição normal	
			20 anos	0,708	0,385	Distribuição normal	
RRRMSE	<0,01	Variâncias não homogêneas	1 ano	0,176	0,142	Distribuição normal	Teste de Kruskal-Wallis e Teste de Dunn
			2 anos	0,180	0,161	Distribuição normal	
			3 anos	0,185	0,108	Distribuição normal	
			4 anos	0,190	0,110	Distribuição normal	
			5 anos	0,196	0,154	Distribuição normal	

			6 anos	0,203	0,148	Distribuição normal	
			7 anos	0,210	0,152	Distribuição normal	
			8 anos	0,213	0,206	Distribuição normal	
			9 anos	0,221	0,204	Distribuição normal	
			10 anos	0,231	0,235	Distribuição não normal	
			11 anos	0,242	0,241	Distribuição normal	
			12 anos	0,254	0,232	Distribuição normal	
			13 anos	0,269	0,260	Distribuição normal	
			14 anos	0,287	0,273	Distribuição normal	
			15 anos	0,309	0,334	Distribuição não normal	
			16 anos	0,338	0,332	Distribuição normal	
			17 anos	0,375	0,329	Distribuição normal	
			18 anos	0,430	0,327	Distribuição normal	
			19 anos	0,519	0,322	Distribuição normal	
			20 anos	0,708	0,302	Distribuição normal	

Tabela B.12. Resultados obtidos na ANOVA de Welch realizada ao grupo da métrica RRMBE e no teste de Kruskal-Wallis ao grupo da métrica RRRMSE

Métrica de erro	ANOVA de Welch		Teste de Kruskal-Wallis	
	Valor de p	Resultado	Valor de p	Resultado
RRMBE	<0,01	Existe pelo menos um grupo que apresenta mediana distinta	-	-
RRRMSE	-	-	<0,01	Existe pelo menos um grupo que apresenta mediana distinta

Tabela B.13. Resultados obtidos no teste de Tukey realizado como análise post-hoc à métrica de erro **RRMBE**: valores de p obtidos para a comparação dos diferentes grupos de anos, sendo que a tabela é lida da direita para a esquerda, ou seja, o valor de p apresentado corresponde à comparação entre o grupo apresentado em linha com o grupo da coluna respectiva. Os valores de p encontram-se assinalados com um sistema de cores: verde quando o valor é inferior ao nível de significância 0,05 e vermelho quando é superior

	1 ano	2 anos	3 anos	4 anos	5 anos	6 anos	7 anos	8 anos	9 anos	10 anos	11 anos	12 anos	13 anos	14 anos	15 anos	16 anos	17 anos	18 anos	19 anos	20 anos	
1 ano	-	0,04	0,01	<0,01	<0,01	<0,01	<0,01	<0,01	<0,01	<0,01	<0,01	<0,01	<0,01	<0,01	<0,01	<0,01	<0,01	<0,01	<0,01	0,02	0,58
2 anos	-	-	1	1	1	0,99	0,89	0,76	0,70	0,49	0,65	0,70	0,77	0,75	0,00	0,03	0,11	0,32	0,77	1	
3 anos	-	-		1	1	1	1	0,98	0,96	0,86	0,93	0,94	0,97	0,95	0,02	0,09	0,26	0,54	0,91	1	
4 anos	-	-	-	-	1	1	1	1	0,99	0,96	0,99	0,99	0,99	0,99	0,05	0,17	0,40	0,68	0,95	1	
5 anos	-	-	-	-	-	1	1	1	1	1	1	1	1	1	0,19	0,42	0,69	0,89	0,99	1	
6 anos	-	-	-	-	-	-	1	1	1	1	1	1	1	1	0,25	0,49	0,75	0,92	0,99	1	
7 anos	-	-	-	-	-	-	-	1	1	1	1	1	1	1	0,52	0,76	0,92	0,98	1	1	
8 anos	-	-	-	-	-	-	-	-	1	1	1	1	1	1	0,70	0,88	0,97	0,99	1	1	
9 anos	-	-	-	-	-	-	-	-	-	1	1	1	1	1	0,79	0,93	0,98	1	1	1	
10 anos	-	-	-	-	-	-	-	-	-	-	1	1	1	1	0,92	0,98	1	1	1	1	
11 anos	-	-	-	-	-	-	-	-	-	-	-	1	1	1	0,90	0,97	1	1	1	1	
12 anos	-	-	-	-	-	-	-	-	-	-	-	-	1	1	0,92	0,98	1	1	1	1	
13 anos	-	-	-	-	-	-	-	-	-	-	-	-	-	1	0,93	0,98	1	1	1	1	
14 anos	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0,97	0,99	1	1	1	1	
15 anos	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	1	1	1	1	

16 anos	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	1	1	1
17 anos	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	1	1
18 anos	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	1
19 anos	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1

Tabela B.14. Resultados obtidos no teste de Dunn realizado como análise post-hoc à métrica de erro RRRMSE: valores de p obtidos para a comparação dos diferentes grupos de anos, sendo que a tabela é lida da direita para a esquerda, ou seja, o valor de p apresentado corresponde à comparação entre o grupo apresentado em linha com o grupo da coluna respectiva. Os valores de p encontram-se assinalados com um sistema de cores: verde quando o valor é inferior ao nível de significância 0,05 e vermelho quando é superior

	1 ano	2 anos	3 anos	4 anos	5 anos	6 anos	7 anos	8 anos	9 anos	10 anos	11 anos	12 anos	13 anos	14 anos	15 anos	16 anos	17 anos	18 anos	19 anos	20 anos	
1 ano	-	0,01	<0,01	<0,01	<0,01	<0,01	<0,01	<0,01	<0,01	<0,01	<0,01	<0,01	<0,01	<0,01	<0,01	<0,01	<0,01	<0,01	<0,01	0,01	0,06
2 anos	-	-	0,19	0,07	<0,01	<0,01	<0,01	<0,01	<0,01	<0,01	<0,01	<0,01	<0,01	<0,01	0,04	0,08	0,09	0,11	0,18	0,28	
3 anos	-	-	-	0,59	0,11	0,05	0,04	0,02	0,02	<0,01	<0,01	<0,01	<0,01	<0,01	0,25	0,36	0,36	0,37	0,44	0,51	
4 anos	-	-	-	-	0,30	0,15	0,12	0,06	0,06	<0,01	0,02	0,02	0,01	0,01	0,44	0,58	0,56	0,55	0,60	0,63	
5 anos	-	-	-	-	-	0,68	0,59	0,37	0,36	0,06	0,15	0,14	0,10	0,08	0,99	0,87	0,94	0,99	0,97	0,89	
6 anos	-	-	-	-	-	-	0,90	0,63	0,61	0,14	0,28	0,26	0,21	0,16	0,76	0,65	0,73	0,83	0,87	0,99	
7 anos	-	-	-	-	-	-	-	0,72	0,70	0,18	0,35	0,32	0,25	0,20	0,69	0,59	0,67	0,77	0,82	0,95	
8 anos	-	-	-	-	-	-	-	-	0,97	0,33	0,55	0,51	0,41	0,32	0,50	0,43	0,51	0,61	0,69	0,85	
9 anos	-	-	-	-	-	-	-	-	-	0,35	0,58	0,53	0,43	0,35	0,49	0,42	0,50	0,60	0,67	0,84	

10 anos	-	-	-	-	-	-	-	-	-	-	0,73	0,80	0,95	0,90	0,15	0,13	0,18	0,27	0,36	0,57
11 anos	-	-	-	-	-	-	-	-	-	-	-	0,94	0,80	0,67	0,26	0,22	0,29	0,38	0,47	0,67
12 anos	-	-	-	-	-	-	-	-	-	-	-	-	0,87	0,73	0,24	0,21	0,27	0,36	0,45	0,65
13 anos	-	-	-	-	-	-	-	-	-	-	-	-	-	0,86	0,19	0,17	0,22	0,31	0,39	0,59
14 anos	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0,15	0,13	0,18	0,26	0,34	0,54
15 anos	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0,89	0,94	0,99	0,98	0,90
16 anos	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0,95	0,90	0,90	0,84
17 anos	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0,95	0,94	0,87
18 anos	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0,99	0,91
19 anos	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0,93

B.2. Parques com 4 anos de dados operacionais

Para o caso dos parques com 4 anos de operação, foram analisados os erros RRMBE e RRRMSE para cada um dos parques sendo que as representações gráficas dos restantes parques para além dos que estão apresentados no capítulo 4.2 estão apresentadas no presente anexo nas Figuras B.1 a B.16 e os dados que originaram as representações gráficas das Figuras 4.6 a 4.7 e Figuras B.1 a B.16 estão apresentados na tabela B.15 e B.16.

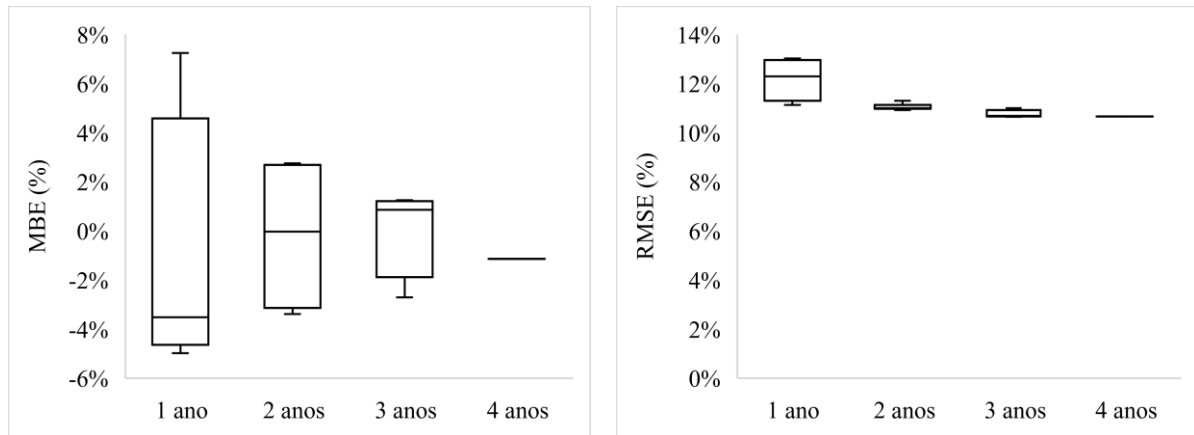


Figura B.1. Representação gráfica do erro **RMBE** (à esquerda) e **RRMSE** (à direita), para todas as combinações realizadas de 1 a 4 anos, para o **parque P_E1**.

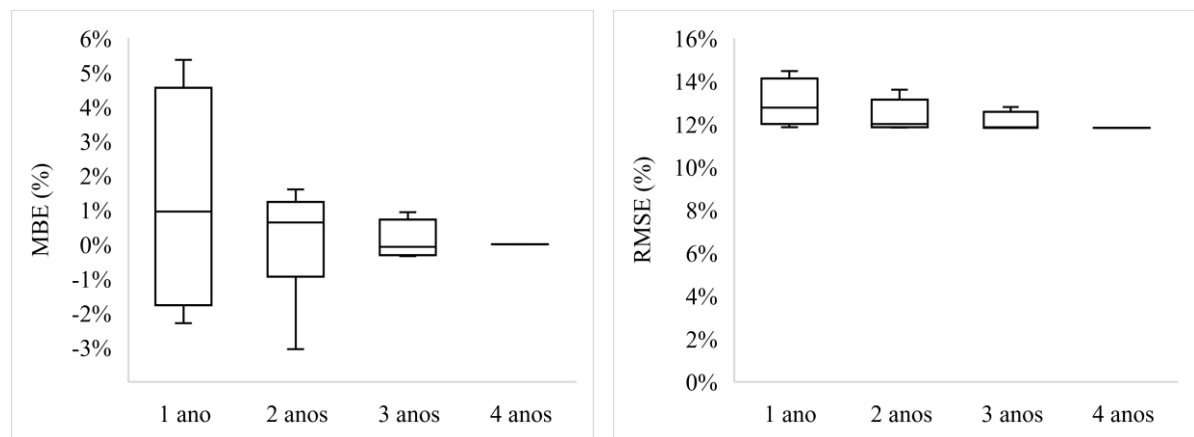


Figura B.2. Representação gráfica do erro **RMBE** (à esquerda) e **RRMSE** (à direita), para todas as combinações realizadas de 1 a 4 anos, para o **parque P_E2**.

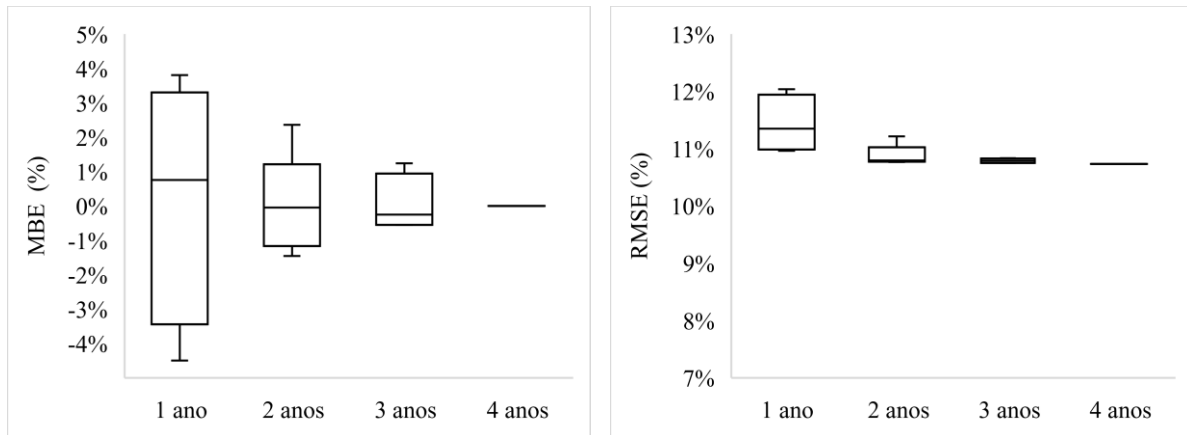


Figura B.3. Representação gráfica do erro **RMBE** (à esquerda) e **RRMSE** (à direita), para todas as combinações realizadas de 1 a 4 anos, para o **parque P_E3**.

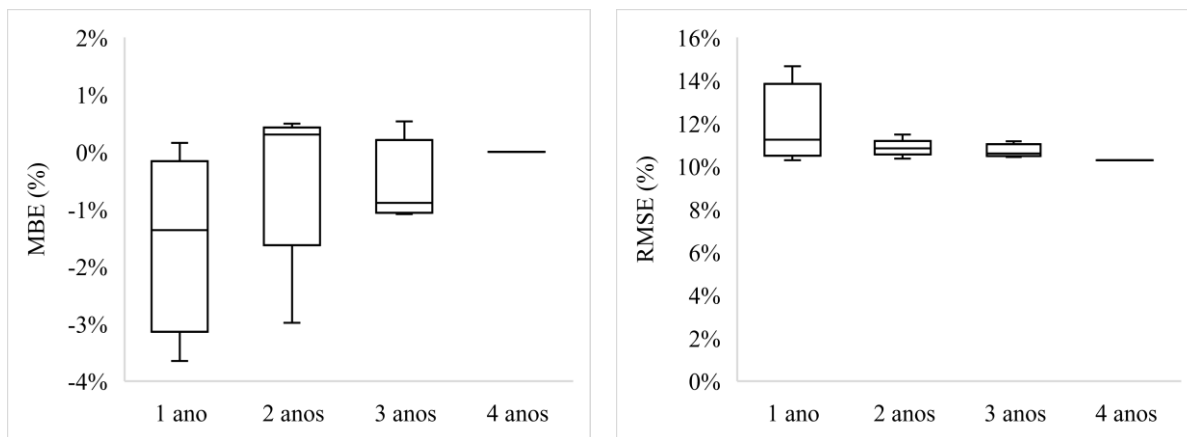


Figura B.4. Representação gráfica do erro **RMBE** (à esquerda) e **RRMSE** (à direita), para todas as combinações realizadas de 1 a 4 anos, para o **parque P_E4**.

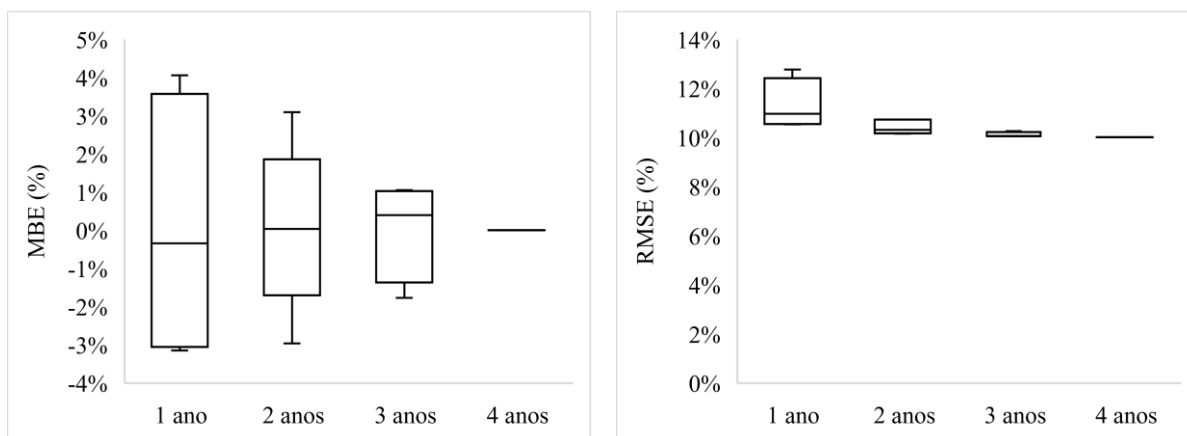


Figura B.5. Representação gráfica do erro **RMBE** (à esquerda) e **RRMSE** (à direita), para todas as combinações realizadas de 1 a 4 anos, para o **parque P_E5**.

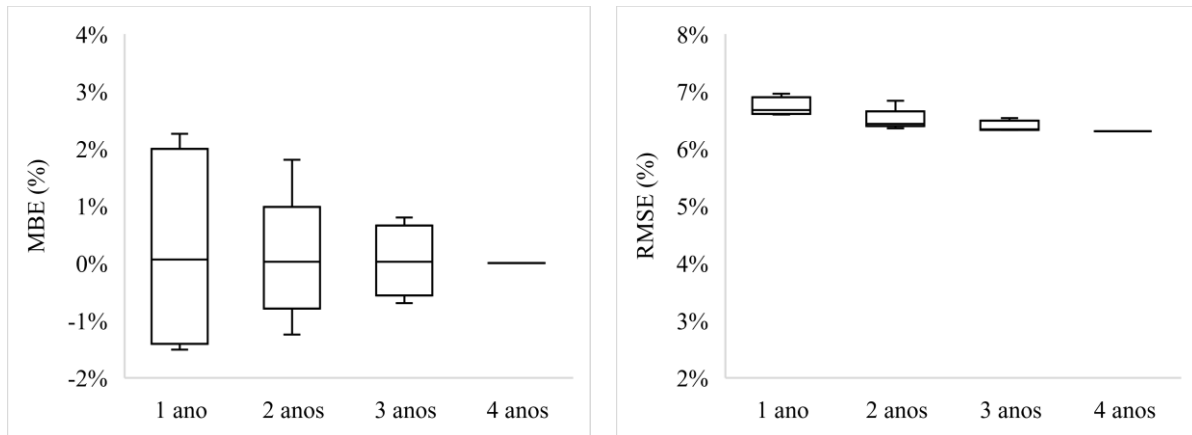


Figura B.6. Representação gráfica do erro **RMBE** (à esquerda) e **RRMSE** (à direita), para todas as combinações realizadas de 1 a 4 anos, para o **parque P_E6**.

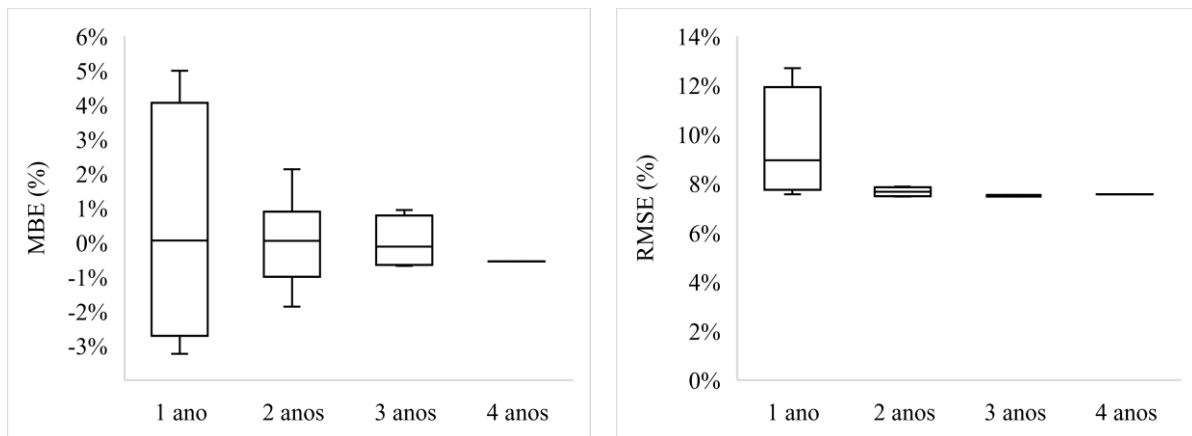


Figura B.7. Representação gráfica do erro **RMBE** (à esquerda) e **RRMSE** (à direita), para todas as combinações realizadas de 1 a 4 anos, para o **parque P_E7**.

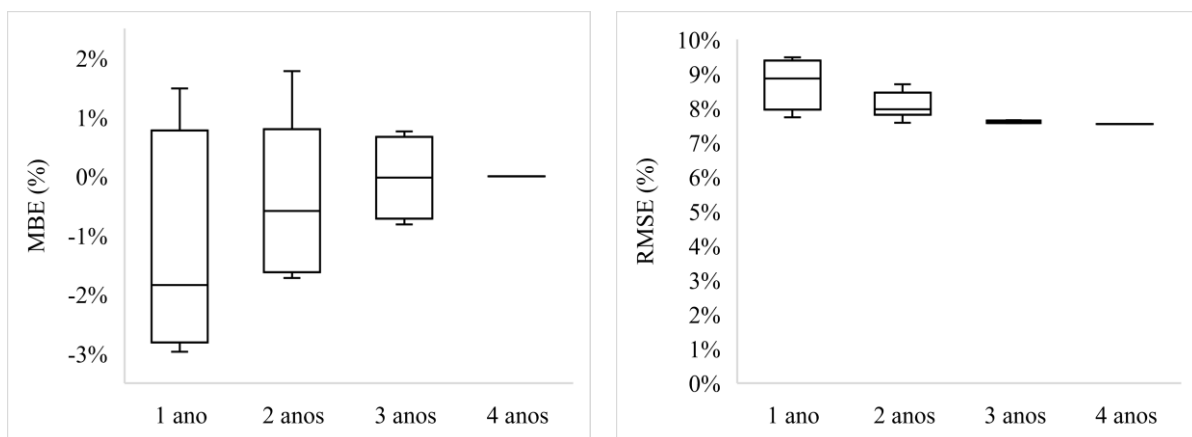


Figura B.8. Representação gráfica do erro **RMBE** (à esquerda) e **RRMSE** (à direita), para todas as combinações realizadas de 1 a 4 anos, para o **parque P_E9**.

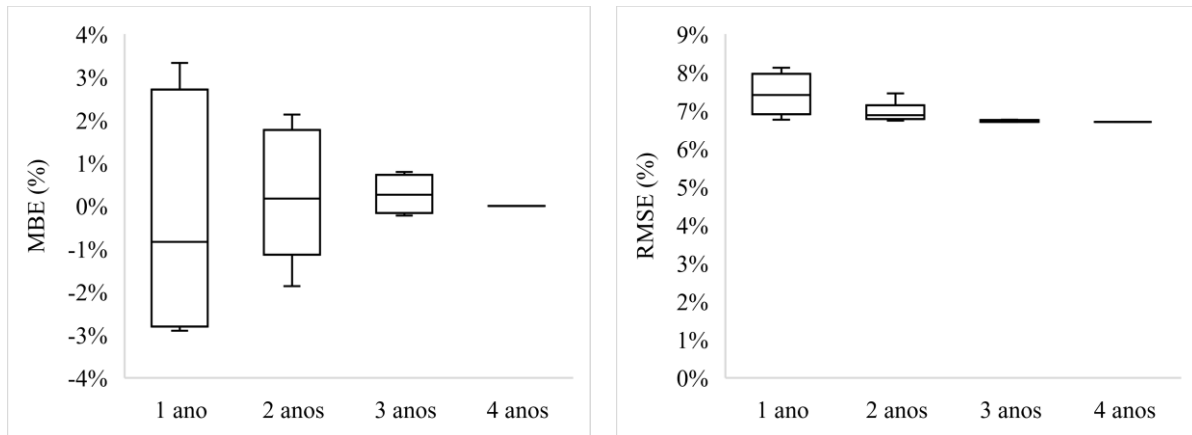


Figura B.9. Representação gráfica do erro **RMBE** (à esquerda) e **RRMSE** (à direita), para todas as combinações realizadas de 1 a 4 anos, para o **parque P_E10**.

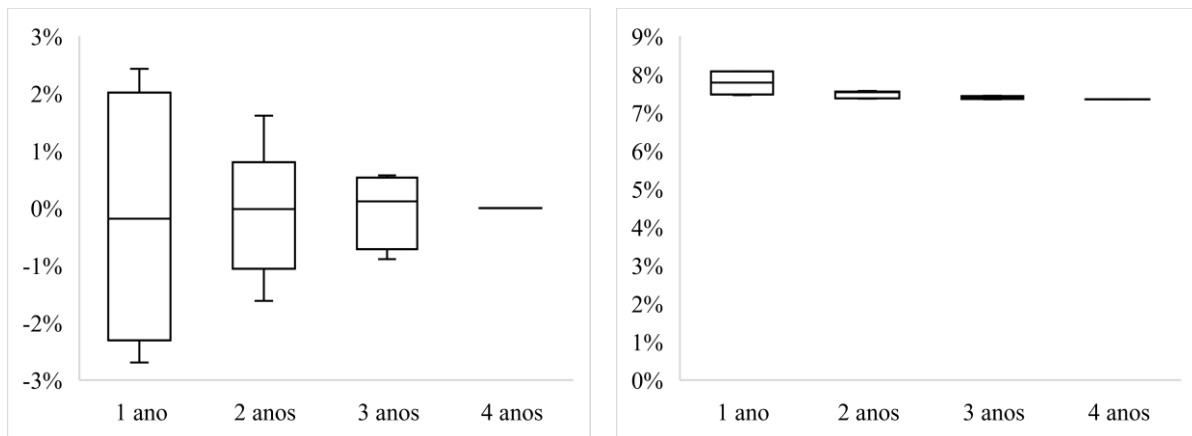


Figura B.10. Representação gráfica do erro **RMBE** (à esquerda) e **RRMSE** (à direita), para todas as combinações realizadas de 1 a 4 anos, para o **parque P_E11**.

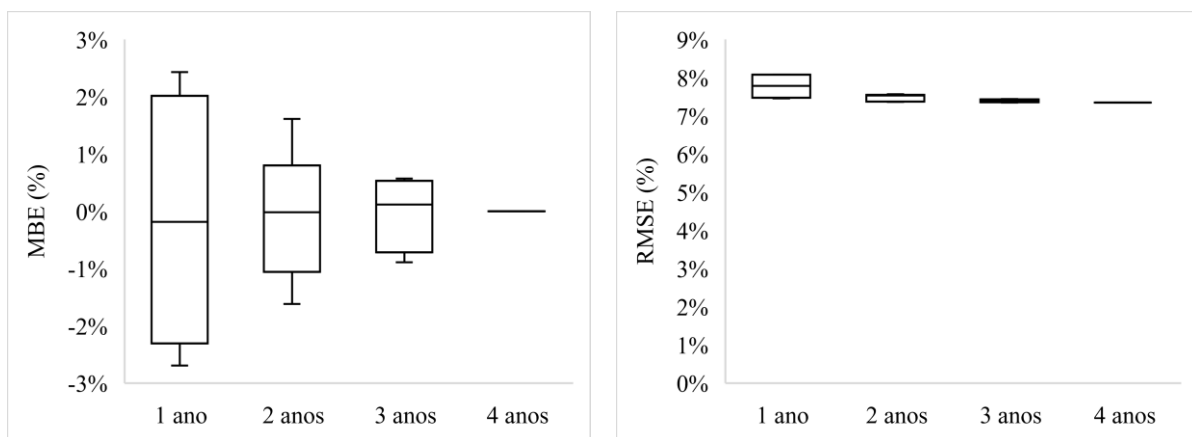


Figura B.11. Representação gráfica do erro **RMBE** (à esquerda) e **RRMSE** (à direita), para todas as combinações realizadas de 1 a 4 anos, para o **parque P_P1**.

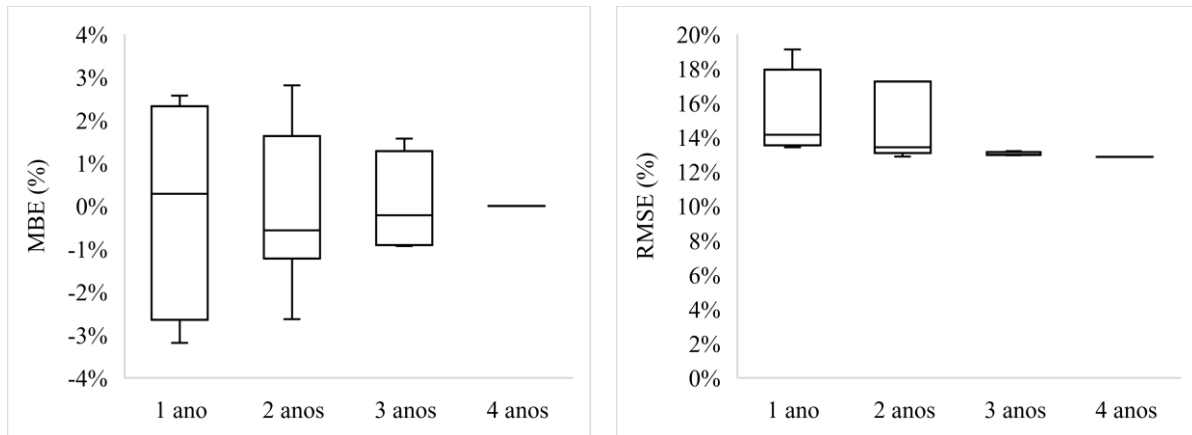


Figura B.12. Representação gráfica do erro **RMBE** (à esquerda) e **RRMSE** (à direita), para todas as combinações realizadas de 1 a 4 anos, para o **parque P_P3**.

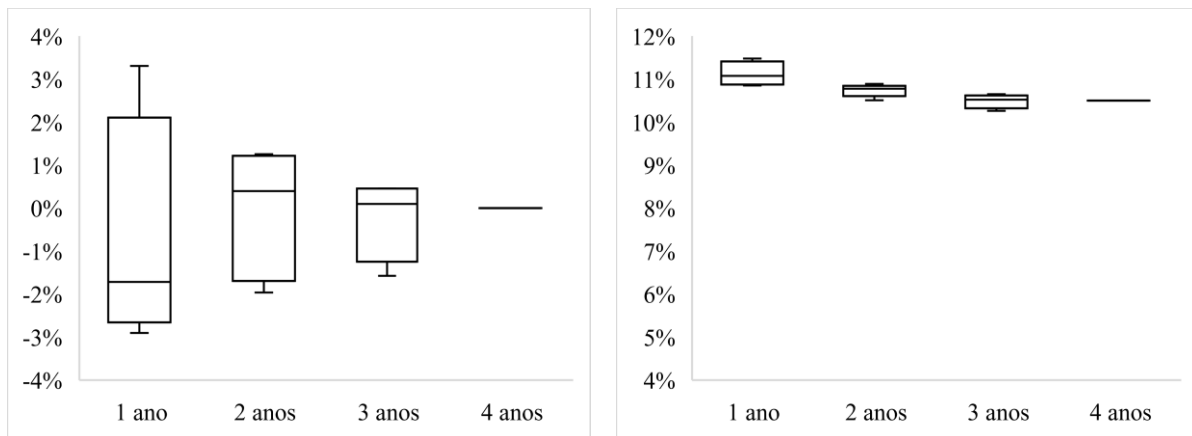


Figura B.13. Representação gráfica do erro **RMBE** (à esquerda) e **RRMSE** (à direita), para todas as combinações realizadas de 1 a 4 anos, para o **parque P_P4**.

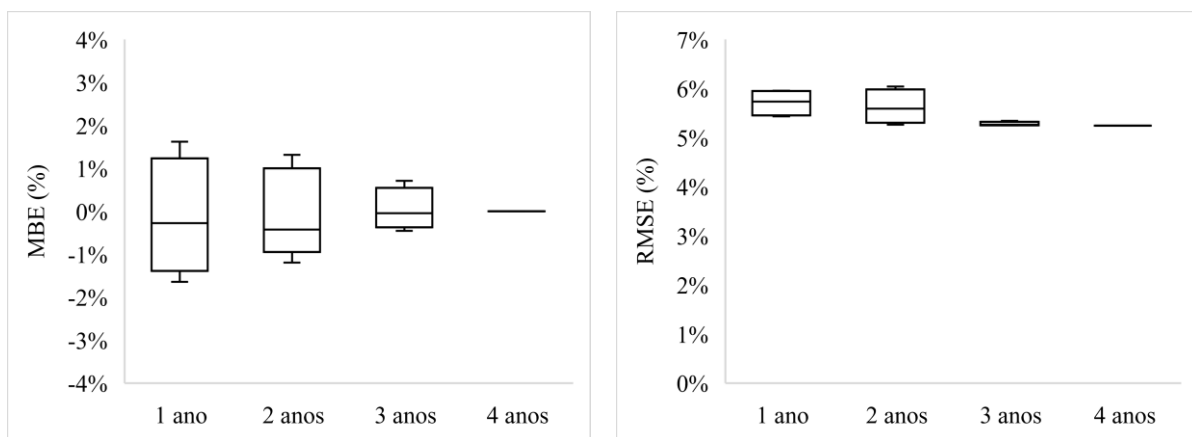


Figura B.14. Representação gráfica do erro **RMBE** (à esquerda) e **RRMSE** (à direita), para todas as combinações realizadas de 1 a 4 anos, para o **parque P_CI**.

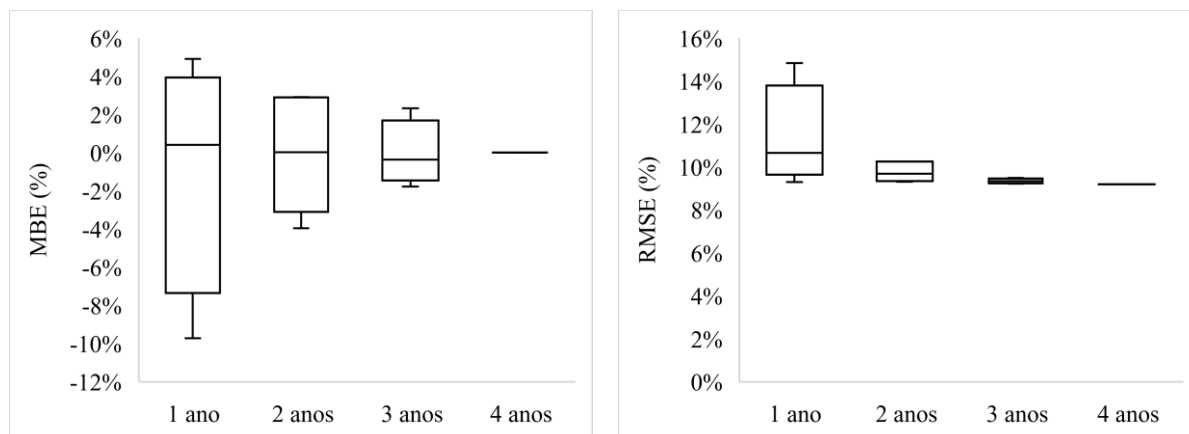


Figura B.15. Representação gráfica do erro **RMBE** (à esquerda) e **RRMSE** (à direita), para todas as combinações realizadas de 1 a 4 anos, para o **parque P_R1**.

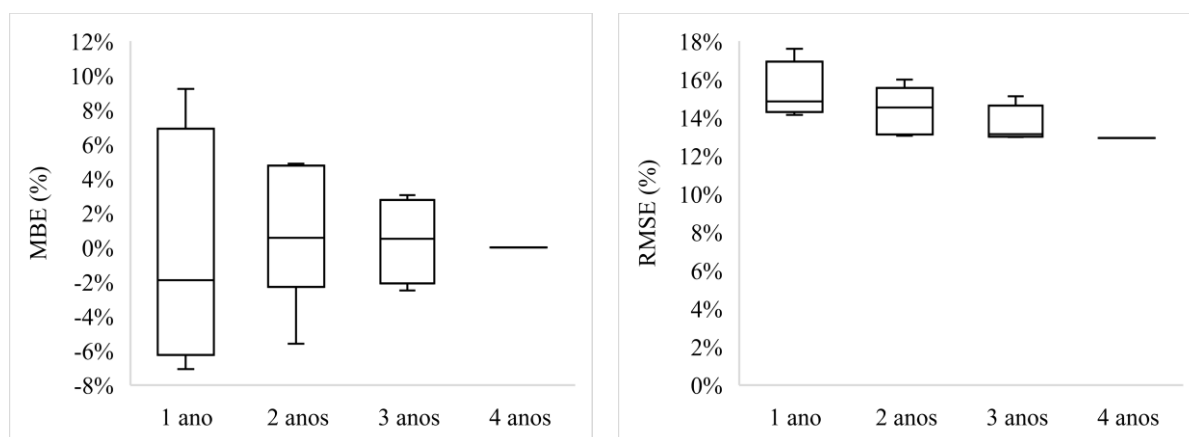


Figura B.16. Representação gráfica do erro **RMBE** (à esquerda) e **RRMSE** (à direita), para todas as combinações realizadas de 1 a 4 anos, para o **parque P_R2**.

Tabela B.15. Dados relativos ao **RMBE**, calculado para todas as combinações de 1 a 4 anos, para o **todos os parques com 4 anos de dados operacionais**

P_E1	1 ano	7,2%	-3,7%	-5,0%	-3,4%	-	-
	2 anos	2,7%	2,7%	2,1%	-2,2%	-3,4%	-3,1%
	3 anos	1,2%	-2,7%	0,6%	1,1%	-	-
	4 anos	-1,1%	-	-	-	-	-
P_E2	1 ano	5,4%	2,1%	-0,2%	-2,3%	-	-
	2 anos	1,1%	0,2%	1,6%	1,0%	-0,2%	-3,1%
	3 anos	0,9%	-0,2%	-0,3%	0,1%	-	-
	4 anos	0%	-	-	-	-	-
P_E3	1 ano	-4,5%	1,8%	-0,3%	3,8%	-	-
	2 anos	-0,8%	-1,5%	-1,1%	0,7%	2,4%	0,8%
	3 anos	-0,5%	1,2%	-0,6%	0,0%	-	-
	4 anos	0%	-	-	-	-	-
	1 ano	-3,6%	0,2%	-1,1%	-1,6%	-	-

P_E4	2 anos	0,4%	-3,0%	0,4%	0,2%	0,5%	-1,2%
	3 anos	-1,0%	-0,8%	-1,1%	0,5%	-	-
	4 anos	0%	-	-	-	-	-
P_E5	1 ano	4,1%	2,1%	-2,8%	-3,1%	-	-
	2 anos	3,1%	1,4%	1,3%	-1,2%	-1,3%	-3,0%
	3 anos	1,1%	-1,8%	-0,2%	1,0%	-	-
	4 anos	0%	-	-	-	-	-
P_E6	1 ano	-1,5%	-1,1%	1,2%	2,3%	-	-
	2 anos	-1,2%	-0,6%	0,1%	0,0%	0,7%	1,8%
	3 anos	-0,7%	0,8%	0,2%	-0,2%	-	-
	4 anos	0%	-	-	-	-	-
P_E7	1 ano	-1,2%	1,3%	5,0%	-3,2%	-	-
	2 anos	0,3%	0,5%	-1,9%	2,1%	-0,7%	-0,3%
	3 anos	0,9%	0,3%	-0,6%	-0,7%	-	-
	4 anos	-0,6%	-	-	-	-	-
P_E8	1 ano	2,9%	-2,0%	-2,7%	-0,5%	-	-
	2 anos	0,7%	0,0%	2,6%	-2,2%	-1,2%	-1,0%
	3 anos	-0,5%	-1,4%	0,7%	1,0%	-	-
	4 anos	0%	-	-	-	-	-
P_E9	1 ano	-1,4%	1,5%	-3,0%	-2,3%	-	-
	2 anos	0,5%	-1,6%	-1,7%	1,8%	-0,4%	-0,8%
	3 anos	0,8%	0,4%	-0,8%	-0,4%	-	-
	4 anos	0%	-	-	-	-	-
P_E10	1 ano	0,8%	-2,5%	-2,9%	3,3%	-	-
	2 anos	-0,9%	-0,5%	2,1%	-1,9%	1,6%	0,8%
	3 anos	0,0%	-0,2%	0,8%	0,5%	-	-
	4 anos	0%	-	-	-	-	-
P_E11	1 ano	-2,7%	-1,1%	0,8%	2,4%	-	-
	2 anos	-1,6%	-0,9%	0,3%	-0,3%	0,5%	1,6%
	3 anos	-0,9%	0,6%	0,4%	-0,2%	-	-
	4 anos	0%	-	-	-	-	-
P_P1	1 ano	-1,7%	2,4%	-2,4%	1,6%	-	-
	2 anos	0,2%	-2,1%	0,2%	-0,2%	2,0%	0,6%
	3 anos	-0,8%	0,6%	-0,6%	0,8%	-	-
	4 anos	0,0%	-	-	-	-	-
P_P2	1 ano	-0,5%	-4,7%	0,3%	-0,6%	-	-
	2 anos	-5,6%	0,9%	1,2%	1,1%	-1,9%	0,7%
	3 anos	0,2%	0,2%	0,3%	-0,6%	-	-
	4 anos	1,8%	-	-	-	-	-
P_P3	1 ano	2,6%	-3,2%	1,6%	-1,0%	-	-
	2 anos	-0,5%	2,8%	-0,7%	-0,8%	-2,6%	1,2%

	3 anos	0,4%	-0,9%	1,6%	-0,8%	-	-
	4 anos	0%	-	-	-	-	-
P_P4	1 ano	3,3%	-1,9%	-1,5%	-2,9%		
	2 anos	0,8%	1,3%	1,2%	0,0%	-1,6%	-2,0%
	3 anos	-0,3%	-1,6%	0,5%	0,5%	-	-
	4 anos	0%	-	-	-	-	-
P_C1	1 ano	1,6%	-0,6%	-1,6%	0,1%	-	-
	2 anos	0,9%	-0,2%	1,3%	-1,2%	-0,7%	-0,9%
	3 anos	0%	-0,5%	-0,1%	0,7%	-	-
	4 anos	0%	-	-	-	-	-
P_R1	1 ano	4,9%	-9,7%	1,0%	-0,2%	-	-
	2 anos	-0,9%	-2,8%	1,0%	2,9%	2,9%	-4,0%
	3 anos	-0,2%	-1,8%	-0,5%	2,3%	-	-
	4 anos	0%	-	-	-	-	-
P_R2	1 ano	9,2%	-0,1%	-7,1%	-3,8%	-	-
	2 anos	4,7%	4,9%	1,7%	-1,2%	-0,6%	-5,6%
	3 anos	3,0%	-0,9%	-2,5%	1,9%	-	-
	4 anos	0%	-	-	-	-	-

Tabela B.16. Dados relativos ao **RRMSE**, calculado para todas as combinações de 1 a 4 anos, para o **todos os parques com 4 anos de dados operacionais**

P_E1	1 ano	13,0%	11,8%	12,8%	11,1%	-	-
	2 anos	10,9%	11,0%	11,0%	11,0%	11,3%	11,1%
	3 anos	10,7%	11,0%	10,7%	10,6%	-	-
	4 anos	10,7%	-	-	-	-	-
P_E2	1 ano	14,5%	12,4%	11,9%	13,1%	-	-
	2 anos	12,1%	11,8%	13,0%	11,9%	11,9%	13,6%
	3 anos	11,9%	11,8%	12,8%	11,8%	-	-
	4 anos	11,8%	-	-	-	-	-
P_E3	1 ano	11,7%	11,1%	11,0%	12,0%	-	-
	2 anos	10,8%	11,0%	10,8%	10,8%	11,2%	10,8%
	3 anos	10,7%	10,8%	10,8%	10,8%	-	-
	4 anos	10,7%	-	-	-	-	-
P_E4	1 ano	11,3%	10,3%	14,7%	11,1%	-	-
	2 anos	10,7%	11,0%	11,1%	10,6%	10,4%	11,5%
	3 anos	10,6%	11,2%	10,6%	10,4%	-	-
	4 anos	10,3%	-	-	-	-	-
P_E5	1 ano	12,8%	11,4%	10,5%	10,6%	-	-
	2 anos	11,6%	10,2%	10,3%	10,4%	10,2%	10,5%
	3 anos	10,1%	10,3%	10,1%	10,1%	-	-
	4 anos	10,0%	-	-	-	-	-

P_E6	1 ano	6,6%	6,6%	7,0%	6,7%	-	-
	2 anos	6,4%	6,4%	6,4%	6,6%	6,4%	6,8%
	3 anos	6,3%	6,5%	6,3%	6,3%	-	-
	4 anos	6,3%	-	-	-	-	-
P_E7	1 ano	7,6%	8,3%	12,7%	9,6%	-	-
	2 anos	7,5%	7,6%	7,7%	7,9%	7,8%	7,5%
	3 anos	7,5%	7,5%	7,5%	7,5%	-	-
	4 anos	7,6%	-	-	-	-	-
P_E8	1 ano	9,1%	9,1%	8,6%	9,8%	-	-
	2 anos	8,2%	8,4%	8,6%	8,5%	9,4%	8,3%
	3 anos	8,2%	8,5%	8,2%	8,2%	-	-
	4 anos	8,1%	-	-	-	-	-
P_E9	1 ano	7,7%	9,1%	9,5%	8,6%	-	-
	2 anos	7,6%	7,9%	8,7%	7,9%	8,0%	8,4%
	3 anos	7,6%	7,6%	7,7%	7,6%	-	-
	4 anos	7,5%	-	-	-	-	-
P_E10	1 ano	6,8%	7,3%	8,1%	7,5%	-	-
	2 anos	6,8%	6,7%	7,0%	7,0%	7,5%	6,8%
	3 anos	6,7%	6,7%	6,8%	6,7%	-	-
	4 anos	6,7%	-	-	-	-	-
P_E11	1 ano	8,1%	7,5%	7,5%	8,1%	-	-
	2 anos	7,5%	7,5%	7,4%	7,4%	7,5%	7,6%
	3 anos	7,4%	7,4%	7,4%	7,3%	-	-
	4 anos	7,3%	-	-	-	-	-
P_P1	1 ano	7,2%	8,4%	7,9%	7,6%	-	-
	2 anos	7,4%	7,3%	7,0%	7,0%	7,3%	7,6%
	3 anos	7,0%	7,0%	7,2%	7,0%	-	-
	4 anos	6,9%	-	-	-	-	-
P_P2	1 ano	13,8%	46,5%	49,1%	13,8%	-	-
	2 anos	50,8%	14,3%	15,2%	47,7%	15,1%	14,5%
	3 anos	13,7%	13,8%	14,1%	13,8%	-	-
	4 anos	14,1%	-	-	-	-	-
P_P3	1 ano	13,4%	14,4%	19,1%	13,9%	-	-
	2 anos	13,6%	13,2%	24,4%	12,9%	13,1%	14,8%
	3 anos	13,0%	13,0%	13,2%	13,0%	-	-
	4 anos	12,9%	-	-	-	-	-
P_P4	1 ano	11,2%	11,0%	10,8%	11,5%	-	-
	2 anos	10,8%	10,9%	10,6%	10,5%	10,8%	10,8%
	3 anos	11,3%	10,6%	10,5%	10,5%	-	-
	4 anos	10,5%	-	-	-	-	-
	1 ano	5,9%	6,0%	5,5%	5,4%	-	-

P_C1	2 anos	5,4%	5,3%	5,8%	6,0%	6,0%	5,3%
	3 anos	5,3%	5,3%	5,2%	5,3%	-	-
	4 anos	5,2%	-	-	-	-	-
P_R1	1 ano	10,7%	14,8%	9,3%	10,7%	-	-
	2 anos	9,3%	9,7%	9,3%	9,7%	9,7%	11,9%
	3 anos	9,3%	9,4%	9,2%	9,5%	-	-
	4 anos	9,2%	-	-	-	-	-
P_R2	1 ano	17,6%	14,9%	14,8%	14,1%	-	-
	2 anos	15,4%	16,0%	13,1%	14,7%	13,0%	14,3%
	3 anos	15,1%	13,0%	13,2%	13,1%	-	-
	4 anos	12,9%	-	-	-	-	-

Tabela B.17. Resultados dos testes de Levene e de Kolmogorov Smirnov, a escolha de teste de hipóteses realizada e respectiva análise post-hoc, para todos os parques com 4 anos de operação.

Métrica de erro	Teste de Levene		Teste de Kolmogorov-Smirnov				Teste de hipóteses e análise post-hoc
	Valor de p	Resultado	Período	Tabelado	Calculado	Resultado	
RRMBE	<0,01	Variâncias homogêneas	1 ano	0,170	0,103	Distribuição normal	Teste de Kruskal-Wallis e teste de Dunn
			2 anos	0,139	0,054	Distribuição normal	
			3 anos	0,170	0,086	Distribuição normal	
			4 anos	0,327	0,439	Distribuição não normal	
RRRMSE	0,39	Variâncias não homogêneas	1 ano	0,170	0,274	Distribuição não normal	Teste de Kruskal-Wallis e teste de Dunn
			2 anos	0,139	0,235	Distribuição não normal	
			3 anos	0,170	0,181	Distribuição não normal	
			4 anos	0,327	0,177	Distribuição normal	

Tabela B.18. Resultados obtidos no teste de Kruskal-Wallis para ambos erros estudados, RMBE e RRMSE, para os parques com 4 anos de operação

Tipo de erro	<i>ANOVA tradicional</i>		Teste de Kruskal-Wallis	
	Valor de p	Resultado	Valor de p	Resultado
RMBE	0,756	As médias dos grupos são estatisticamente semelhantes	-	-
RRMSE	-	-	0,200	As medianas dos grupos são estatisticamente semelhantes

ANEXO C – COMPLEXOS EÓLICOS

Neste anexo são apresentados os dados relativos à análise efetuada aos complexos eólicos que inclui a estimativa de Longo Termo calculada através dos dados do complexo ou dos parques individuais, com a melhor reanálise e com a mesma reanálise utilizada para o complexo, assim como o desvio em relação à estimativa de Longo Termo do complexo como um todo. Inclui ainda os dados das análises estatísticas efetuadas aos complexos eólicos.

Tabela C.1. Resultados obtidos no estudo do complexo C_B1, que inclui a estimativa de Longo Termo dos parques individuais com a melhor reanálise para cada parque e com a melhor série de reanálise para o complexo, assim como o desvio de Longo Termo com dados dos parques individuais com a estimativa realizada com os dados do complexo

	Designação	Estimativa de Longo Termo com dados dos parques individuais (MW)	Estimativa de Longo Termo com dados do complexo (MW)	Desvio (%)
Melhor reanálise parques individuais	B1_P1	108 930	511 028	0,23%
	B1_P2	113 189		
	B1_P3	108 606		
	B1_P4	86 553		
	B1_P5	93 750		
Melhor reanálise complexo	B1_P1	108 930	511 092	0,22%
	B1_P2	113 189		
	B1_P3	108 606		
	B1_P4	86 617		
	B1_P5	93 750		
Complexo	B1		512 231	

Tabela C.2. Resultados obtidos no estudo do **complexo C_B2**, que inclui a estimativa de Longo Termo dos parques individuais com a melhor reanálise para cada parque e com a melhor série de reanálise para o complexo, assim como o desvio de Longo Termo com dados dos parques individuais com a estimativa realizada com os dados do complexo

	Designação	Estimativa de Longo Termo parques individuais	Estimativa de Longo Termo do complexo	Desvio em relação ao complexo
Melhor reanálise parques individuais	B2_P1	127 340	741 641	2,7%
	B2_P2	119 970		
	B2_P3	67 283		
	B2_P4	81 695		
	B2_P5	83 066		
	B2_P6	96 838		
	B2_P7	81 456		
	B2_P8	83 993		
Melhor reanálise complexo	B2_P1	127 340	742 413	2,6%
	B2_P2	119 970		
	B2_P3	67 283		
	B2_P4	82 467		
	B2_P5	83 066		
	B2_P6	96 838		
	B2_P7	81 456		
	B2_P8	83 993		
Complexo	B2		762 200	

Tabela C.3. Resultados obtidos no estudo do **complexo C_B3**, que inclui a estimativa de Longo Termo dos parques individuais com a melhor reanálise para cada parque e com a melhor série de reanálise para o complexo, assim como o desvio de Longo Termo com dados dos parques individuais com a estimativa realizada com os dados do complexo

	Designação	Estimativa de Longo Termo parques individuais	Estimativa de Longo Termo do complexo	Desvio em relação ao complexo
Melhor reanálise parques individuais	B3_P1	105 873	813 106	0,8%
	B3_P2	108 400		
	B3_P3	104 167		
	B3_P4	127 770		
	B3_P5	108 273		
	B3_P6	126 313		
	B3_P7	132 310		
Melhor reanálise complexo	B3_P1	105 873	813 556	0,9%
	B3_P2	108 400		
	B3_P3	104 327		
	B3_P4	127 812		
	B3_P5	108 521		
	B3_P6	126 313		
	B3_P7	132 310		
Complexo	B3		806 689	

Tabela C.4. Resultados obtidos no estudo do **complexo C_B4**, que inclui a estimativa de Longo Termo dos parques individuais com a melhor reanálise para cada parque que é a melhor série de reanálise para o complexo, assim como o desvio de Longo Termo com dados dos parques individuais com a estimativa realizada com os dados do complexo

	Designação	Estimativa de Longo Termo parques individuais	Estimativa de Longo Termo do complexo	Desvio em relação ao complexo
Melhor reanálise complexo	B4_P1	132 633	964 785	0,4%
	B4_P2	109 245		
	B4_P3	119 936		
	B4_P4	123 901		
	B4_P5	113 843		
	B4_P6	120 447		
	B4_P7	117 439		
	B4_P8	127 341		
Complexo	B4		968 439	

Tabela C.5. Resultados obtidos no estudo do **complexo C_B5**, que inclui a estimativa de Longo Termo dos parques individuais com a melhor reanálise para cada parque e com a melhor série de reanálise para o complexo, assim como o desvio de Longo Termo com dados dos parques individuais com a estimativa realizada com os dados do complexo

	Designação	Estimativa de Longo Termo parques individuais	Estimativa de Longo Termo do complexo	Desvio em relação ao complexo
Melhor reanálise parques individuais	B5_P1	100 717	198 277	0,2%
	B5_P2	97 560		
Melhor reanálise complexo	B5_P1	100 717	198 742	0,004%
	B5_P2	98025		
Complexo	B5		198 735	

Tabela C.6. Resultados obtidos no estudo do **complexo C_B6**, que inclui a estimativa de Longo Termo dos parques individuais com a melhor reanálise para cada parque e com a melhor série de reanálise para o complexo, assim como o desvio de Longo Termo com dados dos parques individuais com a estimativa realizada com os dados do complexo

	Designação	Estimativa de Longo Termo parques individuais	Estimativa de Longo Termo do complexo	Desvio em relação ao complexo
Melhor reanálise parques individuais	B6_P1	99 856	263 864	0,42%
	B6_P2	65 421		
	B6_P3	98 587		
Melhor reanálise complexo	B6_P1	99 468	263 476	0,57%
	B6_P2	65 421		
	B6_P3	98587		
Complexo	B6		264 982	

Tabela C.7. Resultados obtidos no estudo do **complexo C_E11**, que inclui a estimativa de Longo Termo dos parques individuais com a melhor reanálise para cada parque que é a melhor série de reanálise para o complexo, assim como o desvio de Longo Termo com dados dos parques individuais com a estimativa realizada com os dados do complexo

	Designação	Estimativa de Longo Termo parques individuais	Estimativa de Longo Termo do complexo	Desvio em relação ao complexo
Melhor reanálise complexo	E11_P1	123 475	225 218	0,1%
	E11_P2	101 743		
Complexo	E11		225 357	

Tabela C.8. Resultados obtidos no estudo do **complexo C_E12**, que inclui a estimativa de Longo Termo dos parques individuais com a melhor reanálise para cada parque e com a melhor série de reanálise para o complexo, assim como o desvio de Longo Termo com dados dos parques individuais com a estimativa realizada com os dados do complexo

	Designação	Estimativa de Longo Termo parques individuais	Estimativa de Longo Termo do complexo	Desvio em relação ao complexo
Melhor reanálise parques individuais	E12_P1	98 362	185 707	0,2%
	E12_P2	77 544		
	E12_P3	9 801		
Melhor reanálise complexo	E12_P1	98 069	185 414	0,3%
	E12_P2	77 544		
	E12_P3	9 801		
Complexo	E12		186 024	

Tabela C.9. Resultados dos testes de Levene e de Kolmogorov Smirnov, a escolha de teste de hipóteses realizada e respectiva análise post-hoc, para todos os complexos em estudo, tendo como base **análise do complexo e da série de reanálise mais adequada para o complexo**

Métrica de erro	Teste de Levene		Teste de Kolmogorov-Smirnov				Teste de hipóteses e análise <i>post-hoc</i>
	Valor de p	Resultado	Fonte dos dados	Tabelado	Calculado	Resultado	
RMBE	0,64	Variâncias não homogêneas	Complexo	0,454	0,294	Distribuição normal	ANOVA de Welch e teste de Tukey
			Parques individuais	0,454	0,175	Distribuição normal	
RRMSE	0,89	Variâncias não homogêneas	Complexo	0,454	0,169	Distribuição normal	
			Parques individuais	0,454	0,171	Distribuição normal	

Tabela C.10. Resultados dos testes de Levene e de Kolmogorov Smirnov, a escolha de teste de hipóteses realizada e respectiva análise post-hoc, para todos os complexos em estudo, tendo como base a **análise do complexo com os dados do próprio complexo e os dados da série de reanálise mais adequada para cada parque**

Métrica de erro	Teste de Levene		Teste de Kolmogorov-Smirnov				Teste de hipóteses e análise <i>post-hoc</i>
	Valor de p	Resultado	Fonte dos dados	Tabelado	Calculado	Resultado	
RMBE	0,82	Variâncias não homogêneas	Complexo	0,454	0,294	Distribuição normal	ANOVA de Welch e teste de Tukey
			Parques individuais	0,454	0,136	Distribuição normal	
RRMSE	0,44	Variâncias não homogêneas	Complexo	0,454	0,169	Distribuição normal	
			Parques individuais	0,454	0,177	Distribuição normal	

Tabela C.11. Resultados obtidos na ANOVA de Welch realizada para ambas métricas de erro, RMBE e RRMSE, para os complexos em estudo, distinguindo as fontes dos dados em “Individuais com melhor série de reanálise para o complexo” e “Individuais com melhor série de reanálise para cada parque”

Métricas de erro	ANOVA de Welch		
	Fonte dos dados	Valor de p	Resultado
RMBE	Individuais com melhor série de reanálise para o complexo	0,64	As medianas dos grupos são estatisticamente semelhantes
	Individuais com melhor série de reanálise para cada parque	0,50	As medianas dos grupos são estatisticamente semelhantes
RRMSE	Individuais com melhor série de reanálise para o complexo	0,93	As medianas dos grupos são estatisticamente semelhantes
	Individuais com melhor série de reanálise para cada parque	0,29	As medianas dos grupos são estatisticamente semelhantes