



**UNIVERSIDADE FEDERAL DO CEARÁ  
CENTRO DE TECNOLOGIA  
DEPARTAMENTO DE ENGENHARIA MECÂNICA  
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA MECÂNICA**

**KEVIN SANTOS GUEDES**

**APLICAÇÃO DE ALGORITMOS DE OTIMIZAÇÃO METAHEURÍSTICOS  
NA ESTIMATIVA DOS PARÂMETROS ÓTIMOS DE DIFERENTES  
DISTRIBUIÇÕES DE VELOCIDADE DO VENTO EM DUAS CIDADES DO  
NORDESTE BRASILEIRO**

**FORTALEZA**

**2020**

KEVIN SANTOS GUEDES

APLICAÇÃO DE ALGORITMOS DE OTIMIZAÇÃO METAHEURÍSTICOS NA  
ESTIMATIVA DOS PARÂMETROS ÓTIMOS DE DIFERENTES DISTRIBUIÇÕES DE  
VELOCIDADE DO VENTO EM DUAS CIDADES DO NORDESTE BRASILEIRO

Dissertação submetida à Coordenação do Curso de Pós-Graduação em Engenharia Mecânica da Universidade Federal do Ceará, como requisito parcial à obtenção do título de mestre em Engenharia Mecânica. Área de Concentração: Processos, Equipamentos e Sistemas para Energias Renováveis.

Orientador(a): Prof<sup>a</sup>. Dr<sup>a</sup>. Carla Freitas de Andrade

Coorientador(a): Prof. Dr. Paulo Alexandre Costa Rocha

FORTALEZA

2020

Dados Internacionais de Catalogação na Publicação  
Universidade Federal do Ceará  
Biblioteca Universitária

Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

---

- G957a Guedes, Kevin Santos.  
Aplicação de Algoritmos de Otimização Metaheurísticos na Estimativa dos Parâmetros Ótimos de Diferentes Distribuições de Velocidade do Vento em Duas Cidades do Nordeste Brasileiro / Kevin Santos Guedes. – 2020.  
70 f. : il. color.
- Dissertação (mestrado) – Universidade Federal do Ceará, Centro de Tecnologia, Programa de Pós-Graduação em Engenharia Mecânica, Fortaleza, 2020.  
Orientação: Profa. Dra. Carla Freitas de Andrade.  
Coorientação: Prof. Dr. Paulo Alexandre Costa Rocha.
1. Energia eólica. 2. Algoritmos de otimização metaheurísticos. 3. Análise estatística de múltiplos critérios.. 4. Modelagem da velocidade do vento. I. Título.

---

CDD 620.1

KEVIN SANTOS GUEDES

APLICAÇÃO DE ALGORITMOS DE OTIMIZAÇÃO METAHEURÍSTICOS NA  
ESTIMATIVA DOS PARÂMETROS ÓTIMOS DE DIFERENTES DISTRIBUIÇÕES DE  
VELOCIDADE DO VENTO EM DUAS CIDADES DO NORDESTE BRASILEIRO

Dissertação submetida à Coordenação do  
Curso de Pós-Graduação em Engenharia Me-  
cânica da Universidade Federal do Ceará,  
como requisito parcial à obtenção do título  
de mestre em Engenharia Mecânica. Área de  
Concentração: Processos, Equipamentos e Sis-  
temas para Energias Renováveis.

Aprovada em: 30/01/2020

BANCA EXAMINADORA

---

Prof<sup>a</sup>. Dr<sup>a</sup>. Carla Freitas de Andrade  
Universidade Federal do Ceará (UFC)

---

Prof. Dr. Paulo Alexandre Costa Rocha  
Universidade Federal do Ceará (UFC)

---

Prof. Dr. Francisco Olimpio Moura Carneiro  
Universidade da Integração Internacional da Lusofonia  
Afro-Brasileira (UNILAB)  
Membro Externo

---

Me. Adão Linhares Muniz  
Secretário Executivo de Energia e Telecomunicações da  
Secretaria de Infraestrutura do Estado do Ceará  
(SEINFRA)  
Membro Externo

A Deus e sua Mãe Maria Santíssima.

## AGRADECIMENTOS

A Deus e sua Mãe Maria Santíssima, pelas graças concedidas durante a minha vida, sem as quais eu jamais seria capaz de fazer algo bom.

À minha família, por sempre estar ao meu lado nas dificuldades e alegrias.

Aos meus amigos Arthur, Júlio, Caio, Paulo, Venício, Ahed, Davi e Nicolas por suportarem as minhas ausências, pelas incontáveis resenhas e por sempre me apoiarem nos meus planos.

À Sofia, pelo seu constante apoio, carinho, alegria e duradoura amizade.

Aos colegas de turma e ao Valdi, pela disposição, ajuda e companheirismo durante o curso.

Aos professores Dr<sup>a</sup>. Carla Freitas de Andrade e Dr. Paulo Alexandre Costa Rocha pelos conselhos e pela excelente orientação.

Às minhas amigas Eduarda e Mariana, pela amizade, brincadeiras e suporte.

Aos colegas Lindemberg, Marcus e Rivanilso, pelo compartilhamento de conteúdos e ferramentas essenciais para a conclusão deste trabalho.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001

*"Chegará o dia em que teremos que provar ao mundo que a grama é verde."  
(G.K. Chesterton)*

## RESUMO

Para um melhor aproveitamento da energia eólica, é imprescindível uma minuciosa análise do recurso eólico. Nesta análise, as distribuições de velocidade do vento desempenham um importante papel, sendo a mais comum a de Weibull. Contudo, esta distribuição não é sempre a mais adequada, o que torna necessária a avaliação de diferentes distribuições para a obtenção de informações mais confiáveis. Outra etapa essencial é a estimativa dos parâmetros que regem estas distribuições, pois a precisão destas estimativas afeta diretamente os cálculos de geração de energia. Nos últimos anos, diferentes métodos de otimização foram utilizados com esta finalidade e, em comparação com os tradicionais métodos numéricos, eles apresentaram melhor desempenho. Contudo, em energia eólica, a aplicação destes métodos é centrada nas distribuições convencionais de dois parâmetros, tais como Weibull, Gamma e Lognormal. Ademais, diferentes autores ressaltam a falta de estudos que utilizam métodos de otimização com este fim. Diante deste cenário, quatro algoritmos de otimização metaheurístico (MOA), nomeadamente, Pássaros Migratórios, Competição Imperialista, Busca Harmônica e Busca do Cuco, foram utilizados, neste estudo, para ajustar 11 distribuições de velocidade do vento. O estudo foi conduzido em duas cidades no Nordeste do Brasil, sendo esta uma das melhores regiões do mundo para geração de energia eólica, uma vez que os ventos são propícios para tal propósito. Foram também analisadas três funções objetivo distintas para determinar qual delas deve ser aplicada aos MOA para que um melhor ajuste seja obtido. Por fim, os ajustes obtidos pelos MOA foram confrontados, através de uma objetiva análise estatística, com os obtidos pelo método numérico Máxima Verossimilhança (MLE). A função objetivo que garantiu melhores resultados foi a maximização do coeficiente de determinação. Em relação ao MLE, os MOA apresentaram ajustes significativamente melhores, visto que os valores de Score Global ( $GS$ ) obtidos foram menores. Tal fato demonstra a capacidade dos MOA de estimarem precisamente os parâmetros dos modelos de distribuição. Nas duas regiões analisadas, os modelos de três parâmetros proporcionaram, em geral, melhor ajuste do que os de dois parâmetros. Em São João do Cariri, o melhor ajuste foi obtido através da distribuição Gamma Generalizada, que apresentou  $GS = 0,005766$ , e a amplamente utilizada distribuição de Weibull ocupou a quarta posição com  $GS = 0,009369$ . Em Petrolina, o melhor ajuste foi obtido através da distribuição Lindley Generalizada Extendida, que apresentou  $GS = 0,005246$ , e a de Weibull ocupou a quinta posição com  $GS = 0,007894$ . Os modelos de Birnbaum-Saunders e Lognormal foram os que apresentaram piores ajustes nas duas regiões.

**Palavras-chave:** Energia eólica, Algoritmos de otimização metaheurísticos, Modelagem da velocidade do vento, Análise estatística de múltiplos critérios.

## ABSTRACT

For a better use of wind energy, a thorough analysis of the wind resource is essential. In this analysis, wind speed distributions play an important role, the most common being the Weibull distribution. However, this distribution is not always the most suitable, which makes it necessary to evaluate different distributions to obtain more reliable information. Another essential step is the estimation of the parameters governing these distributions, as the accuracy of these estimates directly affects the power generation calculations. In the last years, different optimization methods have been used for this purpose and, compared to traditional numerical methods, they have performed better. However, in wind energy, the application of these methods is centered on the conventional two-parameters distributions, such as Weibull, Gamma and Lognormal. Moreover, different authors emphasize the lack of studies that use optimization methods for this purpose. Given this scenario, four metaheuristic optimization algorithms (MOA), namely, Migrating Birds, Imperialist Competition, Harmony Search and Cuckoo Search, were used in this study to fit 11 wind speed distributions. The study was conducted in two cities in Northeastern Brazil, which is one of the best regions in the world for wind energy generation, since winds are favorable for this purpose. Three distinct objective functions were also analyzed to determine which one should be applied to the MOA in order to obtain a better fit. Finally, the fits obtained by the MOA were compared, through an objective statistical analysis, with those obtained by the Maximum Likelihood (MLE) numerical method. The objective function that achieved best results was the maximization of the determination coefficient. Regarding the MLE, the MOA presented significantly better fits, since the Global Score ( $GS$ ) values obtained were lower. This demonstrates the ability of MOA to accurately estimate the parameters of the distribution models. In both analyzed regions, the three-parameter models generally provided better fit than the two-parameter models. In São João do Cariri, the best fit was obtained through the Generalized Gamma distribution, which presented  $GS = 0,005766$ , and the widely used Weibull distribution ranked fourth with  $GS = 0,009369$ . In Petrolina, the best fit was obtained through the Extended Generalized Lindley distribution, which presented  $GS = 0,005246$ , and the Weibull distribution ranked fifth with  $GS = 0,007894$ . Birnbaum-Saunders and Lognormal models presented the worst fits in both regions.

**Keywords:** Wind energy, Metaheuristic optimization algorithms, Wind speed modeling, Multi-criteria statistical analysis.

## LISTA DE FIGURAS

Figura 1 – Aerodinâmica de voo das aves migratórias . . . . .	30
Figura 2 – Processo de Assimilação do método ICA . . . . .	32
Figura 3 – Distribuições de dois (esquerda) e três (direita) parâmetros em SCR-25 .	51
Figura 4 – Distribuições de dois (esquerda) e três (direita) parâmetros ajustadas pelo MLE em SCR-25 . . . . .	52
Figura 5 – Distribuições de dois (esquerda) e três (direita) parâmetros em PTR-11 .	57
Figura 6 – Distribuições de dois (esquerda) e três (direita) parâmetros ajustadas pelo MLE em PTR-11 . . . . .	58

## LISTA DE TABELAS

Tabela 1 – Modelos de distribuição da velocidade do vento . . . . .	26
Tabela 2 – Informações geográficas das estações SCR-25 e PTR-11 . . . . .	38
Tabela 3 – Estatísticas básicas dos dados anemométricos . . . . .	39
Tabela 4 – Configuração dos algoritmos metaheurísticos . . . . .	44
Tabela 5 – Função objetivo geradora do melhor resultado . . . . .	45
Tabela 6 – Análise estatística SCR-25 . . . . .	47
Tabela 7 – Melhores modelos de distribuição em SCR-25 . . . . .	50
Tabela 8 – Análise estatística PTR-11 . . . . .	53
Tabela 9 – Melhores modelos de distribuição em PTR-11 . . . . .	56
Tabela 10 – Parâmetros estimados em SCR-25 . . . . .	67
Tabela 11 – Parâmetros estimados em PTR-11 . . . . .	68

## LISTA DE ABREVIATURAS E SIGLAS

W	Distribuição de Weibull
G	Distribuição Gamma
BS	Distribuição de Birnbaum-Saunders
N	Distribuição de Nakagami
LN	Distribuição Lognormal
GL	Distribuição Lindley Generalizada
GEV	Distribuição Valor Extremo Generalizado ( <i>Generalized Extreme Value</i> )
B	Distribuição Burr
D	Distribuição Dagum
EGL	Distribuição Lindley Generalizada Extendida ( <i>Extended Generalized Lindley</i> )
GG	Distribuição Gamma Generalizada
GI	Gaussiana Inversa
WI	Weibull Inversa
PV	Pearson Tipo V
Norm	Normal
Gum	Gumbel
K	Kappa
R	Rayleigh
MOA	Algoritmo de otimização metaheurístico ( <i>Metaheuristic optimization algorithms</i> )
MBO	Pássaros Migratórios ( <i>Migrating Birds Optimization</i> )
ICA	Competição Imperialista ( <i>Imperialist Competitive Algorithm</i> )

HS	Busca Harmônica ( <i>Harmony Search</i> )
CS	Busca do Cuco ( <i>Cuckoo Search</i> )
PSO	Enxame de Partículas ( <i>Particle Swarm Optimization</i> )
GWO	Otimizador do Lobo Cinzento ( <i>Grey Wolf Optimizer</i> )
GA	Algoritmo Genético ( <i>Genetic Algorithm</i> )
ACO	Método da Colônia de Formigas ( <i>Ant Colony Optimization</i> )
BA	Algoritmo do Morcego ( <i>Bat Algorithm</i> )
SA	Recozimento Simulado ( <i>Simulated Annealing</i> )
FA	Algoritmo do Vaga-lume ( <i>Firefly Algorithm</i> )
MLE	Método da Máxima Verossimilhança ( <i>Maximum Likelihood Estimation</i> )
MMLE	Método da Máxima Verossimilhança Modificado ( <i>Modified Maximum Likelihood Estimation</i> )
GM	Método Gráfico ( <i>Graphical Method</i> )
MM	Método do Momento ( <i>Method of Moment</i> )
EM	Método Empírico ( <i>Empirical Method</i> )
LSM	Método dos Mínimos Quadrados ( <i>Least Squares Method</i> )
EPF	Método da Energia Padrão ( <i>Energy Pattern Factor Method</i> )
SONDA	Sistema de Organização de Dados Ambientais
BSRN	<i>Baseline Surface Radiation Network</i>
SCR-25	Estação de coleta de dados em São João do Cariri - PB
PTR-11	Estação de coleta de dados em Petrolina - PE
SNCDFT	Transformação da Função de Distribuição Cumulativa Normal Padrão ( <i>Standard Normal Cumulative Distribution Function Transformation</i> )
HM	Memória Harmônica ( <i>Harmony Memory</i> )

## LISTA DE SÍMBOLOS

$x$	Dados de velocidade do vento
$n$	Número total de dados de velocidade do vento
$k$	Parâmetro de forma
$c$	Parâmetro de escala
$p$	Segundo parâmetro de forma
$u$	Parâmetro de localização
$f(x)$	Função densidade de probabilidade
$F(x)$	Função de distribuição cumulativa
$\bar{P}_w$	Potência média gerada por um aerogerador
$P_w(x)$	Potência de saída de uma turbina eólica em função da velocidade do vento
$\gamma_s$	Função gamma incompleta superior
$\gamma_i$	Função gamma incompleta inferior
$\Phi$	Função de distribuição cumulativa normal padrão
$erf$	Função erro
$n_{[i,j]}$	Número aleatório ente $i$ e $j$
$N_p$	Número de pássaros na formação em V
$n_v$	Número de soluções vizinhas
$n_{nu}$	Número de soluções não utilizadas
$b$	Pássaro (solução) no método Pássaros Migratórios
$\hat{b}$	Solução vizinha no método Pássaros Migratórios
$m$	Número de iterações para a substituição do pássaro líder
$N_{pop}$	População total de países

$N_{imp}$	Número de imperialistas
$N_{col}$	Número de colônias
$P_{rev}$	Probabilidade de Revolução
$NC_i$	Número de colônias recebidas pelo $i$ -ésimo império
$c_i$	Custo do $i$ -ésimo imperialista
$C_i$	Custo normalizado do $i$ -ésimo imperialista
$d$	Distância percorrida por uma colônia em direção ao seu imperialista
$d_{c-i}$	Distância entre uma colônia e seu imperialista
$\beta_{ICA}$	Número positivo maior que 1
$TC_i$	Custo total do $i$ -ésimo império
$NTC_i$	Custo total normalizado do $i$ -ésimo império
$P_{p_i}$	Probabilidade de posse do $i$ -ésimo império
$N_h$	Número de harmonias
$HCMR$	Taxa de Consideração da Memória Harmônica ( <i>Harmony Memory Considering Rate</i> )
$PAR$	Taxa de Ajuste do Som ( <i>Pitch Adjusting Rate</i> )
$bw$	Comprimento de Banda ( <i>Bandwidth</i> )
$N_n$	Número de ninhos
$p_d$	Probabilidade de detecção
$\beta_{CS}$	Constante no método Busca do Cuco
$\alpha$	Constante no método Busca do Cuco
$\nu$	Número aleatório de uma distribuição normal com média 0 e desvio padrão $\sigma_{nu}$
$\mu$	Número aleatório de uma distribuição normal com média 0 e desvio padrão $\sigma_{mu}$
$y_i^{t+1}$	Nova solução no método Busca do Cuco

$y_i^t$	Solução atual no método Busca do Cuco
$y_{melhor}^t$	Melhor solução atual no método Busca do Cuco
$L$	Função semelhança
$f(\theta x_i)$	Distribuição de probabilidade em função dos seus parâmetros
$\theta$	Conjunto de parâmetros que regem uma distribuição de probabilidade
$RMSE$	Raíz do Erro Quadrático Médio ( <i>Root Mean Squared Error</i> )
$RRMSE$	$RMSE$ relativo
$NRMSE$	$RMSE$ normalizado
$MSE$	Erro quadrático médio ( <i>Mean Squared Error</i> )
$MAE$	Erro médio absoluto ( <i>Mean Absolute Error</i> )
$BIC$	Critério de informação Bayesiano ( <i>Bayesian Information Criterion</i> )
$KS$	Teste de Kolmogorov-Smirnov ( <i>Kolmogorov-Smirnov Test</i> )
$F_i$	Probabilidades cumulativas observadas
$\hat{F}_i$	Probabilidades cumulativas previstas
$AIC$	Critério de informação de Akaike ( <i>Akaike Information Criterion</i> )
$n_p$	Número de parâmetros do modelo de distribuição
$DSK$	Desvio de Assimetria e Curtose ( <i>Deviation of skewness and kurtosis</i> )
$\gamma_1$	Assimetria de uma distribuição de probabilidade
$\gamma_2$	Curtose de uma distribuição de probabilidade
$\bar{x}$	Média dos dados de velocidade do vento
$s$	Desvio padrão dos dados de velocidade do vento
$g_1$	Assimetria empírica
$g_2$	Curtose empírica
$DS$	Desvio de Assimetria ( <i>Deviation of Skewness</i> )

$DK$	Desvio de Curtose ( <i>Deviation of Kurtosis</i> )
$R^2$	Coefficiente de determinação
$w_{ij}$	$i$ -ésimo valor ( $i$ -score) do critério $j$
$\bar{w}_j$	Média do critério $j$
$s_j$	Desvio padrão do critério $j$
$z_{ij}$	Score padronizado
$\Phi(z_{ij})$	SNCDF do score padronizado $z_{ij}$
$GS$	Score Global ( <i>Global Score</i> )

## SUMÁRIO

1	INTRODUÇÃO . . . . .	19
1.1	Objetivos . . . . .	20
2	FUNDAMENTAÇÃO TEÓRICA . . . . .	22
2.1	Modelagem do Recurso Eólico . . . . .	22
2.1.1	<i>Distribuições de Velocidade do Vento</i> . . . . .	24
2.2	Métodos de Estimativa dos Parâmetros . . . . .	25
2.3	Algoritmos de Otimização Metaheurísticos . . . . .	28
2.3.1	<i>Pássaros Migratórios</i> . . . . .	29
2.3.1.1	<i>Algoritmo Pássaros Migratórios</i> . . . . .	30
2.3.2	<i>Competição Imperialista</i> . . . . .	31
2.3.2.1	<i>Algoritmo de Competição Imperialista</i> . . . . .	31
2.3.3	<i>Busca Harmônica</i> . . . . .	33
2.3.3.1	<i>Algoritmo Busca Harmônica</i> . . . . .	33
2.3.4	<i>Busca do Cuco</i> . . . . .	35
2.3.4.1	<i>Algoritmo Busca do Cuco</i> . . . . .	35
2.4	Método da Máxima Verossimilhança . . . . .	36
3	METODOLOGIA . . . . .	38
3.1	Dados de velocidade do vento . . . . .	38
3.2	Linguagem R . . . . .	39
3.3	Análise estatística de múltiplos critérios . . . . .	39
3.4	Funções objetivo . . . . .	42
3.5	Critério de parada . . . . .	43
3.6	Configuração dos métodos de otimização metaheurísticos . . . . .	43
4	RESULTADOS E DISCUSSÕES . . . . .	45
4.1	Seleção da Função Objetivo . . . . .	45
4.2	Estação SCR-25 . . . . .	47
4.2.1	<i>Seleção do modelo de distribuição ótimo para SCR-25</i> . . . . .	50
4.3	Estação PTR-11 . . . . .	53
4.3.1	<i>Seleção do modelo de distribuição ótimo para PTR-11</i> . . . . .	55
5	CONCLUSÃO . . . . .	59

REFERÊNCIAS . . . . .	61
APÊNDICE A – PARÂMETROS ESTIMADOS . . . . .	67

## 1 INTRODUÇÃO

Em meio à crescente demanda de energia e dos impactos ambientais causados pela utilização de combustíveis fósseis, as fontes renováveis recebem especial atenção como meios viáveis para suprir tal necessidade e reduzir as emissões de dióxido de carbono na atmosfera e a dependência dos combustíveis fósseis. Dentre as diversas fontes renováveis, a energia eólica é uma que apresenta acelerado ritmo de desenvolvimento em várias regiões do planeta (MOHAMMADI; ALAVI; MCGOWAN, 2017). Em 2018, no Brasil, foram instalados 75 parques eólicos, somando um total de 1,94MW. Tal valor representa 30,42% de toda a potência instalada no mesmo ano (ABEEÓLICA, 2019a). De 2010 a 2019, a potência de energia eólica *onshore* instalada nas Américas aumentou 221,7%, passando de 46GW para 148GW (GWEC, 2020).

Para um melhor aproveitamento do recurso eólico de uma região, diversos fatores devem ser analisados, sendo um dos mais importantes a distribuição de probabilidade que melhor se ajusta aos dados de vento da região em questão. Uma vez conhecida esta distribuição, o potencial eólico da região e a viabilidade de instalação de um parque eólico podem ser devidamente mensurados. Por muitos anos, a distribuição de Weibull foi a mais utilizada. Contudo, estudos recentes demonstraram que esta distribuição não é sempre a mais adequada para descrever determinados regime de ventos, de maneira tal que, para a obtenção de uma maior precisão na estimativa do potencial eólico, é essencial a análise de diferentes modelos de distribuição. Tal fato pode ser observado nos estudos conduzidos por Masseran (2018), Kantar et al. (2018), Aries, Boudia e Ounis (2018), Mohammadi, Alavi e McGowan (2017), Jung e Schindler (2017), Arslan, Acitas e Senoglu (2017) e Wais (2017).

Outra etapa essencial é a correta estimativa dos parâmetros que regem estas distribuições, pois, ainda que um determinado modelo seja adequado para uma região, se os seus parâmetros não forem estimados de forma correta, o ajuste obtido será pouco preciso, o que pode comprometer as análises futuras. Os métodos que são utilizados para estimar tais parâmetros podem ser divididos em duas categorias: Métodos Numéricos e Métodos de Otimização Metaheurísticos (MOA - *Metaheuristic Optimization Algorithms*). Esta última categoria apresenta aplicações recentes em diversas áreas, o que pode ser visto nos estudos conduzidos por Jaafari et al. (2019), Mansourzadeh et al. (2019), êardhanan et al. (2019) e Jalal e Goharzay (2019). Em energia eólica, foi demonstrado por diferentes pesquisadores, tais como Zhao et al. (2019), Wang et al. (2018), Jiang et al. (2017), Andrade et al. (2019) e Wang, Hu e Ma (2016), as vantagens e o superior desempenho dos métodos de otimização em relação aos métodos numéricos. No entanto, a partir de uma revisão na literatura, foi observado que as aplicações dos métodos de otimização estão centradas principalmente em

modelos de distribuição convencionais de dois parâmetros, como pode ser visto nos recentes trabalhos conduzidos por Zhao et al. (2019), Wang et al. (2018) e Jiang et al. (2017). Além disso, foi ressaltado por diferentes autores a falta de estudos que utilizassem tais métodos na estimativa dos parâmetros desconhecidos das distribuições de velocidade do vento.

Inspirado por estes cenários, foram utilizadas, neste trabalho, quatro MOA, sendo eles, Pássaros Migratórios (MBO - *Migrating Birds Optimization*), Competição Imperialista (ICA - *Imperialist Competitive Algorithm*), Busca Harmônica (HS - *Harmony Search*) e Busca do Cuco (CS - *Cuckoo Search*) para ajustar um grupo de onze modelos de distribuições de velocidade do vento, no qual estão inclusos modelos de distribuição não convencionais e modelos de três parâmetros. Ademais, três funções objetivo distintas foram aplicadas aos métodos de otimização a fim de determinar qual delas proporcionou ajustes mais precisos.

Foram utilizados dados de velocidade do vento coletados em duas cidades no Nordeste do Brasil, sendo elas São João do Cariri - PB e Petrolina - PE. A região do Nordeste brasileiro é uma das melhores do mundo para a geração de energia eólica. Nas regiões Sul e Nordeste do Brasil, os ventos são predominantemente unidirecionais, constantes e de velocidade estável, o que os tornam extremamente propícios para a geração de energia eólica (ABEEÓLICA, 2019c). Durante a temporada da "safra dos ventos", que corresponde aos meses de Junho a Novembro, o fator de capacidade médio mensal no Brasil pode chegar aos 60%, enquanto que, durante o mesmo período, no Nordeste Brasileiro, é comum que os parques eólicos atinjam fatores de capacidade superiores a 80% (ABEEÓLICA, 2019b; ABEEÓLICA, 2019c).

Posteriormente, os resultados obtidos foram avaliados a partir de uma abordagem integrada recentemente proposta por Masseran (2018) que considera mutuamente, em um único valor denominado, neste trabalho, de Score Global, quatro testes estatísticos comumente utilizados neste tipo de estudo, nomeadamente, Coeficiente de Determinação ( $R^2$ ), Teste de Kolmogorov-Smirnov ( $KS$ ), Critério de Informação de Akaike ( $AIC$  - *Akaike Information Criterion*) e Desvio de Assimetria e Curtose ( $DSK$  - *Deviation of skewness and kurtosis*). Por fim, para avaliar a aplicabilidade dos MOA, o método determinístico da Máxima Verossimilhança (MLE - *Maximum Likelihood Estimation*) foi utilizado como referencial, visto que este método é amplamente utilizado e geralmente fornece resultados significativos, como pode ser observado nos estudos conduzidos por Jung e Schindler (2019), Morgan et al. (2011) e Kantar et al. (2018), por exemplo.

## 1.1 Objetivos

O objetivo deste trabalho foi avaliar a aplicabilidade dos algoritmos de otimização metaheurísticos MBO, ICA, HS e CS na determinação dos parâmetros ótimos de onze distribuições de velocidade do vento. Os objetivos específicos são:

1. Implementar os algoritmos de otimização MBO, ICA, HS e CS com três funções objetivo distintas e o método determinístico MLE;
2. Aplicar os métodos de otimização e o método determinístico no ajuste dos dados de velocidade do vento das regiões de São João do Cariri - PB e Petrolina - PE, ambas situadas na região Nordeste do Brasil;
3. Conduzir uma análise estatística de múltiplos critérios para quantificar a qualidade dos ajustes obtidos;
4. Identificar a função objetivo que deve ser aplicada aos métodos de otimização para que os mesmos apresentem ajustes mais precisos;
5. Confrontar os resultados obtidos pelos métodos de otimização configurados com a melhor função objetivo com os obtidos através do MLE, a fim de avaliar a aplicabilidade do MBO, ICA, HS e CS no ajuste das distribuições de probabilidade utilizadas;
6. Determinar os conjuntos (distribuição e método de estimativa dos parâmetros) que proporcionaram maior precisão nas duas regiões analisadas.

## 2 FUNDAMENTAÇÃO TEÓRICA

### 2.1 Modelagem do Recurso Eólico

Ter o conhecimento das distribuições de frequência da velocidade do vento é um fator imprescindível para a caracterização do comportamento da velocidade do vento e para a avaliação do potencial eólico e do desempenho dos sistemas de geração de energia em determinada região (ALAVI; MOHAMMADI; MOSTAFAEIPOUR, 2016). Desse modo, torna-se necessária a determinação do modelo de distribuição que melhor se adequa ao conjunto de dados de velocidade do vento analisados. Se esta distribuição é conhecida, o potencial eólico e a viabilidade econômica do parque eólico podem ser facilmente calculados (ZHAO et al., 2019).

A potência de saída de um aerogerador depende de diversos fatores, tais como altura do rotor, densidade de potência, densidade do ar, distribuição da velocidade do vento, formato da curva de potência e potência nominal da turbina eólica, intensidade de turbulência e velocidade média do vento. Destes fatores, o mais importante é a distribuição de velocidade do vento (AKDAG; GULER, 2015). Em termos matemáticos, a potência média gerada por uma turbina eólica  $\bar{P}_w$  é dada por:

$$\bar{P}_w = \int_0^{\infty} P_w(x)f(x)dx \quad (2.1)$$

na qual  $P_w(x)$  é a potência de saída do aerogerador em função da velocidade do vento  $x$  e  $f(x)$  é a função de distribuição de probabilidade. A partir da Equação 2.1, é possível ver que a seleção incorreta da função  $f(x)$  mais precisa e/ou uma má estimativa dos parâmetros que regem esta distribuição impactam diretamente na potência média gerada pela turbina eólica, o que pode comprometer a análise do recurso eólico e a eficiência na geração de energia em determinada região. De acordo com Alavi, Mohammadi e Mostafaeipour (2016), uma mesma turbina eólica instalada em dois locais com velocidade média do vento semelhante pode apresentar potência de saída completamente distinta devido às diferenças nos regimes de vento dos dois locais, fato que ressalta a importância das distribuições de frequência da velocidade do vento. Ademais, a maior incerteza no cálculo de  $\bar{P}_w$  encontra-se na escolha da distribuição da velocidade do vento  $f(x)$ , visto que o fabricante da turbina eólica conhece com bastante precisão a variável  $P_w(x)$  (MORGAN et al., 2011). As distribuições são também utilizadas no cálculo de importantes indicadores, tais como densidade média de potência, densidade média de potência efetiva, fator de disponibilidade e fator de capacidade. Sendo assim, a precisão destas distribuições é essencial para um bom planejamento de parques eólicos e para a redução das incertezas atreladas ao desenvolvimento do potencial eólico.

Dentre as diversas distribuições, a de Weibull (W) é uma das mais utilizadas, como pode ser visto nos conduzidos por Zhao et al. (2019), Masseran (2018), Aries, Boudia e Ounis (2018), Wang et al. (2018), Jung e Schindler (2017) e Jiang et al. (2017), por exemplo. A facilidade de estimar os dois parâmetros dessa distribuição, a sua flexibilidade e boa precisão em diferentes regiões são as principais características que tornam esta distribuição uma das mais utilizadas (ZHAO et al., 2019). Contudo, nem sempre esta distribuição é a mais adequada para descrever certos regimes de vento (WANG; HU; MA, 2016). Alguns exemplos de outras distribuições que já foram utilizadas com este propósito são:

- Lindley Generalizada Extendida (EGL - *Generalized Lindley Distribution*);
- Rayleigh (R);
- Lognormal (LN);
- Gamma (G);
- Birnbaum-Saunders (BS);
- Nakagami (N);
- Lindley Generalizada (GL - *Generalized Lindley*);
- Valor Extremo Generalizado (GEV - *Generalized Extreme Value*);
- Gaussiana Inversa (GI);
- Weibull Inversa (WI);
- Pearson Tipo V (PV);
- Burr (B);
- Dagum (D);
- Normal (Norm);
- Log-Logística (LL);
- Gumbel (Gum);
- Kappa (K).

Kantar et al. (2018) compararam o desempenho da distribuição EGL com o das distribuições W, R, LN e G em diferentes regiões da Turquia. Os resultados demonstraram que a EGL foi a que apresentou melhor precisão na maioria das regiões examinadas. Estudos similares foram conduzidos por Mohammadi, Alavi e McGowan (2017), Alavi, Mohammadi e Mostafaeipour (2016) e Arslan, Acitas e Senoglu (2017), demonstrando o bom desempenho, respectivamente, das distribuições de BS, N e GL no ajuste dos dados anemométricos em diferentes regiões.

Aries, Boudia e Ounis (2018), com o intuito de reduzir as incertezas atreladas às estimativas do recurso eólico, analisaram a precisão de oito distribuições de probabilidade na modelagem da velocidade do vento em quatro regiões da Argélia. De acordo com os testes estatísticos coeficiente de determinação e raiz quadrada do erro quadrático médio, as distribuições G e GEV foram as que apresentaram melhor desempenho. Em um estudo conduzido por Brano et al. (2011) no sul da Itália, a distribuição de três parâmetros B foi a que apresentou maior precisão quando comparada com as distribuições W, R, G, LN, GI e PV. Jung e Schindler (2017) avaliaram o desempenho de 24 distribuições unimodais e 24 distribuições bimodais em um estudo conduzido em escala global e, a partir dos resultados, foi demonstrado que apenas em poucas regiões a distribuição W foi a que proporcionou maior precisão.

Diante destes fatos, é possível ver que a precisão de uma distribuição está diretamente relacionada com as características do regime de ventos no qual ela foi aplicada, de maneira tal que, se um determinado modelo de distribuição apresentou boa precisão em um região, isto não significa que em todas as regiões ele irá proporcionar bom ajuste. Desse modo, torna-se imprescindível, para a obtenção de um ajuste mais preciso, a análise de diferentes modelos de distribuição.

### ***2.1.1 Distribuições de Velocidade do Vento***

Com o intuito de garantir a seleção de um modelo preciso, onze distribuições de probabilidade foram avaliadas neste trabalho, sendo elas:

- Weibull (W): Uma das mais utilizadas na modelagem dos recurso eólico e que apresenta bom desempenho em diversas regiões;
- Gamma (G) e Lognormal (LN): Ambas muito comuns na análise do recurso eólico;
- Birnbaum-Saunders (BS), Nakagami (N), Lindley Generalizada (GL - *Generalized Lindley*) e Lindley Generalizada Extendida (EGL - *Extended Generalized Lindley*): Modelos cuja aplicação em energia eólica é mais recente;

- Burr (B), Dagum (D), Valor Extremo Generalizado (GEV - *Generalized Extreme Value*) e Gamma Generalizada (GG): Modelos mais flexíveis e que apresentaram elevado desempenho em diferentes regiões.

Os modelos W, G, BS, N, LN e GL possuem dois parâmetros, já os modelos GEV, B, D, EGL e GG possuem três parâmetros. Deste grupo, somente os modelos W, G e LN são caracterizados como convencionais, isto é, que são mais comumente utilizados neste tipo de estudo. Já os modelos BS, N, GL, GEV, B, D, EGL e GG são caracterizados como não convencionais. A função densidade de probabilidade  $f(x)$  e a função de distribuição acumulada  $F(x)$  destes modelos estão apresentadas na Tabela 1, na qual  $k$  é o parâmetro de forma,  $p$  o segundo parâmetro de forma,  $c$  o parâmetro de escala,  $u$  o parâmetro de localização,  $\gamma_s$  e  $\gamma_i$  são, respectivamente, as funções gamma incompleta superior e inferior e  $\Phi$  a função de distribuição cumulativa normal padrão. Foram também apresentadas algumas referências de estudos anteriores que utilizaram estas distribuições em energia eólica.

## 2.2 Métodos de Estimativa dos Parâmetros

Outra etapa importante para a obtenção de um ajuste preciso é a correta estimativa dos parâmetros que regem as funções de distribuição de probabilidade. Ainda que uma determinada distribuição seja capaz de representar precisamente o regime de ventos de uma região, se os seus parâmetros não forem estimados corretamente, o ajuste obtido será impreciso e pouco confiável. Além disso, a análise do potencial eólico de uma região e a viabilidade econômica de um projeto são importantes aspectos que dependem da boa estimativa desses parâmetros (AKDAG; DINLER, 2009).

A fim de obter uma maior precisão nas estimativas dos parâmetros das distribuições, os pesquisadores utilizam diferentes métodos, conforme pode ser observado nos trabalhos conduzidos por Gugliani et al. (2018), Aries, Boudia e Ounis (2018), Usta et al. (2018), Jung e Schindler (2017), Alavi, Sedaghat e Mostafaeipour (2016), Mohammadi et al. (2016), Akdag e Guler (2015) e Rocha et al. (2012). Os métodos de estimativa comumente utilizados são:

- Método da Máxima Verossimilhança (MLE - *Maximum Likelihood Estimation*);
- Método da Máxima Verossimilhança Modificado (MMLE - *Modified Maximum Likelihood Estimation*);
- Método Gráfico (GM - *Graphical Method*)
- Método do Momento (MM - *Method of Moment*);
- Método Empírico (EM - *Empirical Method*);

Tabela 1 – Modelos de distribuição da velocidade do vento

Modelo	$f(x)$	$F(x)$	Referências
W	$\frac{k}{c} \left(\frac{x}{c}\right)^{k-1} e^{-\left(\frac{x}{c}\right)^k}$	$1 - e^{-\left(\frac{x}{c}\right)^k}$	Wang et al. (2018) Jiang et al. (2017) Andrade et al. (2014) Usta et al. (2018)
G	$\frac{x^{k-1} e^{-\frac{x}{c}}}{c^k \Gamma(k)}$	$\frac{\gamma_i(k, \frac{x}{c})}{\Gamma(k)}$	Masseran (2018) Carta, Ramírez e Velázquez (2009) Mazzeo, Oliveti e Labonia (2018)
BS	$\frac{e^{-\frac{1}{2k^2}(\frac{x}{c} + \frac{c}{x} - 2)}}{2\sqrt{2\pi k c}} \left[ \left(\frac{c}{x}\right)^{\frac{1}{2}} + \left(\frac{c}{x}\right)^{\frac{3}{2}} \right]$	$\Phi \left[ \frac{1}{k} \left\{ \left(\frac{x}{c}\right)^{\frac{1}{2}} - \left(\frac{c}{x}\right)^{\frac{1}{2}} \right\} \right]$	Mohammadi, Alavi e McGowan (2017) Soukissian e Karathanasi (2017) Miao et al. (2019)
N	$\frac{2k^k}{\Gamma(k)c^k} x^{2k-1} e^{-\left(\frac{k}{c}x^2\right)}$	$1 - \frac{\gamma_s(k, \frac{k}{c}x^2)}{\Gamma(k)}$	Aries, Boudia e Ounis (2018) Alavi, Mohammadi e Mostafaeipour (2016) Soukissian e Karathanasi (2017)
LN	$\frac{1}{xk\sqrt{2\pi}} e^{\left[-\frac{1}{2}\left(\frac{\ln(x)-c}{k}\right)^2\right]}$	$\frac{1}{2} + \frac{1}{2} \operatorname{erf} \left[ \frac{\ln(x)-c}{k\sqrt{2}} \right]$	Alavi, Sedaghat e Mostafaeipour (2016) Kiss e János (2008) Ouarda et al. (2015)
GL	$\frac{kc^2(1+x)e^{-cx}}{c+1} \left(1 - \frac{1+c+cx}{1+c} e^{-cx}\right)^{k-1}$	$\left(1 - \frac{1+c+cx}{1+c} e^{-cx}\right)^k$	Arslan, Acitas e Senoglu (2017)
GEV	$\frac{\left[1 - \frac{k(x-u)}{c}\right]^{\frac{1}{k}-1}}{c} e^{\left[-\left(1 - \frac{k(x-u)}{c}\right)^{\frac{1}{k}}\right]}$	$e^{\left[-\left(1 - \frac{k(x-u)}{c}\right)^{\frac{1}{k}}\right]}$	Aries, Boudia e Ounis (2018) Mohammadi, Alavi e McGowan (2017) Miao et al. (2019) Soukissian e Karathanasi (2017)
B	$\frac{kp\left(\frac{x}{c}\right)^{p-1}}{c\left[1 + \left(\frac{x}{c}\right)^p\right]^{k+1}}$	$1 - \left[1 + \left(\frac{x}{c}\right)^p\right]^{-k}$	Brano et al. (2011) Mazzeo, Oliveti e Labonia (2018) Qin, Li e Xiong (2011)
D	$\frac{kp\left(\frac{x}{c}\right)^{kp-1}}{c\left[1 + \left(\frac{x}{c}\right)^p\right]^{k+1}}$	$\left[1 + \left(\frac{x}{c}\right)^p\right]^{-k}$	Chiodo e Falco (2016) Jung et al. (2017) Jung e Schindler (2017)
EGL	$\frac{k^2 pc(1+xc)^{2p-1} e^{k-k(1+xc)^p}}{k+1}$	$1 - \frac{e^{k-k(1+xc)^p} [1+k(1+xc)^p]}{k+1}$	Kantar et al. (2018)
GG	$ p  x^{kp-1} \frac{e^{-\left(\frac{x}{c}\right)^p}}{c^{kp}\Gamma(k)}$	$\gamma_i\left(k, \left(\frac{x}{c}\right)^p\right)$	Morgan et al. (2011) Kiss e János (2008) Carta, Ramírez e Velázquez (2009)

- Método dos Mínimos Quadrados (LSM - *Least Squares Method*);
- Método da Energia Padrão (EPF - *Energy Pattern Factor Method*);

Tais métodos são classificados como Métodos Numéricos ou Métodos Determinísticos. Contudo, há outra categoria mais recente, composta por algoritmos de otimização metaheurísticos (MOA - *Metaheuristic Optimization Algorithms*), cuja aplicação na otimização dos parâmetros desconhecidos das distribuições mostrou-se bastante promissora.

Além disso, diferentes estudos demonstraram o superior desempenho destes métodos quando comparados com os tradicionais Métodos Numéricos. Os estudos conduzidos por Wang et al. (2018), Jiang et al. (2017), Andrade et al. (2019) e Wang, Hu e Ma (2016), por exemplo, comprovam tal fato. Alguns exemplos de métodos que compõem essa categoria são:

- Enxame de Partículas (PSO - *Particle Swarm Optimization*);
- Otimizador do Lobo Cinzento (GWO - *Grey Wolf Optimizer*);
- Algoritmo Genético (GA - *Genetic Algorithm*);
- Método da Colônia de Formigas (ACO - *Ant Colony Optimization*);
- Algoritmo do Morcego (BA - *Bat Algorithm*);
- Recozimento Simulado (SA - *Simulated Annealing*);
- Algoritmo do Vaga-lume (FA - *Firefly Algorithm*);
- Busca Harmônica (HS - *Harmony Search*);
- Busca do Cuco (CS - *Cuckoo Search*);

Wang et al. (2018) utilizaram os métodos de otimização PSO, CS e GWO e os métodos numéricos EM, EPF e MLE para estimar os parâmetros da distribuição W na China. A partir da análise estatística, foi demonstrado que os ajustes obtidos através dos métodos de otimização foram superiores aos obtidos pelos métodos numéricos. Jiang et al. (2017) estimaram os parâmetros das distribuições W, R, G e LN a partir dos métodos de otimização PSO, CS e BA e dos métodos numéricos MM, MLE e LSM. Posteriormente, uma análise comparativa foi conduzida e os resultados demonstraram que todos métodos de otimização apresentaram melhor desempenho do que os métodos numéricos. Wang et al. (2015) utilizaram uma variação do PSO e o CS para estimar os parâmetros das distribuições GEV e Gumbel na modelagem da velocidade do vento *offshore*. Ambos os métodos apresentaram bom desempenho. Wang, Hu e Ma (2016) utilizaram o CS para ajustar as distribuições W, G, R, LN, Norm e LL em quatro regiões da China. Em comparação com os métodos numéricos utilizados, o CS foi o que proporcionou maior precisão no ajuste. De maneira similar, Zhao et al. (2019), Chang (2011), Andrade et al. (2019) e Wang, Wang e Wu (2016) estimaram os parâmetros da distribuição W a partir de diferentes métodos de otimização.

A partir destes fatos, fica evidente o desempenho competitivo dos métodos de otimização em relação aos tradicionais métodos numéricos. Além disso, é possível perceber que os métodos de otimização foram empregados majoritariamente em distribuições convencionais

de dois parâmetros (W, R, G, LN, Norm e LL). Tal fato complementa a justificativa de escolha das distribuições apresentadas na Seção 2.1.1.

### 2.3 Algoritmos de Otimização Metaheurísticos

Algoritmos de otimização metaheurísticos (MOA - *Metaheuristic Optimization Algorithms*) são métodos que apresentam em sua estrutura um determinado *tradeoff* entre aleatoriedade e busca local, isto é, quanto maior a aleatoriedade, menor a busca local e vice-versa. Os principais componentes de qualquer MOA são: Diversificação e Intensificação. Diversificação é a capacidade do método de gerar diversas e distintas soluções, garantindo, deste modo, uma busca global no espaço de possíveis soluções para o problema. Já a intensificação é a capacidade do algoritmo de focar em uma determinada região do espaço de busca explorando a informação de que a atual melhor solução encontra-se nesta região. A combinação destes dois componentes garante, geralmente, que o ótimo global seja alcançável (YANG, 2010c).

Os MOA apresentam também um importante mecanismo de compartilhamento de informações, o que, em determinadas circunstâncias, pode acelerar a convergência do algoritmo (SHAYANFAR; GHAREHCHOPOGH, 2018). Este compartilhamento de informações é uma das principais características que distinguem os MOA de uma simples varredura de todas as soluções possíveis para o problema. Os MOA combinam diferentes conceitos de forma inteligente e simulam estratégias de aprendizado a fim de encontrar a solução ótima de forma eficiente e em menor tempo (OSMAN; LAPORTE, 1996).

Nos MOA, o processo de busca pelos parâmetros ótimos do problema se dá através da minimização (ou maximização) de uma função objetivo, sendo esta uma função previamente definida e com capacidade de representar numericamente a qualidade das soluções geradas. Ademais, é importante ressaltar que, em um mesmo problema, diferentes funções objetivo podem ser utilizadas para conduzir o processo de otimização. Desse modo, torna-se imprescindível uma análise minuciosa dos resultados gerados a partir de diferentes funções objetivo a fim de determinar qual delas é a mais adequada na obtenção da solução ótima do problema.

Neste estudo, foram utilizados quatro MOA para determinar os parâmetros ótimos das distribuições de velocidade do vento e proporcionar um ajuste mais preciso, sendo eles:

- Pássaros Migratórios (MBO - *Migrating Birds Optimization*);
- Competição Imperialista (ICA - *Imperialist Competitive Algorithm*);
- Busca Harmônica (HS - *Harmony Search*);

- Busca do Cuco (CS - *Cuckoo Search*);

Estes métodos foram selecionados devido ao seu desempenho em diversas áreas científicas, tais como energia nuclear, previsão espacial de incêndio, robótica, entre outras, como pode ser observado nos estudos conduzidos por Jaafari et al. (2019), Mansourzadeh et al. (2019), êardhanan et al. (2019), Jalal e Goharzay (2019), Gerist e Maheri (2019), Almeida (2019) e Chiranjeevi e Jena (2018); porque não foram encontradas aplicações dos métodos MBO, ICA e HS em distribuições de três parâmetros e em distribuições não convencionais, sendo neste trabalho, portanto, a primeira vez que eles são utilizados com tal propósito; e, por fim, porque de acordo com Jiang et al. (2017), as pesquisas e aplicações de métodos de otimização na estimativa dos parâmetros das distribuições de velocidade do vento são insuficientes. Ademais, como demonstrado anteriormente na Seção 2.2, a aplicação de tais métodos esteve até então centrada principalmente em distribuições convencionais de dois parâmetros. Sendo assim, ao selecionar quatro MOA, este trabalho apresenta uma ampla comparação do desempenho destes métodos na determinação dos parâmetros ótimos das distribuições de velocidade do vento, focando não em um pequeno grupo de modelos convencionais, mas em um amplo grupo composto de onze distribuições, sendo a maioria delas modelos não convencionais, conforme apresentado na Seção 2.1.1.

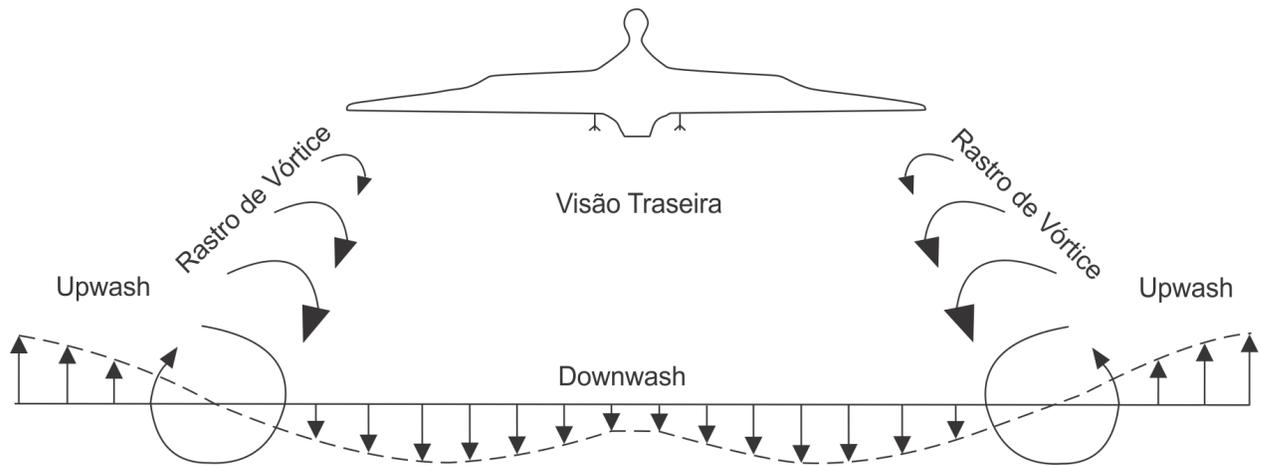
### 2.3.1 *Pássaros Migratórios*

Proposto por Duman, Uysal e Alkaya (2012), o método Pássaros Migratórios (MBO - *Migrating Birds Optimization*) é inspirado na formação de voo em V comumente adotada pelas aves migratórias, na qual o pássaro posicionado na região dianteira, denominado de líder, guia o bando situado nas duas linhas da formação.

Durante o voo das aves, devido à diferença de pressão entre a superfície inferior e superior das asas, um par de vórtices é gerado, conforme a Figura 1. Estes vórtices deixam um rastro composto por uma região de *Downwash*, que aumenta o arrasto induzido nas asas e, conseqüentemente, dificulta o voo, e por uma região de *Upwash*, que reduz o arrasto induzido e proporciona maior economia de energia para as aves localizadas nesta região (MAKAS; YUMUSAK, 2016). A formação em V é organizada de tal maneira que todas as aves, com exceção do líder, são posicionadas na região de *Upwash*. O pássaro líder, por não receber tal benefício, apresenta maior gasto de energia (DUMAN; UYSAL; ALKAYA, 2012).

Em um grupo composto por 25 aves, por exemplo, cada pássaro, com exceção do líder, atinge, em virtude deste benefício aerodinâmico, uma amplitude de voo aproximadamente 70% maior do que a de um pássaro isolado (LISSAMAN; SHOLLENBERGER, 1970).

Figura 1 – Aerodinâmica de voo das aves migratórias



Fonte: Adaptado de Duman, Uysal e Alkaya (2012)

### 2.3.1.1 Algoritmo Pássaros Migratórios

O MBO simula o benefício da economia de energia a partir do compartilhamento de soluções e é conduzido da seguinte maneira (MAKAS; YUMUSAK, 2016):

1. Geração aleatória da população de  $N_p$  pássaros (soluções) e distribuição aleatória destes na formação em  $V$ . O pássaro posicionado na região dianteira é o líder.
2. Geração de  $n_v$  soluções vizinhas para o pássaro líder. Caso uma dessas novas soluções seja melhor, o líder é substituído por ela. As  $2n_{nu}$  melhores soluções não utilizadas são compartilhadas com os próximos dois pássaros da formação.
3. Geração de  $n_v - n_{nu}$  soluções vizinhas para cada um dos dois pássaros subsequentes da formação e combinação destas com as  $n_{nu}$  soluções recebidas do pássaro da frente, totalizando  $n_v$  vizinhos. Caso a solução vizinha seja melhor, o pássaro é substituído por ela e as  $n_{nu}$  melhores soluções não utilizadas são compartilhadas com o próximo pássaro da linha.
4. O passo 3 é executado até que todos os pássaros sejam processados.
5. Neste passo, supõe-se que o pássaro líder, após  $m$  iterações, atinge a exaustão, sendo, portanto, necessária a sua substituição. O líder é então deslocado para o final de umas das duas linhas da formação e o segundo pássaro desta linha assume a posição de líder.
6. Repetição dos passos 2 a 5 até que o critério de parada seja atingido. A melhor solução obtida é retornada.

Neste algoritmo, as soluções vizinhas são geradas de acordo com a Equação 2.2, na qual  $b_{ij}$  é a  $i$ -ésima solução composta por  $j$  dimensões, sendo cada dimensão um dos parâmetros das distribuições de probabilidade,  $\hat{b}_{ij}$  são as soluções vizinhas,  $r \in [1, N_p]$  é um número inteiro diferente de  $i$  escolhido aleatoriamente e  $n_{[-1,1]}$  é um número aleatório entre -1 e 1.

$$\hat{b}_{ij} = b_{ij} + n_{[-1,1]} \cdot (b_{ij} - b_{rj}) \quad (2.2)$$

O pássaro líder é o que gasta mais energia gerando  $n_v$  vizinhos. Já os demais pássaros recebem as  $n_{nu}$  soluções não utilizadas do pássaro da frente e geram apenas  $n_v - n_{nu}$  soluções vizinhas, simulando a economia de energia proporcionada pela formação em V.

### 2.3.2 *Competição Imperialista*

Proposto por Atashpaz-Gargari e Lucas (2007), o método da Competição Imperialista (ICA - *Imperialist Competitive Algorithm*) é um algoritmo de otimização baseado no processo sociopolítico de competição imperialista. Neste método, cada país consiste em uma possível solução e o número de dimensões que compõem cada país, neste trabalho, é igual ao número de parâmetros das distribuições de probabilidade. Os países mais fortes são aqueles que apresentam as melhores soluções, ou seja, melhor valor da função objetivo, também denominada custo neste algoritmo.

#### 2.3.2.1 *Algoritmo de Competição Imperialista*

O ICA é conduzido da seguinte maneira:

1. Geração aleatória da população inicial de  $N_{pop}$  países. Para formação dos impérios, os  $N_{imp}$  países com menor custo (melhor valor da função objetivo) são designados como imperialistas, e os  $N_{col}$  países restantes, denominados colônias, são distribuídos entre os imperialistas. A quantidade de colônias que cada imperialista recebe é diretamente proporcional ao seu poder.
2. Movimentação das colônias em direção aos seus respectivos imperialistas (Assimilação) ou em uma direção aleatória (Revolução). Caso uma das colônias se torne mais poderosa (melhor valor da função objetivo), esta colônia se torna a imperialista do seu respectivo império. A probabilidade de Revolução  $P_{rev}$  é igual à 30% neste trabalho, conforme utilizado por Macêdo (2018) e Santos (2018).
3. Cálculo do custo total dos impérios.
4. A colônia mais fraca do império mais fraco é disputada entre os impérios mais fortes, de maneira tal que o império mais forte é o que apresenta maior probabilidade de posse da colônia mais fraca.

5. Eliminação dos impérios mais fracos. Neste trabalho, a eliminação ocorre quando o império não possui nenhuma colônia.
6. Repetição dos passos 2 a 5 até que reste somente um império, sendo o último imperialista a solução do problema.

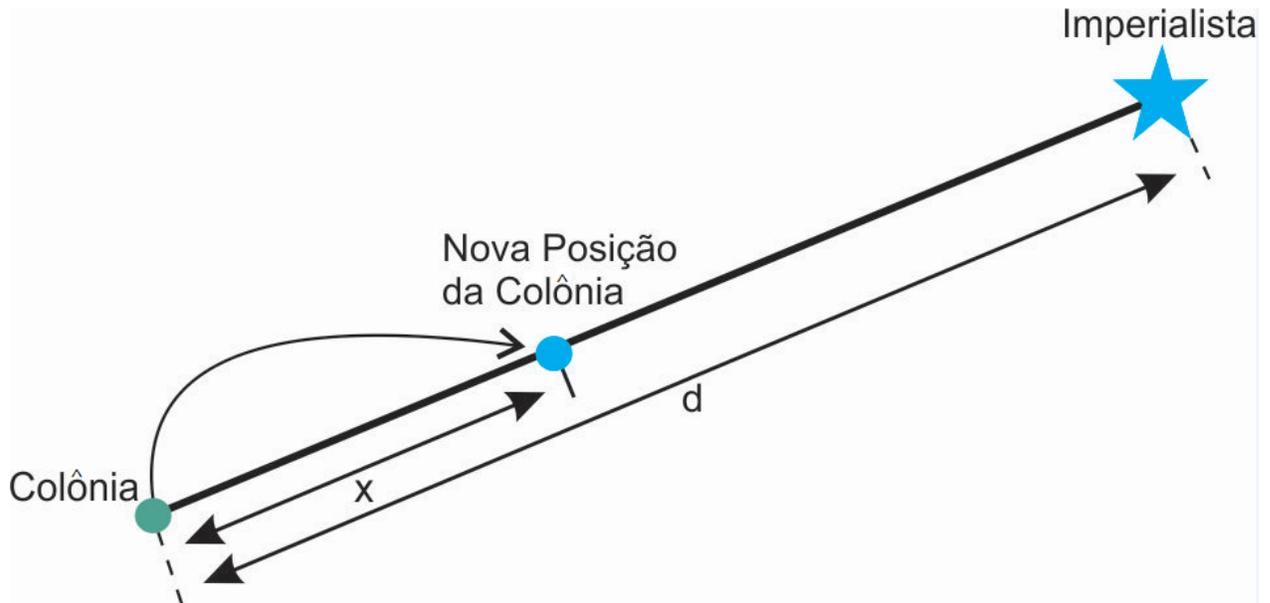
No primeiro passo, o número de colônias  $NC_i$  que o  $i$ -ésimo imperialista recebe é dado pelas Equações 2.4 e 2.3, nas quais  $c_i$  é o custo do  $i$ -ésimo imperialista e  $C_i$  o seu custo normalizado. As colônias que serão distribuídas para cada imperialista são escolhidas aleatoriamente.

$$C_i = c_i - \max(c) \quad (2.3)$$

$$NC_i = \text{arredondamento} \left( \left\lfloor \frac{C_i}{\sum_{n=1}^{N_{imp}} C_n} \cdot N_{col} \right\rfloor \right) \quad (2.4)$$

No segundo passo, a distância  $d$  que a colônia percorre em direção ao seu imperialista (Assimilação) é um valor aleatório entre 0 e  $\beta_{ICA} \cdot d_{c-i}$ , na qual  $\beta_{ICA}$  é um número positivo maior que 1 e  $d_{c-i}$  a distância entre a colônia e seu imperialista. A Figura 2 mostra um esquema deste processo.

Figura 2 – Processo de Assimilação do método ICA



Fonte: Adaptado de Atashpaz-Gargari e Lucas (2007)

No terceiro passo, o custo total  $TC_i$  dos impérios é calculado pela Equação 2.5, na qual  $\xi = 0, 1$ , conforme utilizado por Atashpaz-Gargari e Lucas (2007).

$$TC_i = \text{Custo}(\text{imperialista}_i) + \xi \cdot \text{m\u00e9dia}\{\text{Custo}(\text{col\u00f4nias do imp\u00e9rio}_i)\} \quad (2.5)$$

No quarto passo, a probabilidade de posse  $P_{p_i}$  de cada imp\u00e9rio \u00e9 calculada a partir das Equa\u00e7\u00f5es 2.6 e 2.7, na quais  $NTC_i$  \u00e9 o custo total normalizado do  $i$ -\u00e9simo imp\u00e9rio.

$$NTC_i = TC_i - \max(TC) \quad (2.6)$$

$$P_{p_i} = \left| \frac{NTC_i}{\sum_{n=1}^{N_{imp}} NTC_n} \right| \quad (2.7)$$

Para cada imp\u00e9rio \u00e9 gerado um n\u00famero aleat\u00f3rio  $n_{[0,1]}$  entre 0 e 1. O imp\u00e9rio que apresentar maior valor de  $P_{p_n} - n_{[0,1]}$  recebe a col\u00f4nia que est\u00e1 sendo disputada.

O mecanismo de competi\u00e7\u00e3o imperialista traz, gradualmente, um aumento no poder dos imp\u00e9rios mais fortes e uma diminui\u00e7\u00e3o no poder dos imp\u00e9rios mais fracos (ATASHPAZ-GARGARI; LUCAS, 2007). Com o decorrer das itera\u00e7\u00f5es, todos os imp\u00e9rios, com exce\u00e7\u00e3o do mais forte, entram em colapso e todas as col\u00f4nias s\u00e3o controladas por um \u00fanico imperialista, sendo este a solu\u00e7\u00e3o do problema (MORADI et al., 2014).

### 2.3.3 Busca Harm\u00f4nica

Inicialmente proposto por Geem, Kim e Loganathan (2001) o m\u00e9todo Busca Harm\u00f4nica (HS - *Harmony Search*) \u00e9 baseado no processo artificial de busca pela melhor conson\u00e2ncia de notas musicais por parte de um grupo musical (MANSOURZADEH et al., 2019). Esta combina\u00e7\u00e3o de notas \u00e9 denominada harmonia. A busca se d\u00e1 atrav\u00e9s da experi\u00eancia do grupo ou de um processo aleat\u00f3rio de improvisa\u00e7\u00e3o, e, a partir de um crit\u00e9rio est\u00e9tico, \u00e9 determinada a qualidade da nova harmonia gerada. De forma an\u00e1loga, o HS procura mimetizar este comportamento a partir da otimiza\u00e7\u00e3o das solu\u00e7\u00f5es existentes e da gera\u00e7\u00e3o aleat\u00f3ria de novas solu\u00e7\u00f5es, adotando como crit\u00e9rio de qualidade das solu\u00e7\u00f5es o valor da fun\u00e7\u00e3o objetivo.

#### 2.3.3.1 Algoritmo Busca Harm\u00f4nica

Como proposto por Geem, Kim e Loganathan (2001), o HS \u00e9 conduzido da seguinte maneira:

1. Inicializa\u00e7\u00e3o da Mem\u00f3ria Harm\u00f4nica (HM - *Harmony Memory*) composta por  $N_h$  harmonias geradas de forma aleat\u00f3ria, sendo cada uma delas uma poss\u00edvel solu\u00e7\u00e3o. Neste estudo, cada nota musical que comp\u00f5e uma harmonia \u00e9 um dos par\u00e2metros das distribui\u00e7\u00f5es de velocidade do vento.

2. Improvisação de novas harmonias. Caso a nova harmonia seja melhor do que a pior harmonia presente na HM, a pior é substituída pela nova.
3. Repetição do passo 2 até que o critério de parada seja atingido.

No segundo passo, a geração de cada nota musical que compõe a nova harmonia é regida pelos parâmetros Taxa de Consideração da Memória Harmônica (*HMCR* - *Harmony Memory Considering Rate*), Taxa de Ajuste do Som (*PAR* - *Pitch Adjusting Rate*) e Comprimento de Banda (*bw* - *Bandwidth*). O *HMCR* varia de 0 a 1 e é definido como a probabilidade de um valor existente na HM ser selecionado para gerar a nova nota, enquanto que  $1 - HMCR$  é a probabilidade da nova nota ser gerada aleatoriamente a partir conjunto total de possíveis soluções. Se o *HMCR* for próximo de 1, a diversidade das soluções não será preservada (pior busca global), enquanto que se o *HMCR* for próximo de 0, a diversidade será alta (ASKARZADEH; ZEBARJADI, 2014). Este parâmetro é de suma importância, pois, caso as notas que compõem a solução ótima não estiverem presentes na HM, seria impossível a obtenção desta solução.

Posteriormente, cada componente obtido pela consideração da memória harmônica é examinado para decidir se ele será ou não ajustado. Com esta finalidade, é introduzido o parâmetro *PAR*, que varia de 0 a 1 e é definido como a probabilidade da nota musical gerada ser substituída por um elemento de sua vizinhança dentro de um intervalo  $[-bw, bw]$  (MANSOURZADEH et al., 2019). Os parâmetros *PAR* e *bw* são responsáveis pela busca local e impedem que o algoritmo se prenda a mínimos locais (ASKARZADEH; ZEBARJADI, 2014).

A seleção correta do *HMCR*, *PAR* e *bw* é de suma importância para o bom desempenho do HS. Askarzadeh e Zebarjadi (2014), com o intuito de garantir bom equilíbrio entre busca global e local e simplificar a determinação destes parâmetros, propuseram as Equações 2.8, 2.9 e 2.10, nas quais  $n_{[0,1]}$  é um número aleatório entre 0 e 1, sendo estas as equações utilizadas neste trabalho para implementar o método HS.

$$HMCR = 0,9 + (0,1n_{[0,1]}) \quad (2.8)$$

$$PAR = \frac{1 - n_{[0,1]}}{2} \quad (2.9)$$

$$bw = n_{[0,1]} \quad (2.10)$$

### 2.3.4 Busca do Cuco

Proposto por Yang e Deb (2009), o método Busca do Cuco (CS - *Cuckoo Search*) é baseado no comportamento parasita de algumas espécies de pássaro Cuco em combinação com o voo de Lévy.

Cucos são aves parasitas que geralmente depositam os seus ovos no ninho de outros pássaros (JIANG et al., 2017). Após a deposição dos ovos, é possível que o hospedeiro identifique o ovo parasita, e, caso isto ocorra, o ovo do cuco é removido ou o hospedeiro abandona o ninho e constrói um novo em outro local. Algumas espécies de cuco são capazes até mesmo de mimetizar o padrão e a cor dos ovos do ninho hospedeiro, reduzindo as chances do ovo parasita ser detectado (YANG; DEB, 2009).

#### 2.3.4.1 Algoritmo Busca do Cuco

No algoritmo CS, cada ovo em um ninho representa uma possível solução para o problema, e cada ovo de cuco depositado é uma nova solução. Ademais, o algoritmo obedece três regras idealizadas (YANG, 2010a):

- Cada cuco deposita somente um ovo por vez. O ninho no qual este ovo será depositado é escolhido aleatoriamente.
- Os ninhos com melhor qualidade de ovos (melhores soluções) são transferidos para a próxima geração.
- O número de ninhos hospedeiros é fixo, e a probabilidade do ovo de cuco ser detectado é de  $p_d$ . Caso o ovo parasita seja detectado, ele é removido ou o hospedeiro abandona o ninho para construir um novo em outro local.

As novas soluções são geradas a partir do voo de Lévy, sendo este definido como uma caminhada aleatória cujo comprimento dos passos aleatórios segue a distribuição de Lévy. A direção desta caminhada segue a distribuição uniforme. Neste trabalho, o comprimento dos passos foi modelado de acordo com o algoritmo de Mantegna, que segue a distribuição de Lévy e é dado por (YANG, 2010c):

$$\text{Comprimento dos passos aleatórios} = \frac{\mu}{|\nu|^{\frac{1}{\beta_{CS}}}} \quad (2.11)$$

na qual  $\beta_{CS} = 1.5$  (EMARY; ZAWBAA; SHARAWI, 2019; YANG, 2010b),  $\mu$  e  $\nu$  são números aleatórios de uma distribuição normal:

$$\mu \sim N(0, \sigma_\mu^2), \quad \nu \sim N(0, \sigma_\nu^2) \quad (2.12)$$

onde  $\sigma_\nu = 1$  e  $\sigma_\mu$  é dado pela Equação 2.13.

$$\sigma_\mu = \left[ \frac{\Gamma(1 + \beta_{CS}) \sin\left(\frac{\pi\beta_{CS}}{2}\right)}{\Gamma\left(\frac{1+\beta_{CS}}{2}\right) \cdot \beta_{CS} \cdot 2^{\frac{\beta_{CS}-1}{2}}}\right]^{\frac{1}{\beta_{CS}}} \quad (2.13)$$

As novas soluções são então geradas conforme a Equação 2.14, na qual  $y_i^{t+1}$  é a nova solução do  $i$ th cuco,  $y_i^t$  é a solução atual,  $y_{melhor}^t$  é a melhor solução atual e  $\alpha$  é uma constante, normalmente igual a 1 (YANG, 2010c; YANG; DEB, 2009).

$$y_i^{t+1} = y_i^t + \alpha \frac{\mu}{|\nu|^{\frac{1}{\beta_{CS}}}} (y_i^t - y_{best}^t) \quad (2.14)$$

Através do voo de Lévy, as novas soluções podem apresentar valores bem distintos da melhor solução atual, o que intensifica o processo de busca global e evita que o algoritmo fique estagnado em mínimos locais (EMARY; ZAWBAA; SHARAWI, 2019).

Dito isto, o CS é conduzido da seguinte maneira (JIANG et al., 2017):

1. Geração aleatória da população inicial composta por  $N_n$  ninhos. Neste trabalho, o número de dimensões de cada ninho é equivalente à quantidade de parâmetros dos modelos de distribuição.
2. Cálculo da função objetivo para cada ninho, sendo a melhor solução mantida para próxima geração.
3. Geração de novas soluções a partir do voo de Lévy, conforme a Equação 2.14. Caso a nova solução seja melhor do que a anterior, a anterior é substituída pela nova, caso não, a solução anterior é mantida.
4. Geração de um número aleatório  $n_{[0,1]}$  entre 0 e 1 para cada ninho. Se  $n_{[0,1]} > p_d$ , significa que o ovo de cuco foi descoberto, e então o hospedeiro abandona o ninho para construir um novo em outro local. Os novos ninhos são comparados com os ninhos obtidos no passo 3, de modo que somente os melhores são mantidos.
5. Repetição dos passos 2 a 4 até que o critério de parada seja atingido. A melhor solução obtida é retornada.

## 2.4 Método da Máxima Verossimilhança

Neste trabalho, o Método da Máxima Verossimilhança (MLE - *Maximum Likelihood Estimation*) foi utilizado como referencial para avaliar o desempenho dos métodos metaheurísticos MBO, ICA, HS e CS na determinação dos parâmetros das onze distribuições

de velocidade do vento apresentadas na Seção 2.1.1. De acordo com Alavi, Mohammadi e Mostafaeipour (2016), este método é considerado como uma técnica bastante robusta na determinação dos parâmetros desconhecidos das distribuições de probabilidade e, quando há uma grande quantidade de amostras, este método é mais eficiente do que outros métodos numéricos tradicionais. No estudo conduzido por Kantar et al. (2018), o MLE foi selecionado dentre os diversos métodos numéricos tradicionais devido às suas propriedades estatísticas favoráveis. Nos estudos conduzidos por Mohammadi, Alavi e McGowan (2017) e Morgan et al. (2011), por exemplo, o MLE foi selecionado para estimar os parâmetros de diferentes distribuições de velocidade do vento. Tais fatos justificam a escolha deste método como referencial para análise do desempenho dos MOA.

No MLE, o ajuste do modelo é realizado através da identificação dos parâmetros que maximizam a probabilidade de serem obtidos os dados observados (ARIES; BOUDIA; OUNIS, 2018). Tal processo se dá através da maximização da função semelhança  $L$ , calculada pela seguinte Equação:

$$L = \prod_{i=1}^n f(\theta|x_i) \quad (2.15)$$

na qual  $\theta$  é o conjunto de parâmetros que regem as distribuições de probabilidade. A formulação  $f(\theta|x_i)$  é utilizada para explicitar o fato de que, no MLE, as distribuições de probabilidade são vistas não em função da variável aleatória, como normalmente elas são utilizadas, mas sim em função dos parâmetros que a regem. Sendo assim, para aplicação do MLE, é necessário antes a obtenção do conjunto de amostras que será ajustado.

### 3 METODOLOGIA

#### 3.1 Dados de velocidade do vento

Neste trabalho, os dados anemométricos utilizados foram obtidos nas estações PTR-11, localizada na cidade de Petrolina, e SCR-25, localizada na cidade de São João do Cariri, ambas na região Nordeste do Brasil, e foram disponibilizados pelo Sistema de Organização Nacional de Dados Ambientais (SONDA), projeto do Governo Federal Brasileiro. Informações geográficas das duas regiões são apresentadas na Tabela 2.

O Nordeste é a região do Brasil com melhores condições para geração de energia eólica, motivo pelo qual aproximadamente 86% da capacidade instalada de energia eólica do País encontra-se nesta região. No dia 13 de Novembro de 2018, 104% da demanda de energia do Nordeste veio da fonte eólica, sendo o excedente exportado para as demais regiões do País, e, no mesmo ano, durante a temporada da "safra dos ventos", que corresponde aos meses de Junho a Novembro, o fator de capacidade chegou a 86% nesta região, valor significativamente maior do que a média mundial de 25% (ABEEÓLICA, 2019b). Tais fatos, atrelados ao potencial ainda inexplorado desta região, justificam a escolha dos dados anemométricos das cidades de São João do Cariri e Petrolina para este estudo.

Tabela 2 – Informações geográficas das estações SCR-25 e PTR-11

Local	Latitude (°)	Longitude (°)	Altitude (m)
São João do Cariri (Estação SCR-25)	-7,3817	-36,5272	718
Petrolina (Estação PTR-11)	-9,0689	-40,3197	387

Antes de serem disponibilizados, os dados são submetidos a um processo de validação baseado no controle de qualidade de dados da *Baseline Surface Radiation Network* (BSRN). Tal processo não corrige os dados, mas gera um código de validação que sinaliza os dados caracterizados como suspeitos, ficando a critério do usuário utilizá-los ou não. Dentre os dados selecionados, somente um pequeno percentual foi caracterizado como suspeito. Diante desta situação, os dados foram submetidos a um processo de tratamento para remoção dos valores que apresentaram erros de medição. Em ambas as regiões analisadas, o percentual de aproveitamento foi superior a 99%, demonstrando a alta confiabilidade dos dados. A Tabela 3 apresenta algumas estatísticas básicas sobre os dados coletados pelas estações SCR-25 e PTR-11. Ambos foram medidos a 50m de altura e em intervalos de 10 minutos. O ano de medição foi escolhido com base na disponibilidade de dados, de maneira tal que, em ambas as regiões, foi selecionado o ano mais recente com dados medidos durante todos os meses.

Os dados utilizados das estações SCR-25 e PTR-11 são referentes aos anos de 2008 e 2010, respectivamente.

Tabela 3 – Estatísticas básicas dos dados anemométricos

Station	SCR-25	PTR-11
Ano de medição	2008	2010
Mínima (m/s)	0.002	0.008
Máxima (m/s)	12.98	14.30
Média (m/s)	5.238	4.886
Desvio padrão (m/s)	2.332	1.745
Assimetria	0.178	0.073
Curtose	2.449	3.105
Quantidade de dados	52662	52514

### 3.2 Linguagem R

Para implementação dos algoritmos de otimização metaheurísticos apresentados neste trabalho e do método determinístico máxima verossimilhança, foi utilizada a linguagem de programação R a partir do ambiente de desenvolvimento integrado RStudio. O RStudio é, atualmente, utilizado para a implementação de rotinas computacionais em diferentes linguagens, tais como C, C++, HTML, entre outras. A linguagem R foi escolhida devido à sua ampla utilização entre estatísticos e pela sua versatilidade no processamento e apresentação de dados (R-PROJECT, 2019).

### 3.3 Análise estatística de múltiplos critérios

Para a seleção do modelo de distribuição que melhor se ajusta aos dados de velocidade do vento de uma região, diversos testes estatísticos podem ser utilizados, tais como: Erro quadrático médio (*MSE - Mean Squared Error*) Raíz quadrada do erro quadrático médio (*RMSE - Root Mean Squared Error*), *RMSE* relativo (*RRMSE*), Erro médio absoluto (*MAE - Mean Absolute Error*), Critério de informação de Akaike (*AIC - Akaike Information Criterion*), Critério de informação Bayesiano (*BIC - Bayesian Information Criterion*), Teste de Kolmogorov-Smirnov (*KS - Kolmogorov-Smirnov Test*), Coeficiente de determinação ( $R^2$ ), entre outros.

Contudo, ao conduzir uma análise com diferentes critérios, é possível que os resultados variem significativamente entre os modelos e a análise seja pouco conclusiva (MASSERAN, 2018). Ademais, de acordo com Kantar et al. (2018), não há, atualmente, um consenso de qual critério é o melhor para determinar o modelo de distribuição mais preciso. Em uma análise

conduzida com dois critérios, por exemplo, é possível que um certo modelo seja o melhor em um critério, e outro modelo seja o melhor no outro critério. Esta situação causaria confusão e, conseqüentemente, uma má seleção do modelo mais adequado. Em meio a este cenário, Masseran (2018) propôs uma abordagem de múltiplos critérios baseada em uma agregação multiplicativa com scores padronizados para cada critério estatístico. Neste processo, todos os critérios são simultaneamente considerados e um único resultado é gerado, sendo este o parâmetro adotado na seleção do modelo mais adequado.

Na primeira etapa desta abordagem integrada, são calculados os testes estatísticos  $KS$ ,  $AIC$ , Desvio de Assimetria e Curtose ( $DSK$  - *Deviation of skewness and kurtosis*) e  $1 - R^2$ :

- $KS$ : Este teste calcula a maior distância entre as probabilidades cumulativas observadas ( $F_i$ ) e as probabilidades cumulativas previstas ( $\hat{F}_i$ ) a fim de determinar se este modelo é adequado para descrever os dados de velocidade do vento e é dado pela Equação 3.1 (OUARDA; CHARRON; CHEBANA, 2016). Valores baixos de  $KS$  indicam que um bom ajuste foi obtido.

$$KS = \max_{1 \leq i \leq n} |F_i - \hat{F}_i| \quad (3.1)$$

- $AIC$ : Este teste estatístico é uma medida da informação perdida quando um modelo é ajustado a um conjunto de dados e é dado pela Equação 3.2, na qual  $L$  é a função semelhança, apresentada na Seção 2.4 e calculada pela Equação 2.15, e  $n_p$  é o número de parâmetros do modelo. Quanto menor o  $AIC$ , melhor o ajuste.

$$AIC = -2\log(L) + 2n_p \quad (3.2)$$

- $DSK$ : É utilizado para estimar a similaridade entre as características de forma do modelo ajustado e da distribuição empírica dos dados (MASSERAN, 2018). A assimetria  $\gamma_1$  e a curtose  $\gamma_2$  de uma distribuição podem ser calculadas, respectivamente, através das Equações 3.3 e 3.4:

$$\gamma_1 = \frac{E(X^3) - 3E(X^2)E(X) + 2E^3(X)}{\left[\sqrt{E(X^2) - E^2(X)}\right]^3} \quad (3.3)$$

$$\gamma_2 = \frac{E(X^4) - 4E(X^3)E(X) + 6E(X^2)E^2(X) - 3E^4(X)}{\left[\sqrt{E(X^2) - E^2(X)}\right]^4} \quad (3.4)$$

nas quais:

$$E(X^n) = \int_{-\infty}^{\infty} x^n f(x) dx \quad (3.5)$$

A assimetria empírica  $g_1$  e a curtose empírica  $g_2$  dos dados são calculadas pelas Equações 3.6 e 3.7, nas quais  $\bar{x}$  e  $s$  são, respectivamente, a média e o desvio padrão dos dados de

velocidade do vento (JUNG; SCHINDLER, 2017).

$$g_1 = \frac{\sum_{i=1}^n (x_i - \bar{x})^3}{(n-1)s^3} \quad (3.6)$$

$$g_2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^4}{(n-1)s^4} \quad (3.7)$$

O desvio entre a assimetria teórica e empírica ( $DS$ ) e o desvio da curtose ( $DK$ ) são calculados pelas Equações 3.8 e 3.9, respectivamente.

$$DS = \gamma_1 - g_1 \quad (3.8)$$

$$DK = \gamma_2 - g_2 \quad (3.9)$$

Por fim, o critério  $DSK$  é calculado pela Equação 3.10. Valores baixos de  $DSK$  indicam boa qualidade de ajuste (TOFALLIS, 2008).

$$DSK = \sqrt{(DS)^2 \cdot (DK)^2} \quad (3.10)$$

- $1 - R^2$ : Este critério mede a correlação entre as probabilidades observadas e as probabilidades previstas pelo modelo de distribuição e é dado por (ALAVI; MOHAMMADI; MOSTAFAEIPOUR, 2016):

$$R^2 = \frac{\sum_{i=1}^n (\hat{F}_i - \bar{F})^2}{\sum_{i=1}^n (\hat{F}_i - \bar{F})^2 + \sum_{i=1}^n (F_i - \hat{F}_i)^2} \quad (3.11)$$

onde

$$\bar{F} = \frac{\sum_{i=1}^n \hat{F}_i}{n} \quad (3.12)$$

O valor de  $R^2$  varia entre 0 e 1, sendo os valores próximos a 1 indicativo de bom ajuste. Na abordagem integrada é utilizado o valor  $1 - R^2$ , dessa maneira, a minimização de todos os critérios torna-se indicativo de boa qualidade de ajuste.

Posteriormente, cada critério é submetido a um processo de padronização. Neste processo, um conjunto de dados com uma determinada média e desvio padrão é convertido em um conjunto com média 0 e desvio padrão 1, conforme a Equação 3.13:

$$z_{ij} = \frac{w_{ij} - \bar{w}_j}{s_j} \quad (3.13)$$

onde  $w_{ij}$  é o  $i$ -ésimo valor ( $i$ -score) do critério  $j$ ,  $\bar{w}_j$  e  $s_j$  são, respectivamente, a média e desvio padrão do critério  $j$  e  $z_{ij}$  é o score padronizado.

Durante a padronização, é possível que os  $z$ -scores apresentem sinais positivos e negativos, o que afetaria a agregação multiplicativa. Portanto, para garantir que os scores padronizados sejam passíveis de comparação, é utilizada a Transformação da Função de Distribuição Cumulativa Normal Padrão (SNCDF - *Standard Normal Cumulative Distribution Function Transformation*), dada pela Equação 3.14, para que os  $z$ -scores fiquem entre 0 e 1.

$$\Phi(z_{ij}) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z_{ij}} e^{-\frac{t^2}{2}} dt \quad (3.14)$$

Os  $z$ -scores de cada critério, após serem submetidos à SNCDF são multiplicados e um único valor é gerado, sendo este denominado, neste trabalho, de Score Global ( $GS$  - *Global Score*). A distribuição com menor  $GS$  é a mais adequada para a região e, como demonstrado por Masseran (2018), o modelo com menor  $GS$  apresenta não somente o melhor ajuste em termos numéricos como também o melhor ajuste graficamente. Portanto, a utilização desta abordagem simplifica a análise comparativa entre diferentes modelos e garante a seleção da distribuição mais adequada para uma determinada região de forma objetiva.

### 3.4 Funções objetivo

Como descrito na Seção 2.3, o processo de busca pelos parâmetros ótimos das distribuições nos MOA se dá através da minimização (ou maximização) de uma função objetivo. Neste estudo, foram utilizadas três funções objetivo, sendo elas:

1. Maximização do coeficiente de determinação ( $R^2$ ): O teste  $R^2$  foi apresentado anteriormente na seção 3.3 e é calculado pela Equação 3.11. Esta função objetivo pode também ser interpretada como a minimização do valor  $1 - R^2$  e foi utilizada por Jiang et al. (2017) e Wang et al. (2015).
2. Minimização da Raíz do Erro Quadrático Médio ( $RMSE$ ): O  $RMSE$  é uma medida da precisão do modelo de distribuição ajustado e é dado pela Equação 3.15 (MOHAMMADI; ALAVI; MCGOWAN, 2017).

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (F_i - \hat{F}_i)^2} \quad (3.15)$$

Quanto menor o valor do  $RMSE$ , mais preciso é o ajuste obtido. A minimização do erro quadrático como função objetivo, que é muito similar à minimização do  $RMSE$ , foi utilizada por Wang et al. (2018) e Andrade et al. (2019).

3. Minimização do  $(1 - R^2) + NRMSE$ : Função objetivo híbrida formada pela soma das duas funções anteriores. O termo  $NRMSE$  indica que o  $RMSE$  foi normalizado, de

acordo com a Equação 3.16, para que o  $R^2$  e  $RMSE$  fiquem ambos na mesma escala. Não foram encontrados estudos que utilizassem esta função objetivo para determinar os parâmetros desconhecidos das distribuições de velocidade do vento.

$$NRMSE = \frac{RMSE}{\max(F) - \min(F)} \quad (3.16)$$

É importante ressaltar que o resultado gerado a partir das três funções objetivos não é necessariamente o mesmo. Ademais, não foram encontrados trabalhos que comparassem o desempenho destas funções objetivos na determinação dos parâmetros ótimos das distribuições de velocidade do vento. Portanto, os resultados gerados a partir das três funções objetivo foram comparados através da abordagem integrada descrita na Seção 3.3 a fim de determinar a que proporciona melhores resultados (menor  $GS$ ).

### 3.5 Critério de parada

De acordo com Yang (2010c), quando não há limite de tempo para o processo de otimização, é teoricamente possível que um determinado método encontre o ótimo global. Neste contexto, se caso fosse adotado como critério de parada um valor máximo de iterações, o tempo de processamento seria restringido e haveria possibilidade da melhor solução encontrada pelos métodos ser uma solução prematura e pouco otimizada. Portanto, para evitar esta possibilidade, o critério de parada adotado neste trabalho foi a convergência do método, ou seja, quando a função objetivo sofrer uma variação menor ou igual a  $10^{-6}$ , o algoritmo cessa e retorna a solução obtida.

### 3.6 Configuração dos métodos de otimização metaheurísticos

Como apresentado na Seção 2.3, cada MOA possui um determinado conjunto de parâmetros. A influência destes parâmetros na busca pela solução ótima, como demonstrado por Santos (2018), é praticamente nula, isto é, independente da configuração adotada, é sempre possível que a solução ótima seja encontrada, e a única variável afetada por estes parâmetros é o tempo necessário para a convergência do método. Além disso, de acordo com Macêdo (2018), o ajuste destes parâmetros não influencia o resultado final obtido. Tais afirmações são explicadas pelo fato de que, como os MOA apresentam em sua estrutura uma geração aleatória de soluções, há sempre a possibilidade de que uma solução gerada aleatoriamente seja a solução ótima do problema. Dito isto, neste trabalho, a configuração adotada foi a mesma utilizada em referências que obtiveram bons resultados com os MOA em diferentes aplicações. Os parâmetros adotados e as referências de onde foram extraídos são apresentados na Tabela 4.

Tabela 4 – Configuração dos algoritmos metaheurísticos

Método	Referências	Parâmetros
MBO	Makas e Yumusak (2016)	$N_p = 51$ $m = 10$ $n_v = 3$ $n_{nu} = 1$
ICA	Atashpaz-Gargari e Lucas (2007)	$N_{pop} = 50$ $N_{imp} = 5$
HS	Askarzadeh e Zebarjadi (2014)	$N_h = 50$
CS	Santos (2018), Macêdo (2018)	$N_n = 50$ $P_d = 0,25$

Fonte: Elaborado pelo Autor.

## 4 RESULTADOS E DISCUSSÕES

### 4.1 Seleção da Função Objetivo

Para cada distribuição de probabilidade, foram comparados os resultados obtidos através das três funções objetivo. Em ambas as regiões analisadas, a maximização do  $R^2$  foi a função objetivo que apresentou, na maioria dos casos, menor  $GS$ , ou seja, melhor ajuste.

Em São João do Cariri (SCR-25), 68,18% dos melhores resultados foram obtidos através da maximização do coeficiente de determinação  $R^2$ . As funções objetivo maximização do  $RMSE$  e híbrida proporcionaram o melhor resultado em somente 20,46% e 11,36% dos casos, respectivamente.

Em Petrolina (PTR-11), 43,18% dos melhores resultados foram obtidos através da maximização do coeficiente de determinação  $R^2$ . As funções objetivo maximização do  $RMSE$  e híbrida proporcionaram o melhor resultado em somente 29,55% e 27,27% dos casos, respectivamente. Na Tabela 5 foram apresentadas as funções objetivo que proporcionaram o melhor resultado em cada uma das distribuições e métodos utilizados.

Por tais motivos e por padronização, a discussão conduzida a seguir foi baseada somente nos resultados gerados a partir da função objetivo maximização do  $R^2$ .

Tabela 5 – Função objetivo geradora do melhor resultado

Método	Modelo	SCR-25	PTR-11
MBO	W	$H$	$RMSE$
	G	$R^2$	$R^2$
	BS	$RMSE$	$RMSE$
	N	$R^2$	$H$
	LN	$RMSE$	$RMSE$
	GL	$R^2$	$RMSE$
	GEV	$RMSE$	$R^2$
	B	$R^2$	$H$
	D	$R^2$	$R^2$
	EGL	$R^2$	$R^2$
GG	$R^2$	$RMSE$	
ICA	W	$RMSE$	$RMSE$
	G	$RMSE$	$H$

*Continua na próxima página*

Tabela 5 – *Continuação da página anterior*

Método	Modelo	SCR-25	PTR-11
	BS	$R^2$	$H$
	N	$R^2$	$RMSE$
	LN	$R^2$	$RMSE$
	GL	$H$	$R^2$
	GEV	$R^2$	$R^2$
	B	$R^2$	$RMSE$
	D	$R^2$	$R^2$
	EGL	$R^2$	$R^2$
	GG	$R^2$	$R^2$
HS	W	$H$	$H$
	G	$R^2$	$H$
	BS	$RMSE$	$H$
	N	$RMSE$	$H$
	LN	$R^2$	$RMSE$
	GL	$H$	$H$
	GEV	$R^2$	$R^2$
	B	$RMSE$	$RMSE$
	D	$RMSE$	$RMSE$
	EGL	$R^2$	$R^2$
	GG	$R^2$	$H$
CS	W	$R^2$	$R^2$
	G	$R^2$	$R^2$
	BS	$R^2$	$R^2$
	N	$R^2$	$RMSE$
	LN	$R^2$	$R^2$
	GL	$R^2$	$R^2$
	GEV	$R^2$	$R^2$
	B	$H$	$H$
	D	$R^2$	$R^2$
	EGL	$R^2$	$R^2$
	GG	$R^2$	$H$

É importante ressaltar que o fato da função objetivo  $R^2$ , nesta pesquisa, apresentar

melhores resultados não significa que as outras duas funções não sejam viáveis. Nesta análise em específico, baseada nos critérios  $AIC$ ,  $KS$ ,  $DSK$  e  $1 - R^2$ , a função objetivo  $R^2$  garantiu bons resultados. Contudo, se uma análise for conduzida com outros critérios estatísticos, as outras duas funções apresentadas podem proporcionar melhor desempenho.

## 4.2 Estação SCR-25

Uma vez estimados os parâmetros das distribuições, os testes estatísticos  $1 - R^2$ ,  $KS$ ,  $AIC$  e  $DSK$  referentes à estação SCR-25 foram calculados e apresentados na Tabela 6, assim como o Score Global  $GS$  gerado a partir destes quatro testes. Nesta seção, foi utilizada a nomenclatura  $Distribuição_{Método}$  para fazer referência a um determinado modelo de distribuição e o método utilizado para estimar os parâmetros; por exemplo,  $GG_{ICA}$  é referente à distribuição GG com parâmetros estimados pelo método ICA. Os parâmetros estimados nesta região foram apresentados na Tabela 10, localizada no Apêndice A deste trabalho.

Tabela 6 – Análise estatística SCR-25

Método	Modelo	$1 - R^2$	$KS$	$AIC$	$DSK$	Score Global
MBO	W	0,000574	0,014455	238132,745	0,114126	0,009369
	G	0,003784	0,036525	244618,814	1,474199	0,051648
	BS	0,011968	0,071369	307993,008	3,753875	0,382511
	N	0,000786	0,020145	238968,175	0,251003	0,013540
	LN	0,009629	0,053931	266166,376	9,133399	0,390851
	GL	0,005056	0,043610	245272,189	2,643198	0,083748
	GEV	0,000253	0,011110	237936,878	0,022947	0,007442
	B	0,000534	0,017753	238491,024	0,217354	0,011528
	D	0,001943	0,022985	238187,991	0,096140	0,014818
	<b>EGL</b>	<b>0,000397</b>	<b>0,009580</b>	<b>237020,772</b>	<b>0,006519</b>	<b>0,006453</b>
GG	0,000102	0,007636	237246,008	0,025975	0,005795	
ICA	W	0,000357	0,015399	238130,005	0,128029	0,009732
	G	0,003900	0,035331	244982,699	1,401672	0,050449
	BS	0,009624	0,057409	295427,995	4,705488	0,356582
	N	0,000790	0,020240	238976,293	0,250689	0,013605
	LN	0,009568	0,053470	266863,515	8,798650	0,387863
	GL	0,005062	0,043675	245335,853	2,640323	0,084100
	GEV	0,000251	0,011352	237948,057	0,025259	0,007549

*Continua na próxima página*

Tabela 6 – *Continuação da página anterior*

Método	Modelo	$1 - R^2$	$KS$	$AIC$	$DSK$	$GS$
	B	0,000596	0,018732	238528,885	0,217207	0,012126
	<b>D</b>	<b>0,000796</b>	<b>0,018286</b>	<b>238520,273</b>	<b>1,607747</b>	<b>0,016699</b>
	EGL	0,000385	0,012710	237163,310	0,020883	0,007745
	GG	0,000190	0,007468	237251,611	0,023672	0,005766
HS	W	0,000353	0,014897	238114,779	0,126053	0,009477
	G	0,003787	0,035550	244719,734	1,452952	0,050359
	BS	0,009632	0,057295	295447,595	4,686767	0,355545
	N	0,000786	0,020224	238962,676	0,251681	0,013585
	LN	0,009594	0,054127	266735,745	8,921705	0,392168
	GL	0,005060	0,043604	245332,193	2,640350	0,083958
	GEV	0,000260	0,011895	237880,164	0,009455	0,007708
	B	0,000534	0,017413	238480,578	0,214431	0,011331
	D	0,002234	0,022562	238126,034	0,064479	0,014609
	<b>EGL</b>	<b>0,000433</b>	<b>0,010752</b>	<b>237003,080</b>	<b>0,004846</b>	<b>0,006889</b>
	GG	0,000102	0,007646	237249,452	0,026446	0,005800
CS	W	0,000353	0,014789	238112,183	0,125258	0,009423
	G	0,003783	0,036121	244660,830	1,465260	0,051108
	BS	0,009620	0,056998	296556,512	4,730798	0,356122
	<b>N</b>	<b>0,000786</b>	<b>0,020075</b>	<b>238964,664</b>	<b>0,251052</b>	<b>0,013496</b>
	LN	0,009563	0,053921	267044,953	8,726596	0,390115
	GL	0,005056	0,043399	245303,450	2,641275	0,083494
	GEV	0,000245	0,011140	237918,596	0,016894	0,007429
	B	0,000536	0,018090	238501,912	0,218988	0,011722
	<b>D</b>	<b>0,000658</b>	<b>0,021876</b>	<b>238830,136</b>	<b>3,266149</b>	<b>0,026639</b>
	EGL	0,000209	0,010214	237121,839	0,042366	0,006725
	GG	0,000103	0,007831	237252,272	0,026292	0,005865
MLE	W	0,001317	0,018823	238020,523	0,113605	0,011925
	G	0,015054	0,053450	243627,468	1,801609	0,112394
	BS	0,124033	0,146974	272893,147	10,994830	0,918208
	N	0,004060	0,028407	238798,948	0,263907	0,021988
	LN	0,051467	0,090703	258070,094	25,183177	0,733778
	GL	0,016834	0,054603	243740,888	2,829225	0,145563
	GEV	0,001244	0,017380	237592,360	0,000116	0,010478

*Continua na próxima página*

Tabela 6 – *Continuação da página anterior*

Método	Modelo	$1 - R^2$	$KS$	$AIC$	$DSK$	$GS$
	B	0,001796	0,021748	238526,282	0,232945	0,014821
	<b>D</b>	<b>0,002926</b>	<b>0,028497</b>	<b>238099,968</b>	<b>0,101334</b>	<b>0,019059</b>
	EGL	0,000821	0,014641	236949,896	0,005532	0,008597
	GG	0,000517	0,011197	236940,846	0,001241	0,007054

Dos resultados da Tabela 6, é possível perceber as vantagens da utilização da abordagem integrada de Masseran (2018), que considera todos os critérios mutuamente em um único valor, sendo este o Global Score ( $GS$ ). Ao comparar, por exemplo, os resultados da  $EGL_{MBO}$  e  $EGL_{HS}$ , a variação observada nos testes estatísticos impossibilitaria a decisão objetiva do método que garantiu melhor estimativa dos parâmetros. O resultado obtido a partir do MBO é melhor nos testes  $1 - R^2$  e  $KS$ , enquanto que o resultado obtido pelo HS é melhor nos testes  $AIC$  e  $DSK$ . Nesta situação, seria necessário mensurar a importância de cada um dos critérios e a decisão do melhor modelo seria subjetiva. Contudo, a partir da abordagem integrada, a análise se torna mais objetiva e mais precisa. Neste caso, o  $GS$  da  $EGL_{HS}$  é de 0,006889 e o da  $EGL_{MBO}$  é 0,006453, sendo este o melhor resultado dentre os dois comparados.

Esta abordagem é útil não somente para comparar os resultados obtidos a partir de diferentes métodos como também para comparar os diferentes modelos de distribuição. Por exemplo, quando comparados os resultados das distribuições  $D_{ICA}$  e  $N_{CS}$ , a distribuição  $D_{ICA}$  apresenta valores inferiores de  $KS$  e  $AIC$  e a  $N_{CS}$  apresenta valores inferiores de  $1 - R^2$  e  $DSK$ . Quando comparadas em relação ao  $GS$ , fica evidente o melhor desempenho da distribuição  $N_{CS}$ , com  $GS = 0,013496$ , em relação à distribuição  $D_{ICA}$ , que apresentou  $GS = 0,016699$ . Estas duas situações reafirmam a vantagem e justificam a utilização de uma abordagem integrada neste tipo de estudo.

Comparando o mesmo modelo de distribuição, pode ser visto que os valores de  $GS$  obtidos pelos MOA são significativamente inferiores aos obtidos pelo método determinístico MLE. A distribuição  $D_{CS}$  foi a única que não apresentou valor de  $GS$  inferior quando comparada com a  $D_{MLE}$ . Contudo, ainda que o CS não tenha obtido um resultado superior ao MLE na distribuição D, os outros três métodos utilizados atingiram tal feito, demonstrando a aplicabilidade e elevado desempenho dos MOA na determinação dos parâmetros das distribuições de dois e três parâmetros utilizadas. Ademais, é possível ver que, na maioria das distribuições, os valores de  $GS$  obtidos pelos MOA apresentaram pequenas variações, demonstrando que tais métodos obtiveram soluções semelhantes. Por exemplo, os  $GS$  da  $GL_{MBO}$ ,  $GL_{ICA}$ ,  $GL_{HS}$  e  $GL_{CS}$  são, respectivamente, 0,083748, 0,084100, 0,083958 e

0,083494.

Para o mesmo método de estimativa dos parâmetros, o  $GS$  da distribuição GG foi o menor e o da LN o maior. No caso do MLE, o  $GS$  da distribuição BS foi o maior. Desse modo, dentre as distribuições analisadas, a GG foi a que proporcionou melhor ajuste nesta região, enquanto que a BS e LN apresentaram os piores ajustes.

#### 4.2.1 Seleção do modelo de distribuição ótimo para SCR-25

Para determinar a distribuição de probabilidade e o método de estimativa que garantiram melhor ajuste ao conjunto de dados anemométricos da cidade de São João do Cariri, foi extraído, da Tabela 6, o resultado com menor  $GS$  de cada distribuição, sendo estes apresentados na Tabela 7. Isto é, foram comparados, por exemplo, os resultados das distribuições  $BS_{MBO}$ ,  $BS_{ICA}$ ,  $BS_{HS}$ ,  $BS_{CS}$  e  $BS_{MLE}$ , e a que apresentou menor  $GS$  (melhor ajuste) foi explicitada na Tabela 7. Foi também apresentada a posição de cada distribuição no ranking.

Tabela 7 – Melhores modelos de distribuição em SCR-25

Modelo	Método	$1 - R^2$	$KS$	$AIC$	$DSK$	$GS$	Rank
W	MBO	0,000574	0,014455	238132,745	0,114126	0,009369	4
G	HS	0,003787	0,035550	244719,734	1,452952	0,050359	8
BS	HS	0,009632	0,057295	295447,595	4,686767	0,355545	10
N	CS	0,000786	0,020075	238964,664	0,251052	0,013496	6
LN	ICA	0,009568	0,053470	266863,515	8,798650	0,387863	11
GL	CS	0,005056	0,043399	245303,450	2,641275	0,083494	9
GEV	CS	0,000245	0,011140	237918,596	0,016894	0,007429	3
B	HS	0,000534	0,017413	238480,578	0,214431	0,011331	5
D	HS	0,002234	0,022562	238126,034	0,064479	0,014609	7
EGL	MBO	0,000397	0,009580	237020,772	0,006519	0,006453	2
GG	ICA	0,000190	0,007468	237251,611	0,023672	0,005766	1

A distribuição que proporcionou melhor ajuste foi a  $GG_{ICA}$ , seguida da  $EGL_{MBO}$  e  $GEV_{CS}$ , em segundo e terceiro lugar, respectivamente.

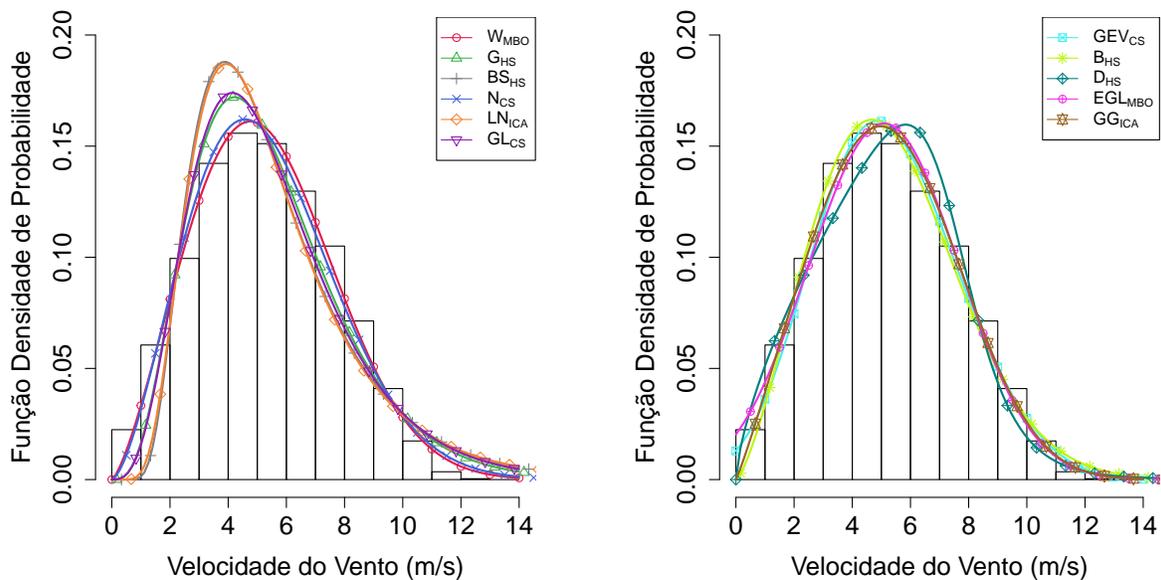
A distribuição  $W_{MBO}$ , que é uma das mais utilizadas devido à sua versatilidade e bom ajuste em diferentes regiões, ocupou a quarta posição, confirmando que esta distribuição não é sempre a mais apropriada. Contudo, dentre as distribuições de dois parâmetros, a  $W_{MBO}$  foi a que apresentou menor  $GS$ .

De todas as onze distribuições, a  $BS_{HS}$  e  $LN_{ICA}$  foram as que apresentaram os piores valores de  $GS$ . No geral, é possível inferir que, para a região de São João do Cariri, as distribuições de três parâmetros, cujo ajuste foi realizado através dos MOA com a função

objetivo  $R^2$ , proporcionaram melhor ajuste em relação às distribuições de dois parâmetros, que foram ajustadas da mesma maneira e ocuparam as posições mais elevadas no ranking. Ademais, é importante ressaltar, mais uma vez, que nenhum dos melhores resultados foi fornecido pelo MLE, comprovando o bom desempenho dos MOA.

Na Figura 3, foram apresentados graficamente os ajustes obtidos através dos modelos de distribuição da Tabela 7. Para melhor visualização, as distribuições de dois e três parâmetros foram apresentadas separadamente.

Figura 3 – Distribuições de dois (esquerda) e três (direita) parâmetros em SCR-25



Fonte: Elaborada pelo Autor.

A partir da Figura 3, é possível ver que, dentre os modelos de dois parâmetros, as distribuições  $W_{MBO}$  e  $N_{CS}$  foram as que apresentaram melhor ajuste. Já as distribuições  $G_{HS}$ ,  $GL_{CS}$ ,  $BS_{HS}$  e  $LN_{ICA}$  foram as que apresentaram piores ajustes. Estes resultados estão perfeitamente de acordo com o resultado numérico obtido a partir da abordagem integrada, visto que, dentre os modelos de dois parâmetros, as distribuições  $W_{MBO}$  e  $N_{CS}$  foram as que apresentaram menor  $GS$  e as distribuições  $G_{HS}$ ,  $GL_{CS}$ ,  $BS_{HS}$  e  $LN_{ICA}$  maior  $GS$ .

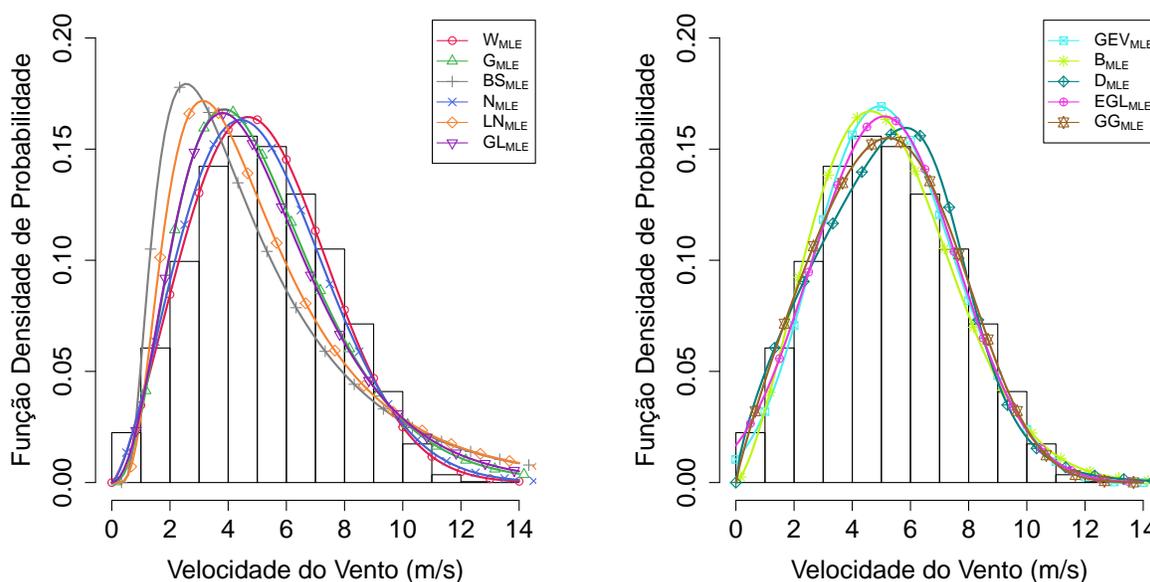
Quanto às distribuições de três parâmetros, todas, com exceção da  $D_{HS}$ , apresentaram desempenho gráfico semelhante. A distribuição  $D_{HS}$  exibiu valor máximo de densidade no intervalo de  $5 - 6\text{m/s}$ , enquanto que a densidade máxima observada nos dados é no intervalo de  $4 - 5\text{m/s}$ . Ademais, também pode ser visto que para as velocidades entre  $9 - 11\text{m/s}$ , esta distribuição apresentou densidade inferior à dos dados observados. Este fraco desempenho da distribuição  $D_{HS}$  concorda com o seu valor de  $GS$ , que foi o pior dentre as distribuições de três parâmetros. Quanto aos demais modelos de três parâmetros, devido à grande similaridade

entre eles, é necessário considerar os valores de  $GS$  para determinar qual foi o mais preciso, sendo, neste caso, a  $GG_{ICA}$  com  $GS = 0,005766$

Quanto ao desempenho dos MOA, os resultados apresentados na Tabela 7 não são conclusivos para afirmar se algum método é superior aos demais. Nenhum método apresentou predominância significativa em relação aos outros, isto é, diferentes métodos garantiram o melhor resultado em diferentes distribuições. No entanto, é válido ressaltar que os métodos MBO, ICA e HS, utilizados no ajuste de distribuições de três parâmetros e distribuições não convencionais pela primeira vez neste estudo, garantiram bons resultados para as distribuições W e EGL, LN e GG, G, BS, B e D, respectivamente. Já o CS garantiu o melhor resultado nas distribuições N, GL e GEV.

Para comparar o desempenho gráfico dos MOA, foi apresentado, na Figura 4, os resultados obtidos pelo método determinístico MLE.

Figura 4 – Distribuições de dois (esquerda) e três (direita) parâmetros ajustadas pelo MLE em SCR-25



Fonte: Elaborada pelo Autor.

Ao comparar os resultados das distribuições de dois parâmetros obtidos pelo método MLE (Figura 4) com os dos MOA (Figura 3), foi possível observar que os modelos W e N se mantiveram muito similares em ambos os casos, enquanto que os demais modelos divergiram substancialmente. Por exemplo, nos intervalos entre  $0 - 3\text{m/s}$  e  $5 - 14\text{m/s}$ , os modelos  $G_{MLE}$ ,  $BS_{MLE}$ ,  $LN_{MLE}$  e  $GL_{MLE}$  pouco se aproximaram dos valores de densidade observados no histograma, enquanto que, nos ajustes obtidos pelos MOA, os mesmos modelos de distribuição exibiram valores de densidade significativamente mais próximos do histograma.

Nas distribuições de três parâmetros, foi possível observar que os modelos D e GG se mantiveram muito similares em ambos os casos (Figura 3 e Figura 4). Já os modelos  $GEV_{MLE}$ ,  $B_{MLE}$  e  $EGL_{MLE}$  exibiram valor máximo de densidade um pouco maior do que valor máximo de densidade do histograma. Estes mesmos modelos de distribuição, quando ajustados pelos MOA (Figura 3), apresentaram valor máximo de densidade estreitamente próximos dos valores máximos do histograma.

Tais comparações reafirmam, em termos gráficos, a superioridade dos MOA em relação ao MLE. No caso das distribuições que se mantiveram similares nesta estação (W, N, D e GG), os valores de  $GS$  apresentados anteriormente (Tabela 7) demonstraram o ajuste superior dos MOA.

### 4.3 Estação PTR-11

Os resultados dos testes estatísticos referentes à estação PTR-11 e os parâmetros estimados foram apresentados, respectivamente, nas Tabelas 8 e 11, sendo esta última localizada no Apêndice A deste trabalho.

Tabela 8 – Análise estatística PTR-11

Método	Modelo	$1 - R^2$	$KS$	$AIC$	$DSK$	$GS$
MBO	W	0,000612	0,017412	209002,163	0,004601	0,008266
	G	0,004228	0,045361	218950,901	0,370237	0,051379
	<b>BS</b>	<b>0,008148</b>	<b>0,062829</b>	<b>270459,619</b>	<b>1,550372</b>	<b>0,279593</b>
	N	0,001783	0,031105	211471,356	0,018213	0,016877
	LN	0,008260	0,063264	243633,937	2,077473	0,300827
	GL	0,007615	0,055561	227299,561	1,966291	0,192398
	GEV	0,000576	0,014806	209148,212	0,055842	0,007755
	B	0,000624	0,018686	208870,487	0,019825	0,008748
	D	0,000310	0,012385	207765,650	0,177590	0,006860
	EGL	0,000426	0,009722	207496,629	0,011904	0,005246
GG	0,000582	0,015121	208995,341	0,015806	0,007544	
ICA	<b>W</b>	<b>0,000650</b>	<b>0,016483</b>	<b>208839,208</b>	<b>0,008137</b>	<b>0,007894</b>
	G	0,004259	0,047060	219403,763	0,357497	0,053866
	BS	0,008156	0,063038	270538,725	1,552027	0,280481
	N	0,001784	0,031222	211483,671	0,018225	0,016951
	LN	0,008031	0,061071	241838,032	2,233963	0,295729

*Continua na próxima página*

Tabela 8 – *Continuação da página anterior*

Método	Modelo	$1 - R^2$	$KS$	$AIC$	$DSK$	$GS$
	GL	0,007623	0,055233	227200,820	1,966589	0,190958
	GEV	0,000603	0,015276	209232,175	0,032188	0,007824
	B	0,000644	0,018596	208887,686	0,019816	0,008733
	D	0,000354	0,009717	207868,938	0,488103	0,007548
	EGL	0,000453	0,013724	207772,467	0,028113	0,006548
	GG	0,000624	0,013973	208932,538	0,023746	0,007194
HS	W	0,000613	0,017577	209029,892	0,004045	0,008336
	G	0,004453	0,050014	220256,394	0,337027	0,058689
	<b>BS</b>	<b>0,008259</b>	<b>0,059861</b>	<b>267403,783</b>	<b>1,643141</b>	<b>0,278916</b>
	N	0,001786	0,031230	211480,466	0,018219	0,016953
	LN	0,008323	0,056429	239133,621	2,484453	0,284167
	GL	0,007629	0,056887	227956,160	1,964457	0,199920
	GEV	0,000587	0,014901	209193,964	0,042432	0,007733
	B	0,000624	0,018845	208889,514	0,019486	0,008814
	D	0,000269	0,011278	207831,574	0,274764	0,007007
	EGL	0,000407	0,010054	207527,216	0,014464	0,005350
	GG	0,000591	0,015355	209062,035	0,026045	0,007721
CS	W	0,000612	0,017412	209002,166	0,004601	0,008266
	G	0,004228	0,045429	218966,785	0,369798	0,051473
	BS	0,008147	0,062742	270332,000	1,554897	0,279734
	N	0,001783	0,031104	211471,207	0,018213	0,016875
	LN	0,008027	0,061102	241806,339	2,239767	0,296050
	GL	0,007615	0,055575	227305,815	1,966273	0,192475
	<b>GEV</b>	<b>0,000573</b>	<b>0,015180</b>	<b>209203,073</b>	<b>0,055597</b>	<b>0,007910</b>
	B	0,000624	0,018747	208880,268	0,019537	0,008774
	D	0,000116	0,008508	207893,067	0,626664	0,007684
	EGL	0,000387	0,011380	207580,049	0,012057	0,005707
	GG	0,000582	0,015183	208992,895	0,015068	0,007559
MLE	W	0,003001	0,027339	208373,240	0,032883	0,012938
	G	0,025066	0,072438	215112,315	0,656764	0,125662
	BS	0,177518	0,175007	244775,917	6,572223	0,926383
	N	0,010131	0,048517	210121,376	0,012957	0,031527

*Continua na próxima página*

Tabela 8 – *Continuação da página anterior*

Método	Modelo	$1 - R^2$	$KS$	$AIC$	$DSK$	$GS$
	LN	0,063472	0,106895	227333,310	7,470845	0,663506
	GL	0,037652	0,083018	217336,642	2,112967	0,310026
	GEV	0,006277	0,036546	208778,585	0,023315	0,019597
	B	0,002752	0,026769	208423,086	0,037634	0,012663
	D	0,000655	0,015492	207675,236	0,145036	0,007761
	EGL	0,001231	0,017127	207336,423	0,006538	0,007471
	GG	0,003224	0,027008	208219,292	0,003414	0,012468

Da Tabela 8, também é possível perceber a vantagem da utilização da abordagem integrada na comparação dos resultados obtidos. Por exemplo, a  $BS_{MBO}$  apresentou valores inferiores de  $1 - R^2$  e  $DSK$ , enquanto que a  $BS_{HS}$  apresentou valores inferiores de  $KS$  e  $AIC$ , no entanto, a partir do  $GS$ , é evidente o melhor desempenho da  $BS_{HS}$  com  $GS = 0,278916$ , enquanto que a  $BS_{MBO}$  apresentou  $GS = 0,279593$ . Outra situação em que tal abordagem mostrou-se útil é ao comparar, por exemplo, a  $W_{ICA}$  e  $GEV_{CS}$ . O resultado fornecido pelo ICA possui valores inferiores de  $AIC$  e  $DSK$ , enquanto que o resultado obtido pelo CS apresentou valores inferiores de  $1 - R^2$  e  $KS$ . Comparando-as em relação ao  $GS$ , a  $GEV_{CS}$  apresentou  $GS = 0,007910$  e a  $W_{ICA}$   $GS = 0,007894$ , sendo este o melhor resultado dentre os dois comparados.

Comparando o mesmo modelo de distribuição, pode ser visto que todos os valores de  $GS$  obtidos pelos MOA são inferiores aos obtidos pelo método determinístico MLE, demonstrando, novamente, a aplicabilidade e elevado desempenho dos MOA na determinação dos parâmetros das distribuições de velocidade do vento utilizadas. Ademais, os valores de  $GS$  obtidos pelos MOA apresentaram pequenas variações, demonstrando que tais métodos obtiveram soluções semelhantes. Por exemplo, os  $GS$  da  $N_{MBO}$ ,  $N_{ICA}$ ,  $N_{HS}$  e  $N_{CS}$  são, respectivamente, 0,016877, 0,016951, 0,016953 e 0,016875.

Para o mesmo método de estimativa dos parâmetros, o  $GS$  da distribuição EGL foi o menor e o da LN o maior. No caso do MLE, o  $GS$  da BS foi o maior. Desse modo, dentre as distribuições analisadas, a EGL foi a que proporcionou melhor ajuste nesta região, enquanto que a BS e LN apresentaram os piores ajustes.

#### 4.3.1 Seleção do modelo de distribuição ótimo para PTR-11

O resultado com menor  $GS$  de cada distribuição utilizada em Petrolina foi extraído da Tabela 8 e apresentado na Tabela 9.

Tabela 9 – Melhores modelos de distribuição em PTR-11

Modelo	Método	$1 - R^2$	$KS$	$AIC$	$DSK$	$GS$	Rank
W	ICA	0,000650	0,016483	208839,208	0,008137	0,007894	5
G	MBO	0,004228	0,045361	218950,901	0,370237	0,051379	8
BS	HS	0,008259	0,059861	267403,783	1,643141	0,278916	10
N	CS	0,001783	0,031104	211471,207	0,018213	0,016875	7
LN	HS	0,008323	0,056429	239133,621	2,484453	0,284167	11
GL	ICA	0,007623	0,055233	227200,820	1,966589	0,190958	9
GEV	HS	0,000587	0,014901	209193,964	0,042432	0,007733	4
B	ICA	0,000644	0,018596	208887,686	0,019816	0,008733	6
D	MBO	0,000310	0,012385	207765,650	0,177590	0,006860	2
EGL	MBO	0,000426	0,009722	207496,629	0,011904	0,005246	1
GG	ICA	0,000624	0,013973	208932,538	0,023746	0,007194	3

A distribuição que proporcionou melhor ajuste foi a  $EGL_{MBO}$ , seguida da  $D_{MBO}$  e  $GG_{ICA}$ , em segundo e terceiro lugar, respectivamente.

A distribuição  $W_{ICA}$  ocupou a quinta posição, confirmando, novamente, que esta distribuição não é sempre a mais apropriada. Contudo, a  $W_{ICA}$  foi a que apresentou menor  $GS$  quando comparada com as demais distribuições de dois parâmetros.

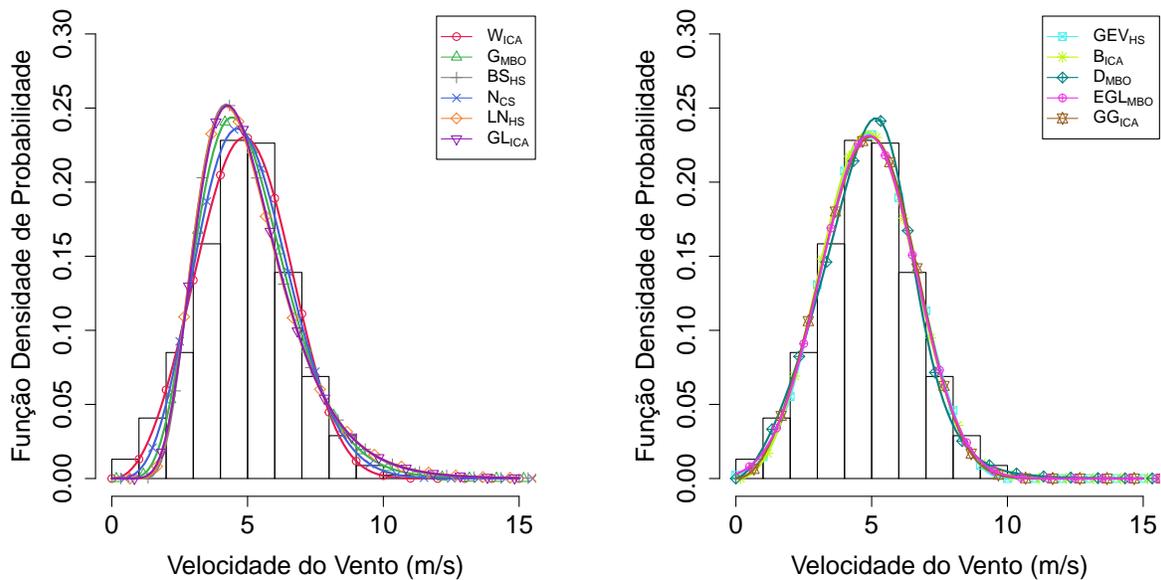
De todas as onze distribuições, a  $BS_{HS}$  e  $LN_{HS}$  foram as que apresentaram os piores valores de  $GS$ . No geral, é possível inferir que, para a região de Petrolina, as distribuições de três parâmetros, cujo ajuste foi realizado através dos MOA com a função objetivo  $R^2$ , proporcionaram melhor ajuste em relação às distribuições de dois parâmetros, que foram ajustadas da mesma maneira e ocuparam as posições mais elevadas no ranking. Ademais, é importante ressaltar, mais uma vez, que nenhum dos melhores resultados foi fornecido pelo MLE, comprovando o bom desempenho dos MOA.

Na Figura 5, foram apresentados graficamente os ajustes obtidos através dos modelos de distribuição da Tabela 9.

A partir da Figura 5, é possível ver que, dentre os modelos de dois parâmetros, os modelos  $W_{ICA}$  e  $N_{CS}$  foram os que apresentaram melhor ajuste. A distribuição  $W_{ICA}$  foi a que melhor se aproximou do valor máximo de densidade observado no histograma. Já as distribuições  $G_{MBO}$ ,  $GL_{ICA}$ ,  $BS_{HS}$  e  $LN_{HS}$  foram as que apresentaram piores ajustes. Estes resultados concordam perfeitamente com os valores de  $GS$  obtidos, visto que, dentre os modelos de dois parâmetros, as distribuições  $W_{ICA}$  e  $N_{CS}$  foram as que apresentaram menor  $GS$  e as distribuições  $G_{MBO}$ ,  $GL_{ICA}$ ,  $BS_{HS}$  e  $LN_{HS}$  maior  $GS$ .

Quanto às distribuições de três parâmetros, todas, com exceção da  $D_{MBO}$ , apresentaram desempenho gráfico semelhante. A distribuição  $D_{MBO}$  exibiu valor máximo de densidade um pouco maior do que o valor máximo observado no histograma. Contudo, é possível

Figura 5 – Distribuições de dois (esquerda) e três (direita) parâmetros em PTR-11



Fonte: Elaborada pelo Autor.

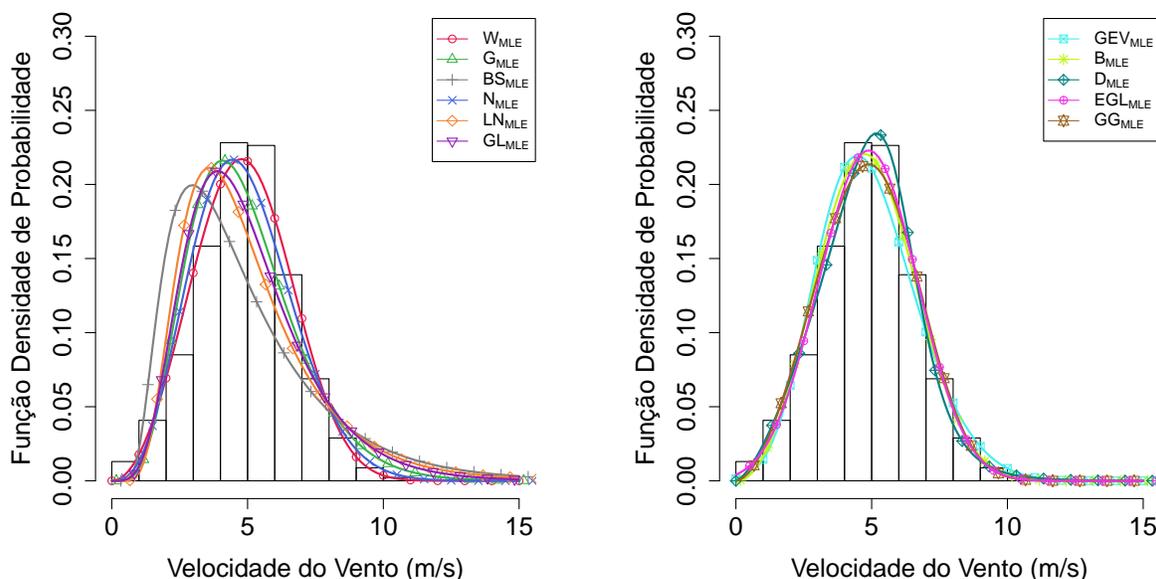
perceber sutis diferenças deste modelo em relação aos demais. Por exemplo, nos intervalos de  $3 - 4\text{m/s}$  e  $7 - 8\text{m/s}$ , a densidade exibida por esta distribuição é levemente inferior à densidade obtida pelas outras distribuições de três parâmetros. Desse modo, para evitar conclusões subjetivas na decisão do modelo mais preciso, é necessário considerar os valores dos testes estatísticos. Neste caso, a  $EGL_{MBO}$  foi a distribuição que apresentou melhor desempenho com  $GS = 0,005246$ . Já o modelo  $D_{MBO}$ , que foi o mais distinto em relação aos outros de três parâmetros, ocupou a segunda posição no ranking com  $GS = 0,006860$ .

Quanto ao desempenho dos MOA, os resultados apresentados na Tabela 9 não são conclusivos para afirmar se algum método é superior aos demais. Assim como na região de São João do Cariri, nenhum método apresentou predominância significativa em relação aos outros. Os métodos MBO e ICA garantiram bons resultados para as distribuições G, D e EGL, e W, GL, B e GG, respectivamente. Já o método HS garantiu bons resultados para os modelos BS, LN e GEV, e o CS para o modelo N.

Para comparar o desempenho gráfico dos MOA, foram apresentados, na Figura 6, os resultados obtidos pelo método determinístico MLE.

Comparando os ajustes obtidos pelo método MLE (Figura 6) com os obtidos pelos MOA (Figura 5), foi possível observar que, para os modelos W e N, as curvas geradas pelos MOA se aproximaram mais do valor máximo de densidade do histograma. Já nos demais modelos de dois parâmetros, os ajustes obtidos pelo MLE se distanciaram dos valores de densidade do histograma, enquanto que os mesmos modelos, quando ajustados pelos MOA, apresentaram

Figura 6 – Distribuições de dois (esquerda) e três (direita) parâmetros ajustadas pelo MLE em PTR-11



Fonte: Elaborada pelo Autor.

valores de densidade mais próximos.

Nas distribuições de três parâmetros, o modelo D se manteve similar em ambos os casos (Figura 5 e Figura 6). No entanto, os demais modelos de três parâmetros, quando ajustados pelos MOA, aproximaram-se mais do valor máximo de densidade observados no histograma do que quando ajustados pelo MLE, motivo pela qual as curvas estão sobrepostas na Figura 5.

Estas comparações reafirmam, graficamente, o melhor desempenho dos MOA no ajuste das distribuições utilizadas. No caso das distribuições que se mantiveram similares nesta estação (W, N e D), os valores de  $GS$  apresentados anteriormente (Tabela 9) comprovaram o ajuste superior dos MOA.

## 5 CONCLUSÃO

Em vista da carência de estudos que aplicam métodos de otimização na estimativa dos parâmetros das distribuições de velocidade do vento, foram utilizados, nesta dissertação, quatro algoritmos de otimização metaheurísticos (MOA - *Metaheuristic Optimization Algorithm*), sendo eles Pássaros Migratórios (MBO - *Migrating Birds Optimization*), Competição Imperialista (ICA - *Imperialist Competitive Algorithm*) e Busca Harmônica (HS - *Harmony Search*), cuja aplicação no ajuste de distribuições de três parâmetros e distribuições não convencionais foi realizada pela primeira vez neste trabalho, e Busca do Cuco (CS - *Cuckoo Search*), para ajustar 11 distribuições aos dados anemométricos das regiões de São João do Cariri e Petrolina, ambas localizadas no Nordeste do Brasil.

O ajuste foi feito a partir de três funções objetivo distintas, e, posteriormente, foi identificada a que gerou melhores resultados. Por fim, os ajustes obtidos foram analisados através da abordagem integrada de Masseran (2018), que considera mutuamente quatro testes estatísticos comumente utilizados neste tipo de estudo, sendo eles o Coeficiente de Determinação ( $1 - R^2$ ), Teste de Kolmogorov-Smirnov (*KS - Kolmogorov-Smirnov Test*), Critério de informação de Akaike (*AIC - Akaike Information Criterion*) e Desvio de Assimetria e Curtose (*DSK - Deviation of skewness and kurtosis*), em um único valor denominado Global Score (*GS*).

Os resultados obtidos a partir dos MOA foram então confrontados com os resultados obtidos a partir do método determinístico Máxima Verossimilhança (*MLE - Maximum Likelihood Estimation*), a fim de avaliar o desempenho dos MOA na estimativa dos parâmetros das distribuições utilizadas.

Ao comparar o resultado obtido a partir das três funções objetivo, a maximização do coeficiente de determinação  $R^2$  foi a que proporcionou melhor ajuste em ambas as regiões, visto que o *GS* obtido através desta função foi o menor na maioria dos modelos de distribuição utilizados. Em São João do Cariri 68,18% dos melhores resultados foram obtidos através da maximização do  $R^2$ , e em Petrolina 43,18% dos melhores resultados foram obtidos através da mesma função.

Posteriormente, ao conduzir a análise comparativa, a abordagem integrada mostrou-se útil. Diante das variações observadas nos testes estatísticos  $1 - R^2$ , *KS*, *AIC* e *DSK*, seria necessário mensurar importância destes testes para determinar o modelo de distribuição mais preciso. Contudo, utilizando o *GS* como critério de decisão, as comparações foram realizadas de forma simples e objetiva. Além disso, os valores de *GS* concordaram perfeitamente com o desempenho gráfico dos modelos ajustados, visto que os modelos com melhor ajuste gráfico apresentaram também os menores valores de *GS*.

Na região de São João do Cariri, as distribuições de velocidade do vento ajustadas a partir dos MOA, com exceção da  $D_{CS}$ , apresentaram maior precisão quando comparadas com a mesma distribuição ajustada pelo MLE, isto é, os valores de  $GS$  obtidos pelos MOA foram inferiores aos valores de  $GS$  obtidos pelo MLE. Na região de Petrolina, todas as distribuições ajustadas a partir dos MOA, sem exceções, apresentaram precisão superior. Tais resultados demonstram o elevado desempenho dos MOA e a sua viabilidade de uso na determinação dos parâmetros ótimos das distribuições de dois e três parâmetros. Quanto ao desempenhos dos MOA, os resultados não foram conclusivos para afirmar se algum método é superior aos demais. Além disso, em ambas as regiões, os resultados obtidos pelos MOA foram muito similares, visto que, para uma mesma distribuição, os quatro MOA forneceram valores de  $GS$  próximos.

Em ambas as regiões, os modelos de três parâmetros proporcionaram, em geral, melhor ajuste (menor  $GS$ ) do que os modelos de dois parâmetros. Foi possível também concluir que a amplamente utilizada distribuição de Weibull ( $W$ ) não é a mais adequada nas duas regiões analisadas. Contudo, dentre as distribuições de dois parâmetros, a de Weibull foi a que proporcionou melhor ajuste. Em São João do Cariri, o melhor ajuste foi obtido através da distribuição  $GG_{ICA}$ , que apresentou  $GS = 0,005766$ , e a distribuição  $W_{MBO}$  ocupou a quarta posição com  $GS = 0,009369$ . Em Petrolina, o melhor ajuste foi obtido através da recentemente proposta distribuição  $EGL_{MBO}$ , que apresentou  $GS = 0,005246$ , e a distribuição de  $W_{ICA}$  ocupou a quinta posição com  $GS = 0,007894$ . Os modelos de Birnbaum-Saunders ( $BS$ ) e Lognormal ( $LN$ ) foram os que apresentaram pior ajuste (maior  $GS$ ) em ambas as regiões. Por fim, é válido ressaltar que, neste trabalho, o desempenho apresentado pelas distribuições de probabilidade é referente aos dados coletados em São João do Cariri (Estação SCR-25) e Petrolina (Estação PTR-11), ou seja, se as mesmas distribuições foram aplicadas a um conjunto de dados diferente, o resultado não será necessariamente o mesmo.

## REFERÊNCIAS

- ABEEÓLICA. *Eólica já é a segunda fonte da matriz elétrica brasileira com 15GW de capacidade instalada*. 2019. Disponível em: <[http://abeeolica.org.br/wp-content/uploads/2019/06/Boletim-Anual\\_2018\\_Ingl%C3%AAs.pdf](http://abeeolica.org.br/wp-content/uploads/2019/06/Boletim-Anual_2018_Ingl%C3%AAs.pdf)>. Acesso em: 26 de fevereiro de 2020.
- ABEEÓLICA. *Eólica já é a segunda fonte da matriz elétrica brasileira com 15GW de capacidade instalada*. 2019. Disponível em: <<http://abeeolica.org.br/noticias/eolica-ja-e-a-segunda-fonte-da-matriz-eletrica-brasileira-com-15-gw-de-capacidade-instalada/>>. Acesso em: 31 de dezembro de 2019.
- ABEEÓLICA. *InfoVento - 14. Energia Eólica - Os bons ventos do Brasil*. 2019. Disponível em: <[http://abeeolica.org.br/wp-content/uploads/2020/02/Infovento-14\\_PT.pdf](http://abeeolica.org.br/wp-content/uploads/2020/02/Infovento-14_PT.pdf)>. Acesso em: 27 de fevereiro de 2020.
- AKDAG, S. A.; DINLER, A. A new method to estimate weibull parameters for wind energy applications. *Energy Conversion and Management*, v. 50, n. 7, p. 1761 – 1766, 2009. ISSN 0196-8904.
- AKDAG, S. A.; GULER, O. A novel energy pattern factor method for wind speed distribution parameter estimation. *Energy Conversion and Management*, v. 106, p. 1124 – 1133, 2015. ISSN 0196-8904.
- ALAVI, O.; MOHAMMADI, K.; MOSTAFAEIPOUR, A. Evaluating the suitability of wind speed probability distribution models: A case of study of east and southeast parts of iran. *Energy Conversion and Management*, v. 119, p. 101 – 108, 2016. ISSN 0196-8904.
- ALAVI, O.; SEDAGHAT, A.; MOSTAFAEIPOUR, A. Sensitivity analysis of different wind speed distribution models with actual and truncated wind data: A case study for kerman, iran. *Energy Conversion and Management*, v. 120, p. 51 – 61, 2016. ISSN 0196-8904.
- ALMEIDA, F. S. Optimization of laminated composite structures using harmony search algorithm. *Composite Structures*, v. 221, p. 110852, 2019. ISSN 0263-8223.
- ANDRADE, C. F.; NETO, H. F. M.; ROCHA, P. A. C.; SILVA, M. E. V. An efficiency comparison of numerical methods for determining weibull parameters for wind energy applications: A new approach applied to the northeast region of brazil. *Energy Conversion and Management*, v. 86, p. 801 – 808, 2014. ISSN 0196-8904.
- ANDRADE, C. F.; SANTOS, L. F.; MACÊDO, M. V. S.; ROCHA, P. A. C.; GOMES, F. F. Four heuristic optimization algorithms applied to wind energy: determination of weibull curve parameters for three brazilian sites. *International Journal of Energy and Environmental Engineering*, v. 10, n. 1, p. 1–12, Mar 2019. ISSN 2251-6832.
- ARIES, N.; BOUDIA, S. M.; OUNIS, H. Deep assessment of wind speed distribution models: A case study of four sites in algeria. *Energy Conversion and Management*, v. 155, p. 78 – 90, 2018. ISSN 0196-8904.

- ARSLAN, T.; ACITAS, S.; SENOGLU, B. Generalized lindley and power lindley distributions for modeling the wind speed data. *Energy Conversion and Management*, v. 152, p. 300 – 311, 2017. ISSN 0196-8904.
- ASKARZADEH, A.; ZEBARJADI, M. Wind power modeling using harmony search with a novel parameter setting approach. *Journal Of Wind Engineering And Industrial Aerodynamics*, v. 135, p. 70–75, 2014.
- ATASHPAZ-GARGARI, E. E.; LUCAS, C. Imperialist competitive algorithm: an algorithm for optimization inspired by imperialistic competition. *Proceedings of IEEE Congress on Evolutionary Computation*, Singapore, p. 4661– 4667, 2007.
- BRANO, V. L.; ORIOLI, A.; CIULLA, G.; CULOTTA, S. Quality of wind speed fitting distributions for the urban area of palermo, italy. *Renewable Energy*, v. 36, n. 3, p. 1026 – 1039, 2011. ISSN 0960-1481.
- CARTA, J.; RAMÍREZ, P.; VELÁZQUEZ, S. A review of wind speed probability distributions used in wind energy analysis: Case studies in the canary islands. *Renewable and Sustainable Energy Reviews*, v. 13, n. 5, p. 933 – 955, 2009. ISSN 1364-0321.
- CHANG, T. P. Wind energy assessment incorporating particle swarm optimization method. *Energy Conversion and Management*, v. 52, n. 3, p. 1630 – 1637, 2011. ISSN 0196-8904.
- CHIODO, E.; FALCO, P. D. Inverse burr distribution for extreme wind speed prediction: Genesis, identification and estimation. *Electric Power Systems Research*, v. 141, p. 549 – 561, 2016. ISSN 0378-7796.
- CHIRANJEEVI, K.; JENA, U. R. Image compression based on vector quantization using cuckoo search optimization technique. *Ain Shams Engineering Journal*, v. 9, n. 4, p. 1417 – 1431, 2018. ISSN 2090-4479.
- DUMAN, E.; UYSAL, M.; ALKAYA, A. F. Migrating birds optimization: A new metaheuristic approach and its performance on quadratic assignment problem. *Information Sciences*, v. 217, p. 65 – 77, 2012. ISSN 0020-0255.
- ÊARDHANAN, M. N.; LI, Z.; BOCEWICZ, G.; BANASZAK, Z.; NIELSEN, P. Metaheuristic algorithms for balancing robotic assembly lines with sequence-dependent robot setup times. *Applied Mathematical Modelling*, v. 65, p. 256 – 270, 2019. ISSN 0307-904X.
- EMARY, E.; ZAWBAA, H. M.; SHARAWI, M. Impact of lévy flight on modern meta-heuristic optimizers. *Applied Soft Computing*, v. 75, p. 775 – 789, 2019. ISSN 1568-4946.
- GEEM, Z. W.; KIM, J. H.; LOGANATHAN, G. A. New heuristic optimization algorithm: Harmony search. *Simulation*, v. 76, n. 2, p. 60–68, 2001.
- GERIST, S.; MAHERI, M. R. Structural damage detection using imperialist competitive algorithm and damage function. *Applied Soft Computing*, v. 77, p. 1 – 23, 2019. ISSN 1568-4946.

- GUGLIANI, G.; SARKAR, A.; LEY, C.; MANDAL, S. New methods to assess wind resources in terms of wind speed, load, power and direction. *Renewable Energy*, v. 129, p. 168 – 182, 2018. ISSN 0960-1481.
- GWEC. *Americas wind installations rise 12% in 2019 to 13.4GW*. 2020. Disponível em: <<https://gwec.net/americas-wind-installations-rise-12-in-2019-to-13-4gw/>>. Acesso em: 07 de fevereiro de 2020.
- JAAFARI, A.; ZENNER, E. K.; PANAHI, M.; SHAHABI, H. Hybrid artificial intelligence models based on a neuro-fuzzy system and metaheuristic optimization algorithms for spatial prediction of wildfire probability. *Agricultural and Forest Meteorology*, v. 266-267, p. 198 – 207, 2019. ISSN 0168-1923.
- JALAL, M.; GOHARZAY, M. Cuckoo search algorithm for applied structural and design optimization: Float system for experimental setups. *Journal of Computational Design and Engineering*, v. 6, n. 2, p. 159 – 172, 2019. ISSN 2288-4300.
- JIANG, H.; WANG, J.; WU, J.; GENG, W. Comparison of numerical methods and metaheuristic optimization algorithms for estimating parameters for wind energy potential assessment in low wind regions. *Renewable and Sustainable Energy Reviews*, v. 69, p. 1199 – 1217, 2017. ISSN 1364-0321.
- JUNG, C.; SCHINDLER, D. Global comparison of the goodness-of-fit of wind speed distributions. *Energy Conversion and Management*, v. 133, p. 216 – 234, 2017. ISSN 0196-8904.
- JUNG, C.; SCHINDLER, D. Wind speed distribution selection – a review of recent development and progress. *Renewable and Sustainable Energy Reviews*, v. 114, p. 109290, 2019. ISSN 1364-0321.
- JUNG, C.; SCHINDLER, D.; LAIBLE, J.; BUCHHOLZ, A. Introducing a system of wind speed distributions for modeling properties of wind speed regimes around the world. *Energy Conversion and Management*, v. 144, p. 181 – 192, 2017. ISSN 0196-8904.
- KANTAR, Y. M.; USTA, I.; ARIK, I.; YENILMEZ, I. Wind speed analysis using the extended generalized lindley distribution. *Renewable Energy*, v. 118, p. 1024 – 1030, 2018. ISSN 0960-1481.
- KISS, P.; JÁNOSI, I. M. Comprehensive empirical analysis of era-40 surface wind speed distribution over europe. *Energy Conversion and Management*, v. 49, n. 8, p. 2142 – 2151, 2008. ISSN 0196-8904.
- LISSAMAN, P. B. S.; SHOLLENBERGER, C. A. Formation flight of birds. *Science*, n. 168, p. 1003–1005, 1970.
- MACÊDO, M. V. S. *Aplicação de Algoritmos de Otimização Heurística à Energia Eólica: Determinação dos Parâmetros da Curva de Weibull para Duas Regiões Brasileiras*. 88 p. Dissertação (Mestrado em Engenharia Mecânica) — Universidade Federal do Ceará, Fortaleza, 2018.

- MAKAS, H.; YUMUSAK, N. System identification by using migrating birds optimization algorithm: a comparative performance analysis. *Information Sciences*, v. 24, p. 1879–1900, 2016.
- MANSOURZADEH, F.; KHAMSEH, A. G.; SAFDARI, J.; NOROUZI, A. Utilization of harmony search algorithm to optimize a cascade for separating multicomponent mixtures. *Progress in Nuclear Energy*, v. 111, p. 165 – 173, 2019. ISSN 0149-1970.
- MASSERAN, N. Integrated approach for the determination of an accurate wind-speed distribution model. *Energy Conversion and Management*, v. 173, p. 56 – 64, 2018. ISSN 0196-8904.
- MAZZEO, D.; OLIVETI, G.; LABONIA, E. Estimation of wind speed probability density function using a mixture of two truncated normal distributions. *Renewable Energy*, v. 115, p. 1260 – 1280, 2018. ISSN 0960-1481.
- MIAO, S.; GU, Y.; LI, D.; LI, H. Determining suitable region wind speed probability distribution using optimal score-radar map. *Energy Conversion and Management*, v. 183, p. 590 – 603, 2019. ISSN 0196-8904.
- MOHAMMADI, K.; ALAVI, O.; MCGOWAN, J. G. Use of birnbaum-saunders distribution for estimating wind speed and wind power probability distributions: A review. *Energy Conversion and Management*, v. 143, p. 109 – 122, 2017. ISSN 0196-8904.
- MOHAMMADI, K.; ALAVI, O.; MOSTAFAEIPOUR, A.; GOUDARZI, N.; JALILVAND, M. Assessing different parameters estimation methods of weibull distribution to compute wind power density. *Energy Conversion and Management*, v. 108, p. 322 – 335, 2016. ISSN 0196-8904.
- MORADI, M. H.; ZEINALZADEH, A.; MOHAMMADI, Y.; ABEDINI, M. An efficient hybrid method for solving the optimal sitting and sizing problem of dg and shunt capacitor banks simultaneously based on imperialist competitive algorithm and genetic algorithm. *Electrical Power and Energy Systems*, v. 54, p. 101–111, 2014.
- MORGAN, E. C.; LACKNER, M.; VOGEL, R. M.; BAISE, L. G. Probability distributions for offshore wind speeds. *Energy Conversion and Management*, v. 52, n. 1, p. 15 – 26, 2011. ISSN 0196-8904.
- OSMAN, I. H.; LAPORTE, G. Metaheuristics: A bibliography. *Annals of Operational Research*, v. 63, p. 513–628, 1996.
- OUARDA, T.; CHARRON, C.; CHEBANA, F. Review of criteria for the selection of probability distributions for wind speed data and introduction of the moment and l-moment ratio diagram methods, with a case study. *Energy Conversion and Management*, v. 124, p. 247 – 265, 2016. ISSN 0196-8904.
- OUARDA, T.; CHARRON, C.; SHIN, J.-Y.; MARPU, P.; AL-MANDOOS, A.; AL-TAMIMI, M.; GHEDIRA, H.; HOSARY, T. A. Probability distributions of wind speed in the uae. *Energy Conversion and Management*, v. 93, p. 414 – 434, 2015. ISSN 0196-8904.

- QIN, Z.; LI, W.; XIONG, X. Estimating wind speed probability distribution using kernel density method. *Electric Power Systems Research*, v. 81, n. 12, p. 2139 – 2146, 2011. ISSN 0378-7796.
- R-PROJECT. 2019. Disponível em: <<https://www.r-project.org/about.html>>. Acesso em: 23 de maio de 2019.
- ROCHA, P. A. C.; SOUSA, R. C.; ANDRADE, C. F.; SILVA, M. E. V. Comparison of seven numerical methods for determining weibull parameters for wind energy generation in the northeast region of brazil. *Applied Energy*, v. 89, n. 1, p. 395 – 400, 2012. ISSN 0306-2619. Special issue on Thermal Energy Management in the Process Industries.
- SANTOS, L. F. *Parametrização e Aplicação de Métodos Heurísticos no Ajuste da Curva de Distribuição de Weibull em Energia Eólica*. 82 p. Dissertação (Mestrado em Engenharia Mecânica) — Universidade Federal do Ceará, Fortaleza, 2018.
- SHAYANFAR, H.; GHAREHCHOPOGH, F. S. Farmland fertility: A new metaheuristic algorithm for solving continuous optimization problems. *Applied Soft Computing*, v. 71, p. 728 – 746, 2018. ISSN 1568-4946.
- SOUKISSIAN, T. H.; KARATHANASI, F. E. On the selection of bivariate parametric models for wind data. *Applied Energy*, v. 188, p. 280 – 304, 2017. ISSN 0306-2619.
- TOFALLIS, C. Selecting the best statistical distribution using multiple criteria. *Computers Industrial Engineering*, v. 54, n. 3, p. 690 – 694, 2008. ISSN 0360-8352.
- USTA, I.; ARIK, I.; YENILMEZ, I.; KANTAR, Y. M. A new estimation approach based on moments for estimating weibull parameters in wind power applications. *Energy Conversion and Management*, v. 164, p. 570 – 578, 2018. ISSN 0196-8904.
- WAIS, P. Two and three-parameter weibull distribution in available wind power analysis. *Renewable Energy*, v. 103, p. 15 – 29, 2017. ISSN 0960-1481.
- WANG, J.; HU, J.; MA, K. Wind speed probability distribution estimation and wind energy assessment. *Renewable and Sustainable Energy Reviews*, v. 60, p. 881 – 899, 2016. ISSN 1364-0321.
- WANG, J.; HUANG, X.; LI, Q.; MA, X. Comparison of seven methods for determining the optimal statistical distribution parameters: A case study of wind energy assessment in the large-scale wind farms of china. *Energy*, v. 164, p. 432 – 448, 2018. ISSN 0360-5442.
- WANG, J.; QIN, S.; JIN, S.; WU, J. Estimation methods review and analysis of offshore extreme wind speeds and wind energy resources. *Renewable and Sustainable Energy Reviews*, v. 42, p. 26 – 42, 2015. ISSN 1364-0321.
- WANG, Z.; WANG, C.; WU, J. Wind energy potential assessment and forecasting research based on the data pre-processing technique and swarm intelligent optimization algorithms. *Sustainability*, v. 8, p. 1191, 11 2016.
- YANG, S. D. X. S. Engineering optimisation by cuckoo search. *International Journal of Mathematical Modelling and Numerical Optimisation*, v. 1, n. 4, p. 330–343, 2010.

YANG, X.-S. *Firefly Algorithm, Lévy Flights and Global Optimization*. London: Springer London, 2010. 209–218 p.

YANG, X.-S. *Nature-inspired metaheuristic algorithms*. [S.l.]: Luniver press, 2010.

YANG, X. S.; DEB, S. Cuckoo search via lévy flights. *In: Proceedings of World Congress on Nature Biologically Inspired Computing (NaBIC 2009)*, USA, p. 210–214, 2009.

ZHAO, X.; WANG, C.; SU, J.; WANG, J. Research and application based on the swarm intelligence algorithm and artificial intelligence for wind farm decision system. *Renewable Energy*, v. 134, p. 681 – 697, 2019. ISSN 0960-1481.

## APÊNDICE A – PARÂMETROS ESTIMADOS

Tabela 10 – Parâmetros estimados em SCR-25

Método	Modelo	$k$	$c$	$p$	$u$
MBO	W	2,370687	6,009842	-	-
	G	4,423883	1,228761	-	-
	BS	0,444477	5,006488	-	-
	N	1,283610	34,355130	-	-
	LN	0,487203	1,600586	-	-
	GL	3,596458	0,551307	-	-
	GEV	-0,211519	2,332288	-	4,314224
	B	22,438320	21,772042	2,387160	-
	D	0,200500	7,803006	8,517126	-
	EGL	0,201729	0,196448	3,050599	-
GG	0,697487	6,976942	2,961900	-	
ICA	W	2,335101	5,983722	-	-
	G	4,595746	1,179445	-	-
	BS	0,484721	4,972712	-	-
	N	1,284759	34,389554	-	-
	LN	0,482657	1,602548	-	-
	GL	3,612720	0,551714	-	-
	GEV	-0,209794	2,331139	-	4,310956
	B	22,453135	21,854862	2,387253	-
	D	0,268928	7,555532	6,670501	-
	EGL	0,198401	0,222972	2,816565	-
GG	0,707058	6,942840	2,972689	-	
HS	W	2,339947	5,979219	-	-
	G	4,472600	1,214841	-	-
	BS	0,483976	4,962833	-	-
	N	1,281132	34,350025	-	-
	LN	0,484346	1,605399	-	-
	GL	3,612823	0,551828	-	-
	GEV	-0,223638	2,341085	-	4,306569

*Continua na próxima página*

Tabela 10 – *Continuação da página anterior*

Método	Modelo	$k$	$c$	$p$	$u$
	B	22,537475	21,760738	2,391053	-
	D	0,194671	7,907900	8,693193	-
	EGL	0,200836	0,194442	3,079340	-
	GG	0,699304	6,968861	2,956382	-
CS	W	2,341917	5,979464	-	-
	G	4,444221	1,222963	-	-
	BS	0,485725	4,968315	-	-
	N	1,283428	34,336305	-	-
	LN	0,481657	1,603572	-	-
	GL	3,608271	0,552018	-	-
	GEV	-0,216553	2,337735	-	4,310159
	B	22,414137	21,805577	2,384687	-
	D	0,300131	7,376135	6,176464	-
	EGL	0,156743	0,307449	2,511320	-
	GG	0,697903	6,980167	2,958908	-
MLE	W	2,372090	5,893431	-	-
	G	3,807273	1,375982	-	-
	BS	0,682202	4,213422	-	-
	N	1,238562	32,881914	-	-
	LN	0,614066	1,518876	-	-
	GL	2,810022	0,520792	-	-
	GEV	-0,231151	2,238004	-	4,360931
	B	16,458330	18,212635	2,447453	-
	D	0,201844	7,919395	8,492404	-
	EGL	0,170376	0,229321	2,937391	-
	GG	0,499297	7,706802	3,760511	-

Tabela 11 – Parâmetros estimados em PTR-11

Método	Modelo	$k$	$c$	$p$	$u$
MBO	W	3,276481	5,458617	-	-
	G	8,475742	0,591915	-	-

*Continua na próxima página*

Tabela 11 – *Continuação da página anterior*

Método	Modelo	$k$	$c$	$p$	$u$
	BS	0,345531	4,792380	-	-
	N	2,290860	27,374349	-	-
	LN	0,337510	1,564538	-	-
	GL	10,113503	0,804863	-	-
	GEV	-0,299510	1,671291	-	4,301603
	B	23,600301	13,984260	3,334848	-
	D	0,300047	6,323836	8,733032	-
	EGL	0,049584	0,404030	3,233919	-
	GG	0,880405	5,701942	3,554275	-
ICA	W	3,244604	5,466702	-	-
	G	8,648446	0,579729	-	-
	BS	0,345650	4,798061	-	-
	N	2,292683	27,384632	-	-
	LN	0,344199	1,566764	-	-
	GL	10,090962	0,805051	-	-
	GEV	-0,281225	1,651682	-	4,299259
	B	23,266382	13,891365	3,337686	-
	D	0,333515	6,257556	8,119210	-
	EGL	0,054811	0,329600	3,559815	-
	GG	0,835337	5,809061	3,663966	-
HS	W	3,281529	5,457621	-	-
	G	8,946073	0,560364	-	-
	BS	0,352121	4,783253	-	-
	N	2,291748	27,399873	-	-
	LN	0,354176	1,563574	-	-
	GL	10,347821	0,809629	-	-
	GEV	-0,289099	1,654311	-	4,305860
	B	23,502013	13,945897	3,339704	-
	D	0,315376	6,272039	8,507307	-
	EGL	0,050899	0,389756	3,282426	-
	GG	0,834345	5,799145	3,685888	-
CSO	W	3,276481	5,458616	-	-

*Continua na próxima página*

Tabela 11 – *Continuação da página anterior*

Método	Modelo	$k$	$c$	$p$	$u$
	G	8,481551	0,591523	-	-
	BS	0,345858	4,793218	-	-
	N	2,290842	27,374076	-	-
	LN	0,344440	1,567352	-	-
	GL	10,115592	0,804898	-	-
	GEV	-0,299348	1,666135	-	4,303301
	B	23,818638	14,016271	3,336542	-
	D	0,341943	6,211641	7,933295	-
	EGL	0,046391	0,417870	3,227522	-
	GG	0,883640	5,695251	3,545757	-
MLE	W	3,026071	5,453699	-	-
	G	6,048018	0,807804	-	-
	BS	0,569060	4,181806	-	-
	N	1,873844	26,918002	-	-
	LN	0,469644	1,501393	-	-
	GL	4,875211	0,667052	-	-
	GEV	-0,167510	1,701846	-	4,166856
	B	21,687596	14,489571	3,114514	-
	D	0,284671	6,416491	8,740200	-
	EGL	0,060063	0,379754	3,178275	-
	GG	0,796360	5,937828	3,476265	-