



**UNIVERSIDADE FEDERAL DO PIAUÍ**

**CENTRO DE TECNOLOGIA**

**PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA**

**MESTRADO ACADÊMICO EM CONTROLE E AUTOMAÇÃO DE SISTEMAS**

**IGOR SOARES BRASIL**

**ESTIMAÇÃO DE CURVAS DE POTÊNCIA DE AEROGERADORES USANDO O  
MODELO FUZZY DE TAKAGI-SUGENO-KANG**

**TERESINA**

**2018**

IGOR SOARES BRASIL

ESTIMAÇÃO DE CURVAS DE POTÊNCIA DE AEROGERADORES USANDO O MODELO  
FUZZY DE TAKAGI-SUGENO-KANG

Dissertação apresentada ao Curso de Mestrado Acadêmico em Controle e Automação de Sistemas do Programa de Pós-Graduação em Engenharia Elétrica do Centro de Tecnologia da Universidade Federal do Piauí, como requisito parcial à obtenção do título de mestre em Engenharia Elétrica. Área de Concentração: Controle e Automação de Sistemas

Orientador: Prof. Dr. Guilherme de Alencar Barreto

Co-Orientador: Prof. Dr. Luís Gustavo Mota Souza

TERESINA

2018

IGOR SOARES BRASIL

ESTIMAÇÃO DE CURVAS DE POTÊNCIA DE AEROGERADORES USANDO O MODELO  
FUZZY DE TAKAGI-SUGENO-KANG

Dissertação apresentada ao Curso de Mestrado Acadêmico em Controle e Automação de Sistemas do Programa de Pós-Graduação em Engenharia Elétrica do Centro de Tecnologia da Universidade Federal do Piauí, como requisito parcial à obtenção do título de mestre em Engenharia Elétrica. Área de Concentração: Controle e Automação de Sistemas

Aprovada em: 14 de Dezembro de 2018

BANCA EXAMINADORA

---

Prof. Dr. Guilherme de Alencar Barreto (Orientador)  
Universidade Federal do Ceará (UFC)

---

Prof. Dr. Luís Gustavo Mota Souza (Co-Orientador)  
Universidade Federal do Piauí (UFPI)

---

Prof. Dr. Marcos Antônio Tavares Lira  
Universidade Federal do Piauí (UFPI)

---

Prof. Dr. José Medeiros de Araújo Júnior  
Universidade Federal do Piauí (UFPI)

---

Prof. Dr. Ajalmar Rêgo da Rocha Neto  
Instituto Federal do Ceará (IFCE)

## **AGRADECIMENTOS**

Agradeço inicialmente aos meus pais Francisco Brasil e Adélia Soares e ao meu irmão Ítalo pelo apoio incondicional durante toda minha jornada pessoal e profissional.

Agradeço ao meu orientador Prof. Dr. Guilherme Alencar Barreto pela oportunidade de aprendizado, amizade e colaboração durante a execução deste trabalho e ao meu Coorientador Prof. Dr. Luís Gustavo Mota Souza pela disponibilidade, amizade e dedicação para a realização desse projeto.

Agradeço aos professores do programa de pós-graduação de Engenharia Elétrica da UFPI, pela oportunidade de absorver e compartilhar conhecimentos.

Por último, mas não menos importante, agradeço aos meus amigos e família que criei em Teresina-PI e Açailândia-MA.



“Que nada na minha vida seja fácil e de graça.  
Esse tipo de conquista não vale de nada.”

(Helder Thug)

## RESUMO

A curva de potência de um aerogerador é a representação teórica da relação entre velocidade do vento e potência elétrica gerada. A estimação ou obtenção de curvas de potência a partir de dados medidos é fundamental para predição de potência elétrica gerada por um aerogerador isolado ou por um parque eólico como um todo. Em geral, curvas de potência são usadas para dimensionamento do aerogerador ou do parque eólico e também no monitoramento da condição de operação dos aerogeradores. Costuma-se utilizar modelos de regressão polinomial para estimação da curva de potência, mas há também soluções baseadas em outros modelos não lineares, tais como redes neurais artificiais e modelos logísticos cujos parâmetros são estimados via algoritmos de otimização meta-heurísticas. Nesta dissertação, contudo, é introduzida uma solução alternativa baseada no sistema de inferência *fuzzy* de Takagi-Sugeno-Kang (TSK). A principal vantagem da abordagem proposta está em sua simplicidade e adequação ao problema de interesse. As funções de saída das regras do modelo TSK são lineares (mais exatamente equações da reta), cujos os parâmetros são estimados pelo método dos mínimos quadrados ordinário (MQO). A fim de lidar com a presença de *outliers* nos dados, o modelo TSK proposto é estendido pela substituição do método MQO por estimadores- $M$ , que é um arcabouço de estimação robusta de parâmetros. A metodologia proposta é comparada com o estado da arte em estimação da curva de potência de aerogeradores e os resultados indicam um desempenho consistente da metodologia proposta em relação aos métodos avaliados.

**Palavras-chave:** Curva de potência. *Outliers*. Regressão não-linear. Aerogeradores. Takagi-Sugeno-Kang. Estimador- $M$

## ABSTRACT

The power curve of a wind turbine is the theoretical representation of the relationship between wind speed and the generated electric power. The power curves estimate from measured data is fundamental to predict the electric power generated by an isolated wind turbine or by the wind farm as a whole. In general, power curves are used to design wind turbines or wind farms and also wind turbine operation condition monitoring. It is usually used polynomial regression models to estimate the power curve, but there are also solutions based on other nonlinear models such as artificial neural networks and logistic models whose parameters are estimated via metaheuristic optimization algorithms. In this masters thesis, an alternative solution is introduced based on the Takagi-Sugeno-Kang (TSK) fuzzy inference system. The main advantage of the proposed approach lies in its simplicity and adequacy to the problem of interest. The output functions of the rules of the TSK model are linear (more precisely equations of the line), the parameters are estimated by the ordinary least squares (OLS) method. In order to deal with the presence of outliers in the data, the proposed TSK model is extended by replacing the MQO method with estimators-  $M$ , which is a robust parameter estimation framework. The proposed methodology is compared with the state of the art in estimation of the power curve of wind turbines and the results indicate a consistent performance of the methodology proposed in relation to the evaluated methods.

**Keywords:** Power curve. Outliers. Non-linear regression. Wind turbines. Takagi-Sugeno-Kang. M-estimator

## LISTA DE ILUSTRAÇÕES

Figura 1 – Fluxo de ar através da área circular de varredura das pás do aerogerador. . . . .	26
Figura 2 – Curva de potência típica de uma aerogerador com controle de potência por ângulo de passo. . . . .	28
Figura 3 – Gráfico das funções peso para os estimadores Bisquare e de Huber. . . . .	40
Figura 4 – Representação em blocos de um neurônio artificial $\alpha$ . . . . .	41
Figura 5 – Rede MLP com duas camadas ocultas. . . . .	42
Figura 6 – Componentes de um Sistema de Inferência Fuzzy (do inglês, FIS). . . . .	46
Figura 7 – Diagrama de blocos do ANFIS. Os blocos retangulares dispõem de parâmetros que são adaptados durante a fase de treinamento. . . . .	48
Figura 8 – Tipos de função de pertinência <i>fuzzy</i> . . . . .	53
Figura 9 – Exemplo de variável linguística e seus atributos. . . . .	54
Figura 10 – Curva de potência da turbina eólica WT1. . . . .	58
Figura 11 – Diagrama de dispersão para os dados da Tabela 4. . . . .	60
Figura 12 – Gráfico das retas de regressão de acordo com os dados da Tabela 5 com e sem os dados referentes aos EUA. . . . .	61
Figura 13 – Curva de potência e regressão para o modelo polinomial de ordem 8. . . . .	65
Figura 14 – Curva de potência e regressão para o modelo de Rede MLP com 5 neurônios na camada oculta. . . . .	66
Figura 15 – Curva de potência e regressão para o modelo ANFIS com 5 funções de pertinência gaussianas. . . . .	67
Figura 16 – Curvas de potência estimados pelos modelos TSK-OLS e TSK-M (bisquare) para 5 funções de pertinência gaussianas. . . . .	68
Figura 17 – Curvas de potência estimadas pelos modelos TSK-OLS e TSK-M (bisquare) para funções de pertinência triangulares. a) 3 funções; b) 4 funções; e c) 5 funções. . . . .	70
Figura 18 – Curvas de potência estimados para o modelo TSK-OLS e TSK-M(bisquare) para a heurística 1 para 3 funções de pertinência gaussianas. . . . .	71
Figura 19 – Curvas de potência estimados pelos modelos TSK-OLS e TSK-M(bisquare) para a heurística 2 para 3 funções de pertinência gaussianas. . . . .	71
Figura 20 – Curvas de potência estimados pelos modelos TSK-OLS e TSK-M(bisquare) para a heurística 3 para 3 funções de pertinência gaussianas. . . . .	72

Figura 21 – Curvas de potência estimados para o modelo TSK-OLS e TSK-M(bisquare) para a heurística 3 para 5 funções de pertinência gaussianas. . . . .	72
Figura 22 – Boxplot dos resíduos para os modelos apresentados na Tabela 12. . . . .	74
Figura 23 – Comparação das curvas de regressão para os modelos TSK-OLS, TSK-M e MLP para o aerogerador da Prainha-CE. . . . .	75

## LISTA DE TABELAS

Tabela 1 – Os 5 estados com maior geração de energia eólica no ano de 2017. . . . .	21
Tabela 2 – Função peso e objetivo para alguns dos estimadores-M encontrados na literatura.	40
Tabela 3 – Especificações do parque eólico em estudo. . . . .	57
Tabela 4 – Consumo per capita de cigarro em vários países em 1930 e as mortes devido a câncer de pulmão em 1950. . . . .	60
Tabela 5 – Estimação dos parâmetros da reta de regressão com e sem os dados dos Estados Unidos(EUA) . . . . .	61
Tabela 6 – Valores obtidos para os índices SSE e $R^2$ para diferentes ordens do modelo polinomial. . . . .	64
Tabela 7 – Coeficiente de regressão e SSE para o modelo baseado em rede MLP . . . .	65
Tabela 8 – Valores obtidos para os índices SSE e $R^2$ para o modelo ANFIS. . . . .	66
Tabela 9 – Coeficiente de regressão e SSE para o modelo TSK-OLS e TSK-M para funções gaussianas. . . . .	67
Tabela 10 – Coeficiente de regressão e SSE para o modelo TSK-OLS e TSK-M para funções triangulares/trapezoidais. . . . .	68
Tabela 11 – Coeficiente de regressão e SSE para o modelo TSK-OLS e TSK-M para os teste das heurísticas para 3 funções de pertinência gaussianas. . . . .	69
Tabela 12 – Comparação do desempenho dos modelos quanto ao SSE e $R^2$ . . . . .	73
Tabela 13 – Especificações do aerogerador da Prainha-CE. . . . .	74

## LISTA DE ABREVIATURAS E SIGLAS

ANFIS	<i>Adaptive-Network-Based Fuzzy Inference System</i>
FDP	Função de Densidade de Probabilidade.
IRLS	<i>Iteratively Reweighted Least Squares.</i>
MAD	<i>Median Absolute Deviation.</i>
MQO	Mínimos Quadrados Ordinários.
RNA	Rede Neural Artificial.
SSE	<i>Sum of Squared Errors.</i>
WLS	<i>Weighted Least Squares.</i>

## LISTA DE SÍMBOLOS

$A$	Área circular varrida pelas pás de um aerogerador de eixo horizontal.
$b$	Limiar de ativação do neurônio.
$c$	Limiar de identificação de um <i>outlier</i> utilizado na regressão robusta.
$C$	Conjunto de dados para obtenção de centróide.
$C_p$	Coefficiente de potência ( $P_a/P_w$ ).
$D$	Dimensão dos dados de entrada de um neurônio artificial.
$E_c$	Energia cinética.
$f$	Uma função genérica qualquer.
$g$	Função de densidade de probabilidade da potência elétrica.
$J$	Função de custo da regressão cujo objetivo é reduzir seu valor de resposta.
$k$	Cardinalidade de $\theta$ .
$k_\beta$	Cardinalidade do vetor $\beta$ .
$m$	Massa.
$n$	Cardinalidade do conjunto amostral.
$p$	Potência elétrica fornecida pelo rotor.
$\mathbf{p}$	Vetor de $p_i \quad i \in \{1 \dots n\}$ .
$P$	Potência mecânica genérica.
$P_a$	Potência eólica absorvida.
$P_w$	Potência eólica disponível.
$t$	Tempo.
$u$	Residual normalizado pelo seu desvio padrão estimado ( $\frac{\epsilon_i}{\hat{\sigma}_\epsilon}$ ).
$v$	Velocidade do vento.



$V$	Matriz de Vandermonde de $v$ .
$x$	Variável aleatória qualquer.
$y$	Variável dependente qualquer.
$w_i$	Componente de ponderação de um determinado dado de índice $i$ em que $i \in \{1 \dots n\}$ .
$W$	Matriz diagonal $n \times n$ com $W_{ii} = w_i$ .
$X$	FDP de $x$ .
$z$	Dados de entrada de um neurônio artificial.
$\alpha$	Identificação de um neurônio qualquer.
$\beta$	Vetor dos coeficientes de uma regressão.
$\gamma$	Constante utilizada para estimar o desvio padrão de uma distribuição através do MAD. Seu valor depende do tipo de distribuição em análise.
$\varepsilon$	Residual da regressão.
$\eta$	Eficiência de conversão de energia mecânica em elétrica do rotor.
$\theta$	Vetor de parâmetros de dependência da FDP $g$ .
$\theta_x$	Vetor de parâmetros de dependência da FDP $X$ .
$\lambda$	Constante que regulariza a intensidade da penalização $\ell_d$ .
$\mu$	Valor esperado de potência elétrica.
$\mu_A(x)$	Grau de pertinência que o elemento $x$ pertence ao universo $X$ no conjunto <i>fuzzy</i> $A$ .
$\rho$	Densidade do ar.
$\sigma^2$	Variância da potência elétrica.
$\sigma_\varepsilon^2$	Variância de $\varepsilon$ .
$\hat{\sigma}$	Desvio padrão estimado.

$\varphi(\cdot)$	Função de ativação do neurônio.
$\psi$	Derivada da função $\Psi$ .
$\Psi$	Função contínua, derivável em $\mathbb{R}$ e com um mínimo global.

## SUMÁRIO

<b>1</b>	<b>CONCEITOS GERAIS</b> . . . . .	17
<b>1.1</b>	<b>Introdução</b> . . . . .	17
<b>1.2</b>	<b>Motivação</b> . . . . .	18
<b>1.3</b>	<b>Justificativa</b> . . . . .	20
<b>1.4</b>	<b>Objetivos Geral e Específicos</b> . . . . .	22
<b>1.5</b>	<b>Produção Científica</b> . . . . .	22
<b>1.6</b>	<b>Estrutura do Trabalho</b> . . . . .	22
<b>2</b>	<b>FUNDAMENTOS DE GERAÇÃO EÓLICA</b> . . . . .	24
<b>2.1</b>	<b>Necessidade da Modelagem da Curva de Potência</b> . . . . .	24
<i>2.1.1</i>	<i>Avaliação e Previsão da Energia Produzida</i> . . . . .	24
<i>2.1.2</i>	<i>Estimativa do Fator de Capacidade</i> . . . . .	25
<i>2.1.3</i>	<i>Seleção de Aerogeradores</i> . . . . .	25
<i>2.1.4</i>	<i>Monitoramento Online de Curvas de Potência</i> . . . . .	25
<b>2.2</b>	<b>Fundamentos Teóricos da Modelagem de um Aerogerador de Eixo Horizontal</b> . . . . .	26
<b>2.3</b>	<b>Problemas de Modelagem</b> . . . . .	29
<i>2.3.1</i>	<i>Diferença nos Modelos de Aerogeradores</i> . . . . .	29
<i>2.3.2</i>	<i>Comportamento nas Velocidades de Corte</i> . . . . .	29
<i>2.3.3</i>	<i>Aerogerador Único versus Grupo de Aerogeradores</i> . . . . .	30
<i>2.3.4</i>	<i>Fatores de Influência</i> . . . . .	30
<b>2.4</b>	<b>Aquisição de Dados para Modelagem</b> . . . . .	31
<i>2.4.1</i>	<i>Análise Estatística da Velocidade do Vento</i> . . . . .	32
<i>2.4.2</i>	<i>Levantamento da Curva de Potência</i> . . . . .	32
<b>2.5</b>	<b>Conclusões</b> . . . . .	33
<b>3</b>	<b>ESTIMAÇÃO DA CURVA DE POTÊNCIA VIA MODELOS POLINOMIAIS E NEURAIS</b> . . . . .	35
<b>3.1</b>	<b>Modelo de regressão polinomial</b> . . . . .	35
<b>3.2</b>	<b>Modelo polinomial ponderado</b> . . . . .	36
<i>3.2.1</i>	<i>Estimadores-M</i> . . . . .	38
<b>3.3</b>	<b>Rede Perceptron Multicamadas (MLP)</b> . . . . .	41

3.3.1	<i>Levenberg-Marquardt</i> . . . . .	44
3.4	<b>Sistemas Neuro-Fuzzy</b> . . . . .	45
3.4.1	<i>Adaptive-Network-Based Fuzzy Inference System (ANFIS)</i> . . . . .	45
3.5	<b>Conclusão do Capítulo</b> . . . . .	49
4	<b>ESTIMAÇÃO DA CURVA DE POTÊNCIA VIA MODELO DE TAKAGI-SUGENO-KANG</b> . . . . .	51
4.1	<b>Fundamentos de Modelagem Fuzzy</b> . . . . .	51
4.1.1	<i>Conjuntos Fuzzy</i> . . . . .	51
4.1.2	<i>Funções de Pertinência</i> . . . . .	52
4.1.3	<i>Operações com Conjuntos Fuzzy</i> . . . . .	52
4.1.4	<i>Variáveis Linguísticas</i> . . . . .	53
4.2	<b>Modelo fuzzy Takagi-Sugeno-Kang (TSK)</b> . . . . .	54
4.2.1	<i>Modelo TSK para estimação de Curva de potência</i> . . . . .	55
4.3	<b>Metodologia de Ajuste de Hiperparâmetros do Modelo TSK</b> . . . . .	56
4.3.1	<i>Algoritmo K-Médias</i> . . . . .	56
4.4	<b>Conjunto de Dados</b> . . . . .	57
4.5	<b>Ocorrência de Amostras Discrepantes (<i>Outliers</i>)</b> . . . . .	58
4.5.1	<i>Modelo TSK Robusto</i> . . . . .	61
4.6	<b>Índices de Desempenho</b> . . . . .	62
4.6.1	<i>Sum of Squared Errors (SSE)</i> . . . . .	62
4.6.2	<i>Coefficiente de Determinação</i> . . . . .	62
4.7	<b>Softwares de Simulação</b> . . . . .	62
4.8	<b>Conclusão do capítulo</b> . . . . .	63
5	<b>RESULTADOS E DISCUSSÃO</b> . . . . .	64
5.1	<b>Resultados do Modelo Polinomial</b> . . . . .	64
5.2	<b>Resultados da Rede MLP</b> . . . . .	64
5.3	<b>Resultados do Modelo ANFIS</b> . . . . .	65
5.4	<b>Resultados do Modelo TSK</b> . . . . .	66
5.4.1	<i>Uma abertura comum a todas as funções de pertinência</i> . . . . .	67
5.4.2	<i>Raios distintos por pertinência gaussiana</i> . . . . .	69
5.4.2.1	<i>Quatro e Cinco Funções de Pertinência</i> . . . . .	69
5.5	<b>Comparação entre os modelos apresentados</b> . . . . .	73

<b>6</b>	<b>CONCLUSÕES E TRABALHOS FUTUROS</b> . . . . .	<b>76</b>
<b>6.1</b>	<b>Conclusões</b> . . . . .	<b>76</b>
<b>6.2</b>	<b>Trabalhos Futuros</b> . . . . .	<b>77</b>
	<b>REFERÊNCIAS</b> . . . . .	<b>78</b>
	<b>APÊNDICES</b> . . . . .	<b>83</b>
	<b>APÊNDICE A – Distribuição de Probabilidade da Potência Elétrica Ge-</b> <b>rada</b> . . . . .	<b>83</b>
<b>A.1</b>	<b>Função de Densidade de Probabilidade (FDP) de <math>p</math></b> . . . . .	<b>83</b>

## 1 CONCEITOS GERAIS

Este capítulo se inicia com uma breve introdução a respeito da importância de gerar energia elétrica a partir de fontes renováveis. Dentre os principais tipos de produção de energia renovável, é enfatizada a de origem eólica, pois trata-se da base do nosso objeto de estudo. Partindo disso, é salientada a importância do estudo de modelagem, com foco em aerogeradores, bem como as principais dificuldades em sua aplicação, para que se tenha uma previsão adequada do fornecimento elétrico. E então, é apresentada, de forma sucinta, proposta da dissertação, a qual será abordada detalhadamente no decorrer deste texto dissertativo.

Além do supracitado, este capítulo também aborda a motivação da realização deste trabalho, quais os objetivos almejados e é anunciada a publicação, por hora, oriunda do estudo desenvolvido.

Por fim, é especificada como esta dissertação será organizada nos capítulos posteriores, sendo feita uma breve contextualização a respeito de cada um deles.

### 1.1 Introdução

Uma matriz energética que proporcione o mínimo possível de agressão ambiental tem sido de fundamental interesse para a maioria dos países desenvolvidos e em desenvolvimento no mundo. Assim, há uma forte tendência mundial para que a produção energética se volte a fontes renováveis, como a solar e eólica.

Com o crescimento da indústria eólica, aerogeradores estão sendo instalados em diversos tipos de ambientes distintos, que influenciam significativamente a produção energética local. Em face de tais dificuldades, faz-se necessário uma modelagem adequada da relação entre a velocidade do vento e a potência gerada (conhecida genericamente como *curva de potência*) pelos aerogeradores, com o intuito de prever corretamente sua produção energética, bem como monitorar o seu desempenho.

Dada sua importância, a modelagem (ou estimação) de curvas de potência é um tema que desperta o interesse da comunidade de energias renováveis, haja visto o grande número de artigos disponíveis na literatura especializada, principalmente na forma de surveys (MARCIOU-KAITIS *et al.*, 2017; WANG *et al.*, 2018; SOHONI; GUPTA; NEMA, 2016; LEE *et al.*, 2015; SHOKRZADEH; JOZANI; BIBEAU, 2014; LYDIA *et al.*, 2014; CLIFTON *et al.*, 2013; LYDIA *et al.*, 2013; LI *et al.*, 2001).

Dentre os métodos encontrados na literatura, destacam-se o modelo polinomial, o polinomial ponderado e redes neurais artificiais. Em particular, o modelo polinomial tem seu uso disseminado entre os praticantes da área devido não somente à sua formulação mais simples quando comparado a outros modelos, tais como redes neurais, mas principalmente porque encontra-se disponível em planilhas de cálculo populares (e.g. Excel<sup>TM</sup> e LibreOffice<sup>TM</sup> Calc).

## 1.2 Motivação

Historicamente, o ser humano tem como desafio descrever matematicamente sistemas experimentais, com objetivo de entendê-los e resolver problemas relacionados aos mesmos. Em meados dos anos 90, houve um aumento na necessidade de obtenção de modelos matemáticos a partir de dados observados. Isso aconteceu devido a disponibilidade de computadores acessíveis e com bom desempenho, junto ao fato do aumento da complexidade dos sistemas, o que dificultava a *modelagem baseada na física do processo* ou modelagem fenomenológica, que é a obtenção das equações básicas do sistema (AGUIRRE, 2015).

Modelos matemáticos podem assumir formas muito diferentes, dependendo do sistema em estudo, que pode variar entre sistemas social, econômico, ambiental, mecânico ou elétrico. Usualmente, os sistemas mecânicos e elétricos possuem experimentos que podem ser facilmente reproduzidos. O formato do modelo depende fortemente do objetivo final do procedimento de modelagem. Por exemplo, um modelo para projeto ou operação deve conter bem mais detalhes que um modelo para previsão a longo prazo (KEESMAN, 2011).

Geralmente, modelos são desenvolvidos para:

- Obter ou ampliar a visão em diferentes fenômenos, por exemplo, modelar leis físicas ou relações econômicas;
- Analisar o comportamento do processo, usando ferramentas de simulação, por exemplo treinamento de processos ou previsões meteorológicas;
- Controlar processos de uma planta ou robô;
- Estimar variáveis de estado que não podem ser facilmente mensuradas (KEESMAN, 2011).

Um modelo que representa um sistema pode ser determinado pela descrição de todas as suas características ou somente algumas delas, destinado a aplicações específicas (HSIA, 1977). Dentre as técnicas de modelagem de sistemas, é possível trabalhar com a Lógica *Fuzzy* (Nebulosa), que é a lógica que suporta métodos de raciocínio aproximados em vez de exatos,

baseada na teoria de conjuntos *fuzzy* e constitui a base para o desenvolvimento de técnicas e algoritmos de modelagem e controle de processos. Ela permite a redução na complexidade e implementação de projetos, tornando-se uma solução para problemas de controle que eram, até então, intratáveis por técnicas clássicas (GOMIDE; GUDWIN, 1994).

O modelo fuzzy Takagi Sugeno Kang (TSK) atraiu muita atenção da comunidade que estuda modelagem de sistemas *fuzzy*, devido a sua boa performance em várias aplicações (CHU-ANG; HSIAO; JENG, 2003), entretanto essa abordagem não leva em conta dados de treinamento com *outliers*. De acordo com Barnett e Lewis (1974), *outlier* é uma observação, ou conjunto de observações, que aparenta inconsistência com o restante do conjunto de dados. Os autores sugerem que para amostras que contenham *outliers* o modelo matemático ou probabilístico utilizado seria diferente se comparado com sistemas sem a presença de *outliers*.

A utilização de modelos diferentes implicará em diferentes valores de parâmetros a serem estimados, portanto os erros tendem a ser maiores ao se comparar estimações com amostras discrepantes e amostras comuns (BARROS, 2013). Destarte, surge o questionamento: "O que deve ser feito ao se estimar modelos quando há a presença de *outliers* nos dados analisados?"

Para responder a pergunta acima, temos:

- De acordo com Legendre (1805), quando existe a possibilidade de visualização dos dados, o que ocorre para modelos unidimensionais, pode ser feita uma inspeção visual afim de buscar as amostras com maiores erros e eliminá-las. Para modelos multidimensionais essa análise pode ser feita para cada saída ou entrada individualmente;
- Utilizar estimadores que podem lidar automaticamente com *outliers*, ponderando a influência das amostras discrepantes para diminuir a sua ação no resultado final de estimação do modelo (BARROS, 2013). Este tipo de estimação é conhecida por *estimação robusta* e compõe uma área da estatística robusta.

Apesar do conhecimento sobre a presença de *outliers* em dados não ser algo recente, ainda existe incerteza sobre como trabalhar na presença destes dados. Uma gama de autores argumenta que eles podem conter informações importantes, como por exemplo a presença de uma nova classe, e descartam que tais informações tenham sido obtidas por erros grosseiros. Em casos extremos, pode-se optar também por uma abordagem híbrida, onde mesmo utilizando estimadores robustos, resulta-se em erros que destacam-se dos demais, o que justificaria uma limpeza de dados (BARROS, 2013).

O estudo de dados com *outliers* é algo constante em problemas de *Modelagem de*



*Sistemas*. Portanto, o interesse desta dissertação reside na análise de dados com presença de amostras discrepantes, onde serão utilizadas técnicas de modelagem *fuzzy* TSK para sistemas na presença de *outliers*.

### 1.3 Justificativa

As energias renováveis são um grande passo para o desenvolvimento de um mundo mais sustentável. Historicamente, as atividades humanas ocasionam mudanças no meio ambiente (SILVA; LIMA, 2011). Tendo como consequência aquecimento global, chuvas ácidas, destruição da camada de ozônio, entre outros, ligaram o sinal de alerta das grandes nações. Vários países estabeleceram metas e criaram programas com objetivo de criar um ambiente mais favorável à utilização de fontes alternativas de geração de energia elétrica (RAMPINELLI; JUNIOR, 2013).

Os objetivos de desenvolvimento sustentável (ODS), da Organização das Nações Unidas são um conjunto de 17 metas globais estabelecidas pela Assembleia Geral das Nações Unidas. De especial interesse para esse trabalho, temos os ODS números 7 e 13, descritos a seguir:

ODS nº 7 - Assegurar o acesso confiável, sustentável, moderno e a preço acessível à energia para todos.

ODS nº 13 - Tomar medidas urgentes para combater a mudança do clima e seus impactos.

Estes objetivos promovem mecanismos para a criação de capacidades para o planejamento relacionados à mudança do clima, assim como estabelecem a busca por expansão, estrutura e modernização da tecnologia para o fornecimento de serviços de energia modernos e sustentáveis, pilares norteadores dessa dissertação.

Ao fim do ano de 2013, a capacidade eólica mundial atingiu 318 GW e a potência instalada anualmente tem sido, desde 2009, da ordem de 40 GW (COUNCIL, 2014). A matriz energética brasileira é composta por cerca de 85% de fontes providas de energia renovável, enquanto a média mundial é de cerca de 20% (BRASIL, 2013).

De acordo com a INTERNATIONAL ENERGY AGENCY- IEA (2015), em 2015 a energia eólica foi a grande responsável pelo crescimento na geração de fontes de energia por fontes renováveis no mundo. Em 2016, 487 GW de potência de eólica foi instalada em todo o mundo. Segundo a Empresa de Pesquisa Energética, a geração de energia eólica no Brasil atingiu 33,5 TWh, crescimento de 54,9% e a potência eólica atingiu 10.124 MW, expansão de

32,6%. No Boletim Anual de Geração da Associação Brasileira de Energia Eólica (ABEEólica), os estados do nordeste brasileiro aparecem como os maiores produtores de energia eólica do país. Na Tabela 1 são apresentados os 5 estados que possuem geração de energia eólica no ano de 2017.

Tabela 1 – Os 5 estados com maior geração de energia eólica no ano de 2017.

Posição	Estado	TWh
1º	Rio Grande do Norte	13,24
2º	Bahia	7,79
3º	Rio Grande do Sul	5,58
4º	Ceará	5,10
5º	Piauí	4,59

Fonte – (ABEEÓLICA, 2017)

Os estados do Piauí e Ceará estão entre os 5 maiores produtores de energia eólica no Brasil. A geração eólica no estado do Ceará saltará 44% nos próximos anos, o que ilustra entusiasmo com a região nordeste quando o tema é energia eólica (BRUNO CABRAL, 2017), implicando em maiores investimentos relacionados à pesquisa e inovação nesta área. Já quando se fala em Piauí, o estado apresenta potencial para instalação de duas mil torres de energia eólica, algo que representa cerca de 4 GW de potência instalada (CIDADE VERDE, 2017).

A natureza incerta do vento e a alta penetração de energia eólica nos sistemas de potência são um grande desafio para a confiabilidade e estabilidade desses sistemas. Para tornar a energia eólica uma fonte confiável, são necessários modelos precisos para prever a potência e monitorar o desempenho das turbinas eólicas (SOHONI; GUPTA; NEMA, 2016). A criação de modelos matemáticos que estimem a curva de aerogeradores têm diversas aplicações, tais como:

1. Avaliação e previsão de energia eólica;
2. Monitoramento *online* de curvas de potência;
3. Estimativa do fator de capacidade da turbina eólica;
4. Seleção de turbinas eólicas (SOHONI; GUPTA; NEMA, 2016).

Não-linearidades, incertezas paramétricas e a presença de *outliers* são desafios presentes no trabalho que gerou esta dissertação. Quando tais questões se tornam relevantes, sistemas *fuzzy* Takagi-Sugeno são muito úteis devido a sua capacidade de modelar, aproximadamente ou com exatidão, dinâmicas não-lineares sobre domínios compactos por meio da combinação de modelos lineares, localmente válidos (JOHANSEN; SHORTEN; MURRAY-SMITH, 2000; TAKAGI; SUGENO, 1985a).

## 1.4 Objetivos Geral e Específicos

O problema-alvo desta dissertação é o problema de regressão de curva de potência de turbinas eólicas, utilizando o modelo *fuzzy* de Takagi-Sugeno-Kang. Isto posto, o objetivo geral deste trabalho é propor e avaliar uma extensão do modelo *fuzzy* TSK na presença de *outliers*, utilizando estimador robusto.

Em face do objetivo geral apresentado no parágrafo anterior, objetivos específicos foram propostos ao longo do desenvolvimento da dissertação, de tal modo a permitir que o objetivo geral fosse cumprido. Tais objetivos específicos são listados a seguir.

- Implementação de técnicas de estimação de parâmetros utilizando o modelo neurofuzzy ANFIS.
- Desenvolver e avaliar técnicas de estimação-M para o modelo *fuzzy* proposto.
- Utilizar técnicas para determinar os melhores parâmetros do modelo *fuzzy*.
- Avaliar e comparar modelos usuais de estimação de curvas de potência de turbinas eólicas na presença de *outliers*.

## 1.5 Produção Científica

Parte dos resultados produzidos ao longo da pesquisa desenvolvida nesta dissertação foram publicados no seguinte artigo:

BRASIL, I. S., BARRETO, G. A. & SOUZA, L. G. M - Estimação da Curva de Potência de Turbinas Eólicas Usando Modelo de Takagi-Sugeno-Kang, Anais do V Congresso Brasileiro de Sistemas Fuzzy (CBSF 2018), 4 a 6 de Julho de 2018, Fortaleza-CE, p. 447-458.

O artigo foi premiado como um dos 3 melhores artigos aceitos, em cerimônia realizada no dia 07/07/2018, durante o jantar de gala do evento.

## 1.6 Estrutura do Trabalho

A estrutura do trabalho utiliza como suporte principal as orientações do regulamento do Trabalho de Conclusão de Curso do Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal do Piauí. O trabalho é composto de seis capítulos:

- No capítulo 2, chamado “Fundamentos de Geração Eólica”, são feitas as fundamentações teóricas a respeito da modelagem da curva de potência de aerogeradores, bem como o estudo sobre as influências mecânicas do vento em aerogeradores de eixo horizontal, a

transformação da energia mecânica em elétrica e a análise das distribuições estatísticas pertinentes ao desenvolvimento de uma modelagem coerente.

- No capítulo 3 são apresentados outros modelos de regressão utilizados para descrever a geração da energia elétrica de aerogeradores a partir da velocidade do vento, bem como fundamentos acerca de estimadores-M e *outliers*. O modelo neurofuzzy ANFIS também será apresentado, para fins de comparação com o modelo proposto nessa dissertação.
- O capítulo 4 explicará o modelo proposto, assim como as métricas utilizadas para comparar os modelos utilizados.
- O capítulo 5 apresenta os resultados de testes comparativos entre todos os modelos elencados neste trabalho.
- O capítulo 6 finaliza este texto dissertativo com as conclusões obtidas e trabalhos futuros.

## 2 FUNDAMENTOS DE GERAÇÃO EÓLICA

O objetivo principal proposto nesta dissertação é comparar modelos matemáticos que representem a curva de potência de um aerogerador. Para tal, este capítulo destina-se a fazer um levantamento teórico sobre os pontos de principal relevância para a compreensão da problemática que leve à conclusão do objetivo almejado.

### 2.1 Necessidade da Modelagem da Curva de Potência

A curva de potência reflete a resposta da potência elétrica gerada por um aerogerador em função da velocidade de vento local. Elas são fornecidas pelos fabricantes em forma tabular ou gráfica. No entanto, essas curvas são criadas sob condições padronizadas, de forma que podem não representar adequadamente as condições realistas do local em que o aerogerador está instalado. O desempenho do aerogerador nos parques eólicos também não é ideal devido ao seu desgaste e envelhecimento. Dessa forma, a derivação de uma função apropriada para descrever a forma real da curva de potência local é necessária em sistemas de energia eólica.

Modelos precisos dessas curvas são úteis para diversas aplicações em parques eólicos. Os principais são elencados a seguir (SOHONI; GUPTA; NEMA, 2016).

#### 2.1.1 Avaliação e Previsão da Energia Produzida

A avaliação dos recursos eólicos de uma região é feita para identificar áreas adequadas para a instalação de parques eólicos (MATHEW, 2006). Nesse processo, a estimativa da energia é feita usando os dados do vento disponíveis e curva de potência de aerogeradores. Prever a potência de saída do aerogerador em um local candidato também é necessário no dimensionamento e estudos de otimização de custos durante o estágio de projeto de um sistema energético eólico. A precisão na previsão de potência é importante, pois uma superestimação pode resultar em baixa confiabilidade e uma subestimação pode levar ao sobredimensionamento do sistema. Operadores de aerogeradores que comercializam energia diretamente para o mercado de eletricidade também especificam a potência de saída de seus aerogeradores com alta acurácia, de modo que eles sejam capazes de entregar a quantidade correta de energia comercializada (SCHLECHTINGEN; SANTOS; ACHICHE, 2013; MAYA, 2018).

### **2.1.2 Estimativa do Fator de Capacidade**

O fator de capacidade de um aerogerador é um indicador de sua eficiência definido pela razão entre a potência média de saída e a potência nominal de saída do gerador (JOWDER, 2009). Essa informação é essencial para estudos de dimensionamento e otimização de custos, definir qual o aerogerador ideal para uma determinada região e classificar locais em potencial para produção de energia eólica (JOWDER, 2009; POWELL, 1981). Modelos de curva de potência do aerogerador e a estimativa da velocidade de vento média local são usados para estimar o fator de capacidade de um aerogerador. Uma análise comparativa de quatro métodos de modelagem de curva de potência na estimativa do fator de capacidade de um aerogerador é apresentada em Chang *et al.* (2014).

### **2.1.3 Seleção de Aerogeradores**

A curva de potência pode ser usada para comparar genericamente os modelos e pode ajudar na escolha do aerogerador a partir das opções disponíveis. A seleção das características do aerogerador que combinam com o regime de ventos do local ajuda a otimizar a eficiência do sistema de energia eólica (USTUNTAS; SAHIN, 2008).

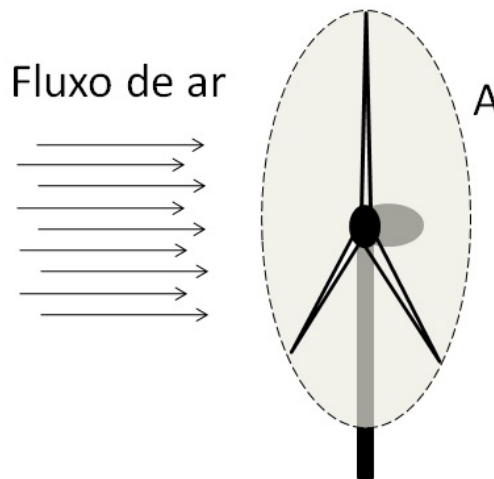
### **2.1.4 Monitoramento Online de Curvas de Potência**

Curvas de potência podem ser usadas para monitorar o desempenho de aerogeradores. Para tal, é necessária uma curva de referência que represente o desempenho esperado desse aerogerador de forma que se possa comparar com a curva real gerada *in loco*. Desvios dos valores reais da saída esperada podem indicar mau desempenho ou falhas (KUSIAK; ZHENG; SONG, 2009). Diferentes tipos de falhas afetam o sistema do aerogerador diferentemente, e causarão o desvio da curva de potência em relação valor esperado de maneiras diferentes. Ferramentas que podem caracterizar e quantificar esses desvios podem auxiliar na identificação precoce de falhas e a análise estatística de *outlier* podem dar indicações da razão específica da anomalia. Monitoramento das condições do aerogerador pelo uso da modelagem de cópula da curva de potência é sugerida em Stephen *et al.* (2011), Gill e Galloway (2012). O reconhecimento precoce das falhas emergentes no equipamento, bem como seu reparo e manutenção em tempo oportuno, podem ajudar significativamente na melhoria do desempenho de aerogeradores.

## 2.2 Fundamentos Teóricos da Modelagem de um Aerogerador de Eixo Horizontal

Seja um aerogerador de eixo horizontal qualquer. Denotaremos por  $A$  a área circular (em  $m^2$ ) varrida por suas pás, tal como esboçado na Figura 1. Considere que o ar do local onde o aerogerador está instalado tenha densidade  $\rho$  constante (em  $kg/m^3$ ) e que a componente de velocidade do vento (em  $m/s$ ) que atravessar perpendicularmente a área  $A$  é constante e igual a  $v$ .

Figura 1 – Fluxo de ar através da área circular de varredura das pás do aerogerador.



Fonte: (SOHONI; GUPTA; NEMA, 2016).

Segundo a mecânica clássica, sabe-se que a potência  $P$  (em Watts) pode ser calculada como

$$P = \frac{\Delta E_c}{\Delta t}, \quad (2.1)$$

em que  $t$  é a medida do tempo (em segundos) e  $E_c$  é a energia cinética do movimento (em Joules).

Ademais, sabe-se que a energia cinética é dada por

$$E_c = \frac{1}{2}mv^2, \quad (2.2)$$

em que  $m$  é a massa (em  $kg$ ). Com  $v$  constante, pode-se escrever a variação da energia cinética como

$$\Delta E_c = \frac{1}{2}\Delta mv^2. \quad (2.3)$$

Portanto, a potência eólica disponível associada a  $v$  em uma área  $A$ , denotada

doravante por  $P_w$ , equivale a energia cinética por unidade de tempo e pode ser obtida por

$$\begin{aligned}
 P_w &= \frac{\Delta E_c}{\Delta t} \\
 &= \frac{1}{2} \frac{\Delta m}{\Delta t} v^2 \\
 &= \frac{1}{2} (\rho A v) v^2 \\
 &= \frac{1}{2} \rho A v^3.
 \end{aligned} \tag{2.4}$$

Porém, apenas uma parte da potência eólica disponível será aproveitada para o movimento do rotor, sendo esta chamada de potência absorvida ( $P_a$ ).

O coeficiente de potência,  $C_p$ , definido por

$$C_p = \frac{P_a}{P_w}, \tag{2.5}$$

é uma grandeza adimensional que caracteriza a capacidade de um aerogerador absorver a potência eólica disponível (SHOKRZADEH; JOZANI; BIBEAU, 2014; CLIFTON *et al.*, 2013).

O valor do coeficiente de potência, que pode ser obtido a partir dos dados do fabricante, é influenciado por aspectos morfológicos e aerodinâmicos do aerogerador, como por exemplo o desenho de suas pás. O seu máximo teórico, conhecido como limite de Betz, é de  $16/27 \approx 0,5926$  (BETZ, 1920). Ademais, ainda deve-se contar com a eficiência do rotor,  $\eta$ , para se obter a potência elétrica gerada a partir do fluxo de vento, chegando assim à seguinte equação teórica:

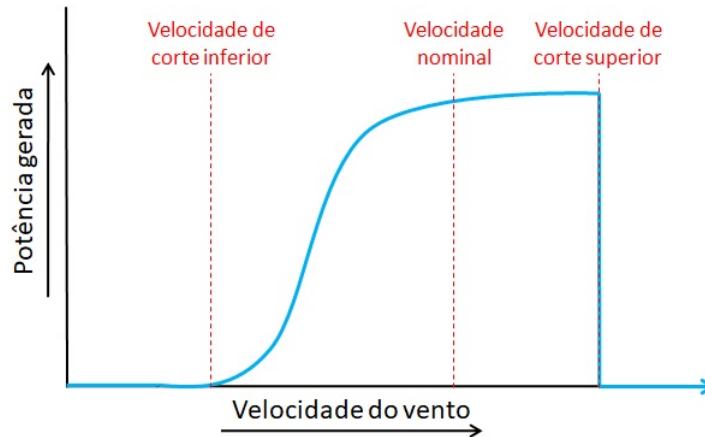
$$p = \frac{1}{2} \eta C_p \rho A v^3. \tag{2.6}$$

Como a potência elétrica gerada será a principal potência analisada neste trabalho, a denotaremos simplesmente por  $p$ , sem nenhum sufixo que a identifique como tal e em caixa baixa, já que ao utilizá-la para desenvolver equações na forma matricial, o que ocorrerá em capítulos à frente, isso torna mais claro que se trata de um valor escalar.

Aerogeradores reais são projetados para fornecer, no máximo, a sua potência nominal, podendo sofrer avarias caso a exceda. A potência elétrica fornecida tem comportamento não linear proporcional ao cubo da velocidade do vento para baixas velocidades, semelhante ao desenvolvimento teórico previamente apresentado. Porém, a medida que se aproxima da potência nominal, os aerogeradores comerciais dispõem de meios para limitar a energia eólica absorvida e, conseqüentemente, controlar a potência fornecida para que não exceda a potência nominal do rotor (MANWELL; MCGOWAN; ROGERS, 2009).



Figura 2 – Curva de potência típica de uma aerogerador com controle de potência por ângulo de passo.



Fonte:(SHOKRZADEH; JOZANI; BIBEAU, 2014).

Existem vários métodos que visam controlar a potência absorvida pelo rotor, dentre os quais podemos destacar dois como os principais utilizados atualmente:

- *controle do ângulo de passo*, que consiste no controle ativo do ângulo de ataque das pás do aerogerador. Dessa forma, pode-se alterar o valor do seu  $C_p$ , que influencia na potência eólica absorvida; e
- *controle por estol*, que é um tipo de controle passivo, que consiste em projetar as pás de forma que o efeito de estol<sup>1</sup> ocorra em velocidades superiores à velocidade eólica nominal.

Além disso, há uma velocidade de vento mínima para o funcionamento do aerogerador, conhecida como velocidade de corte inferior, ou velocidade de entrada (*cut-in speed*), na qual a potência gerada é zero para velocidades iguais ou inferiores. Bem como há um limite de funcionamento superior, conhecido como velocidade de corte superior (*cut-out speed*), em que a potência gerada também é zero para velocidades iguais ou superiores, pois o funcionamento do aerogerador é interrompido com a finalidade de proteger a integridade de seus componentes por conta da alta força do vento (SOHONI; GUPTA; NEMA, 2016). Um gráfico típico da curva de potência de um aerogerador comercial com controle do ângulo de passo está mostrado na Figura 2, em que fica evidenciado o comportamento não linear da curva de potência.

<sup>1</sup> Este efeito ocorre quando o fluxo de ar na superfície da pá deixa, abruptamente, de fixar-se à mesma, passando a girar de forma turbulenta, diminuindo assim a força de sustentação e aumentando a força de arrasto.

## 2.3 Problemas de Modelagem

Alguns aspectos requerem atenção durante a modelagem das curvas de potência dos aerogeradores. A seleção do modelo e a metodologia adotada depende do objetivo da modelagem, dos dados disponíveis, do impacto de vários parâmetros nessas curvas e de outras questões relacionadas. Os seguintes aspectos importantes devem ser considerados durante a modelagem da curva de potência (SOHONI; GUPTA; NEMA, 2016).

### 2.3.1 Diferença nos Modelos de Aerogeradores

As curvas de potência variam com diferentes fabricantes e modelos de aerogeradores. Portanto, o modelo de regressão utilizado para descrevê-los também deve se adequar as suas condições de funcionamento (DIAF *et al.*, 2008). Por exemplo, há diferença entre aerogeradores regulados por ângulo de passo e aerogeradores regulados por estol. Os controlados por ângulo de passo mantêm a produção constante da velocidade nominal até a velocidade de corte superior, enquanto o aerogerador regulado por estol possui uma redução da potência de saída acima da velocidade nominal do vento.

### 2.3.2 Comportamento nas Velocidades de Corte

O comportamento do aerogerador perto de velocidades de corte superior e inferior pode ser difícil de modelar (USTUNTAS; SAHIN, 2008). Estes limites são diferentes para diferentes modelos de aerogerador. Quando a curva de potência é derivada usando os dados medidos, alguns valores negativos e não-nulos de potências abaixo da velocidade de corte inferior podem ser obtidos. Além disso, a histerese de velocidade de corte superior que ocorre durante o período entre o desligamento e a reinicialização do aerogerador afeta a produtividade da mesma (HORVATH; PANZA; KARADZA, 2007). Efeitos da histerese podem ser mais significativos com certos padrões de vento e terrenos, como, por exemplo, ventos instáveis e tempestuosos, que exigem frequentes partida e parada, resultando em perda considerável de produção de energia. A correção da curva de potência que leva em conta este comportamento do aerogerador na velocidade de corte superior pode reduzir erros de previsão de potência.

### 2.3.3 *Aerogerador Único versus Grupo de Aerogeradores*

As curvas fornecidas pelos fabricantes são adequadas para prever a potência de saída de um único aerogerador de um tipo específico. Em um grande parque eólico, alguns aerogeradores estão espalhados por uma área ampla. A produção de energia eólica envolve incertezas devido à natureza estocástica do vento e da variação da curva de potência (JIN; TIAN, 2010). A velocidade e direção do vento encontrados pelos aerogeradores em um parque eólico podem não ser os mesmos devido à variação do vento. Consequentemente, em um parque eólico, a energia produzida por aerogeradores com especificações idênticas também pode diferir, mesmo se a velocidade do vento for a mesma. O efeito de sombreamento, também chamado de efeito esteira dos aerogeradores causa essa diferença, como os aerogeradores que operam em posição sequencial com outros aerogeradores, que podem obter velocidades de vento reduzidas (MCLEAN, 2008). Essa diferença também pode acontecer devido a fatores como desgaste, envelhecimento e sujeira ou depósito de gelo nas lâminas.

Com o crescimento dos projetos de energia eólica, tornou-se essencial o desenvolvimento de métodos para monitorar o desempenho não só de um único aerogerador, mas também o parque eólico como um todo. Portanto, desenvolver modelos para obter a relação entre velocidade do vento e potência de saída quando um grupo de aerogeradores é implantado em um parque eólico é necessário.

### 2.3.4 *Fatores de Influência*

Vários fatores podem causar o desvio da curva de potência do valor teórico (KUSIAK; ZHENG; SONG, 2009; SCHLECHTINGEN; SANTOS; ACHICHE, 2013). Os fatores de influência mais importantes são mostrados a seguir e precisam atenção durante a modelagem.

1. **Condições de vento no local.** O vento é altamente estocástico por natureza. A velocidade e direção do vento mudam continuamente. O vento em um determinado local é afetado por fenômenos climáticos e topologia da região. A turbulência do vento em um dado local afeta a produção de energia (HEDEVANG, 2014). Obstáculos como árvores, edifícios e outras estruturas altas influenciam o vento.
2. **Densidade do Ar.** A pressão, temperatura e umidade local afetam a densidade do ar, afetando, portanto, a potência produzida. O efeito da variação da densidade do ar foi considerado para o desenvolvimento de curvas específicas em Olaofe e Folly (2013), por

exemplo. É mostrado em Schlechtingen, Santos e Achiche (2013) que a temperatura tem a maior influência na densidade do ar e a consideração de seu efeito junto com a direção do vento resultou no melhor desempenho dos modelos.

3. **Extrapolação da Velocidade do Vento.** A velocidade do vento muda com a altura. Esse efeito de cisalhamento do vento é afetado pela rugosidade de terreno. A curva de potência usa a velocidade do vento medida à altura do cubo do aerogerador, mas essa altura varia com diferentes modelos e fabricantes, e nem sempre é possível medir a velocidade do vento a esta altura. Alguns métodos têm sido utilizados na literatura para expressar a variação de velocidade do vento com altura (DIAF *et al.*, 2008). Além disso, a velocidade do vento medida nos mastros é diferente da velocidade no local onde o aerogerador se encontra e, às vezes quando os valores de velocidade do vento neste local específico não estão disponíveis, as medições de velocidade do vento em uma localização próxima são usadas para a determinação do perfil do vento. A precisão da conversão da velocidade do vento medido para a velocidade do vento na altura do cubo e na localização do aerogerador depende de fatores como o perfil vertical do vento no local, posição dos mastros em relação ao aerogerador, e o método utilizado para extrapolção.
4. **Condição do Aerogerador.** A curva de potência é afetada pela condição do aerogerador e equipamento associado. Envelhecimento e desgaste do aerogerador, anomalias e falhas, condições das pás, configurações do controlador e assim por diante, fazem com que a curva de potência se desvie dos valores reais (KUSIAK; ZHENG; SONG, 2009; GILL; GALLOWAY, 2012).

## 2.4 Aquisição de Dados para Modelagem

Os dados necessários para modelar uma curva de potência são a velocidade do vento e a potência de saída. Esses dados devem ser medidos muitas vezes em intervalos periódicos de tempo. Uma vez que os dados necessários sejam coletados, a produção de energia do aerogerador pode ser analisada usando quatro abordagens distintas, ou seja, o uso direto da média dos dados de um curto intervalo de tempo, o método de caixas, o desenvolvimento da velocidade e curvas de potência a partir dos dados e a análise estatística usando medidas sumárias (MANWELL; MCGOWAN; ROGERS, 2009).

### 2.4.1 Análise Estatística da Velocidade do Vento

A distribuição de probabilidade da velocidade do vento descreve a possibilidade de ocorrência de alguns valores da velocidade do vento. As distribuições de probabilidade são geralmente caracterizadas pela função densidade de probabilidade  $f(v)$  ou pela função de densidade acumulada  $F(v)$  (MANWELL; MCGOWAN; ROGERS, 2009). As duas distribuições de probabilidade comumente usadas na análise de dados de vento são a distribuição de Rayleigh e Weibull. A distribuição de Rayleigh requer apenas o conhecimento da velocidade média do vento  $\bar{v}$  e, portanto, é a distribuição de probabilidade de velocidade mais simples. O  $f(v)$  e  $F(v)$  da distribuição de Rayleigh é dada a seguir:

$$f(v) = \frac{\pi}{2} \left( \frac{v}{\bar{v}^2} \right) \exp \left[ -\frac{\pi}{4} \left( \frac{v}{\bar{v}} \right)^2 \right] \quad (2.7)$$

$$F(v) = 1 - \exp \left[ -\frac{\pi}{4} \left( \frac{v}{\bar{v}} \right)^2 \right] \quad (2.8)$$

E as funções  $f(v)$  e  $F(v)$  da distribuição de Weibull são apresentadas a seguir

$$f(v) = \left( \frac{k}{c} \right) \left( \frac{v}{c} \right)^{k-1} \exp \left[ -\left( \frac{v}{c} \right)^k \right] \quad (2.9)$$

$$F(v) = 1 - \exp \left[ -\left( \frac{v}{c} \right)^k \right] \quad (2.10)$$

onde  $k$  é o fator de forma e  $c$  é o fator de escala. Quanto maior o valor de  $k$ , menor é a variação da velocidade do vento observada.

### 2.4.2 Levantamento da Curva de Potência

O IEC 61400-12-1 (IEC, 2005), é o padrão internacional comumente adotado para avaliar o desempenho de produção energética de aerogeradores. A norma descreve a metodologia de medição da curva de potência, que é determinada pela medição simultânea da velocidade do vento no local do teste e da potência eólica gerada. Uma calibração prévia do local é necessária para determinadas condições do terreno. A produção anual de energia é calculada aplicando à curva de potência as distribuições de frequência da velocidade do vento complementadas por

fontes de incerteza e seus efeitos. A norma prescreve a derivação da curva de potência usando a velocidade do vento na altura do cubo e sendo medida com um anemômetro de copo, mas se a velocidade do vento tiver uma grande variação sobre a área varrida pelo rotor, pode haver uma diferença significativa entre a velocidade do vento na altura do cubo e a velocidade do vento média em toda a área de varredura do rotor. Os métodos de medição e a precisão dos instrumentos de medição podem causar variações nas medições e podem levar a grandes erros de previsão.

O padrão IEC usa dados médios de dez minutos agrupados em intervalos de velocidade do vento de  $0,5 \text{ m/s}$  (método de caixas). O vento em um local específico pode ser afetado por vários fatores, como a topologia do local, os obstáculos e fenômenos climáticos. A curva de potência do IEC considera a condição de vento do local atual. Portanto, pode nem sempre ser apropriado aplicar-se às condições de vento de outros locais. Portanto, esforços de pesquisa são necessários para desenvolver curvas de energia específicas do local. Estas curvas podem incorporar as condições de vento do local específico, dando melhores resultados (OLAOFÉ; FOLLY, 2013; IEC, 2005).

A seleção apropriada do método de modelagem é um requisito importante para o planejamento e operação do sistema baseado em energia eólica. Os métodos que consideram apenas a velocidade do vento como entrada podem não levar em consideração a variação causada por vários parâmetros de influência. Métodos que consideram a influência desses parâmetros na curva de potência podem resultar em modelos mais precisos. É mostrado em Olaofé e Folly (2013) que o uso de curvas de potência de locais desenvolvidos, que usaram o conhecimento de parâmetros de aerogeradores para modelagem, resultou em uma avaliação de energia mais precisa do que a curva de potência do aerogerador. Os problemas discutidos acima, se abordados adequadamente, podem resultar em modelos eficientes de curvas de potência.

## 2.5 Conclusões

Neste capítulo foi apresentado inicialmente a importância da modelagem da curva de potência de um aerogerador. Isso é a base de estudo para prever e planejar a produção energética de parques eólicos, monitorar o bom funcionamento do aerogerador, dentre outros.

Na sequência foi feito um estudo do comportamento físico teórico a respeito da transformação da energia cinética eólica em energia elétrica através de um aerogerador de eixo horizontal. Foi considerada a mecânica clássica e as limitações físicas mensuráveis conhecidas

para que seja desenvolvido este esboço teórico. Com isso, chegamos a uma função que relaciona a velocidade do vento com a potência elétrica gerada. O gráfico dessa função é conhecido como curva de potência do aerogerador.

Em seguida foram apresentados os principais desafios para desenvolver um bom modelo para a curva de potência. Que podem ocasionar distorções significativas no formato da curva de potência a depender, por exemplo, de características ambientais ou estruturais do aerogerador.

Por fim, foram apresentadas as metodologias de aquisição dos dados que são utilizados para desenvolver a curva de potência de um aerogerador. Iniciou-se falando sobre a análise estatística da velocidade do vento e então findou na norma IEC 61400-12-1, comumente utilizada por produtores de aerogeradores para formar a curva de potência.

Aos passos desses conceitos preambulares sobre a natureza da problemática terem sido apresentados, o capítulo seguinte tratará de apresentar modelos comumente utilizados para regressão da curva de potência do aerogerador.

### 3 ESTIMAÇÃO DA CURVA DE POTÊNCIA VIA MODELOS POLINOMIAIS E NEURAIAS

Existe um grande número de métodos regressivos para estimação da curva de potência, cada qual com suas características que os tornam mais ou menos apropriados para uma determinada aplicação. Nesta seção são apresentados três modelos convencionais para estimação de curva de potência, começando pelo modelo polinomial clássico e o respectivo estimador de mínimos quadrados ordinários (*ordinary least squares*, OLS), passando por uma versão ponderada deste estimador, chegando ao modelo baseado em redes neurais artificiais. O intuito disso é de avaliá-los e compará-los com o novo modelo de regressão *fuzzy* TSK proposto nessa dissertação.

#### 3.1 Modelo de regressão polinomial

Regressão polinomial muito provavelmente é a técnica mais utilizada na prática para estimar a curva de potência de aerogeradores (LYDIA *et al.*, 2013; SHOKRZADEH; JOZANI; BIBEAU, 2014). Este modelo tenta ajustar um polinômio de ordem  $k$  a um conjunto de  $n$  observações do par entrada-saída  $\{(v_i, p_i)\}_{i=1}^n$ , sendo a  $p_i$  a  $i$ -ésima medição da potência elétrica gerada e  $v_i$  é a velocidade do vento correspondente.

Do ponto de vista da modelagem estocástica, o ajuste do polinômio aos dados assume que o processo gerador dos dados pode ser descrito pela seguinte expressão:

$$p_i = \beta_0 + \beta_1 v_i + \beta_2 v_i^2 + \dots + \beta_k v_i^k + \varepsilon_i, \quad (3.1)$$

em que  $\beta_j$ ,  $j = 0 \dots l$ , são os coeficientes do polinômio e  $\varepsilon_i$  é uma variável aleatória normal, com média nula e variância  $\sigma_\varepsilon^2$ .

Pode-se escrever a Eq. 3.1 na forma matricial como

$$\mathbf{p} = \mathbf{V}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \quad (3.2)$$

em que  $\mathbf{p} = [p_1 \ p_2 \ \dots \ p_n]^\top$ ,  $\boldsymbol{\beta} = [\beta_1 \ \beta_2 \ \dots \ \beta_k]^\top$ ,  $\boldsymbol{\varepsilon} = [\varepsilon_1 \ \varepsilon_2 \ \dots \ \varepsilon_n]^\top$ , com a matriz  $\mathbf{V}$ , de dimensão  $n \times k$ , sendo definida como

$$\mathbf{V} = \begin{bmatrix} 1 & v_1 & \dots & v_1^{k-1} \\ 1 & v_2 & \dots & v_2^{k-1} \\ \vdots & \vdots & & \vdots \\ 1 & v_n & \dots & v_n^{k-1} \end{bmatrix}.$$



A estimativa de  $\beta$  é obtida via método *Ordinary Least Squares* (OLS), que consistem em minimizar o funcional  $J(\beta) = \|\varepsilon\|^2 = \varepsilon^\top \varepsilon = \sum_{i=1}^n \varepsilon_i^2$ , cuja derivada em relação a  $\beta$  é dada por

$$\begin{aligned} \frac{\partial J(\beta)}{\partial \beta} &= \frac{\partial [(\mathbf{p} - \mathbf{V}\beta)^\top (\mathbf{p} - \mathbf{V}\beta)]}{\partial \beta} \\ &= -2\mathbf{V}^\top (\mathbf{p} - \mathbf{V}\beta). \end{aligned} \quad (3.3)$$

E a segunda derivada em relação a  $\beta$  é dada por

$$\begin{aligned} \frac{\partial^2 J(\beta)}{\partial \beta^2} &= \frac{\partial [-2\mathbf{V}^\top (\mathbf{p} - \mathbf{V}\beta)]}{\partial \beta} \\ &= 2\mathbf{V}^\top \mathbf{V}. \end{aligned} \quad (3.4)$$

Perceba que  $v_i \geq 0, \forall i \in \{1, \dots, n\}$ , pois a velocidade de vento mensurada sempre é positiva. Então podemos assumir que  $\mathbf{V}^\top \mathbf{V} \succeq 0$ . Consequentemente, podemos afirmar que  $J(\beta)$  tem a concavidade voltada para cima, portanto há um valor de mínimo global para a função. Igualando a Eq. 3.3 ao vetor de zeros de dimensão  $k$  chega-se à equação do estimador OLS para os parâmetros do modelo polinomial:

$$\begin{aligned} -2\mathbf{V}^\top (\mathbf{p} - \mathbf{V}\hat{\beta}) &= \mathbf{0} \\ \therefore \mathbf{V}^\top (\mathbf{p} - \mathbf{V}\hat{\beta}) &= \mathbf{0} \\ \therefore \mathbf{V}^\top \mathbf{V}\hat{\beta} &= \mathbf{V}^\top \mathbf{p} \\ \therefore \hat{\beta} &= (\mathbf{V}^\top \mathbf{V})^{-1} \mathbf{V}^\top \mathbf{p}. \end{aligned} \quad (3.5)$$

em que  $\hat{\beta}$  é uma estimaco de  $\beta$  e  $\mathbf{0}$  é um vetor de zeros com dimenso  $k$ .

A matriz  $\mathbf{V}$  é uma matriz de Vandermonde<sup>1</sup>. Isto faz com que uma escolha inadequada da ordem do polinmio (e.g.  $k > 8$ ) produza valores elevados para as algumas de suas colunas. Por exemplo, com valores tpicos de velocidade do vento na faixa de 4 a 12 m/s, este fato conduz em geral a erros numricos durante a inverso existente na Eq. (3.5).

### 3.2 Modelo polinomial ponderado

Os dados para estimaco de curva de potncia costumam ser obtidos via sistemas SCADA<sup>2</sup>, estando assim susceptveis a erros de medio, falha em sensores e erros nos sistemas de comunicao (SOHONI; GUPTA; NEMA, 2016; SHOKRZADEH; JOZANI; BIBEAU,

<sup>1</sup> Matriz em que os termos de cada linha esto em progresso geomtrica.

<sup>2</sup> *Supervisory control and data acquisition.*

2014). Os dados também são afetados por turbinas que não estão em produção pois foram desligadas pelo sistema de controle por outra razão que não a operação anômala. Assim, sistemas SCADA podem ter entradas nulas ou dados errôneos, resultando em modelos inexatos da curva de potência.

Pode-se optar por uma inspeção visual dos dados e a remoção manual de observações supostamente anômalas antes de realizar a estimação de parâmetros do modelo polinomial. Contudo, a remoção manual pode levar a modelos inexatos uma vez que o volume de dados oriundos de sistemas SCADA é muito grande, tornando difícil a diferenciação entre observações corretas e anômalas.

Uma alternativa é usar alguma técnica de estimação de parâmetros que seja inerentemente robusta a *outliers*, tal como *estimação-M* (HUBER, 1964a; BARRETO; BARROS, 2015). Grosso modo, esta técnica consiste em reescrever o funcional do método OLS em uma forma mais geral dada por  $\sum_{i=1}^n w_i^2(\varepsilon_i)\varepsilon_i^2$ , em que  $w_i(\varepsilon)$  é o peso associado ao erro  $\varepsilon_i$ . Note, porém, que os pesos dependem dos erros (ou resíduos), os resíduos dependem dos coeficientes estimados, e os coeficientes estimados dependem dos pesos. Esta dependência faz com que um método iterativo seja necessário para estimar os parâmetros, sendo o *iteratively reweighted least-squares*<sup>3</sup> (IRLS)(FOX, 1997) o mais usado para este fim, cujos passos estão descritos a seguir.

### Algoritmo IRLS

- 1º Passo - Fornecer uma estimativa inicial  $\hat{\beta}_0$  usando os mínimos quadrados, como descrito na Eq. 3.5;
- 2º Passo - A cada iteração  $t$ , calcular os resíduos das iterações anteriores  $\varepsilon_i(t-1)$  e seus pesos correspondentes  $w_i(t-1) = w[\varepsilon_i(t-1)]$ ;
- 3º Passo - Obter uma nova estimativa de mínimos quadrados ponderados para  $\beta(t)$ :

$$\hat{\beta}(t) = [\mathbf{V}^T \mathbf{W}(t-1) \mathbf{V}]^{-1} \mathbf{V}^T \mathbf{W}(t-1) \mathbf{p}, \quad (3.6)$$

em que  $\mathbf{W}(t-1) = \text{diag}\{w_i(t-1)\}$  é a matriz  $n \times n$  de pesos da iteração atual.

- Repetir o 2º e 3º passo até a convergência do vetor  $\hat{\beta}_i(t)$ .

<sup>3</sup> Este método é usado na função `robustfit` implementada no Matlab®.

### 3.2.1 Estimadores-M

Estimação-M representa um conhecimento teórico que cobre uma classe muito flexível de estimadores que exercem papel importante no desenvolvimento da estatística robusta e na construção de procedimentos de estimação robustos (HUBER, 1964a; BARROS, 2013).

Estimadores-M são conhecidos por tal nome, devido serem do tipo de Máxima Verossimilhança. Tal conceito foi introduzido por Huber (1964a), e sugere, que para a obtenção de estimativas mais robustas dos parâmetros, a minimização de outra função de erro que não seja a soma de seus quadrados seja utilizada.

O princípio dessa técnica consiste em atribuir um componente multiplicativo de peso a cada dado amostral em função do seu residual. Esse peso tem como finalidade minimizar a influência de *outliers* tanto quanto possível (RUCKSTUHL, 2016). Portanto, considere a seguinte estimação, conhecida como Weighted Least Squares (WLS), que atribui um peso indexado para cada dado amostral

$$\begin{aligned}
 J(\beta) &= \sum_{i=1}^n w_i \left( p_i - \sum_{j=1}^k \beta_j v_i^{j-1} \right)^2 \\
 &= \sum_{i=1}^n w_i (p_i - \mathbf{v}_i^T \beta)^2 \\
 &= \sum_{i=1}^n w_i \varepsilon_i^2,
 \end{aligned} \tag{3.7}$$

em que  $\mathbf{v}_i^T = [1 \quad v_i \quad v_i^2 \quad \dots \quad v_i^k]$  e  $w_i$  é o componente de ponderação para o dado amostral de índice  $i$ . A Eq. 3.7 pode ser escrita na forma matricial como

$$J(\beta) = (\mathbf{p} - \mathbf{V}\beta)^T \mathbf{W} (\mathbf{p} - \mathbf{V}\beta), \tag{3.8}$$

em que  $\mathbf{W}$  é uma matriz diagonal  $n \times n$  com  $\mathbf{W}_{ii} = w_i$ . O mínimo da Eq. 3.8, caso exista, pode ser obtido através da sua primeira derivada em relação a  $\beta$

$$\begin{aligned}
 \frac{\partial J(\beta)}{\partial \beta} &= -2\mathbf{V}^T \mathbf{W} (\mathbf{p} - \mathbf{V}\hat{\beta}) = \mathbf{0} \\
 \therefore \mathbf{V}^T \mathbf{W} (\mathbf{p} - \mathbf{V}\hat{\beta}) &= \mathbf{0} \\
 \therefore \mathbf{V}^T \mathbf{W} \mathbf{V} \hat{\beta} &= \mathbf{V}^T \mathbf{W} \mathbf{p} \\
 \therefore \hat{\beta} &= (\mathbf{V}^T \mathbf{W} \mathbf{V})^{-1} \mathbf{V}^T \mathbf{W} \mathbf{p},
 \end{aligned} \tag{3.9}$$

A segunda derivada em relação a  $\beta$ , dada por

$$\frac{\partial^2 J(\beta)}{\partial \beta^2} = 2\mathbf{V}^T \mathbf{W} \mathbf{V}, \tag{3.10}$$

pode nos informar a respeito da concavidade da função custo. Como visto na seção anterior,  $v_i \geq 0 \forall i \in \{1, \dots, n\}$ , portanto a concavidade da Eq. 3.8 depende do valor de  $w_i$ .

Até então obtemos um método de otimização dos parâmetros  $\beta$ . Entretanto, além desses, faz-se necessário ajustar apropriadamente cada  $w_i$  de forma a mitigar a influência dos *outliers* como mencionado anteriormente. Para tal, é desenvolvida a problemática sob a ótica do estimador-M, que define-se como a solução que minimiza a função objetivo  $J(\beta)$  definida por (SUSANTI *et al.*, 2014; RUCKSTUHL, 2016)

$$J(\beta) = \sum_{i=1}^n \Psi \left( \frac{p_i - v_i^T \beta}{\sigma_\varepsilon} \right), \quad (3.11)$$

em que  $\Psi(\cdot) : \mathbb{R} \rightarrow \mathbb{R}$  é uma função contínua, derivável em  $\mathbb{R}$  e com um mínimo global.

A normalização do residual por seu desvio padrão,  $\sigma_\varepsilon$ , é feita com o intuito de assegurar a invariância em relação à escala utilizada. A estimação do desvio padrão da população a partir de uma amostra pode ser feita de forma robusta através da *Median Absolute Deviation* (MAD) (ROUSSEEUW; CROUX, 1993)

$$\hat{\sigma} = \gamma \cdot MAD \quad (3.12)$$

em que  $\gamma$  é uma constante dependente da distribuição. Para o caso de distribuições normais, ao qual se enquadram as aplicações desenvolvidas neste trabalho, temos  $\gamma = 1,4826$ .

A derivada da Eq. 3.11 em relação a  $\beta$  é

$$\begin{aligned} \frac{\partial J(\beta)}{\partial \beta} &= \sum_{i=1}^n \left[ -\frac{1}{\sigma_\varepsilon} \psi \left( \frac{p_i - v_i^T \beta}{\sigma_\varepsilon} \right) v_i \right] \\ &\therefore \sum_{i=1}^n \left[ \psi \left( \frac{p_i - v_i^T \beta}{\sigma_\varepsilon} \right) v_i \right] \\ &\therefore \sum_{i=1}^n \left[ \psi \left( \frac{\varepsilon_i}{\sigma_\varepsilon} \right) v_i \right], \end{aligned} \quad (3.13)$$

em que  $\psi$  é derivada de  $\Psi$ . A finalidade da função  $\psi$  é de atribuir um peso em relação ao residual, de sorte que há várias funções que podem caracterizá-la. Para o trabalho desenvolvido, foi utilizada a função bponderada de Tukey, ou Bisquare, e a função de Huber, definidas juntas com o MQO na Tabela 2.

Existem alguns benefícios ao se trabalhar com estimadores-M, um deles é que a escolha da função objetivo é flexível, onde cada escolha leva a diferentes estimadores. Os estimadores-M mais comumente encontrados na literatura são Andrews, Bisquare, Cauchy, Fair, Huber, Logistic, MQO, Talwar e Welsh. As funções objetivo e peso dos estimadores de

Bisquare, Huber e MQO estão apresentadas na Tabela 2, em que  $c$  representa um limiar, em que  $u_i = \frac{\varepsilon_i}{\hat{\sigma}_e}$ , e  $c$  é um limiar, de tal forma que todo valor absoluto de  $u_i$  que for superior a  $c$  é considerado como *outlier*. O valor de  $c$  para os estimadores de Huber e Bisquare é chamado de constante de *tuning*, esse valor representa a resistência a *outliers*. Valores menores de  $c$  produzem maior resistência a *outliers*, mas ao custo de menor eficiência quando os erros são normalmente distribuídos.

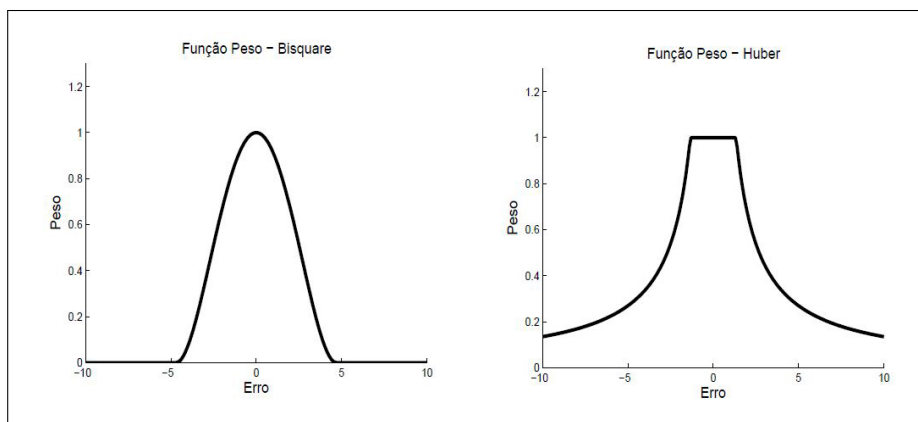
Tabela 2 – Função peso e objetivo para alguns dos estimadores-M encontrados na literatura.

Nome	Função Objetivo	Função Peso
Bisquare	$\begin{cases} \frac{c^2}{6} \left\{ 1 - \left[ 1 - \left( \frac{u_i}{c} \right)^2 \right]^3 \right\}, &  u_i  \leq c \\ \frac{c^2}{6}, &  u_i  > c \end{cases}$	$\begin{cases} \left[ 1 - \left( \frac{u_i}{c} \right)^2 \right]^2, &  u_i  \leq c \\ 0, &  u_i  > c \end{cases}$
Huber	$\begin{cases} \frac{1}{2} u_i^2, &  u_i  \leq c \\ c u_i  - \frac{1}{2} c^2, &  u_i  > c \end{cases}$	$\begin{cases} 1, &  u_i  \leq c \\ \frac{c}{ u_i }, &  u_i  > c \end{cases}$
MQO	$u_i^2$	1

Fonte – (CLEVELAND, 1979; BARROS, 2013; HUBER, 1964b)

Observa-se que o estimador MQO atribui valores de pesos iguais a cada erro, o que faz com que erros causados por *outliers*<sup>4</sup> tenham a mesma importância de erros causados por outras amostras. Para o estimador de Huber, os pesos caem quando  $|u_i| > c$  e para Bisquare os pesos caem quando  $u_i$  se afasta de 0 e são iguais a zero para  $|u_i| > c$  (FOX, 1997). Portanto, quando os *outliers* produzem erros maiores que outras amostras, como apresentado na Figura 3, os estimadores de Huber e Bisquare reduzem a influência desses pontos.

Figura 3 – Gráfico das funções peso para os estimadores Bisquare e de Huber.



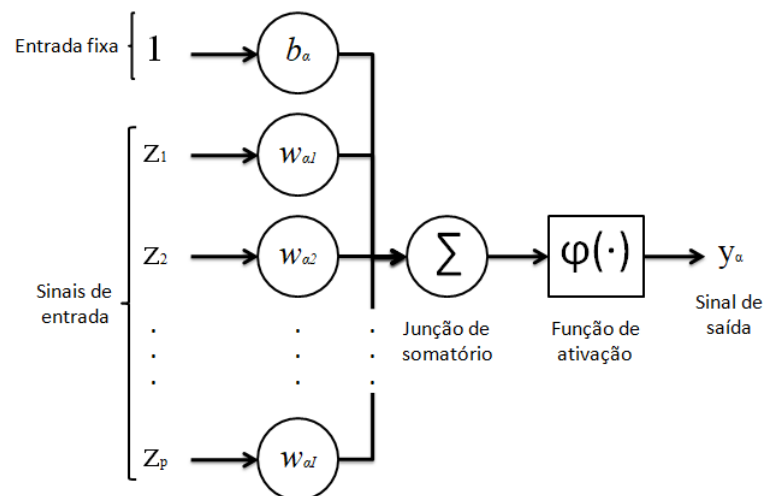
Fonte – (CLEVELAND, 1979; BARROS, 2013; HUBER, 1964b).

<sup>4</sup> O conceito acerca de *outliers* será mais aprofundado na seção 4.5.

### 3.3 Rede Perceptron Multicamadas (MLP)

O cérebro é um computador complexo, não linear e paralelo. Ele tem a capacidade de executar tarefas como reconhecimento de padrões, percepção e controle. Além dessas características, outras, como a capacidade de aprender, memorizar e ainda generalizar, estimularam pesquisas em modelagem algorítmica utilizando o paradigma de sistemas neurais biológicos (ENGELBRECHT, 2007).

Figura 4 – Representação em blocos de um neurônio artificial  $\alpha$ .



Fonte: (HAYKIN, 2007).

Um neurônio artificial, cuja representação em blocos pode ser vista na Figura 4, é um modelo de um neurônio biológico. Podemos identificar três elementos básicos nesse modelo (HAYKIN, 2007):

1. Um conjunto de *sinapses* no qual cada um de seus elementos é caracterizado por um peso  $w$ . Por exemplo, um sinal  $z_j$  na entrada da sinapse  $j$  conectado ao neurônio  $\alpha$  é multiplicado pelo peso sináptico  $w_{\alpha j}$ . O peso sináptico de um neurônio artificial pode estar em um intervalo que inclui valores negativos e positivos.
2. Um somador que faz a combinação linear dos dados de entrada após a sua respectiva ponderação sináptica. Aqui soma-se também o limiar, denotado por  $b_\alpha$ , que atua diminuindo ou aumentando o resultado final da combinação linear.
3. Uma função de ativação, denotado por  $\varphi(\cdot)$ , que, em geral, é usada para limitar a amplitude da saída de um neurônio.

Em termos matemáticos, sendo  $f : \mathbb{R}^D \rightarrow [0, 1]$  (ou  $f : \mathbb{R}^D \rightarrow [-1, 1]$ ) a função que

representa um neurônio artificial e  $D$  a dimensão dos dados de entrada, temos que

$$f(z) = \varphi(\mathbf{w}^\top \mathbf{z} + b_\alpha) \quad (3.14)$$

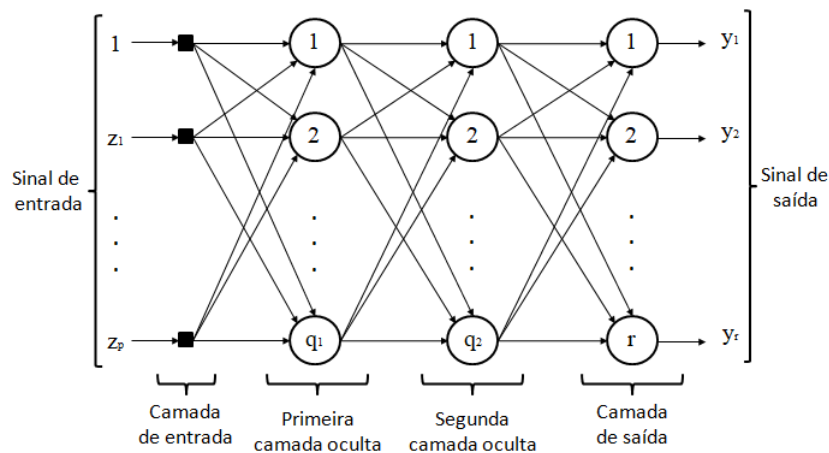
em que  $\mathbf{z} = [z_1 \ z_2 \ \dots \ z_D]^\top$  e  $\mathbf{w} = [w_1 \ w_2 \ \dots \ w_D]^\top$ . Alternativamente, podemos considerar  $b_\alpha = w_0$  e reescrever a Eq. 3.14 como

$$f(z) = \varphi(\mathbf{w}'^\top \mathbf{z}') \quad (3.15)$$

em que  $\mathbf{z}' = [1 \ z_1 \ z_2 \ \dots \ z_D]^\top$  e  $\mathbf{w}' = [b_\alpha \ w_1 \ w_2 \ \dots \ w_D]^\top$ .

Uma Rede Neural Artificial (RNA), cujo paradigma é inspirado no sistema nervoso biológico, é uma cadeia de neurônios artificiais conectados entre si. As entradas de cada neurônio da RNA ou são dados oriundos de informações sobre variáveis externas (e.g. temperatura, pressão, umidade, velocidade do vento etc) ou são a saída de outros neurônios da rede. E a saída também apresenta duas condições de finalidade: ou é o valor de entrada para outro(s) neurônio(s) ou é o valor de saída da própria RNA. O treinamento da RNA é feito através do ajuste dos pesos a fim de que os valores de saída sejam tão próximos quanto possível de um valor esperado. Uma das principais formas de fazer isso é através do método *Back-propagation* (HAYKIN, 2007). Dessa forma, as RNAs podem aprender a relação não-linear entre os conjuntos de dados de entrada e saída.

Figura 5 – Rede MLP com duas camadas ocultas.



Fonte:(HAYKIN, 2007).

A rede *Multilayer Perceptron* (MLP) é uma RNA disposta em camadas de neurônios em que a saída de cada um deles se conecta à entrada de todos neurônios da camada subsequente, caso exista. A MLP é dividida em camadas, das quais podemos distinguir como:

- Camada de entrada - é a interface de entrada de dados da MLP. Essa camada é caracterizada pela quantidade de dados simultâneos recebidos pela MLP, com exceção, caso exista, do 1 fixo para o limiar.
- Camada oculta - São todas as camadas que não estão conectadas diretamente com a saída da MLP. Esta camada é caracterizada pela quantidade de neurônios contidos.
- Camada de saída - é a camada de neurônios diretamente ligadas a saída. Assim como a camada oculta, essa camada é caracterizada pela quantidade de neurônios contidos, ou, se preferir, pela quantidade de saídas da MLP, já que a saída de cada neurônio dessa camada representa a saída da própria MLP.

A Figura 5 ilustra um exemplo da rede MLP em que temos  $D$  variáveis de entrada,  $q_1$  na primeira camada oculta,  $q_2$  na segunda e uma camada de saída com  $r$  neurônios.

As funções de ativação dos neurônios ocultos devem ser não-lineares (e.g. sigmoidais). Neurônios de saída podem usar tanto funções sigmoidais quanto lineares, a depender da aplicação de interesse. Em classificação de padrões, geralmente funções sigmoidais são usadas, enquanto funções lineares são usadas em problemas de regressão ou ajuste de curvas.

A rede MLP com uma camada oculta pode ser considerada como um aproximador universal de função (HAYKIN, 2007). Assim, desde que possua uma quantidade suficiente de neurônios ocultos, ela pode aproximar qualquer função contínua com uma certa precisão (SANTOS, 2010).

Considere uma MLP com apenas uma camada oculta e um único neurônio linear de saída (i.e.  $r = 1$ ). Sua saída é dada por

$$y_n = \mathbf{w}^T \mathbf{h}_n + b_r, \quad (3.16)$$

onde  $\mathbf{w} = [w_1 \cdots w_{q_1}]^T$  é o vetor de pesos que conecta os neurônios ocultos ao neurônio de saída e  $b_r$  é o limiar deste neurônio. O vetor  $\mathbf{h}_n \in \mathbb{R}^{q_1}$ , que contém as saídas atuais dos neurônios ocultos, é calculado como

$$\mathbf{h}_n = [\varphi(\mathbf{m}_1^T \mathbf{x}_n + b_1), \dots, \varphi(\mathbf{m}_{q_1}^T \mathbf{x}_n + b_{q_1})]^T, \quad (3.17)$$

tal que  $\mathbf{x}_n = [x_1 \cdots x_D]^T$  é o vetor de entrada atual,  $\mathbf{m}_j = [m_{j1} \cdots m_{jD}]^T$  é o vetor de pesos do  $j$ -ésimo neurônio oculto,  $b_j$  denota seu limiar, e  $\varphi(\cdot)$  é a função de ativação não-linear associada.

Visto que estamos lidando com um problema de uma única entrada (velocidade do vento,  $v_n$ ) e uma única saída (potência gerada,  $p_n$ ), é preciso treinar apenas uma rede MLP(1,  $q$ , 1), i.e. uma rede com uma única entrada (além do limiar) e um neurônio de saída. Neste caso,



$y_n = \hat{p}_n$  e  $\mathbf{z}_n = [-1 \ v_n]^T$ , com  $\hat{p}_n$  denotando a potência predita para a velocidade do vento atual  $v_n$ . Neste trabalho, os pesos e limiares da rede MLP aplicada ao problema de estimação da curva de potência de aerogeradores são computados via algoritmo de Levenberg-Marquardt (HAGAN; MENHAJ, 1994).

### 3.3.1 Levenberg-Marquardt

O método de treinamento Levenberg-Marquardt é uma alternativa ao treinamento da rede MLP (LEVENBERG, 1944; MARQUARDT, 1963). O algoritmo Levenberg-Marquardt (L.M.) é uma combinação do método de Gauss-Newton e gradiente descendente e tornou-se uma técnica padrão utilizada para problemas não-lineares.

A popularidade do algoritmo L.M. é tamanha devido a existência de diversas implementações disponíveis, mas o fato que a minimização não-linear produz uma matriz Jacobiana esparsa, reduz a complexidade do algoritmo (FRANÇA *et al.*, 2009).

Este método envolve a utilização da derivada segunda da função a ser minimizada, o que acelera a sua convergência. Dada a função  $\mathbf{V}(x)$  a ser minimizada em relação ao vetor de parâmetros  $x$ , o método de Newton seria

$$\Delta x = -[\nabla^2 \mathbf{V}(x)]^{-1} \nabla \mathbf{V}(x) \quad (3.18)$$

onde  $\nabla^2 \mathbf{V}(x)$  é a matriz hessiana e  $\nabla \mathbf{V}(x)$  é o gradiente. Ao assumir que  $\mathbf{V}(x)$  é a soma dos erros quadrados, tem-se

$$\mathbf{V}(x) = \sum_{i=1}^N \varepsilon_i^2(x) \quad (3.19)$$

então, pode ser mostrado que:

$$\begin{aligned} \nabla \mathbf{V}(x) &= \mathbf{J}^T(x) \boldsymbol{\varepsilon}(x) \\ \nabla^2 \mathbf{V}(x) &= \mathbf{J}^T(x) \mathbf{J}(x) + \mathbf{S}(x) \end{aligned} \quad (3.20)$$

onde  $\mathbf{J}(x)$  é a matriz jacobiana e  $\mathbf{S}(x) = \sum_{i=1}^N \varepsilon_i(x) \nabla^2 \varepsilon_i(x)$ . Para o método de Gauss-Newton, assume-se que  $\mathbf{S}(x) \approx 0$  e a Eq. 3.18 se torna:

$$\Delta x = [\mathbf{J}^T(x) \mathbf{J}(x)]^{-1} \mathbf{J}^T(x) \boldsymbol{\varepsilon}(x) \quad (3.21)$$

A modificação de Levenberg-Marquardt para o método Gauss-Newton é,

$$\Delta x = [\mathbf{J}^T(x) \mathbf{J}(x) + \mu \mathbf{I}]^{-1} \mathbf{J}^T(x) \boldsymbol{\varepsilon}(x) \quad (3.22)$$

onde é proposto uma solução adicionando à matriz hessiana a parcela  $\lambda \mathbf{I}$ , sendo  $\lambda$  um escalar, conhecido como constante de regularização e  $\mathbf{I}$  é a matriz identidade (HAGAN; MENHAJ, 1994; JAZAYERI; ROSEHART; WESTWICK, 2007).

### 3.4 Sistemas Neuro-Fuzzy

De acordo com Jang e Sun (1996), os sistemas Neuro-Fuzzy são aqueles que unem a adaptabilidade das redes neurais nas tarefas de reconhecimento de padrões com a capacidade dos sistemas de inferência fuzzy de incorporar conhecimento humano e realizar tomadas de decisões. Este capítulo apresenta o ANFIS (Adaptive Network-based Fuzzy Inference System), uma classe de redes adaptativas que funcionam de maneira equivalente a um sistema de inferência fuzzy.

#### 3.4.1 Adaptive-Network-Based Fuzzy Inference System (ANFIS)

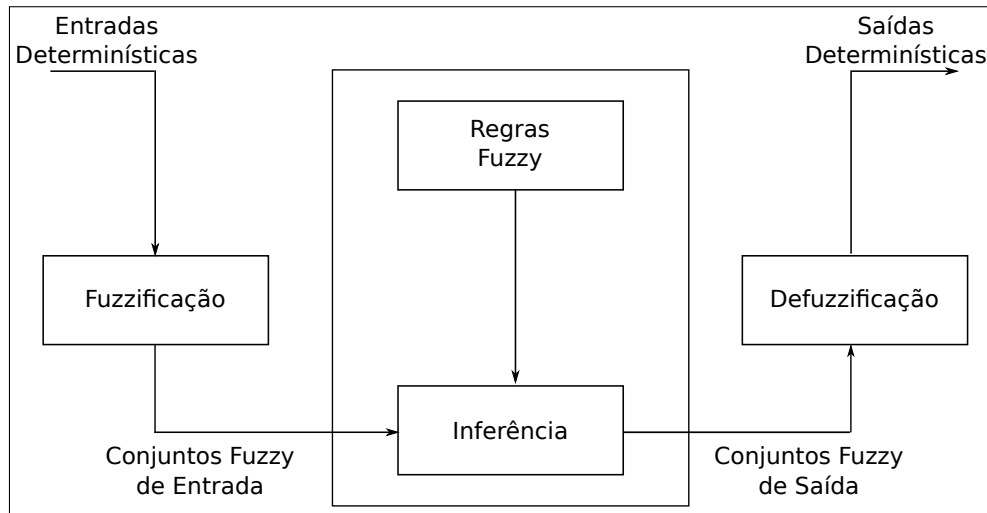
Muitos problemas computacionais podem ser bem caracterizados e resolvidos através de conjuntos que representam a pertinência de seus elementos de forma binária. Estas pertinências são tratadas de forma precisa e isenta de incertezas, ou seja, pode-se afirmar com exatidão se um elemento pertence ou não a determinado conjunto através de valores verdadeiros ou falsos. Por exemplo, pode-se considerar que um homem é alto quando sua altura é maior ou igual a 1,75 m. Neste caso, todos os homens que possuir uma altura maior ou igual a 1,75 m serão ditos altos, caso contrário serão ditos baixos. Tais conjuntos formam a base daquilo que se denomina *raciocínio exato*.

O raciocínio exato se demonstra ineficiente ao tentar caracterizar grandezas imprecisas ou vagas devido ao fato de nem sempre estas grandezas podem ser atribuídas a valores verdadeiros ou falsos. Além do mais, necessita-se tratar formalmente expressões linguísticas complexas que relacionem as grandezas envolvidas. Neste contexto, a *Lógica Fuzzy* e a *Teoria dos Conjuntos Fuzzy* fornecem todos os mecanismos formais para descrever estes tipos de variáveis. Todo raciocínio envolvido nestes casos é dito aproximado, devido a sua similaridade ao modo de raciocínio humano. Para o exemplo citado anteriormente, a Lógica Fuzzy definirá graus de pertinência aos elementos do conjunto de homens altos. Por exemplo, um homem com 1,70 m pertence ao conjunto dos homens altos com grau de pertinência 0,9.

Uma das aplicações da Lógica Fuzzy é a de se estimar saídas através de um conjunto de entradas fornecidas. Sistemas que realizam estes mapeamentos são denominados *Sistemas*

de *Inferência Fuzzy* e suas saídas podem ter diversas finalidades, desde auxílio na tomada de decisões, em problemas de classificação, dentre outros. A Figura 6 ilustra um sistema de inferência fuzzy que é constituído das seguintes partes:

Figura 6 – Componentes de um Sistema de Inferência Fuzzy (do inglês, FIS).



Fonte – (JANG, 1993).

- **Fuzzificador ou Processo de Fuzzificação:** realiza a transformação de uma variável não-fuzzy em variáveis fuzzy através da aplicação de funções de pertinência. Além do mais, este estágio também é responsável pela remoção de ruído que contamina as variáveis de entrada.
- **Regras Fuzzy:** são relações entre variáveis linguísticas que definem, em conjunto com a base de dados, o comportamento dinâmico do sistema de inferência fuzzy. Elas são baseadas no conhecimento e experiência de especialistas no domínio do problema. Estas regras possuem o formato:

$$\text{SE premissa(s) ENTÃO consequência(s)} \quad (3.23)$$

A parte das premissas normalmente envolve mais de uma variável linguística que são combinadas através de operadores lógicos. Estas variáveis podem vir acompanhadas de modificadores de vários tipos cuja função é a de intensificar determinado aspecto destas variáveis. Por exemplo, pode-se aplicar um modificador de contraste para aumentar a diferença de grau de pertinência entre os valores mais baixos e mais altos do espaço de discurso de um conjunto fuzzy. Engelbrecht (2007) traz maiores detalhes acerca de

modificadores linguísticos. Já a parte consequente das regras normalmente utilizam apenas uma variável linguística porém múltiplas variáveis podem estar presentes.

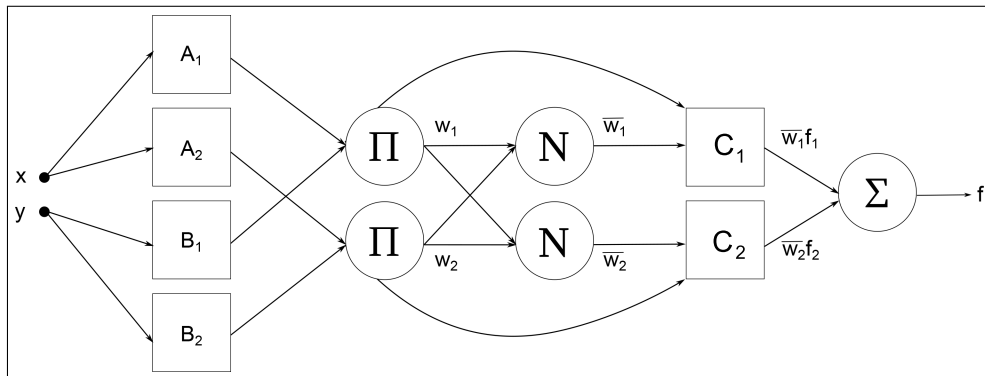
- **Banco de Dados:** define as funções de pertinência dos conjuntos fuzzy utilizados no conjunto de regras.
- **Mecanismo de Inferência:** processo pelo qual as entradas fuzzyficadas são utilizadas para gerar as saídas para cada regra. Esta saída é chamada de *força de disparo da regra*. Após o cálculo de todas as forças de disparo, estas são acumuladas gerando apenas uma saída fuzzy. Nesta fase é possível atribuir pesos que *a priori* quantificam o grau de confiança da regra. Esta análise de confiança também é feita por um especialista no domínio do problema.
- **Defuzzificador ou Processo de Defuzzificação:** realiza a operação inversa do estágio fuzzyficador, ou seja, transforma a saída fuzzy do FIS em um ponto pertencente a um conjunto ordinário (também conhecido como *crisp*). As referências Engelbrecht (2007) e Jang e Sun (1996) apresentam alguns dos tipos de defuzzificadores mais utilizados, suas vantagens e desvantagens.

Como exposto em Lee (1990), Sistemas de Inferência Fuzzy foram aplicados com sucesso em diversas áreas tais como controle, predição e inferência. Porém, existem algumas questões que são inerentes à aplicação destes sistemas. Uma delas é a inexistência de um método padronizado de se mapear o conhecimento e experiência de um especialista em um banco de dados de regras e seus conjuntos Fuzzy associados. Além do mais, faz-se necessário um método efetivo de ajustar os parâmetros das funções de pertinência de forma a minimizar o erro de mapeamento do Sistema de Inferência Fuzzy. Isto posto, Jang (1993) propôs uma arquitetura *feedforward* híbrida denominada ANFIS (*Adaptive-Neural-Based Fuzzy Inference System*), cujas características são endereçadas à solução dos entraves citados anteriormente.

Como uma forma de simplificar a descrição da arquitetura do ANFIS, serão utilizado como exemplo um sistema de duas entradas, denominadas  $x$  e  $y$ , e uma saída  $z$ . Além do mais, serão utilizadas regras fuzzy do tipo Takagi e Sugeno (TK) (TAKAGI; SUGENO, 1985b), onde as saídas de cada regra são determinadas por uma combinação linear das variáveis de entrada mais uma constante. A Figura 7 ilustra a arquitetura do ANFIS, onde os sistemas retangulares são sistemas adaptativos, ou seja, possuem parâmetros ajustáveis que determinam suas relações entrada-saída. As funções de cada nó na mesma camada são descritas em seguida:

- **Camada 1:** tem como função encontrar uma representação fuzzy dos valores de entrada

Figura 7 – Diagrama de blocos do ANFIS. Os blocos retangulares dispõem de parâmetros que são adaptados durante a fase de treinamento.



Fonte – (JANG, 1993).

não fuzzy. Isto é realizado através da aplicação da função de pertinência a um valor obtido do universo de discurso. Qualquer função que siga as restrições básicas de uma função de pertinência pode ser aplicada desde que esta seja diferenciável por partes. Em Jang (1993), funções com formato de sino foram utilizadas. À medida que o treinamento supervisionado se desenrola, os parâmetros são atualizados e, conseqüentemente, o formato das funções é alterado. Por exemplo, a função:

$$\mu(x) = \frac{1}{1 + \left[\left(\frac{x-c_i}{a_i}\right)^2\right]^{b_i}} \quad (3.24)$$

altera o seu formato à medida que  $a_i$ ,  $b_i$  e  $c_i$  são atualizados. Como tais parâmetros definem a interface entre o espaço de entrada não-fuzzy e o espaço fuzzy, ou seja, a parte *SE* de uma regra *SE-ENTÃO* fuzzy, ele são chamados de *parâmetros de premissa*, onde  $i$  denota uma regra fuzzy.

- **Camada 2:** realiza uma operação  $E$  lógica caso uma das regras fuzzy possua mais de uma variável linguística associada à parte da premissa de uma regra fuzzy. A operação de interseção de duas variáveis fuzzy  $A$  e  $B$  pode ser realizada através de operações matemáticas chamadas  $T$ -normas, porém a mais comumente utilizada é a multiplicação:

$$w_i = \mu_A(x) \times \mu_B(x) \quad (3.25)$$

Este bloco tem como saída o que se chama *força de disparo da regra*.

- **Camada 3:** esta camada realiza uma normalização com respeito ao somatório das forças de ativação de todas as regras:

$$\bar{w}_i = \frac{w_i}{\sum_{i=1}^N w_i} \quad (3.26)$$

onde  $N$  é o número de regras. Os novos valores produzidos são chamados de *forças de disparo normalizadas*.

- **Camada 4:** após o processo de normalização esta camada aplica a parte consequente da regra *SE-ENTÃO* fuzzy. Em (JANG, 1993) a parte consequente é do tipo Takagi-Sugeno (TS), ou seja, a saída de cada regra é uma combinação linear das variáveis de entrada mais um termo constante:

$$O_i = \bar{w}_i f_i = \bar{w}_i (p_i x + q_i y + \dots + r_i) \quad (3.27)$$

onde o conjunto  $p_i, q_i, r_i$  é chamado de *parâmetros consequentes*.

- **Camada 5:** camada de nó único que realiza a acumulação das saídas da camada anterior:

$$O = \sum_i^N \bar{w}_i f_i \quad (3.28)$$

A partir da descrição da arquitetura fica clara a necessidade de ajustar os parâmetros dos nós adaptativos. Para isso, um algoritmo de treinamento híbrido é empregado, onde a primeira fase é dita *forward* e a segunda é dita *backward*, de forma similar àquela apresentadas para as redes neurais MLP. Na primeira fase o fluxo de sinais se propaga da camada de entrada até a camada de saída, enquanto os parâmetros das premissas são mantidos fixos e os parâmetros da parte consequente são ajustados através do algoritmo *Least Mean Squares* (LMS). Na segunda fase, após a determinação dos parâmetros da parte consequente, o fluxo de sinais se propaga da camada de saída em direção a camada de entrada e os parâmetros da premissa são ajustados através de *error backpropagation*. Devido à fixação de parte dos parâmetros, o algoritmo converge mais rapidamente, pois reduz a dimensão do espaço de busca do algoritmo de retropropagação do erro (JANG; SUN, 1996).

### 3.5 Conclusão do Capítulo

A princípio, neste capítulo, foi feita uma breve revisão a respeito de alguns dos modelos comumente utilizados para regressão da curva de potência de aerogeradores. Iniciando pelo modelo polinomial clássico, em que seus parâmetros, são estimados pelo método dos

mínimos quadrados ordinários (*ordinary least squares*, OLS). Em seguida, foi apresentado o modelo baseado em redes neurais artificiais seguido pelo modelo neurofuzzy ANFIS.

No próximo capítulo será tratado especificamente acerca do modelo proposto para a execução deste trabalho, do conjunto de dados e da importância da análise de dados com *outliers*.

## 4 ESTIMAÇÃO DA CURVA DE POTÊNCIA VIA MODELO DE TAKAGI-SUGENO-KANG

Neste trabalho são reportados os resultados de um amplo estudo comparativo envolvendo modelos convencionais descritos no capítulo anterior com uma nova proposta baseada no modelo *fuzzy* Takagi-Sugeno-Kang (TSK). Sistemas baseados em regra *fuzzy* podem ser úteis para realizar tarefas de classificação, simulação de processos e diagnóstico, ferramentas de suporte a decisões online, identificação de sistemas e controle de processos.

### 4.1 Fundamentos de Modelagem Fuzzy

Sistemas *fuzzy* têm capacidade de processar informações semelhantes aos seres humanos. De acordo com Zimmermann (2010), as pessoas são capazes de lidar com processos bastante complexos, baseados em informações imprecisas e a estratégia de decisão adotada por elas, também é, de natureza imprecisa. Devido a isto, a tomada de decisão, por parte dos seres humanos, não se resume a um "sim" ou "não", a "verdadeiro" ou "falso", ou a "pertence" ou "não pertence", o que expressa que ela não pode ser expressa com exatidão numérica. A teoria de conjuntos *fuzzy* (ou nebulosos) foi proposta por Zadeh (1965) e possibilita o tratamento de informações imprecisas ou vagas a sistemas de computadores.

#### 4.1.1 Conjuntos Fuzzy

A lógica clássica está definida sobre a concepção de conjuntos. Portanto um valor pode pertencer ou não a um determinado conjunto. Dado conjunto  $A$ , definido em um universo  $X$ , os elementos deste universo pertencem ou não a  $A$ . Esta relação pode ser mapeada pela função característica  $\mu_A(x)$ :

$$\mu_A(x) = \begin{cases} 1, & \text{se } x \in A, \\ 0, & \text{se } x \notin A. \end{cases} \quad (4.1)$$

Para a lógica *fuzzy*, qualquer valor pode pertencer a esse conjunto, de modo que a pertinência pode assumir um número infinito de valores no intervalo  $[0, 1]$ . Essa função de pertinência é representada por:

$$\mu_A(x) : X \rightarrow [0; 1]; x \in X, \quad (4.2)$$

onde  $\mu_A(x)$  representa o grau com que o elemento  $x$ , pertencente ao universo de discurso  $X$ , inserido no conjunto *fuzzy*  $A$ .



A representação de um conjunto *fuzzy* depende da natureza e da dimensão do universo de discurso em que os elementos pertencem. Se o universo de discurso for discreto e finito, o seu conjunto *fuzzy* pode ser representado por um vetor contendo os graus de pertinência de  $X$ , representada pela notação abaixo (MENDEL, 1995):

$$A = \sum_{i=1}^N \mu_A(x_i)/x_i, \quad (4.3)$$

em que  $\mu_A(x_i)/x_i$  é o grau de pertinência do elemento  $x_i$  em relação ao conjunto *fuzzy* A. Já quando o universo de discurso for contínuo, substitui-se o  $\sum$  da Eq.( 4.3) por  $\int_{i=1}^N$ .

#### 4.1.2 Funções de Pertinência

Um conjunto *fuzzy* pode ter funções de pertinência de diversos tipos. A escolha do tipo ideal da função de pertinência é uma tarefa árdua, sendo necessário a presença de um especialista para determinar o melhor formato para cada aplicação. Os formatos das funções de pertinência são divididos em triangular, trapezoidal e gaussiana e estão ilustradas abaixo na Figura 8

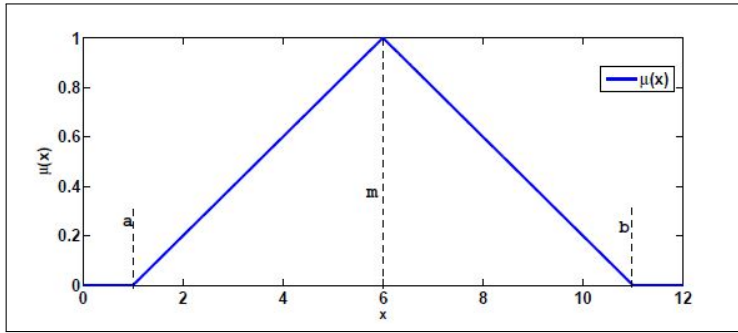
#### 4.1.3 Operações com Conjuntos Fuzzy

Os principais operadores utilizados em conjuntos *fuzzy* são a união, intersecção e complemento. Os operadores *T-norma* e *S-norma* representam, respectivamente, as operações de intersecção e união. O operador *T-norma* é representado por  $\mu_{A \cap B}(x_i) = \min[\mu_A(x_i), \mu_B(x_i)]$  e deve satisfazer as propriedades abaixo:

1. Contorno:  $T(0, 0) = 0$  e  $T(\mu_A(x), 1) = \mu_A(x)$ ;
  2. Comutativa:  $T(\mu_A(x), \mu_B(x)) = T(\mu_B(x), \mu_A(x))$ ;
  3. Associativa:  $T(\mu_A(x), T(\mu_B(x), \mu_C(x))) = T(T(\mu_A(x), \mu_B(x)), \mu_C(x))$ ;
  4. Monotonicidade:  $T(\mu_A(x), \mu_B(x)) \leq T(\mu_C(x), \mu_D(x))$  se  $\mu_A(x) \leq \mu_C(x)$  e  $\mu_B(x) \leq \mu_D(x)$
- (ZIMMERMANN, 2010; BRANCO, 2009)

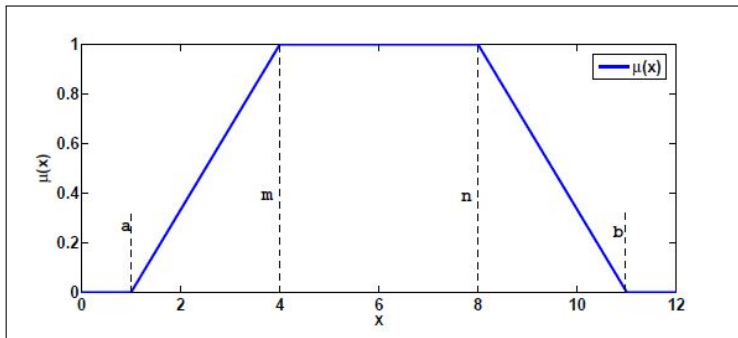
O principal operador de união é o máximo, representado por  $\mu_{A \cup B}(x_i) = \max[\mu_A(x_i), \mu_B(x_i)]$ . No entanto, existem outros operadores que também são *S-normas* e elas devem obedecer às propriedades a seguir:

1. Contorno:  $S(1, 1) = 1$  e  $S(\mu_A(x), 0) = \mu_A(x)$ ;
2. Comutativa:  $S(\mu_A(x), \mu_B(x)) = S(\mu_B(x), \mu_A(x))$ ;
3. Associativa:  $S(\mu_A(x), S(\mu_B(x), \mu_C(x))) = S(S(\mu_A(x), \mu_B(x)), \mu_C(x))$ ;

Figura 8 – Tipos de função de pertinência *fuzzy*

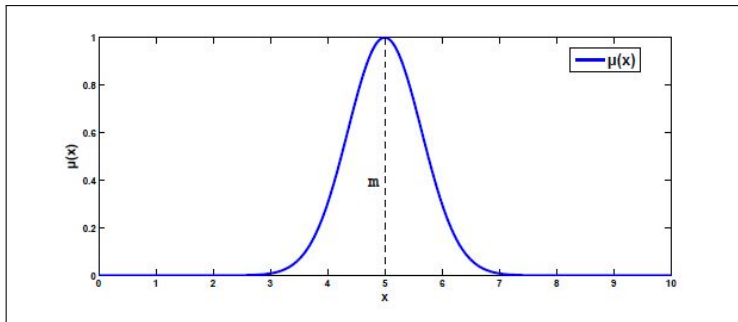
(a)

$$\mu(x) = \begin{cases} 0, & \text{se } x \leq a, \\ \frac{x-a}{m-a}, & \text{se } x \in [a, m] \\ \frac{b-x}{b-m}, & \text{se } x \in [m, b] \\ 0, & \text{se } x \geq b \end{cases}$$



(b)

$$\mu(x) = \begin{cases} 0, & \text{se } x \leq a, \\ \frac{x-a}{m-a}, & \text{se } x \in [a, m] \\ 1, & \text{se } x \in [m, n] \\ \frac{b-x}{b-n}, & \text{se } x \in [n, b] \\ 0, & \text{se } x \geq b \end{cases}$$



(c)

$$\mu(x) = \exp^{-\sigma(x-m)^2}$$

Fonte – Branco (2009) - Adaptado

4. Monotonicidade:  $S(\mu_A(x), \mu_B(x)) \leq S(\mu_C(x), \mu_D(x))$  se  $\mu_A(x) \leq \mu_C(x)$  e  $\mu_B(x) \leq \mu_D(x)$

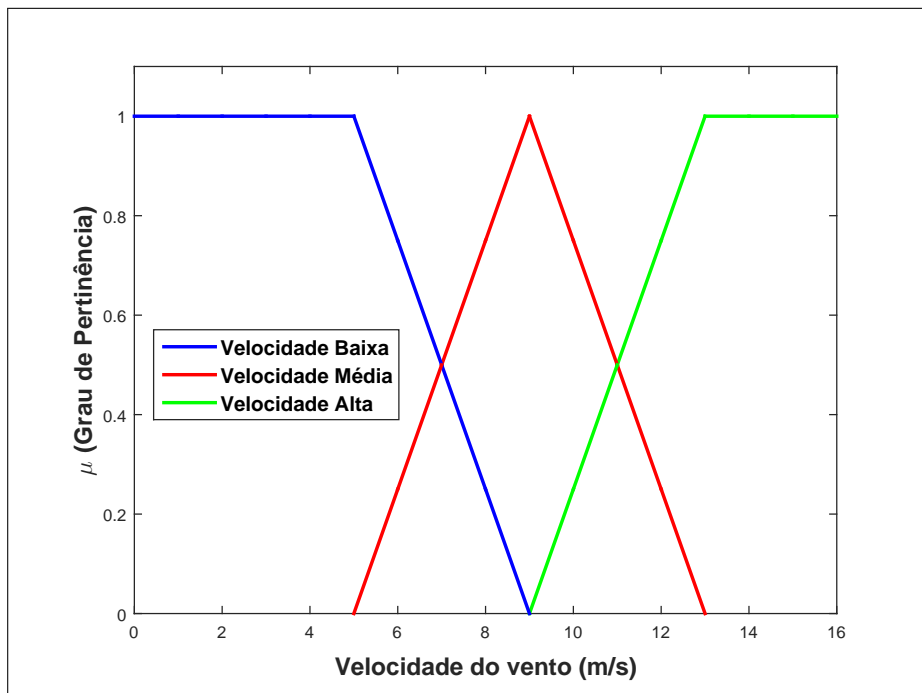
O complemento é outra operação básica sobre conjuntos *fuzzy*. Ele é representado por  $\bar{A}$  e é definido como  $\mu_{\bar{A}}(x) = 1 - \mu_A(x_i)$  (ZADEH, 1965).

#### 4.1.4 Variáveis Linguísticas

O trabalho com conjuntos *fuzzy* possibilita o tratamento de informações para fenômenos complexos ou mal definidos. Sistemas de inferência são compostos por variáveis linguísticas associadas através de conectivos para formar regras *fuzzy*. Elas implicam em variáveis de saída, que são combinadas, no processo de inferência, para fornecer uma única saída *fuzzy*.

Variáveis linguísticas são caracterizadas e construídas por termos linguísticos. Os valores das variáveis são formados a partir de termos próprios, tais como: "BAIXO", "ALTO" e "MÉDIO"; de conectivos lógicos "E", "OU" e "NÃO"; de modificadores "MUITO" e "POUCO" e de delimitadores. A Figura 9 exemplifica a variável linguística do vento e seus atributos (SANDRI; CORREA, 1999).

Figura 9 – Exemplo de variável linguística e seus atributos.



Fonte – Autor.

## 4.2 Modelo fuzzy Takagi-Sugeno-Kang (TSK)

O modelo *fuzzy* Takagi-Sugeno-Kang (TSK) consiste em um sistema de inferência, capaz de descrever sistemas dinâmicos não-lineares a partir de sistemas dinâmicos lineares (MOZELLI, 2008). Apesar de ter sido proposto originalmente por Takagi e Sugeno (1985a), o modelo também é conhecido por Takagi-Sugeno-Kang devido a trabalhos subsequentes de Sugeno e Kang (1988).

Em um sistema de inferência TSK, a  $l$ -ésima regra apresenta uma expressão funcional dada por

$$\text{Se } f_l \left( x_1 \text{ é } \mu_1^{(l)}, \dots, x_k \text{ é } \mu_k^{(l)} \right), \text{ então } y_l = \beta_0^{(l)} + \beta_1^{(l)} x_1 + \dots + \beta_k^{(l)} x_k \quad (4.4)$$

onde  $\{x_1, \dots, x_k\}$  são as variáveis de entrada,  $\{\mu_1^{(l)}, \dots, \mu_k^{(l)}\}$  são as funções de pertinência das

entradas para a  $l$ -ésima regra e  $(\beta_0^{(l)}, \dots, \beta_k^{(l)})$  são os parâmetros do modelo de regressão linear da  $l$ -ésima regra (TAKAGI; SUGENO, 1985a; SUGENO; KANG, 1988). A função  $f_l(\cdot)$  denota a função lógica que conecta as premissas da  $l$ -ésima regra:

$$f_l(x_1 \text{ é } \mu_1^{(l)}, \dots, x_k \text{ é } \mu_k^{(l)}) = \mu_1^{(l)}(x_1) \wedge \dots \wedge \mu_k^{(l)}(x_k), \quad (4.5)$$

em que  $\wedge$  denota a operação lógica de conjunção/interseção (AND). Neste trabalho, usaremos a função *mínimo* para o operador lógico AND, de tal forma que a ativação da  $l$ -ésima regra será dada pela seguinte expressão:

$$\alpha_l = \min(\mu_1^{(l)}(x_1), \mu_2^{(l)}(x_2), \dots, \mu_k^{(l)}(x_k)). \quad (4.6)$$

Em um sistema composto de  $L$  regras, a saída  $y$  agregada correspondente é obtida por

$$y = \frac{\sum_{l=1}^L \alpha_l (\beta_0^{(l)} + \beta_1^{(l)} x_1 + \dots + \beta_k^{(l)} x_k)}{\sum_{l=1}^L \alpha_l}. \quad (4.7)$$

#### 4.2.1 Modelo TSK para estimação de Curva de potência

Para o problema de estimação proposto nesta dissertação, tem-se apenas uma variável de entrada (velocidade do vento,  $v$ ) e uma única variável de saída (potência gerada,  $p$ ), o que simplifica o modelo TSK consideravelmente.

A primeira simplificação recai logo sobre a ativação da  $l$ -ésima regra, que passa a ser  $\alpha_l = \mu^{(l)}(v)$ , com a função de pertinência  $\mu^{(l)}$  sendo representada matematicamente por um dentre três possíveis subconjuntos fuzzy (baixa, média, alta). A segunda simplificação recai sobre as funções dos consequentes da  $l$ -ésima regra. Com uma variável apenas, a função de saída da  $l$ -ésima regra reduz-se a uma reta:  $p_l = \beta_0^{(l)} + \beta_1^{(l)} v$ , com  $p_l$  sendo a potência estimada pela  $l$ -ésima regra.

A potência final predita pelo modelo TSK para o problema de interesse é então calculada por:

$$y = \frac{\sum_{l=1}^L \alpha_l p_l}{\sum_{l=1}^L \alpha_l} = \frac{\sum_{l=1}^L \alpha_l (\beta_0^{(l)} + \beta_1^{(l)} v)}{\sum_{l=1}^L \alpha_l}. \quad (4.8)$$

Um total de  $2L$  parâmetros (dois por regra) deve ser estimado para o modelo TSK. Os coeficientes das  $L$  regras podem ser pelo método OLS, apresentado na Eq. (3.5). Neste caso, o modelo TSK será denotado por TSK-OLS.

### 4.3 Metodologia de Ajuste de Hiperparâmetros do Modelo TSK

Para a obtenção de resultados adequados no modelo proposto, é necessário que os hiperparâmetros do modelo TSK estejam especificados corretamente, em particular a abertura e alcance das funções de pertinência. Os seguintes cenários foram considerados com o objetivo de se obter melhores resultados. Eles foram testados para 3, 4 e 5 funções de pertinência gaussianas.

- Cenário 1 - Abertura comum a todas as funções gaussianas. Variou-se as aberturas da função gaussiana entre metade do desvio-padrão da velocidade  $v$  até o triplo deste valor.
- Cenário 2 - Uma abertura diferente para cada uma das funções gaussianas. Para esse caso, utilizou-se algumas heurísticas:

Heurística 1: Igual ao dobro da distância entre o centro da gaussiana atual e o centro da gaussiana vizinha mais distante;

Heurística 2: Igual à metade da distância entre o centro da gaussiana atual e o centro da gaussiana vizinha mais distante;

Heurística 3: Igual a 1/3 da distância entre o centro da gaussiana atual e o centro da gaussiana vizinha mais distante;

Heurística 4: Igual a 1/4 da distância entre o centro da gaussiana atual e o centro da gaussiana vizinha mais distante;

Heurística 5: Igual a 1/5 da distância entre o centro da gaussiana atual e o centro da gaussiana vizinha mais distante.

Além dos cenários supracitados, também foram testadas as mesmas heurísticas para funções de pertinência triangulares e trapezoidais. Em ambos os casos, a determinação dos centros das funções de pertinência foram obtidos por clusterização, via algoritmo  $k$ -médias ( $k$ -means).

#### 4.3.1 Algoritmo $K$ -Médias

Para um conjunto de velocidade do vento  $v_i^T = [1 \quad v_i \quad v_i^2 \quad \dots \quad v_i^k]$ , com  $v_i \in \mathbb{R}$ , o algoritmo de clusterização  $k$ -means em  $K$  subconjuntos distintos, chamados de *clusters*. O critério de clusterização mais utilizado é a soma das distâncias euclidianas quadradas entre cada ponto do conjunto de dados  $C_i$  e o centroide  $k_i$  (centro do *cluster*) de cada subconjunto de dados. Esse critério é chamado de erro de clusterização e depende de cada centroide.

O algoritmo  $k$ -means encontra soluções ótimas locais com relação ao erro de clusteri-

zação. É um algoritmo iterativo rápido que tem sido usado em muitos problemas de clusterização. Esse método é baseado em pontos aleatoriamente alocados como centros dos *clusters* e move-se em cada iteração, afim de minimizar o erro de clusterização. A principal desvantagem do método reside na sua sensibilidade às posições iniciais dos centros dos *clusters*. Portanto, para a obtenção de melhores resultados neste algoritmo, se faz necessário várias execuções com variações nas posições iniciais dos centros dos *clusters* (LIKAS; VLASSIS; VERBEEK, 2003).

#### 4.4 Conjunto de Dados

O conjunto de dados para validação da metodologia proposta foi utilizado primeiramente por Lee *et al.* (2015) e está disponível para uso público<sup>1</sup>. A Tabela 3 apresenta as especificações dos aerogeradores do parque eólico em estudo. São disponibilizados dados de velocidade do vento e potência elétrica de 6 estações meteorológicas (4 *in-land* e 2 *offshore*), denotadas WT1 a WT6. Para a elaboração desta dissertação, foram utilizados os dados referentes à estação WT1, apresentados abaixo na Figura 10

Tabela 3 – Especificações do parque eólico em estudo.

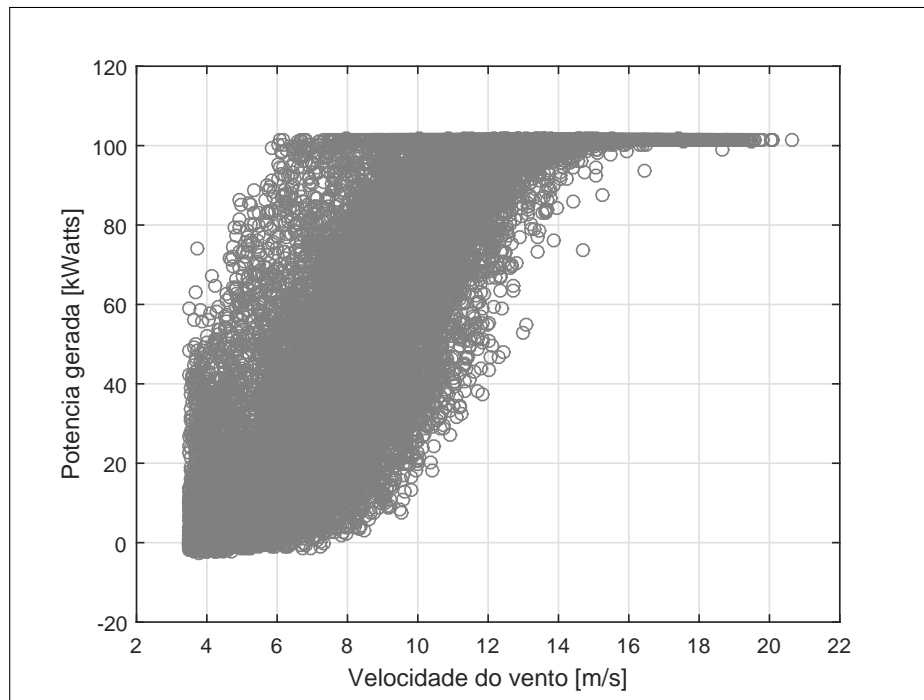
Parque eólico	
Altura ( <i>m</i> )	120
Diâmetro do rotor ( <i>m</i> )	80
Velocidade de corte inferior ( <i>m/s</i> )	3,5
Velocidade de corte superior ( <i>m/s</i> )	20
Velocidade nominal ( <i>m/s</i> )	13
Potência nominal ( <i>MW</i> )	1,5 – 2,0

Fonte – (LEE *et al.*, 2015).

Os dados ambientais da turbina WT1, dentre eles a velocidade do vento, são coletados por sensores em uma torre meteorológica, enquanto os dados de potência são medidos na própria turbina. Cada torre meteorológica está associada a 2 turbinas eólicas, o que significa que cada medida de velocidade feita naquela torre é pareado com as de potência das 2 torres associadas. Os dados relativos a essa turbina (no total de  $N = 47.542$  medidas) foram coletados de 30/07/2010 a 31/07/2011. É prática corrente arranjar os dados obtidos em fazendas eólicas em blocos de 10 minutos, período em que velocidades do vento são consideradas estacionárias (i.e. aproximadamente constantes), conseqüentemente, tomam-se as médias das leituras da potência de saída e da velocidade do vento a cada intervalo de 10 minutos.

<sup>1</sup> <https://aml.engr.tamu.edu/2001/09/01/publications/>

Figura 10 – Curva de potência da turbina eólica WT1.



Fonte – Autor.

#### 4.5 Ocorrência de Amostras Discrepantes (*Outliers*)

Para a amostra de dados apresentada na Figura 10, é possível obter modelos estatísticos que representem a sua criação. A partir desses modelos é possível avaliar diversas observações, oriundas ou não da mesma distribuição de dados. As observações que não podem ser explicadas pelos modelos propostos geralmente podem ser consideradas como *outliers*.

De acordo com Ben-Gal (2005), existem algumas definições de *outliers* apresentadas ao longo do tempo por autores, a seguir:

- Uma observação que desvia tanto das outras observações, que levanta suspeitas de ter sido gerada por um mecanismo diferente (HAWKINS, 1980);
- Uma observação que parece desviar claramente dos outros membros da amostra na qual ela ocorre (BARNETT; LEWIS, 1974);
- Uma observação em um conjunto de dados que parece ser inconsistente com o restante daquele conjunto de dados (JOHNSON; GEISSER, 1983).

Ainda que existam essas definições, não existe uma definição matemática do que é um *outlier*. De acordo com Hodge e Austin (2004), pode-se resumir as definições comuns de todos autores a:

"Um outlier é uma observação (escalar ou vetorial) que difere marcadamente dos outros membros da amostra a que ela pertence, tal como indicado pelo modelo escolhido para representar esta amostra."

No que se refere a presença de *outliers* nos dados, as causas podem ser bastante variadas:

- Erros de medição durante a coleta dos dados;
- Erros de impressão ou gravação dos dados;
- Falhas elétricas ou mecânicas no sistemas;
- Comportamento imprevisível do sistema;
- Mudança do ponto de operação do sistema;
- Flutuações estatísticas inerentes ao conjunto de dados (BARROS, 2013).

Dados gerados artificialmente não apresentam valores equivocados, ou grosseiros. Já dados reais, no caso deste estudo, geralmente possuem erros, que podem se mostrar em forma de *outliers*. Apesar de *outliers* geralmente serem considerados erros, eles podem possuir informações relevantes e, em alguns casos, podem apresentar as informações mais relevantes de determinada amostra (BEN-GAL, 2005). De acordo com Ripley (2004), não é suficiente examinar os dados e eliminar os *outliers*, alguns aspectos devem ser considerados, enumerados a seguir:

1. Localizar *outliers* em dados multivariados ou altamente estruturados pode ser tarefa muito laboriosa ou até mesmo impossível;
2. Rejeitar *outliers* afeta a teoria da distribuição. Em particular, variâncias serão subestimadas a partir dos dados "limpos";
3. Decidir entre manter ou rejeitar uma observação pode ser perda de tempo. O melhor a fazer seria diminuir o peso dessas informações em vez de rejeitá-las.

Ainda no que tange a retirada de *outliers*, há o risco da remoção de amostras simplesmente pelo fato delas não se ajustarem ao restante dos dados. Assim, existe o risco de escolher dados com base no desejo de melhorar os resultados, o que pode levar a obtenção de resultados enviesados, em vez de resultados cientificamente significativos (BARROS, 2013).

Um exemplo ilustrado da tese de BARROS (2013) traz um caso de regressão linear simples com *outliers*. O conjunto de dados da Tabela 4 corresponde ao número de mortes (por milhões de pessoas) por câncer de pulmão no mundo em 1950 devido ao uso per capita em um dado país no ano de 1930.



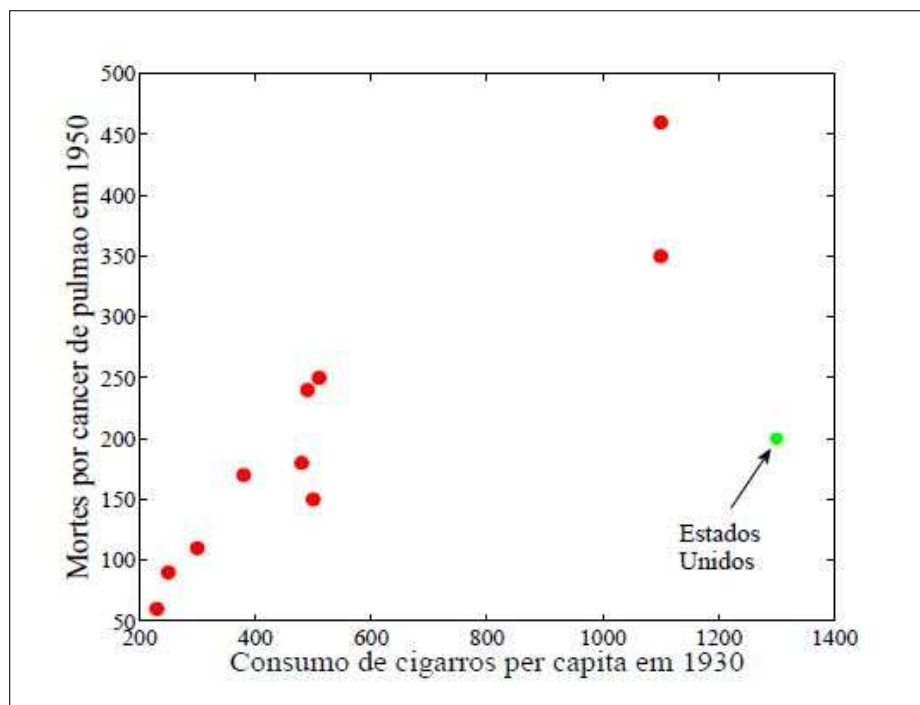
Tabela 4 – Consumo per capita de cigarro em vários países em 1930 e as mortes devido a câncer de pulmão em 1950.

País	Cigarro per capita	Mortes por milhões de pessoas
Austrália	480	180
Canadá	500	150
Dinamarca	380	170
Finlândia	1100	350
Grã Bretanha	1100	460
Islândia	230	60
Holanda	490	240
Noruega	250	90
Suécia	300	110
Suíça	510	250
Estados Unidos	1300	200

Fonte – (BARROS, 2013)

O diagrama de dispersão dos dados, está ilustrado na Figura 11, com o ponto referente aos Estados Unidos em destaque. Com a presença dos dados referentes aos Estados Unidos, percebe-se a dificuldade de representar os dados por uma tendencia linear. Ao se aplicar o método dos mínimos quadrados ordinário para a obtenção das retas de regressão resultantes, obtêm-se os parâmetros apresentados na Tabela. 5. Observa-se que os parâmetros da reta de regressão foram alterados com e sem a presença dos dados dos Estados Unidos.

Figura 11 – Diagrama de dispersão para os dados da Tabela 4.



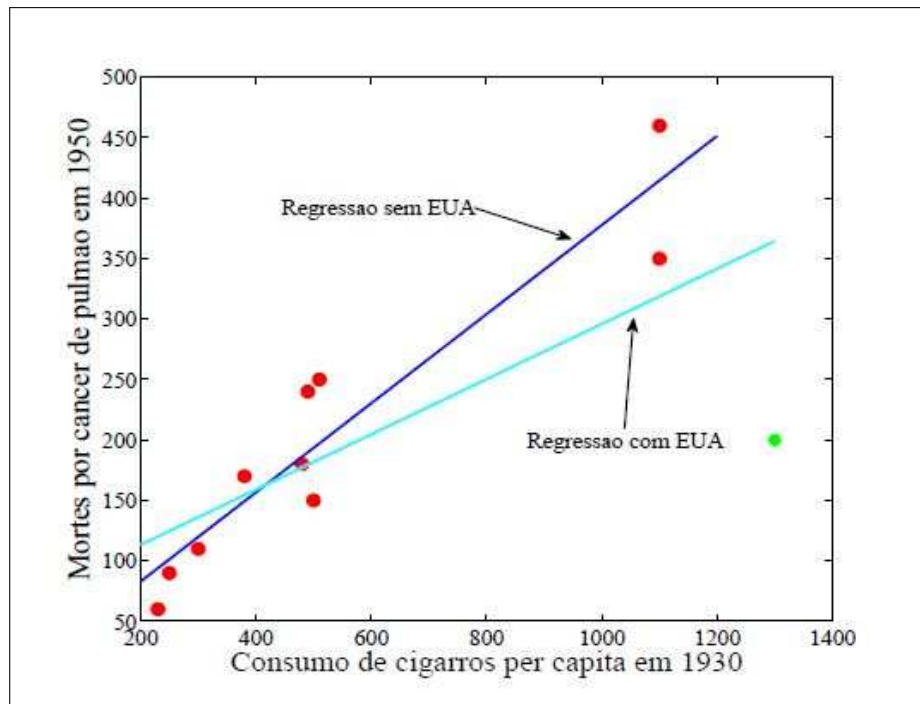
Fonte – (BARROS, 2013).

Tabela 5 – Estimação dos parâmetros da reta de regressão com e sem os dados dos Estados Unidos(EUA)

Dados	Parâmetro $\beta_0$	Parâmetro $\beta_1$
Sem EUA	9,14	0,37
Com EUA	67,56	0,23

Fonte – (BARROS, 2013)

Figura 12 – Gráfico das retas de regressão de acordo com os dados da Tabela 5 com e sem os dados referentes aos EUA.



Fonte – (BARROS, 2013).

Na Fig 12 observa-se as retas de regressão resultantes para os dados com e sem os Estados Unidos, obtidos por mínimos quadrados e apresentados na Tabela 5. Percebe-se que a análise das duas retas de regressão levariam a conclusões diferentes, a importância da investigação dos dados é muito relevante para que não se obtenha resultados tendenciosos, visto que o dado referente aos Estados Unidos é considerado um *outlier* para o problema apresentado.

#### 4.5.1 Modelo TSK Robusto

Na Seção 4.2.1 vimos que os parâmetros dos consequentes do modelo TSK podem ser estimados via método OLS. Podemos substituir o método OLS pelo algoritmo IRLS descrito no Capítulo 3, de modo a tornar o modelo TSK robusto à presença de outliers nos dados. Neste

caso, denotaremos a versão robusta do modelo TSK como TSK- $M$ .

## 4.6 Índices de Desempenho

Em problemas de regressão, uma importante tarefa é a de quantificar a qualidade obtida. Dito isto, é necessário comparar diversos algoritmos e estruturas utilizando índices de desempenho. A seguir são apresentados os índices de desempenho utilizados para o problema de regressão proposto. Juntamente com esses índices, também serão analisados figuras de mérito para observar como os modelos se comportam.

### 4.6.1 *Sum of Squared Errors (SSE)*

A Soma dos Erros Quadrados, do inglês *Sum of Squared Errors*, apresentado na Eq.( 4.9) abaixo é um índice que descreve a relação entre o conjunto de dados esperados e observados, utilizado para determinar como os valores de algum modelo predito estão relacionados com os valores reais utilizados para a determinação deste modelo.

$$SSE = \sum_{i=1}^N (y_i - \hat{f}(x_i))^2 \quad (4.9)$$

### 4.6.2 *Coefficiente de Determinação*

O coeficiente de determinação, conhecido também por  $R^2$  ou por coeficiente de correlação múltipla, que tem sua fórmula descrita na Eq.( 4.10) abaixo, é um parâmetro que mede o ajuste de um modelo estatístico em relação aos dados observados.

$$R^2 = 1 - \frac{\sum_{i=1}^N (y_i - \hat{f}(x_i))^2}{\sum_{i=1}^N (y_i - \bar{y})^2}. \quad (4.10)$$

O  $R^2$  varia entre 0 e 1 e indica, em porcentagem, o quanto o modelo consegue explicar os valores observados (OZER, 1985).

## 4.7 *Softwares de Simulação*

Durante esse trabalho de mestrado todos os algoritmos foram desenvolvidos em ambientes de programação tais como Matlab<sup>©</sup> e Octave, escolhidos pela facilidade de implementação dos códigos e interatividade com o usuário.

#### 4.8 Conclusão do capítulo

Neste capítulo foram apresentados os modelos alternativos baseados no sistema de inferência *fuzzy* TSK, afim de compará-los com os modelos mais utilizados, os dados para realização do trabalho também foram explicados, tal como o que é e como lher dar com *outliers* e os índices de desempenho utilizados para avaliar os modelos propostos. O modelo neurofuzzy anfis também foi apresentado, para fins de comparação com o o modelo proposto. No capítulo a seguir, serão expostos os resultados desta comparação.

## 5 RESULTADOS E DISCUSSÃO

Neste capítulo iremos apresentar os resultados oriundos de uma comparação abrangente e sistemática do desempenho dos modelos descritos no Capítulo 3, 4 e 5. Por sistemática, entende-se que avaliaremos a sensibilidade de cada modelo em relação aos seus principais parâmetros, tais como a ordem do polinômio (modelo polinomial) e número de neurônios ocultos (rede MLP). Para o modelo TSK e ANFIS será avaliada a sensibilidade dos modelos em relação ao número de funções de pertinência (granularidade da fuzzificação) para variável velocidade do vento e em relação ao raio de abertura no caso da função de pertinência gaussiana.

### 5.1 Resultados do Modelo Polinomial

Variou-se a ordem  $k$  do modelo polinomial de 1 até 8, com o melhor desempenho alcançado nos testes pelo modelo de ordem  $k = 8$ , como exposto na Tabela 6, com valor de  $SSE = 8,0629 \times 10^6$  e  $R^2 = 0,8529$ .

Tabela 6 – Valores obtidos para os índices SSE e  $R^2$  para diferentes ordens do modelo polinomial.

Ordem $k$	SSE	$R^2$
1	$1,0471 \times 10^7$	0,8089
2	$9,5108 \times 10^6$	0,8264
3	$8,3965 \times 10^6$	0,8468
4	$8,1183 \times 10^6$	0,8519
5	$8,0952 \times 10^6$	0,8523
6	$8,0814 \times 10^6$	0,8525
7	$8,0711 \times 10^6$	0,8527
8	$8,0629 \times 10^6$	0,8529

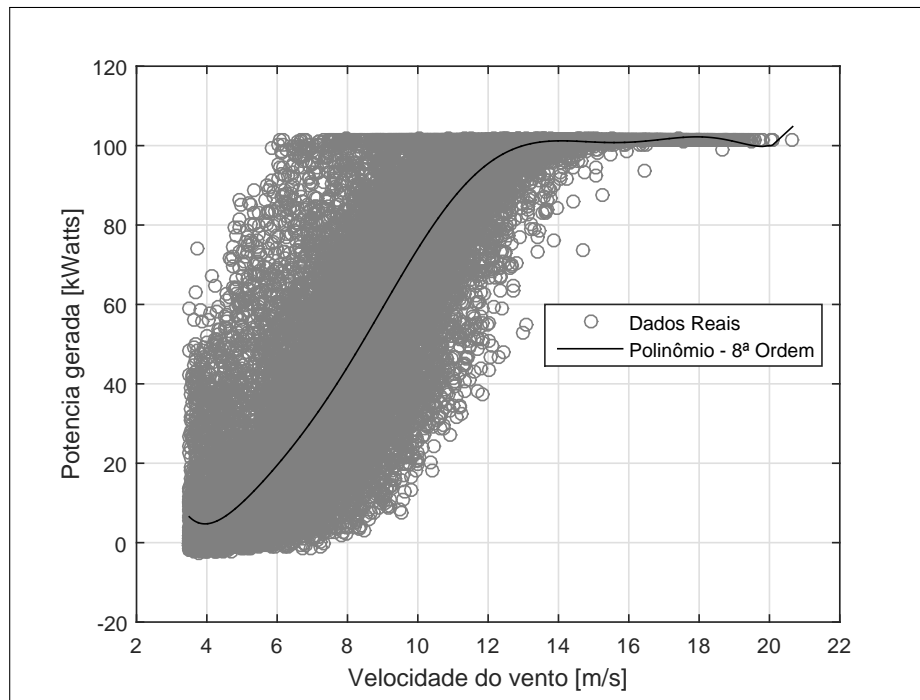
Fonte – o autor.

A curva de potência para o melhor modelo polinomial pode ser visto na Fig. 13. Observa-se que a curva obtida replica o comportamento dos dados, com exceção quando a velocidade  $v$  aumenta muito, onde a potência gerada começa a aumentar juntamente com a velocidade do vento.

### 5.2 Resultados da Rede MLP

No modelo baseado em rede MLP, utilizou-se 90% dos dados para treinamento da rede e 10% para teste e validação do modelo obtido. O algoritmo de treinamento utilizado, como

Figura 13 – Curva de potência e regressão para o modelo polinomial de ordem 8.



Fonte – o autor.

exposto anteriormente, foi o de Levenberg-Marquardt (L.M), com variação dos neurônios da camada escondida de  $q = 1$  até  $q = 5$ . A Tabela 7 apresenta os valores dos índices SSE e  $R^2$ .

Percebe-se certa semelhança entre as curvas da Fig. 13 e Fig. 14, assim como nos valores obtidos dos SSE's. Ambas as curvas foram capazes de representar bem o comportamento das amostras, com certa dificuldade quando a velocidade de entrada aumenta muito.

Tabela 7 – Coeficiente de regressão e SSE para o modelo baseado em rede MLP

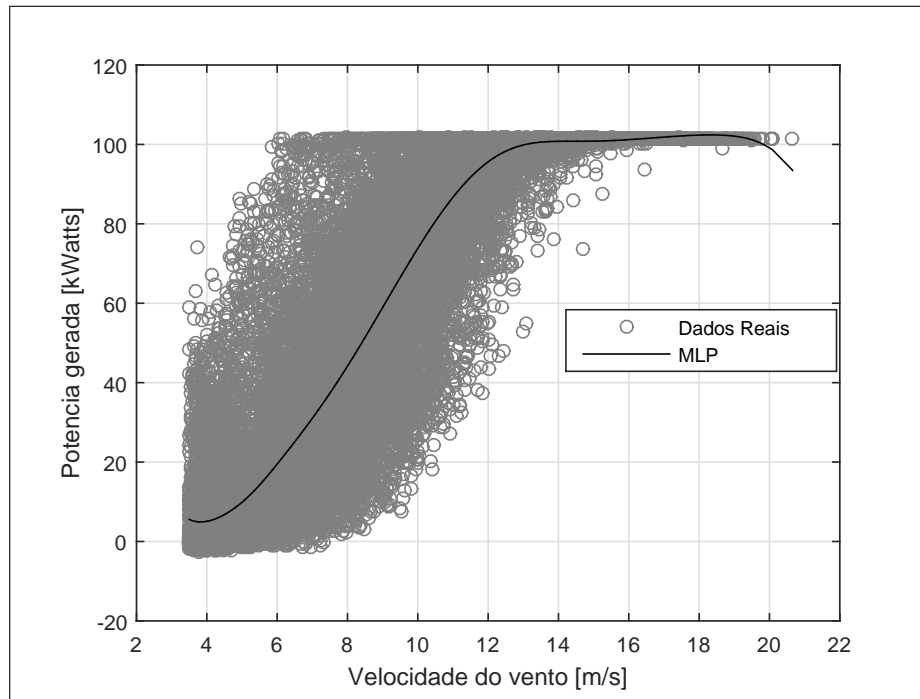
Nº de neurônios ocultos	SSE	$R^2$
1	$8,1139 \times 10^6$	0,8263
2	$8,0769 \times 10^6$	0,8254
3	$8,0796 \times 10^6$	0,8253
4	$8,0664 \times 10^6$	0,8258
5	$8,0618 \times 10^6$	0,8271

Fonte – o autor.

### 5.3 Resultados do Modelo ANFIS

Variou-se a quantidade de funções de pertinência gaussianas do modelo ANFIS de 3 até 5, com melhor desempenho apresentado no modelo com 5 funções de pertinência, como exposto na Tabela 8, com valor de  $SSE = 8,0598 \times 10^6$  e  $R^2 = 0,8529$ . A curva de potência

Figura 14 – Curva de potência e regressão para o modelo de Rede MLP com 5 neurônios na camada oculta.



Fonte – o autor.

para o melhor modelo ANFIS obtido pode ser visto na Fig. 15.

Tabela 8 – Valores obtidos para os índices  $SSE$  e  $R^2$  para o modelo ANFIS.

Qtd de Funções de pertinência	$SSE$	$R^2$
3	$8,0610 \times 10^6$	0,8529
4	$8,0612 \times 10^6$	0,8528
5	$8,0598 \times 10^6$	0,8529

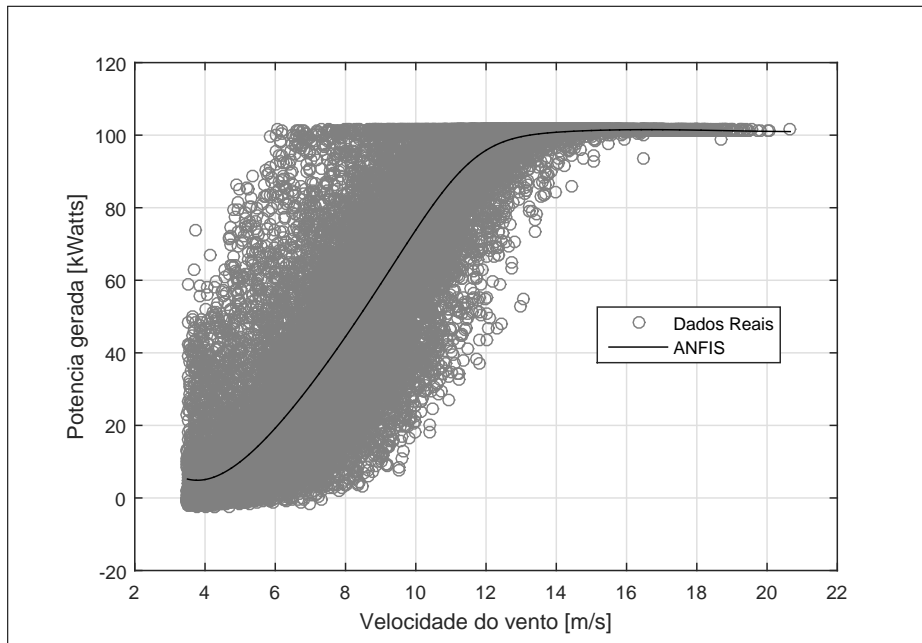
Fonte – o autor.

Observa-se que o modelo representa bem os dados testados, com valores obtidos de  $SSE$  inferiores aos modelos MLP e polinomial.

#### 5.4 Resultados do Modelo TSK

Para as variantes do modelo TSK avaliadas foram testadas tanto funções de pertinência triangulares/trapezoidais, quanto gaussianas. As posições dos centros das funções de pertinência foram encontradas por clusterização dos vetores de medidas  $\{(v_i, p_i)\}_{i=1}^N$  usando o algoritmo k-médias, onde o valor de  $K$  variou de 3 a 5, quantidade igual ao número de funções de pertinência testadas. Posteriormente foram testados dois cenários, como explicado na Seção 4.3

Figura 15 – Curva de potência e regressão para o modelo ANFIS com 5 funções de pertinência gaussianas.



Fonte – o autor.

para a obtenção de melhores resultados. Para a realização desses testes, foram testados os dois estimadores-M apresentados na Subseção 3.2.1.

#### 5.4.1 Uma abertura comum a todas as funções de pertinência

Foram testados dois cenários para cada estimador, o primeiro onde aconteceu a variação comum da abertura de todas as funções, variando entre metade do valor do desvio-padrão ( $\gamma = 2,87m/s$ ) da velocidade do vento até três vezes ( $\gamma = 8,61m/s$ ) esse valor e para o segundo cenário foi considerado uma abertura diferente para cada função.

Tabela 9 – Coeficiente de regressão e SSE para o modelo TSK-OLS e TSK-M para funções gaussianas.

Nº de funções	$SSE-OLS$	$R^2 - OLS$	$SSE-M$	$R^2 - M$	Melhor $Spread (m/s)$
3	$8,0619 \times 10^6$	0,8529	$8,2939 \times 10^6$	0,8486	2,90
4	$8,0724 \times 10^6$	0,8527	$8,2874 \times 10^6$	0,8488	8,54
5	$8,060 \times 10^6$	0,8529	$8,2976 \times 10^6$	0,8488	1,44

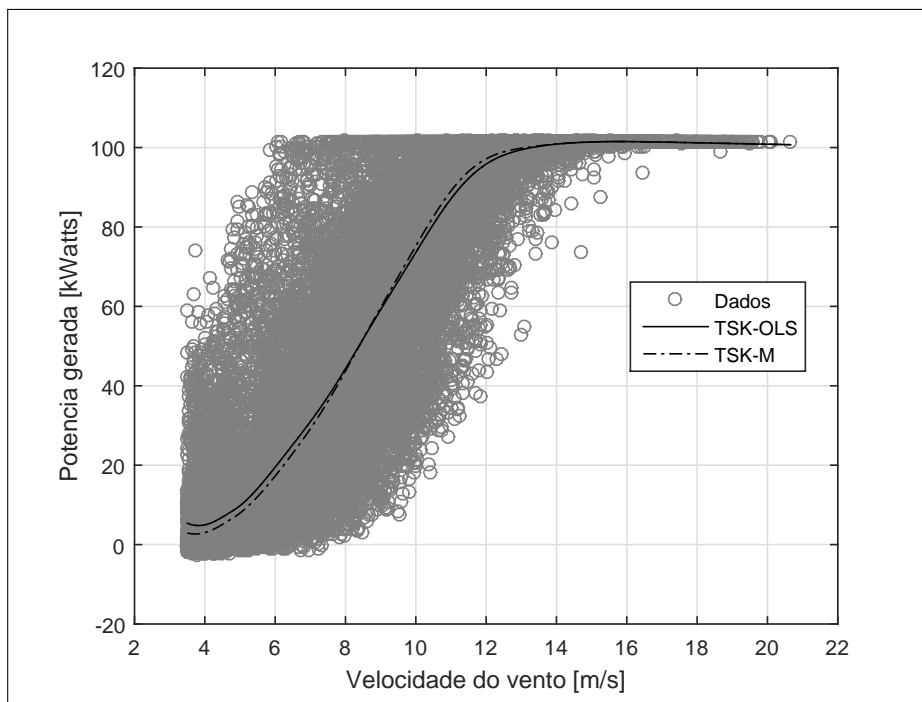
Fonte – o autor.

Ao analisar a Tabela 9, percebe-se que os valores de  $SSE$  e  $R^2$  para cada tipo de estimação são bem similares, porém para o 3 e 4 funções de pertinência gaussianas, a curva não



conseguiu se adequar a última parcela dos dados. No entanto, ao se analisar o cenário com 5 funções as duas curvas foram bastante fiéis aos dados, como apresentado na Fig. 16 e os centros encontrados foram  $c_1 = 4,8167m/s$ ,  $c_2 = 6,6958m/s$ ,  $c_3 = 9,0543m/s$ ,  $c_4 = 11,3130m/s$ ,  $c_5 = 14,5690m/s$ .

Figura 16 – Curvas de potência estimados pelos modelos TSK-OLS e TSK-M (bisquare) para 5 funções de pertinência gaussianas.



Fonte – o autor.

Os resultados para os testes realizados com funções de pertinência triangulares e trapezoidais estão expostos na Tabela 10. Em todos os casos a melhor abertura para as funções triangulares foi igual a  $4,5m/s$  para cada lado além dos centros obtidos, com obtenção dos melhores resultados com a utilização de 5 funções de pertinência.

Tabela 10 – Coeficiente de regressão e SSE para o modelo TSK-OLS e TSK-M para funções triangulares/trapezoidais.

Nº de funções	$SSE-OLS$	$R^2 - OLS$	$SSE-M$	$R^2 - M$
3	$8,1122 \times 10^6$	0,8520	$8,3595 \times 10^6$	0,8474
4	$8,1577 \times 10^6$	0,8511	$8,3919 \times 10^6$	0,8469
5	$8,1109 \times 10^6$	0,8520	$8,3331 \times 10^6$	0,8479

Fonte – o autor.

Na Fig. 17 são apresentados os resultados para os testes com funções de pertinências

Tabela 11 – Coeficiente de regressão e SSE para o modelo TSK-OLS e TSK-M para os teste das heurísticas para 3 funções de pertinência gaussianas.

heurística	$SSE-OLS$	$R^2 - OLS$	$SSE-M$	$R^2 - M$
H1	$8,0761 \times 10^6$	0,8528	$8,3595 \times 10^6$	0,8491
H2	$8,0754 \times 10^6$	0,8527	$8,3919 \times 10^6$	0,8484
H3	$8,0837 \times 10^6$	0,8529	$8,3331 \times 10^6$	0,8485
H4	$8,1478 \times 10^6$	0,8528	$8,3331 \times 10^6$	0,8485
H5	$8,2129 \times 10^6$	0,8527	$8,3331 \times 10^6$	0,8484

Fonte – o autor.

triangulares. Observa-se que as três curvas foram capazes de captar a dinâmica dos dados, apresentando dificuldade para replicá-los quando  $v > 14m/s$ , região em que com o aumento da velocidade  $v$  a potência permanece constante. Em Fig. 17.c, o modelo TSK-M proposto foi capaz de representar bem essa 3ª região da curva.

#### 5.4.2 Raios distintos por pertinência gaussiana

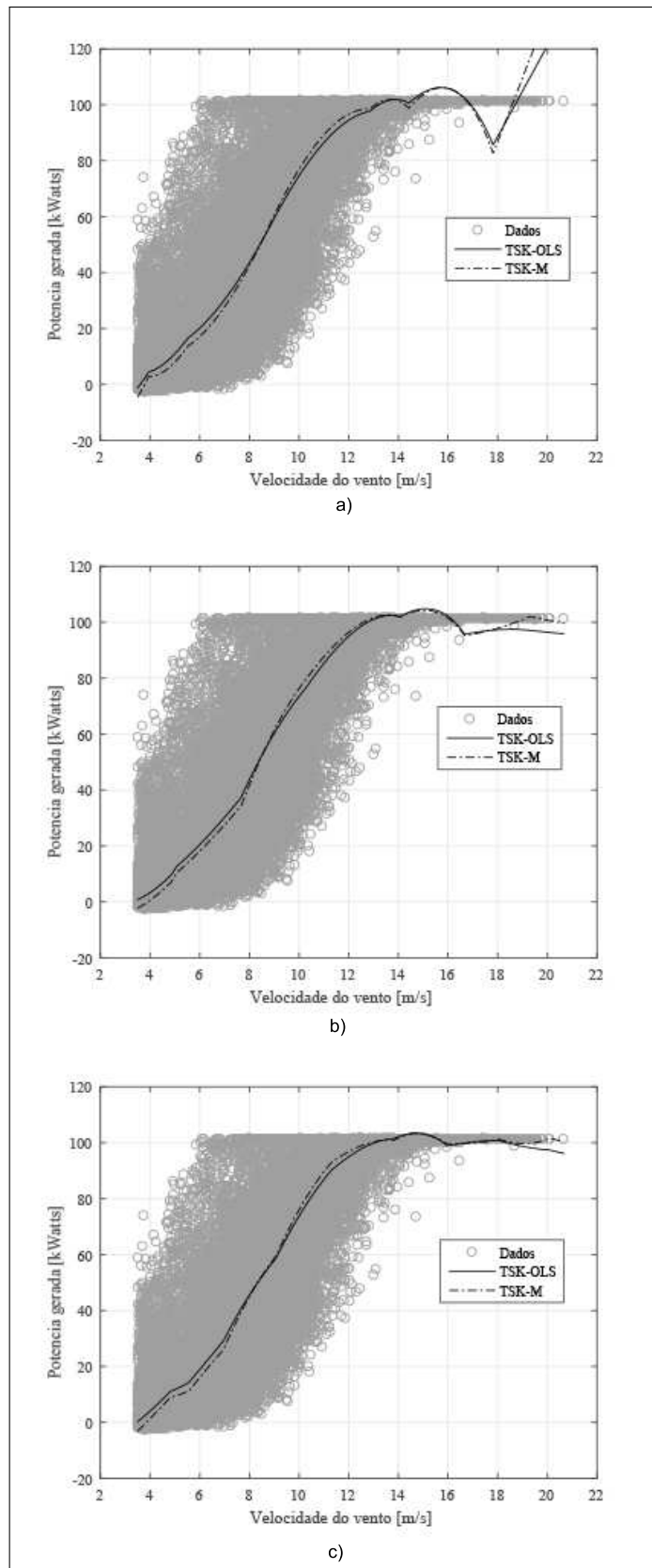
Para esse cenário, são testadas cinco heurísticas, assim como especificado na Seção 4.3, por simplificação cada heurística será denotada pela letra H.

Para 3 funções de pertinência, quando o valor da distância entre as gaussianas foi mais alto, como no caso de H.1, ao analisar graficamente, observa-se que as curvas se comportaram de maneira inusitada e variaram muito na parte do gráfico em que a velocidade é maior que  $V_{cutoff}$ , apresentado na Fig. 18. Porém para os casos de H.2 e H.3, nas Figuras 19 e 20 o estimador robusto se mostrou bem fiel aos dados, com rendimento superior ao estimador convencional. Quando ocorre a diminuição da abertura das funções gaussianas, para H.4 e H.5, houve uma piora na curva de regressão, apesar do baixo erro. A curva apresentou uma configuração anormal em relação aos outros casos para  $v > V_{cutoff}$ , onde um aumento contínuo na velocidade do vento levou ao aumento da potência de saída, apresentando característica similar a uma reta.

##### 5.4.2.1 Quatro e Cinco Funções de Pertinência

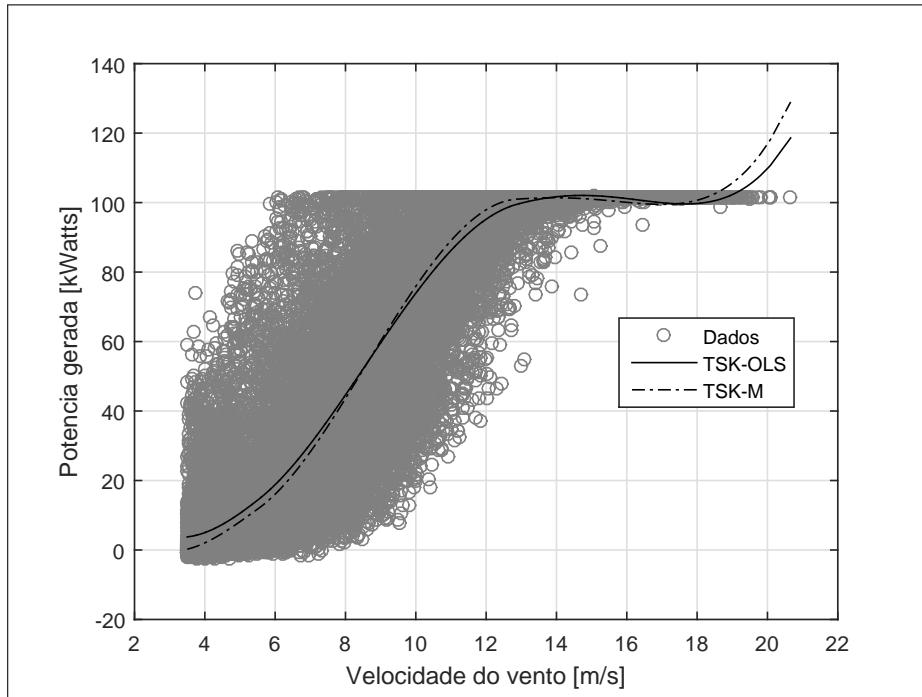
Para quatro funções de pertinência, a diminuição das distâncias entre as gaussianas também provocou uma piora nas curvas de regressão, sem apresentar diferenças significativas no valores de  $SSE$  e  $R^2$ . Já para cinco funções de pertinência, essa diminuição provocou uma melhora na representação dos dados, a Fig. 21 apresenta o cenário testado para H.3, com resultados bem similares a H.4 e H.5.

Figura 17 – Curvas de potência estimadas pelos modelos TSK-OLS e TSK-M (bisquare) para funções de pertinência triangulares. a) 3 funções; b) 4 funções; e c) 5 funções.



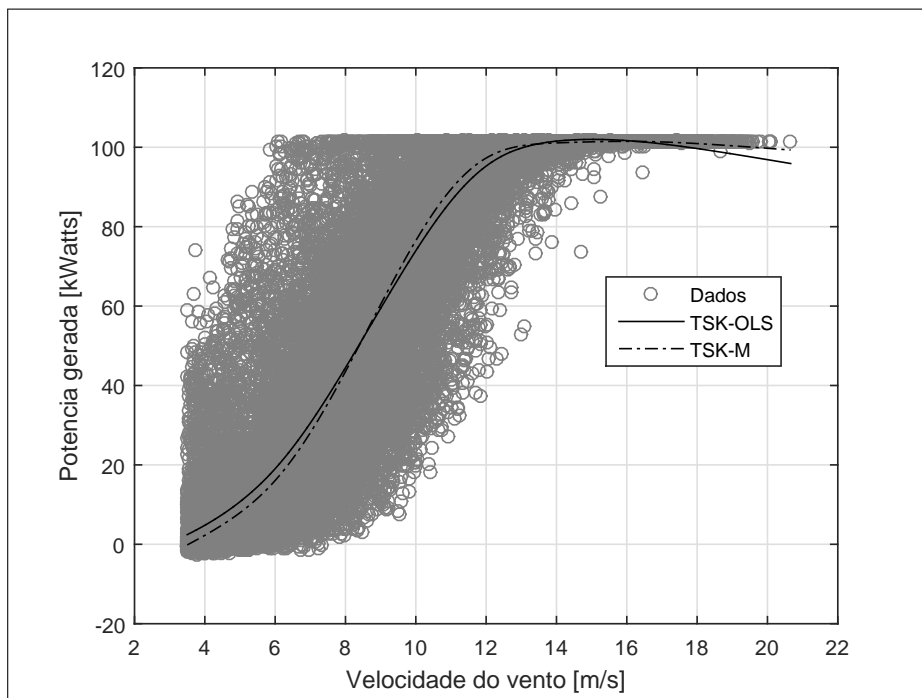
Fonte – o autor.

Figura 18 – Curvas de potência estimados para o modelo TSK-OLS e TSK-M(bisquare) para a heurística 1 para 3 funções de pertinência gaussianas.



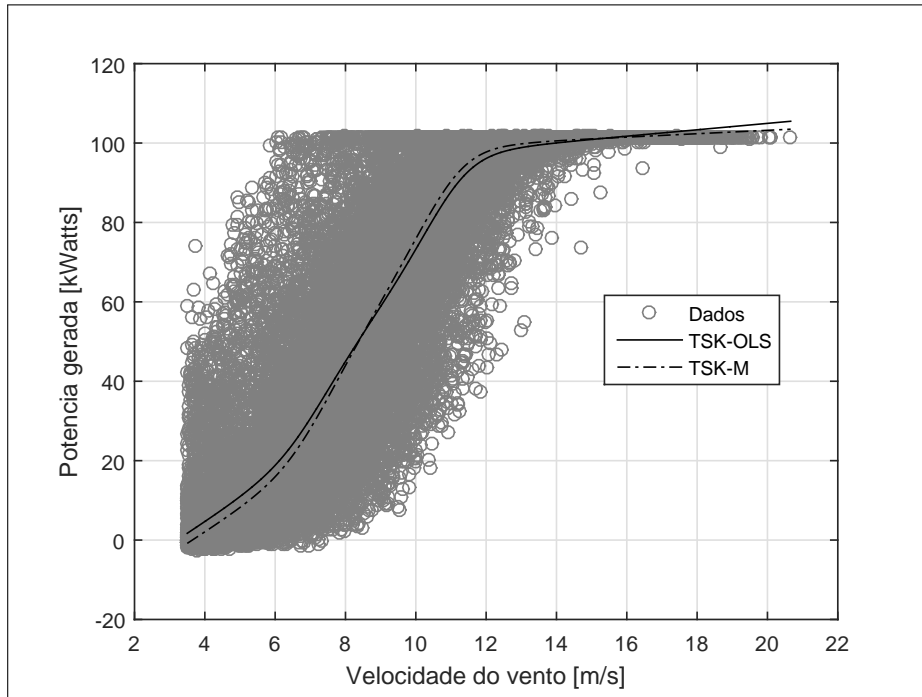
Fonte – o autor.

Figura 19 – Curvas de potência estimados pelos modelos TSK-OLS e TSK-M(bisquare) para a heurística 2 para 3 funções de pertinência gaussianas.



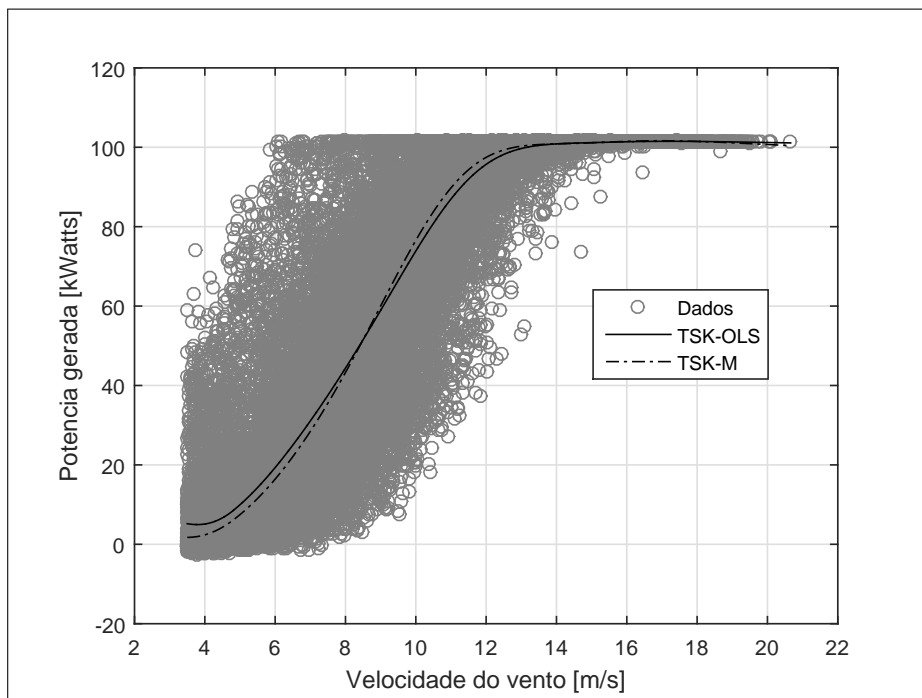
Fonte – o autor.

Figura 20 – Curvas de potência estimados pelos modelos TSK-OLS e TSK-M(bisquare) para a heurística 3 para 3 funções de pertinência gaussianas.



Fonte – o autor.

Figura 21 – Curvas de potência estimados para o modelo TSK-OLS e TSK-M(bisquare) para a heurística 3 para 5 funções de pertinência gaussianas.



Fonte – o autor.

Não houve diferenças significativas nas curvas de regressão utilizando o estimador bisquare ou de Huber. Nos casos de aberturas diferentes para cada função, ao se considerar funções triangulares e trapezoidais, os resultados não foram superiores ao das funções de pertinência gaussianas.

### 5.5 Comparação entre os modelos apresentados

Ao fazer a comparação dos modelos avaliados, é necessário avaliá-los não só matematicamente, como graficamente, visto que a curva possui certa dificuldade em ser replicada na parte final dos dados. Os melhores desempenhos dos modelos avaliados, segundo o SSE e o coeficiente de regressão, estão dispostos na Tabela 12. Para fins de notação, o T representa que a estimação veio a partir de funções triangulares/trapezoidais.

Tabela 12 – Comparação do desempenho dos modelos quanto ao SSE e  $R^2$ .

Modelo	SSE	$R^2$
Polinomial	$8,0629 \times 10^6$	0,8529
MLP	$8,0618 \times 10^6$	0,8271
ANFIS	$8,0598 \times 10^6$	0,8529
TSK-OLS	$8,0605 \times 10^6$	0,8529
TSK-M	$8,2976 \times 10^6$	0,8488
TSK-OLS-T	$8,1109 \times 10^6$	0,8520
TSK-M-T	$8,3331 \times 10^6$	0,8479
TSK-OLS-H.3	$8,0612 \times 10^6$	0,8529
TSK-M-H.3	$8,2998 \times 10^6$	0,8485

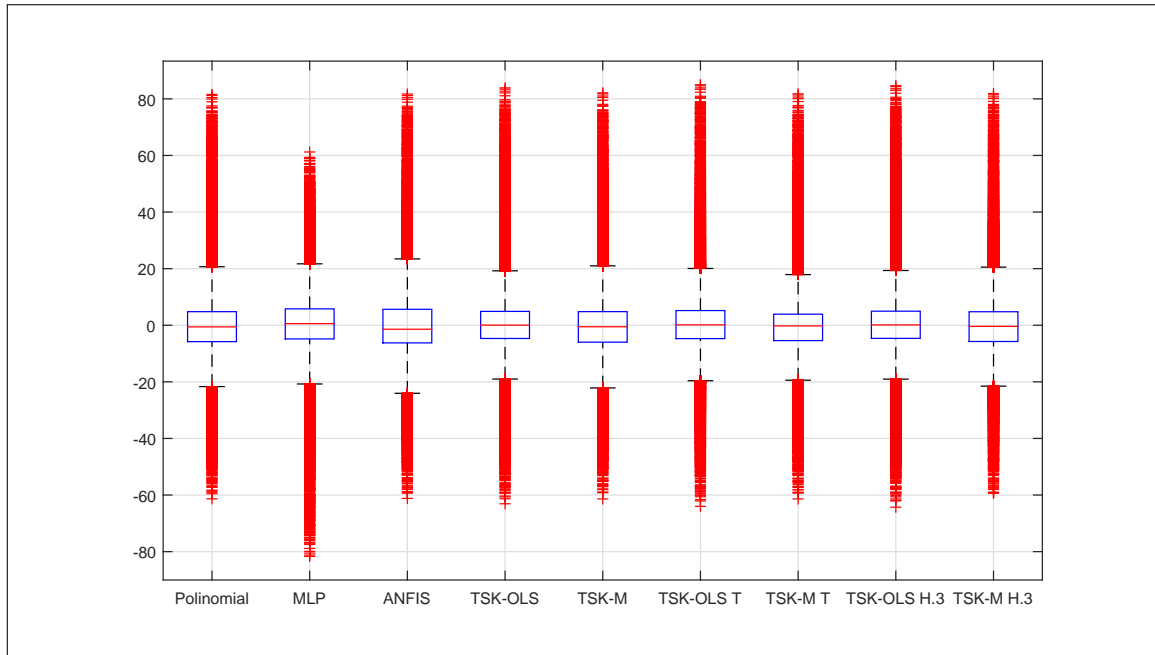
Fonte – o autor.

No quesito SSE, observa-se que o modelo ANFIS apresentou o melhor rendimento, seguido do modelo de MLP, polinomial e dos modelos TSK-OLS. Contudo, o modelo ANFIS e o modelo MLP apresentam tempo de processamento e gasto computacional elevados, tornando-os mais trabalhosos computacionalmente. Em relação aos coeficientes de determinação obtidos, o modelos Polinomial, ANFIS, TSK-OLS e TSK-OLS-H.3 obtiveram melhor rendimento, com coeficiente de determinação igual a 0,8529. O modelo baseado em MLP apresentou o pior resultado, com  $R^2 = 0,8271$ .

A título de completude, mostramos na Fig. 22 os boxplots dos resíduos para os modelos apresentados na Tabela 12. Como era de se esperar, para o tipo de aplicação que estamos abordando, há muitos *outliers* presentes nos dados (SOHONI; GUPTA; NEMA, 2016; SHOKRZADEH; JOZANI; BIBEAU, 2014). A presença de *outliers* acaba interferindo nega-

tivamente no processo de estimação de parâmetros de todos os modelos e gerando assim uma distribuição de resíduos que também apresenta *outliers*, conforme pode ser verificado na referida figura.

Figura 22 – Boxplot dos resíduos para os modelos apresentados na Tabela 12.



Fonte – o autor.

Com o objetivo de mostrar a robustez do modelo TSK proposto a *outliers*, outros testes foram realizados com dados obtidos de um aerogerador na Prainha, Ceará. As especificações do aerogerador citado são apresentadas na Tabela 13.

Tabela 13 – Especificações do aerogerador da Prainha-CE.  
Aerogerador Prainha

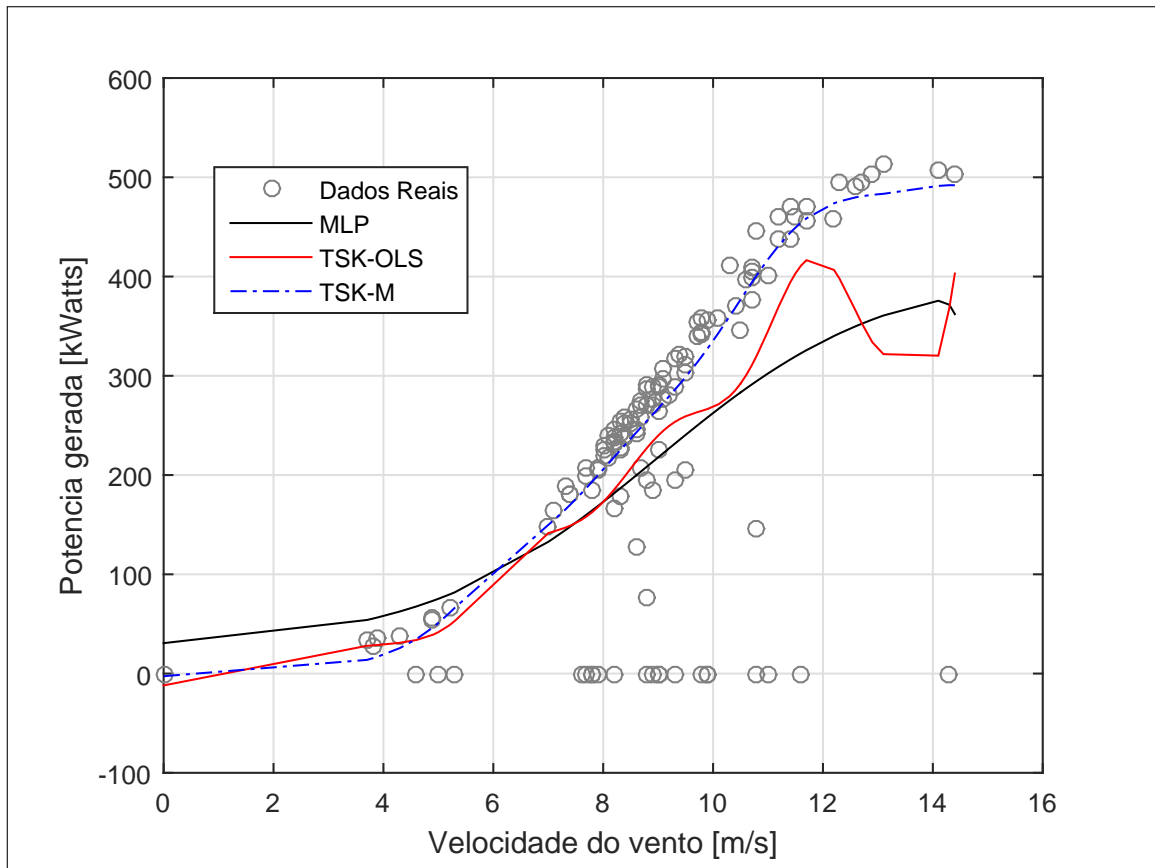
Altura ( <i>m</i> )	60
Diâmetro do rotor ( <i>m</i> )	40
Velocidade de corte inferior ( <i>m/s</i> )	3,0
Velocidade de corte superior ( <i>m/s</i> )	13
Velocidade nominal ( <i>rpm</i> )	28 – 36
Potência nominal ( <i>kW</i> )	500 – 600

Fonte – (MOURA A. P., 2003; SILVA, 2003).

A Fig. 23 apresenta as curvas de regressão dos modelos MLP, TSK-OLS e TSK-M para os dados do aerogerador da Prainha. Observa-se que tanto o modelo TSK-OLS quanto o modelo baseado em rede MLP apresentam um péssimo rendimento e não conseguem replicar os

dados. Já o modelo proposto nesta dissertação replica bem a dinâmica dos dados do aerogerador, ainda que existam *outliers* nas amostras analisadas.

Figura 23 – Comparação das curvas de regressão para os modelos TSK-OLS, TSK-M e MLP para o aerogerador da Prainha-CE.



Fonte – (MOURA A. P., 2003; SILVA, 2003).

Pelos resultados observados nos testes, conclui-se que o modelo proposto apresenta vantagens ao compará-lo com o modelo de rede MLP. O modelo TSK-M responde melhor em situações com poucos dados, por isso o conjunto de dados utilizado para a validação da proposta apresenta resultados melhores com o modelo de rede MLP do que com alguns dos modelos TSK. O conjunto de dados da estação WT1 apresenta 47.542 pares de entrada-saída, com diversos *outliers*. Esta situação beneficia o modelo MLP, o que justifica o fato dos modelos comparados apresentarem resultados parecidos.



## 6 CONCLUSÕES E TRABALHOS FUTUROS

### 6.1 Conclusões

O trabalho desenvolvido nessa dissertação tratou essencialmente de apresentar modelos de regressão aplicados à estimação da curva de potência de aerogeradores que visassem estabelecer uma solução mais parcimoniosa que modelos comumente utilizados. Além disso, é importante que os modelos apresentados ainda sejam competitivos quanto a capacidade de minimizar o residual quadrático em relação aos demais modelos.

De todos os casos analisados, o modelo ANFIS com 5 funções de pertinências foi o melhor avaliado, seguido pelo modelo TSK-OLS com 5 funções de pertinências, modelo utilizando a heurística 3, modelo de rede MLP e polinomial. O modelo TSK-M, apesar de apresentar erros maiores que os outros, por vezes mostrou-se mais fiel a curva de potência, como nas Figuras 19 e 20, apesar de erros maiores ao compará-lo com o modelo TSK-OLS. Na comparação com outros modelos restantes, o modelo MLP apresenta erros similares, ou em alguns casos valores melhores do que o modelo TSK, porém se faz necessário a análise do comportamento da curva e do custo computacional como um todo, dito isto, o modelo baseado em MLP é falho ao tentar reproduzir a dinâmica da curva, quando a velocidade de entrada ultrapassa a velocidade de corte superior e em relação ao alto custo computacional. O algoritmo utilizado para treinamento da rede foi o de L.M., visto que a rede não converge ao se utilizar o gradiente descendente, justificando o elevado custo computacional do modelo.

O modelo ANFIS apresentou curva bem fiel ao conjunto de dados testados, porém o seu custo computacional e complexidade elevados não o torna viável, visto que modelos mais simples apresentaram  $R^2$  igual ao modelo ANFIS. Segundo os resultados obtidos, não houve uma melhora significativa ao comparar o modelo ANFIS com o modelo proposto, isto evidencia a qualidade do modelo proposto, pois a aplicação da metodologia para ajuste do modelo TSK é simples ao comparar com um modelo que usa aprendizagem neural.

Os modelos TSK são bastante úteis, pois são capazes de replicar a curva dos aerogeradores por meio de combinação de modelos lineares locais, o que representa menor custo computacional. Estes modelos, juntamente com o modelo polinomial, destacam-se devido a sua velocidade, sendo especialmente interessante para aplicações de monitoramento *on-line*.

Diante dos resultados obtidos e dos objetivos propostos, este texto dissertativo apresentou um amplo estudo bibliográfico acerca do tema tratado, apresentando técnicas capazes

de melhorar os parâmetros dos modelos *fuzzy* testados. O modelo TSK com estimador-M mostrou-se eficaz ao se considerar a dinâmica dos dados, ainda que possua valores de erros maiores que os outros modelos. Logo, o mesmo pode servir como uma alternativa a modelos tradicionais, como o de MLP, assim como representam uma ferramenta que propicie a tomada de decisões.

Todos os códigos que foram desenvolvidos nesta dissertação serão disponibilizados a quem interessar através de uma solicitação por e-mail<sup>1</sup>, mediante as devidas referências no caso de se tornar útil.

## 6.2 Trabalhos Futuros

Esta dissertação possui desdobramentos que podem ser investigados como uma continuação do trabalho de pesquisa aqui apresentado. São listados a seguir algumas potenciais linhas de pesquisa que podem ser desenvolvidas com o objetivo de colocar em prática esta continuidade:

- Realizar um estudo comparativo entre outras técnicas existentes de estimação de curva de potência e o modelo proposto neste trabalho.
- Utilizar técnicas de otimização, tais como *Particle Swarm Optimization*, PSO (KENNEDY, 2011), para determinação dos centros das funções de pertinência.
- Desenvolver modelos de curva de potência que leve em consideração outras variáveis além da velocidade do vento.
- Produzir um método de monitoramento *on-line* de turbinas eólicas, com dados colhidos de aerogeradores em operação.

---

<sup>1</sup> e-mail:igor.brasil@ifma.edu.br

## REFERÊNCIAS

- ABEÉOLICA. Dados mensais associação brasileira de energia eólica. **Associação Brasileira de Energia Eólica**, 2017.
- AGUIRRE, L. A. **Introdução à Identificação de Sistemas – Técnicas Lineares e Não-Lineares Aplicadas a Sistemas Reais**. Minas Gerais: Editora UFMG, 2015. 20–59 p.
- BARNETT, V.; LEWIS, T. **Outliers in statistical data**. New York: Wiley, 1974. 329–400 p.
- BARRETO, G. A.; BARROS, A. L. B. P. On the design of robust linear pattern classifiers based on m-estimators. **Neural Processing Letters**, v. 42, n. 1, p. 119–137, 2015.
- BARROS, A. L. B. P. **Revisitando o problema de classificação de padrões na presença de outliers usando técnicas de regressão robusta**. Tese (Doutorado) — UFC, 2013.
- BEN-GAL, I. Outlier detection. In: **Data mining and knowledge discovery handbook**. Tel-Aviv: Springer, 2005. p. 131–146.
- BETZ, A. Das maximum der theoretisch möglichen ausnützung des windes durch windmotoren. **Zeitschrift für das gesamte Turbinenwesen**, v. 26, p. 307–309, 1920.
- BRANCO, H. M. G. C. **Uma Estratégia para a Detecção e Classificação de Transitórios em Transformadores de Potência pela Utilização de Transformada Wavelet e da Lógica Fuzzy**. Dissertação (Mestrado) — Universidade de São Paulo, 2009.
- BRASIL, E. d. P. E. Balanço energético nacional 2013–ano base 2012: Relatório síntese. **Rio de Janeiro, EPE**, 2013.
- BRUNO CABRAL. **Geração eólica no Ceará vai saltar 44% nos próximos anos**. 2017. Disponível em: <<http://diariodonordeste.verdesmares.com.br/cadernos/negocios/geracao-eolica-no-ceara-vai-saltar-44-nos-proximos-anos-1.1828648#email>>. Acesso em: 15 mai. 2018.
- CHANG, T. P. *et al.* Comparative analysis on power curve models of wind turbine generator in estimating capacity factor. **Energy**, v. 73, p. 88–95, 2014.
- CHUANG, C.-C.; HSIAO, C.-C.; JENG, J.-T. A novel approach for tsf fuzzy modeling with outliers. In: IEEE. **Machine Learning and Cybernetics, 2003 International Conference on**. China, 2003. v. 4, p. 2158–2163.
- CIDADE VERDE. 2017. Disponível em: <<https://cidadeverde.com/noticias/238505/estudo-indica-potencial-para-instalacao-de-duas-mil-torres-de-energia-eolica-no-piaui/>>. Acesso em: 12 set. 2018.
- CLEVELAND, W. S. Robust locally weighted regression and smoothing scatterplots. **Journal of the American statistical association**, Taylor & Francis, v. 74, n. 368, p. 829–836, 1979.
- CLIFTON, A. *et al.* Using machine learning to predict wind turbine power output. **Environmental Research Letters**, n. 8, p. 1–8, 2013.
- COUNCIL, G. W. E. Global wind report. **Annual market update**, 2014.

- DIAF, S. *et al.* Technical and economic assessment of hybrid photovoltaic/wind system with battery storage in corsica island. **Energy Policy**, v. 36, n. 2, p. 743–754, 2008.
- ENGELBRECHT, A. P. **Computational Intelligence: An Introduction**. 2nd. ed. New York: Wiley, 2007.
- FOX, J. **Applied regression analysis, linear models, and related methods**. New York: Sage Publications, Inc, 1997.
- FRANÇA, J. A. de *et al.* Uma implementação do algoritmo levenberg-marquardt dividido para aplicações em visão computacional. **Semina: Ciências Exatas e Tecnológicas**, v. 30, n. 1, p. 51–62, 2009.
- GILL, B. S. S.; GALLOWAY, S. Wind turbine condition assessment through power curve copula modeling. **IEEE Transactions on Sustainable Energy**, v. 3, n. 1, p. 94–101, 2012.
- GOMIDE, F. A. C.; GUDWIN, R. R. Modelagem, controle, sistemas e lógica fuzzy. **SBA Controle & Automação**, Vol.4, 1994.
- HAGAN, M. T.; MENHAJ, M. Training feed-forward networks with the Marquardt algorithm. **IEEE Transactions on Neural Networks**, v. 5, n. 6, p. 989–993, 1994.
- HAWKINS, D. M. **Identification of outliers**. South Africa: Springer, 1980. v. 11.
- HAYKIN, S. **Redes neurais: princípios e prática**. California: Bookman Editora, 2007.
- HEDEVANG, E. Wind turbine power curves incorporating turbulence intensity. **Wind Energy**, v. 17, n. 2, p. 173–195, 2014.
- HODGE, V.; AUSTIN, J. A survey of outlier detection methodologies. **Artificial intelligence review**, Springer, v. 22, n. 2, p. 85–126, 2004.
- HORVATH, L.; PANZA, T.; KARADZA, N. The influence of high wind hysteresis effect on wind turbine power production at bura-dominated site. **European Wind Energy Conference and Exhibition**, p. 1017–1022, 2007.
- HSIA, T. C. System identification: least-squares methods. p. 89–111, 1977.
- HUBER, P. J. Robust estimation of a location parameter. **Annals of Mathematical Statistics**, v. 35, n. 1, p. 73–101, 1964.
- HUBER, P. J. Robust estimation of a location parameter. **The annals of mathematical statistics**, JSTOR, p. 73–101, 1964.
- IEC. **Wind turbines - Power performance measurement of electricity producing wind turbines**. Genebra, Suíça, 2005.
- INTERNATIONAL ENERGY AGENCY- IEA. **World Energy**. 2015. Disponível em: <<https://www.iea.org/>>. Acesso em: 15 mai. 2018.
- JANG, J.-S. Anfis: adaptive-network-based fuzzy inference system. **Systems, Man and Cybernetics, IEEE Transactions on, IEEE**, v. 23, n. 3, p. 665–685, 1993.
- JANG, J.-S. R.; SUN, C.-T. **Neuro-fuzzy and soft computing: a computational approach to learning and machine intelligence**. Taiwan: Prentice-Hall, Inc., 1996.

- JAZAYERI, P.; ROSEHART, W.; WESTWICK, D. T. A multistage algorithm for identification of nonlinear aggregate power system loads. **IEEE Transactions on Power Systems**, IEEE, v. 22, n. 3, p. 1072–1079, 2007.
- JIN, T.; TIAN, Z. Uncertainty analysis for wind energy production with dynamic power curves. **IEEE 11th International Conference on Probabilistic Methods Applied to Power Systems**, p. 745–750, 2010.
- JOHANSEN, T. A.; SHORTEN, R.; MURRAY-SMITH, R. On the interpretation and identification of dynamic takagi-sugenofuzzy models. **IEEE Transactions on Fuzzy systems**, Institute of Electrical and Electronics Engineers (IEEE), v. 8, n. 3, p. 297–313, 2000.
- JOHNSON, W.; GEISSER, S. A predictive view of the detection and characterization of influential observations in regression analysis. **Journal of the American Statistical Association**, Taylor & Francis Group, v. 78, n. 381, p. 137–144, 1983.
- JOWDER, F. A. L. Wind power analysis and site matching of wind turbine generators in kingdom of bahrain. **Applied Energy**, v. 86, n. 4, p. 538–545, 2009.
- KEESMAN, K. **System Identification: An Introduction**. Springer London, 2011. (Advanced Textbooks in Control and Signal Processing). ISBN 9780857295224. Disponível em: <[https://books.google.com.br/books?id=gHssIP\\_dDwUC](https://books.google.com.br/books?id=gHssIP_dDwUC)>.
- KENNEDY, J. Particle swarm optimization. In: **Encyclopedia of machine learning**. Frankfurt: Springer, 2011. p. 760–766.
- KUSIAK, A.; ZHENG, H.; SONG, Z. On-line monitoring of power curves. **Renewable Energy**, v. 34, n. 6, p. 1487–1493, 2009.
- LEE, C. Fuzzy logic in control systems: fuzzy logic controller. i. **Systems, Man and Cybernetics, IEEE Transactions on**, v. 20, n. 2, p. 404–418, Mar 1990. ISSN 0018-9472.
- LEE, G. *et al.* Power curve estimation with multivariate environmental factors for inland and offshore wind farms. **Journal of the American Statistical Association**, Taylor & Francis, v. 110, n. 509, p. 56–67, 2015.
- LEGENDRE, A. M. **Nouvelles méthodes pour la détermination des orbites des comètes**. Paris: F. Didot, 1805.
- LEVENBERG, K. A method for the solution of certain non-linear problems in least squares. **Quarterly of applied mathematics**, v. 2, n. 2, p. 164–168, 1944.
- LI, S. *et al.* Comparative analysis of regression and artificial neural network models for wind turbine power curve estimation. **Journal of Solar Energy Engineering**, v. 123, n. 4, p. 327–332, 2001.
- LIKAS, A.; VLASSIS, N.; VERBEEK, J. J. The global k-means clustering algorithm. **Pattern Recognition**, Elsevier, v. 36, n. 2, p. 451–461, 2003.
- LYDIA, M. *et al.* A comprehensive review on wind turbine power curve modeling techniques. **Renewable and Sustainable Energy Reviews**, v. 30, p. 452–460, 2014.
- LYDIA, M. *et al.* Advanced algorithms for wind turbine power curve modeling. **IEEE Transactions on Sustainable Energy**, v. 4, n. 3, p. 827, 2013.

MANWELL, J.; MCGOWAN, J.; ROGERS, A. **Wind energy explained : theory, design, and application**. 2nd. ed. Massachusetts: John Wiley & Sons, Ltd., 2009.

MARCIUKAITIS, M. *et al.* Non-linear regression model for wind turbine power curve. **Renewable Energy**, v. 113, p. 732–741, 2017.

MARQUARDT, D. W. An algorithm for least-squares estimation of nonlinear parameters. **Journal of the society for Industrial and Applied Mathematics**, SIAM, v. 11, n. 2, p. 431–441, 1963.

MATHEW, S. **Wind Energy-Fundamentals, Resource Analysis and Economics**. Noruega: Springer, 2006.

MAYA, H. C. **Revisitando o Problema de Estimação de Curvas de Potência de Aerogeradores via Modelos Polinomiais Esparsos**. Dissertação (Mestrado) — Universidade Federal do Ceará, 2018.

MCLEAN, J. R. **WP2.6-Equivalent wind power curves**. Canada, 2008. Tech. Rep. EIE/06/022/SI2.442659.

MENDEL, J. M. Fuzzy logic systems for engineering: a tutorial. **Proceedings of the IEEE**, Ieee, v. 83, n. 3, p. 345–377, 1995.

MOURA A. P., S. T. M. . F. A. R. D. Impacto na energia hidroelétrica e nos níveis de tensão em um sistema de transmissão-distribuição conectado com parques eólicos sob regime de ventos diferentes. In: SNPTEE'2003. **Anais do XVII Seminário Nacional de Produção e Transmissão de Energia Elétrica**. Brasil, 2003. v. 1, p. 1–6.

MOZELLI, L. A. **CONTROLE FUZZY PARA SISTEMAS TAKAGI-SUGENO: CONDIÇÕES APRIMORADAS E APLICAÇÕES**. Dissertação (Mestrado) — UNIVERSIDADE FEDERAL DE MINAS GERAIS - UFMG, 2008.

OLAOFE, Z. O.; FOLLY, K. A. Wind energy analysis based on turbine and developed site power curves: a case-study of darling city. **Renewable Energy**, v. 53, p. 306–318, 2013.

OZER, D. J. Correlation and the coefficient of determination. **Psychological Bulletin**, American Psychological Association, v. 97, n. 2, p. 307, 1985.

POWELL, W. R. An analytical expression for the average output power of a wind machine. **Solar Energy**, v. 26, n. 1, p. 77–80, 1981.

RAMPINELLI, G. A.; JUNIOR, C. G. da R. Análise da geração eólica na matriz brasileira de energia elétrica. **RECEN-Revista Ciências Exatas e Naturais**, v. 14, n. 2, p. 273–302, 2013.

RIPLEY, B. Robust statistics. **Course notes**, 2004.

ROUSSEEUW, P. J.; CROUX, C. Alternatives to the median absolute deviation. **Journal of the American Statistical Association**, v. 88, n. 424, p. 1273–1283, 1993.

RUCKSTUHL, A. **Robust Fitting of Parametric Models Based on M-Estimation**. Maryland, 2016.

SANDRI, S.; CORREA, C. Lógica nebulosa. **Instituto Tecnológico da Aeronáutica-ITA, V Escola de Redes Neurais**, pp. C073-c090, São José dos Campos, 1999.

- SANTOS, J. D. de A. **Sobre a estimação do número de neurônios ocultos da rede MLP: uma nova técnica baseada em PCA e SVD**. Tese (Doutorado) — Universidade Federal do Ceará, 2010.
- SCHLECHTINGEN, M.; SANTOS, I. F.; ACHICHE, S. Using datamining approaches for wind turbine power curve monitoring: a comparative study. **IEEE Transactions on Sustainable Energy**, v. 4, n. 3, p. 671–679, 2013.
- SHOKRZADEH, S.; JOZANI, M. J.; BIBEAU, E. Wind turbine power curve modeling using advanced parametric and nonparametric methods. **IEEE Transactions on Sustainable Energy**, v. 5, n. 4, p. 1262–1269, 2014.
- SILVA, R. R. da; LIMA, J. M. de. Análise do tema energia e meio ambiente em livros didáticos de física: um norteador para a elaboração de projetos de sustentabilidade no eja. 2011.
- SILVA, T. M. V. e. **Impacto de Usinas Eólio-Elétricas em um Sistema de Subtransmissão - Distribuição em Regime Permanente**. Dissertação (Mestrado) — Universidade Federal do Ceará, 2003.
- SOHONI, V.; GUPTA, S.; NEMA, R. A critical review on wind turbine power curve modelling techniques and their applications in wind based energy systems. **Journal of Energy**, Hindawi, v. 2016, 2016.
- STEPHEN, B. *et al.* A copula model of wind turbine performance. **IEEE Transactions on Power Systems**, v. 26, n. 2, p. 965–966, 2011.
- SUGENO, M.; KANG, G. Structure identification of fuzzy model. **Fuzzy Sets and Systems**, v. 28, n. 1, p. 15–33, 1988.
- SUSANTI, Y. *et al.* M estimation, s estimation and mm estimation in robust regression. **International Journal of Pure and Applied Mathematics**, v. 91, n. 3, p. 349–360, 2014.
- TAKAGI, T.; SUGENO, M. Fuzzy identification of systems and its applications to modeling and control. **IEEE transactions on Systems, Man, and Cybernetics**, v. 15, n. 1, p. 116–132, 1985.
- TAKAGI, T.; SUGENO, M. Fuzzy identification of systems and its applications to modeling and control. **Systems, Man and Cybernetics, IEEE Transactions on**, IEEE, n. 1, p. 116–132, 1985.
- USTUNTAS, T.; SAHIN, A. D. Wind turbine power curve estimation based on cluster center fuzzy logic modeling. **Journal of Wind Engineering and Industrial Aerodynamics**, v. 96, n. 5, p. 611–620, 2008.
- WANG, J. *et al.* Analyzing the research subjects and hot topics of power system reliability through the Web of Science from 1991 to 2015. **Renewable and Sustainable Energy Reviews**, v. 82, n. 1, p. 700–713, 2018.
- ZADEH, L. A. Fuzzy sets. In: **Fuzzy Sets, Fuzzy Logic, And Fuzzy Systems: Selected Papers by Lotfi A Zadeh**. California: World Scientific, 1965. p. 394–432.
- ZIMMERMANN, H.-J. Fuzzy set theory. **Wiley Interdisciplinary Reviews: Computational Statistics**, Wiley Online Library, v. 2, n. 3, p. 317–332, 2010.

## APÊNDICE A – DISTRIBUIÇÃO DE PROBABILIDADE DA POTÊNCIA ELÉTRICA GERADA

### A.1 Função de Densidade de Probabilidade (FDP) de $p$

Seja  $f$  uma função de regressão<sup>1</sup> qualquer que relaciona a variável dependente  $p$  à variável independente  $v$  de acordo com os  $k_\beta$  coeficientes de regressão do vetor-coluna<sup>2</sup>  $\beta = [\beta_1 \ \beta_2 \ \dots \ \beta_{k_\beta}]^\top$ , por tanto

$$p = f(v|\beta) + \varepsilon \quad (\text{A.1})$$

onde  $\varepsilon$  é o residual da regressão que assumimos ser uma variável aleatória independente, de distribuição normal<sup>3</sup> com média nula e variância fixa igual a  $\sigma_\varepsilon^2$ .

Como  $f$  é uma função determinística e  $\varepsilon$  tem distribuição normal,  $p$  também terá distribuição normal e independente. Assim, o valor esperado de  $p$  é

$$\begin{aligned} E[p] &= E[f(v|\beta) + \varepsilon] \\ &= E[f(v|\beta)] + E[\varepsilon] \\ &= E[f(v|\beta)] + 0 \\ &= \mu \end{aligned} \quad (\text{A.2})$$

em que  $\mu$  é o valor esperado de potência elétrica, que se dá em função da velocidade esperada do vento. E a sua variância é

$$\begin{aligned} \text{Var}[p|v] &= \text{Var}[f(v|\beta) + \varepsilon] \\ &= \text{Var}[f(v|\beta)] + \text{Var}[\varepsilon] \\ &= 0 + \sigma_\varepsilon^2 \\ &= \sigma_\varepsilon^2 \\ &= \sigma^2 \end{aligned} \quad (\text{A.3})$$

Como variância de uma equação determinística é zero, a variância de  $p$ , denotada por  $\sigma^2$ , é igual a do residual.

<sup>1</sup> Trata-se de uma função que tem como saída o valor esperado de uma variável, conhecida como variável dependente, e como argumento uma ou mais variáveis, conhecidas como variáveis independentes.

<sup>2</sup> Assumimos que todos os vetores são vetores-coluna, salvo indicação contrária.

<sup>3</sup> É uma distribuição de probabilidade contínua, parametrizada pela sua esperança matemática  $\mu$  e seu desvio padrão  $\sigma$ . Sua PDF é definida por  $f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]$ .



A partir do que foi obtido preteritamente nessa sessão, podemos definir a FDP de  $p$ , que denotaremos por  $g$ , como

$$g(p|\boldsymbol{\theta}) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(p-\mu)^2}{2\sigma^2}\right] \quad (\text{A.4})$$

onde  $\boldsymbol{\theta} = (\boldsymbol{\beta}, \sigma^2)$  é seu vetor de parâmetros de dependência com cardinalidade  $k = k_\beta + 1$ .

Perceba que não temos como parâmetro de  $g$  o valor esperado da potência  $\mu$ . Isso ocorre pois  $\mu = E[f(v|\boldsymbol{\beta})]$ , assim é dependente de  $\boldsymbol{\beta}$ . Portanto, além dos coeficientes de regressão,  $g$  também depende da variância do residual para ser bem definido.