



UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE CENTRO DE TECNOLOGIA E GEOCIÊNCIAS
DEPARTAMENTO DE ENGENHARIA ELÉTRICA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

JONATA CAMPELO DE ALBUQUERQUE

**DESENVOLVIMENTO DE MODELOS DE CURVAS DE POTÊNCIA VIA IA PARA
PREVISÃO DE GERAÇÃO EÓLICA EM CURTO PRAZO**

Recife

2021

JONATA CAMPELO DE ALBUQUERQUE

**DESENVOLVIMENTO DE MODELOS DE CURVAS DE POTÊNCIA VIA IA PARA
PREVISÃO DE GERAÇÃO EÓLICA EM CURTO PRAZO**

Tese apresentada ao Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal de Pernambuco, Centro de Tecnologia e Geociências, como requisito parcial para a obtenção do título de Doutor em Engenharia Elétrica.

Área de concentração: Processamento de Energia.

Orientador: Prof. Dr. Ronaldo Ribeiro Barbosa de Aquino.

Coorientador: Profa. Dra. Milde Maria da Silva Lira.

Recife

2021

Catálogo na fonte:
Bibliotecária Sandra Maria Neri Santiago

- A345d Albuquerque, Jonata Campelo de.
Desenvolvimento de modelos de curvas de potência via IA para previsão de
geração eólica em curto prazo / Jonata Campelo de Albuquerque. – 2021.
206 f.: il., figs., tabs.
- Orientador: Prof. Dr. Ronaldo Ribeiro Barbosa de Aquino.
Coorientadora: Profa. Dra. Milde Maria da Silva Lira.
Tese (Doutorado) – Universidade Federal de Pernambuco. CTG. Programa de
Pós-Graduação em Engenharia Elétrica, Recife, 2021.
Inclui referências.
1. Engenharia elétrica. 2. Curva de potência. 3. Melhoria da previsão. 4.
Inferência pré-estabelecida. 5. Redes iniciais de aprendizagem profunda. 6.
Modelos de geração eólica. 7. Curto prazo. I. Aquino, Ronaldo Ribeiro Barbosa
de (Orientador). II. Lira, Milde Maria da Silva (Coorientadora). III. Título.
- UFPE
- 621.3 CDD (22. ed.) BCTG/2021-214

JONATA CAMPELO DE ALBUQUERQUE

**DESENVOLVIMENTO DE MODELOS DE CURVAS DE POTÊNCIA VIA IA PARA
PREVISÃO DE GERAÇÃO EÓLICA EM CURTO PRAZO**

Tese apresentada ao Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal de Pernambuco, Centro de Tecnologia e Geociências, como requisito parcial para a obtenção do título de Doutor em Engenharia Elétrica. Área de concentração: Processamento de Energia.

Aprovada em: 12/07/2021.

BANCA EXAMINADORA

Prof. Dr. Ronaldo Ribeiro Barbosa Aquino (Orientador)
Universidade Federal de Pernambuco

Prof. Dr. Pedro Andre de Carvalho Rosas (Examinador Interno)
Universidade Federal de Pernambuco

Profa. Dra. Milde Maria da Silva Lira (Examinadora Externa)
Universidade Federal de Pernambuco

Prof. Dr. Benemar Alencar de Souza (Examinador Externo)
Universidade Federal de Campina Grande

Prof. Dr. José Maurício de Barros Bezerra (Examinador Externo)
Universidade Federal de Pernambuco

Prof. Dr. Manoel Afonso de Carvalho Júnior (Examinador Externo)
Universidade Federal de Pernambuco

Dedico esse trabalho a todos aqueles que ao vislumbrar suas metas e sonhos, são gratos por isto, persistindo e acreditando até conseguirem a obtenção destes sem que para isso mitiguem ou cerceiem as oportunidades dos outros em seu meio de convivência. Abençoados são aqueles que quando crescem ensinam e desenvolvem todos em seu entorno.

AGRADECIMENTOS

Agradeço primeiramente a Deus pela humildade, paciência e inteligência concedidas, fatores primordiais para atingir qualquer meta descriminada na vida.

Agradeço imensamente a meus pais Ezequias e Sônia, por toda condição, amparo, amor, apoio e compreensão prestados a min desde o momento que nasci até hoje. “Posso dizer que sou muito privilegiado pelos pais que tenho”.

Agradeço aos meus dois Orientadores, Prof. Ronaldo Aquino e Profa. Milde Lira pela oportunidade, por todo apoio, parceria, orientação, intuição, direcionamento, paciência e motivação, cedidos para o desenvolvimento satisfatório deste trabalho. Além da Amizade e confiança dividida de forma coletiva entre todos nós. Eles que sem sombra de dúvidas são os arquitetos da minha trajetória.

Agradeço ao professor Otoni Nóbrega Neto pelos bons e práticos ensinamentos com o MATLAB, os quais têm sido bastante utilizados e necessários nas análises e arquiteturas de problemas de previsão. Aprendi a programar muito com o senhor Professor, e por isso, lhe sou muito grato.

Agradeço muito ao Prof. Manoel Afonso, pela primeira oportunidade em trabalhar no grupo de pesquisa Laboratório Digital de Sistemas de Potência (LDSP), fomentando o interesse pela heurística no desenvolvimento de Técnicas e artifícios que contribuem para o avanço da Engenharia Elétrica na Universidade Federal de Pernambuco. Posso dizer que tive o privilégio em trabalhar, ser orientado e aprender muito com o Mestre Afonso.

Agradeço ao Grupo do Laboratório Digital de Sistemas de Potência (LDSP) e a todos os seus professores membros, contribuindo sempre com o repasse de conhecimento ao longo da minha vida acadêmica.

Agradeço a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pela bolsa de estudos oferecida.

Agradeço a todos os meus amigos que me apoiaram de alguma forma ao longo do desenvolvimento deste trabalho.

“Se pudéssemos primeiro saber onde estamos e para onde estamos sendo levados, poderíamos julgar melhor o que fazer, e como fazê-lo” (LINCOLN *apud* ANSOFF, 1983, p. 54).

RESUMO

A modelagem de uma curva de potência é um fator crucial para o desempenho e operação de uma fonte eólica, contribuindo desde o controle até a melhora no desempenho de uma turbina ou parque. Atualmente existem várias técnicas para o ajuste deste fator, cujos métodos utilizados são conhecidos como paramétricos e não paramétricos. Estes últimos criam uma heurística própria, dependendo de característica ou tendência específica de como os dados de (ventos x Potência) de um parque ou turbina variam. Esta Tese visa demonstrar que é possível aumentar a exatidão da previsão de geração eólica em curto prazo com a utilização de algumas técnicas originais que utilizam a inferência pré-estabelecida em sistemas *fuzzy* e técnicas de *deep Learning* como *autoencoders* e redes de aprendizagem profunda do tipo LSTM. Foram desenvolvidas duas novas heurísticas explicadas no decorrer do trabalho e as mesmas foram comparadas com os modelos já desenvolvidos pelo autor em trabalhos anteriores. Os resultados mostram-se significativos tanto no que diz respeito a previsão de dados componentes do vento, como velocidade e direção, quanto a previsão de potência atrelada à aplicação dos modelos de curva de potência desenvolvidos. Desta forma os novos modelos e heurísticas desenvolvidos foram eficazes e inovadores contribuindo para a melhora notável do desempenho das previsões em modelos de geração eólica no curto prazo.

Palavras-chave: curva de potência; melhoria da previsão; inferência pré-estabelecida; redes iniciais de aprendizagem profunda; modelos de geração eólica; curto prazo.

ABSTRACT

The modeling of a power curve is a crucial factor for the performance and operation of a wind source, contributing from the control to the improvement in the performance of a turbine or park. There are currently several techniques for adjusting this factor, whose methods are known as parametric and non-parametric. The latter create their own heuristics, depending on a specific characteristic or trend of how the data (winds x power) of a wind farm or turbine varies. This Thesis aims to demonstrate that it is possible to increase the accuracy of wind generation forecast in the short term with the use of some original techniques that use pre-established inference in fuzzy systems and deep learning techniques such as autoencoders and deep learning networks of the LSTM type. Two new heuristics explained in the course of the work were developed and they were compared with the models already developed by the author in previous works. The results are significant both with regard to forecasting wind component data, such as speed and direction, as well as the power forecast linked to the application of the developed power curve models. Thus, the new models and heuristics developed were effective and innovative, contributing to a notable improvement in the performance of forecasts in wind generation models in the short term.

Keywords: power curve; prediction improvement; pre-established inference; initial deep learning networks; wind generation models; short term.

LISTA DE ILUSTRAÇÕES

Figura 1 – Empreendimentos de Geração em operação na Matriz Energética Brasileira.	23
Figura 2 – Empreendimentos de Geração em construção não iniciada na Matriz Energética.	23
Figura 3 – Empreendimentos de Geração em construção na Matriz Energética Brasileira.	23
Figura 4 – Discrepância entre amplitude do valor previsto em relação ao histórico.....	25
Figura 5 – Discrepância entre amplitude e posição no valor previsto em relação ao histórico.	26
Figura 6 – Demonstração qualitativa de uma curva de potência.	43
Figura 7 – Característica velocidade x Potência para um parque eólico.	44
Figura 8 – Inferência de regras em um sistema Fuzzy (Sugeno).	48
Figura 9 – Estrutura de Aprendizado do ANFIS.	50
Figura 10 – Fluxograma conceitual do ANFIS	53
Figura 11 – Ilustração de uma Restricted Boltzman Machine.....	55
Figura 12 – Ilustração de uma Deep Belief Network com duas camadas.	58
Figura 13 – Ilustração de uma Sigmoid Belief Network.	58
Figura 14 – Estrutura diabolô componente da arquitetura de um Autoencoder.....	59
Figura 15 – Ilustração da primeira camada do stacked autoencoder.	62
Figura 16 – Ilustração da segunda camada do stackedautoencoder.	63
Figura 17 – Ilustração da terceira camada do stacked autoencoder.	63
Figura 18 – Ilustração stacked autoencoder resultante do processo de aprendizagem.....	64
Figura 19 – Diagrama de uma rede neural recorrente.	65
Figura 20 – Significado de uma RNN (Rede Neural Recorrente).....	66
Figura 21 – Módulo de repetição padrão de uma RNN.....	68
Figura 22 – Módulo de repetição padrão de uma LSTM com 4 camadas.....	68
Figura 23 – Simbologias importantes do modelo de uma LSTM.	69
Figura 24 – Caminho de transporte de informação em uma LSTM.	69
Figura 25 – Estrutura do portão (Gate) de permissão da passagem de informação de uma LSTM.	71
Figura 26 – Estrutura do portão (Gate) de permissão da passagem de informação de uma LSTM.	72
Figura 27 – Adição de informação ao estado presente da Célula em uma LSTM.	73
Figura 28 – Passo de atualização da informação em uma LSTM.	74
Figura 29 – Processo de Decisão na construção da saída em uma LSTM	74

Figura 30 – Diagrama relativo ao processo de recorrência independente por entrada, inerente ao processo de decisão na construção da saída.	75
Figura 31 – Diagrama relativo ao processo de recorrência em uma LSTM GRU com acoplamento de entradas, inerente ao processo de decisão na construção da saída.....	76
Figura 32 – Diagrama relativo ao processo de recorrência em uma LSTM CGRU com acoplamento de entradas, inerente ao processo de decisão na construção da saída.....	77
Figura 33 – As várias abordagens de previsão podem ser classificadas de acordo com o tipo de entrada.	82
Figura 34 – O erro vem das previsões meteorológicas numéricas.	87
Figura 35 – Arquitetura da melhor RNA1 para o Parque 01 escolhida com a execução dos treinamentos.	95
Figura 36 – Arquitetura da melhor RNA1 para o Parque 02 escolhida com a execução dos treinamentos.	96
Figura 37 – Comportamento gráfico do MAE referente às simulações para o período de aprendizagem da RNA1 (a) aplicado ao Parque 01 e (b) aplicado ao Parque 02.	96
Figura 38 – Arquitetura da melhor RNA2 para o Parque 01 escolhida com a implementação dos treinamentos.....	98
Figura 39 – Arquitetura da melhor RNA2 para o Parque 02 escolhida com a implementação dos treinamentos.....	99
Figura 40 – Comportamento gráfico do MAE da velocidade média referente às simulações para o período de aprendizagem da RNA2 aplicado em (a) para Parque 01 e em (b) para o Parque 02.	99
Figura 41 – Comportamento gráfico do MAE da direção média referente às simulações para o período de aprendizagem da RNA2 aplicado em (a) para Parque 01 e em (b) para o Parque 02.....	100
Figura 42 – Comportamento gráfico de uma sequência ruidosa.	102
Figura 43 – Comportamento gráfico da tendência de média móvel (momentum) em uma sequência ruidosa.	103
Figura 44 – Variação comportamental da tendência do momentum para três diferentes valores do hiperparâmetro beta.....	103

Figura 45 – Comportamento gráfico do MAE da direção média referente às simulações para o período de aprendizagem da RNA2 aplicado em (a) para Parque 01 e em (b) para o Parque 02.....	106
Figura 46 – Comportamento gráfico do movimento de convergência do gradiente em um processo iterativo de otimização.	107
Figura 47 – Comportamento planar variacional do movimento de convergência do gradiente em um processo iterativo de otimização.	107
Figura 48 – Comportamento variacional otimizado do movimento de convergência do gradiente em um processo iterativo com atualização dos pesos regulada pelo algoritmo backpropagation.....	108
Figura 49 – Arquitetura genérica das LSTMs criadas para os Parques sobre estudo.....	109
Figura 50 – Arquitetura da melhor LSTM1 para o Parque 01 escolhida com a execução dos treinamentos.	110
Figura 51 – Arquitetura da melhor LSTM1 para o Parque 02 escolhida com a execução dos treinamentos.	110
Figura 52 – Comportamento gráfico do MAE referente às simulações para o período de aprendizagem da LSTM1 (a) aplicado ao Parque 01 e (b) aplicado ao Parque 02.	111
Figura 53 – Arquitetura da melhor LSTM2 para o Parque 01 escolhida com a implementação dos treinamentos.....	112
Figura 54 – Arquitetura da melhor LSTM2 para o Parque 02 escolhida com a implementação dos treinamentos.....	112
Figura 55 – Comportamento gráfico do MAE da velocidade média referente às simulações para o período de aprendizagem da LSTM2 aplicado em (a) para Parque 01 e em (b) para o Parque 02.	113
Figura 56 – Comportamento gráfico do MAE da direção média referente às simulações para o período de aprendizagem da LSTM2 aplicado em (a) para Parque 01 e em (b) para o Parque 02.....	113
Figura 57 – Arquitetura da heurística de reconstrução proposta.	116
Figura 58 – Arquitetura das Topologias AET1 e AET2 para os Parques 01 e 02 escolhidas durante o período de aprendizagem.	117
Figura 59 – Comparação entre histórico do conjunto de aprendizagem e a sua reconstrução proposta pela Topologia do AET1 para os parques 01 (a) e 02 em (b).....	119

Figura 60 – Arquitetura da melhor RNA1_AET1 para o Parque 01 escolhida com a execução dos treinamentos.....	120
Figura 61 – Arquitetura da melhor RNA1_AET1 para o Parque 02 escolhida com a execução dos treinamentos.....	120
Figura 62 – Comportamento gráfico do MAE referente às simulações para o período de aprendizagem da RNA1 (a) aplicado ao Parque 01 e (b) aplicado ao Parque 02.	121
Figura 63 – Arquitetura da melhor RNA1_AET2 para o Parque 01 escolhida com a execução dos treinamentos.....	122
Figura 64 – Arquitetura da melhor RNA1_AET2 para o Parque 02 escolhida com a execução dos treinamentos.....	122
Figura 65 – Comportamento gráfico do MAE referente às simulações para o período de aprendizagem da RNA1_AET2 (a) aplicado ao Parque 01 e (b) aplicado ao Parque 02.....	122
Figura 66 – Arquitetura da melhor RNA2 AET2 para o Parque 01 escolhida com a implementação dos treinamentos.	124
Figura 67 – Arquitetura da melhor RNA2 AET2 para o Parque 02 escolhida com a implementação dos treinamentos.	124
Figura 68 – Comportamento gráfico do MAE da velocidade média referente às simulações para o período de aprendizagem da RNA2 AET2 aplicado em (a) para Parque 01 e em (b) para o Parque 02.....	125
Figura 69 – Comportamento gráfico do MAE da direção média referente às simulações para o período de aprendizagem da RNA2 AET2 aplicado em (a) para Parque 01 e em (b) para o Parque 02.	125
Figura 70 – Composição das arquiteturas das LSTM adequadas pelas topologias dos autoencoders para os Parques 01 e 02.....	127
Figura 71 – Arquitetura da melhor LSTM1 AET1 para o Parque 01 escolhida com a execução dos treinamentos.....	128
Figura 72 – Arquitetura da melhor LSTM1 AET1 para o Parque 02 escolhida com a execução dos treinamentos.....	128
Figura 73 – Comportamento gráfico do MAE referente às simulações para o período de aprendizagem da LSTM1 AET1 (a) aplicado ao Parque 01 e (b) aplicado ao Parque 02.....	129

Figura 74 – Arquitetura da melhor LSTM2 AET2 para o Parque 01 escolhida com a execução dos treinamentos.....	130
Figura 75 – Arquitetura da melhor LSTM2 AET2 para o Parque 02 escolhida com a execução dos treinamentos.....	131
Figura 76 – Comportamento gráfico do MAE referente às simulações para o período de aprendizagem da LSTM2_AET2 (a) aplicado ao Parque 01 e (b) aplicado ao Parque 02.....	131
Figura 77 – Comportamento gráfico do MAE da direção média referente às simulações para o período de aprendizagem da LSTM2_AET2 aplicado em (a) para Parque 01 e em (b) para o Parque 02.	132
Figura 78 – Aproximações polinomiais relativas às curvas de potências médias no (a) Parque 01 e no (2) Parque 02.	135
Figura 79 – Arquitetura da CP1RNA para o Parque 01 escolhida no período de aprendizagem.	136
Figura 80 – Arquitetura da CP1RNA para o Parque 02 escolhida no período de aprendizagem.	136
Figura 81 – Comparação entre CP Real e a curva CPRNA1 treinada para o (a) Parque 01e (b) para o Parque 02.....	137
Figura 82 – Arquitetura da CP2RNA para o Parque 01 escolhida no período de aprendizagem.	138
Figura 83 – Arquitetura da CP2RNA para o Parque 02 escolhida no período de aprendizagem.	138
Figura 84 – Comparação entre CP Real e a curva CP2RNA treinada para o (a) Parque 01e (b) para o Parque 02.....	138
Figura 85 – Arquitetura da CP1RNA AET1 para o Parque 01 escolhida no período de aprendizagem.	140
Figura 86 – Arquitetura da CP1RNA AET1 para o Parque 02 escolhida no período de aprendizagem.	140
Figura 87 – Comparação entre CP Real e a curva CPRNA1_AET1 treinada para o (a) Parque 01e (b) para o Parque 02.	140
Figura 88 – Arquitetura da CP1RNA AET2 para o Parque 01 escolhida no período de aprendizagem.	141
Figura 89 – Arquitetura da CP1RNA AET2 para o Parque 02 escolhida no período de aprendizagem.	142

Figura 90 – Comparação entre CP Real e a curva CPRNA1_AET2 treinada para o (a) Parque 01e (b) para o Parque 02.	142
Figura 91 – Arquitetura da CP2RNA AET2 para o Parque 01 escolhida no período de aprendizagem.	143
Figura 92 – Arquitetura da CP2RNA AET2 para o Parque 02 escolhida no período de aprendizagem.	143
Figura 93 – Comparação entre CP Real e a curva CP2RNA AET2 treinada para o (a) Parque 01e (b) para o Parque 02.	144
Figura 94 – Arquitetura do SIF Fuzzy, Mamdani, referente à CP1Fuzzy.....	145
Figura 95 – CP1Fuzzy resultante do processo de aprendizagem relativa a o (a) Parque 01 e ao (b) Parque 02.	145
Figura 96 – Arquitetura do SIF Fuzzy (Mamdani), referente a CP2Fuzzy.	146
Figura 97 – CP2Fuzzy resultante do processo de aprendizagem relativa ao (a) Parque 01 e ao (b) Parque 02.	146
Figura 98 – CP1Fuzzy (Sugeno-ANFIS) resultante do processo de aprendizagem relativa aos Parques 01 e 02.	148
Figura 99 – CP1Fuzzy (Sugeno ANFIS) resultante do processo de aprendizagem relativa a o (a) Parque 01 e ao (b) Parque 02.	148
Figura 100 – CP2Fuzzy (Sugeno-ANFIS) resultante do processo de aprendizagem relativa a o (a) Parque 01 e ao (b) Parque 02.	149
Figura 101 – CP2Fuzzy (Sugeno ANFIS) resultante do processo de aprendizagem relativa a o (a) Parque 01 e ao (b) Parque 02.	150
Figura 102 – Arquitetura Genérica de composição dos Modelos Previsores de Potência definida na Tabela 5 adotada para os Parques 01 e 02.....	153
Figura 103 – Arquitetura da RNA1em (a) e RNA2 em (b) para os parques 01 e 02.	157
Figura 104 – Arquitetura da RNA1 AET1 ou RNA AET2 em (a) e RNA2 AET2 em (b) para os parques 01 e 02.	157
Figura 105 – Arquitetura da LSTM1 AET1 (a) e LSTM2 AET2 em (b) para os parques 01 e 02.....	158
Figura 106 – Alimentação de variáveis no fluxograma genérico de ação na arquitetura do modelo do Previsor geral de potências adotado para os parques 01 e 02.	158
Figura 107 – Comparação do MAE de velocidade em (m/s) do Conjunto de Simulação para os modelos previsores de dados de vento com o modelo da Persistência atuantes (a) no Parque 01 e (b) no Parque 02.	168

Figura 108 – Comparação do MAE de direção em (°) do Conjunto de Simulação para os modelos previsores de dados de vento com o modelo da Persistência atuantes (a) no Parque 01 e (b) no Parque 02.	168
Figura 109 – Comparação do MAE de velocidade em (m/s) do Conjunto de Simulação para os modelos previsores de dados de vento sem o modelo da Persistência atuantes (a) no Parque 01 e (b) no Parque 02.	169
Figura 110 – Comparação do MAE de direção em (°) do Conjunto de Simulação para os modelos previsores de dados de vento sem o modelo da Persistência atuantes (a) no Parque 01 e (b) no Parque 02.	169
Figura 111 – MAE do Conjunto de Simulação para os modelos de curvas de Potência atuantes (a) no Parque 01 e (b) no Parque 02.	174
Figura 112 – NMAE do Conjunto de Simulação para os modelos de curvas de Potência atuantes (a) no Parque 01 e (b) no Parque 02.	174
Figura 113 – RMSE do Conjunto de Simulação para os modelos de curvas de Potência atuantes (a) no Parque 01 e (b) no Parque 02.	175
Figura 114 – Comparação do MAE de Potência em (MW) do Conjunto de Simulação para os modelos previsores de dados de vento com o modelo da Persistência, atuantes (a) no Parque 01 e (b) no Parque 02.	190
Figura 115 – Comparação do MAE de Potência em (MW) do Conjunto de Simulação para os modelos previsores de dados de vento sem o modelo da Persistência, atuantes (a) no Parque 01 e (b) no Parque 02.	191
Figura 116 – Comparação do NMAE(%) de Potência do Conjunto de Simulação para os modelos previsores de dados de vento com o modelo da Persistência atuantes (a) no Parque 01 e (b) no Parque 02.	191
Figura 117 – Comparação do NMAE(%) de Potência do Conjunto de Simulação para os modelos previsores de dados de vento sem o modelo da Persistência atuantes (a) no Parque 01 e (b) no Parque 02.	192

LISTA DE TABELAS

Tabela 1 – Codificação de legenda para as respectivas usinas e centrais energéticas.....	22
Tabela 2 – Parâmetros escolhidos por parque de acordo com esta rotina de aprendizagem, para o melhor modelo de AE de acordo com a topologia adotada AET1 ou AET2 nos Parques 01 e 02.....	118
Tabela 3 – Desempenho das previsões no período de Simulação das redes (LSTM1, LSTM2, LSTM1AET1, LSTM2AET2) para o Parque 01.	154
Tabela 4 – Desempenho das previsões no período de Simulação das redes (LSTM1, LSTM2, LSTM1AET1, LSTM2AET2) para o Parque 02.	155
Tabela 5 – Composição de formação dos Modelos previsores de potência desenvolvidos. ..	156
Tabela 6 – Codificação dos Modelos previsores de dados de vento desenvolvidos.	160
Tabela 7 – Desempenho no período de Simulação dos modelos previsores de dados de vento para as 12h iniciais do horizonte para o Parque 01.	162
Tabela 8 – Desempenho no período de Simulação dos modelos previsores de dados de vento para as 12h finais do horizonte para o Parque 01.	163
Tabela 9 – Desempenho no período de Simulação dos modelos previsores de dados de vento para as 12h iniciais do horizonte para o Parque 02.	165
Tabela 10 – Desempenho no período de Simulação dos modelos previsores de dados de vento para as 12h finais do horizonte para o Parque 02.	166
Tabela 11 – Tabela resumo das médias de MAE para 12h e 24h referentes a análise do Conjunto de Simulação para os modelos previsores de dados de vento no Parque 01.	170
Tabela 12 – Tabela resumo das médias de MAE para 12h e 24h referentes a análise do Conjunto de Simulação para os modelos previsores de dados de vento no Parque 02.	170
Tabela 13 – Aferição dos Ganhos percentuais relativos aos modelos previsores de dados de vento atuantes no Parque 01 em relação ao modelo da persistência.	171
Tabela 14 – Aferição dos Ganhos percentuais relativos aos modelos previsores de dados de vento atuantes no Parque 02 em relação ao modelo da persistência.	171
Tabela 15 – Codificação dos Modelos de curvas de potência desenvolvidos.	172
Tabela 16 – Aferição dos Desempenhos relativos às curvas de potência atuantes no parque 01.	172

Tabela 17 – Aferição dos Desempenhos relativos às curvas de potência atuantes no parque 02.	172
Tabela 18 – Aferição dos Ganhos percentuais relativos aos modelos de curvas de potência atuantes no parque 01 em relação ao modelo de referência.	175
Tabela 19 – Aferição dos Ganhos percentuais relativos aos modelos de curvas de potência atuantes no parque 02 em relação ao modelo de referência.	175
Tabela 20 – Desempenho em MW dos Modelos previsores nas primeiras 12h do período de Simulação para o Parque 01	177
Tabela 21 – Desempenho em MW dos Modelos previsores nas primeiras 12h do período de Simulação para o Parque 02	178
Tabela 22 – Desempenho em MW dos Modelos previsores nas 12h finais do período de Simulação para o Parque 01	179
Tabela 23 – Desempenho em MW dos Modelos previsores nas 12h finais do período de Simulação para o Parque 02	180
Tabela 24 – Desempenho percentual dos Modelos previsores nas primeiras 12h do período de Simulação para o Parque 01	181
Tabela 25 – Desempenho percentual dos Modelos previsores nas primeiras 12h do período de Simulação para o Parque 02	182
Tabela 26 – Desempenho percentual dos Modelos previsores nas 12h finais do período de Simulação para o Parque 01	183
Tabela 27 – Desempenho percentual dos Modelos previsores nas 12h finais do período de Simulação para o Parque 02	184
Tabela 28 – Aferição dos Ganhos percentuais relativos aos índices de MAE dos modelos previsores de potência atuantes nos parques 01 e 02 em relação ao modelo da persistência.	188
Tabela 29 – Aferição dos Ganhos percentuais relativos aos índices de NMAE dos modelos previsores de potência atuantes nos parques 01 e 02 em relação ao modelo da persistência.	189

SUMÁRIO

1	INTRODUÇÃO.....	21
1.1	Descrição do problema	24
1.2	Justificativa	26
1.3	Objetivos.....	27
1.3.1	OBJETIVOS GERAIS.....	27
1.3.2	OBJETIVOS ESPECÍFICOS	28
1.4	Trabalhos publicados	29
1.5	Estrutura do Trabalho	29
2	REVISÃO BIBLIOGRÁFICA	31
3	CURVA DE POTÊNCIA	42
4	FERRAMENTAS DE INTELIGÊNCIA ARTIFICIAL	45
4.1	Redes Neurais Artificiais.....	45
4.2	Sistemas de Inferência Fuzzy.....	46
4.2.1	SISTEMA DE INFERÊNCIA FUZZY MAMDANI	46
4.2.2	SISTEMAS DE INFERÊNCIA FUZZY SUGENO.....	47
4.3	Adaptative Neuro Fuzzy Inference System (ANFIS)	49
4.4	Deep Learning.....	53
4.4.1	REDES USADAS EM ARQUITETURAS DE APRENDIZAGEM PROFUNDA	54
4.4.2	REDES NEURAIIS RECORRENTES	64
4.4.3	REDES LSTM (LONG SHORT-TERM MEMORY).....	67
5	REVISÃO DO ESTADO DA ARTE SOBRE A PREVISÃO DE GERAÇÃO EÓLICA.....	79
5.1	Horizontes de Previsão	80
5.2	Modelos Típicos Utilizados	81
5.3	Modelos de séries temporais	83
5.4	Modelos de séries temporais para previsão de prazos muito curtos.....	85
5.5	Modelagem meteorológica para energia eólica	86
5.6	Modelos de previsão de curto prazo com NWPs.....	87
5.7	Modelagem da velocidade do vento versus energia eólica	89
5.8	Os projetos ANEMOS e outros grandes projetos empresariais de atividades P & D	90
5.9	Conclusões	92

6	MODELOS PREVISORES DE VELOCIDADE DO VENTO E VELOCIDADE E DIREÇÃO DO VENTO, MODELOS DE CURVAS DE POTÊNCIA E MODELO PREVISOR DE GERAÇÃO EÓLICA PROPOSTOS.....	93
6.1	Modelos previsores de velocidade do vento e modelos previsores de velocidade e direção do vento	93
6.1.1	HEURÍSTICA MULTI LAYER PERCEPTRON (MLP)	93
6.1.2	HEURÍSTICA DEEP LEARNING LONG SHORT-TERM MEMORY (LSTM).....	101
6.1.3	HEURÍSTICA MULTI LAYER PERCEPTRON ADEQUADA POR AUTOENCODER (AE-MLP).....	114
6.1.4	HEURÍSTICA DEEP LEARNING LONG SHORT-TERM MEMORY ADEQUADA POR AUTOENCODER (LSTM-AE).....	126
6.2	Metodologia utilizada para os Modelos de curvas de Potência Eólica (WTPCs) descritos nesta tese	132
6.3	Modelo de Referência - Curva de potência Média (Polinomial)	134
6.4	Modelos não paramétricos de Curvas de Potência utilizando RNAs.....	135
6.4.1	MODELO DA CP1RNA	136
6.4.2	MODELO DA CP2RNA:	137
6.5	Modelos não paramétricos de Curvas de Potência utilizando RNAs adequadas por Autoencoder	139
6.5.1	MODELO DA CP1-RNA RETREINADA COM DADOS FILTRADOS PELO AE TOPOLOGIA AET1.....	139
6.5.2	MODELO DA CP1-RNA RETREINADA COM DADOS FILTRADOS PELO AE TOPOLOGIA AET2.....	141
6.5.3	MODELO DA CP2-RNA RETREINADA COM DADOS FILTRADOS PELO AE TOPOLOGIA AET2.....	142
6.6	Modelos não paramétricos de Curvas de Potência utilizando Lógica Fuzzy .	144
6.6.1	MODELO DA CP1FUZZY (MAMDANI)	144
6.6.2	MODELO DA CP2FUZZY (MAMDANI)	145
6.6.3	MODELO DA CP1FUZZY (SUGENO-ANFIS)	147
6.6.4	MODELO DA CP2FUZZY (SUGENO-ANFIS)	149
6.7	Modelos Previsores de Potência	150
6.7.1	ESCOLHA DOS PREVISORES LSTM UTILIZADOS NA COMPOSIÇÃO DA ARQUITETURA DOS MODELOS PREVISORES DE POTÊNCIA	151
7	ANÁLISE GERAL E RESULTADOS.....	159

7.1	Discussão dos resultados	159
7.1.1	DISCUSSÃO DOS RESULTADOS RELATIVOS AOS PREVISORES DE DADOS DE VENTO	160
7.1.2	DISCUSSÃO DOS RESULTADOS RELATIVOS AOS MODELOS APROXIMADORES DE CURVAS DE POTÊNCIA	172
8	CONCLUSÕES E TRABALHOS FUTUROS	193
8.1	Conclusões	193
8.2	Propostas para trabalhos futuros.....	195

1 INTRODUÇÃO

Devido ao crescimento da demanda energética no Brasil, a fonte eólica tem sido uma saída bastante plausível, porém, esta possui uma característica base, a variabilidade dos ventos, o que estimula a concepção de novas heurísticas voltadas para a previsão de energia.

Numa análise global o Brasil melhorou uma posição, passando o Canadá, ocupando o oitavo lugar no ranking mundial no que se refere a capacidade instalada de produção de energia eólica, de acordo com o (Global Wind Statistic ,2017), um documento anual que visa resumir de forma estatística os dados mundiais de energia eólica produzido pelo *Global Wind Energy Council* (GWEC).

Em 2017, o gigante sul-americano adicionou cerca de 52,57 GW de potência eólica à produção mundial, totalizando no globo 539,58 GW de capacidade instalada”, de acordo com a Associação Brasileira de Energia Eólica (ABEEOLICA, 2018), que reúne empresas do setor. Já no ano anterior, ultrapassou a Itália chegando a 9ª posição neste ranking. Em 2017, o país projetou contar com 12,76 GW de capacidade de energia instalada, contra os 12,39 GW do Canadá (Global Wind Statistic, 2017).

Liderando este ranking está a China, com 188,23 GW; seguida por: Estados Unidos, com 89,07 GW, Alemanha, com 56,132 GW de capacidade instalada. Além da Índia, da Espanha, do Reino Unido e da França, sendo estes países citados os sete primeiros do ranking.

Para alcançar a perspectiva aproximada de 12,8 GW, o número de Parques eólicos cresceu consideravelmente nos últimos anos, passando de uma capacidade eólica fiscalizada de 6.629MW em 2015 para patamares de 12.509,74MW em 2018 e conseguindo ultrapassar o objetivo logrado em 2017, atualmente 17.247,26MW. Outro fator importante é o crescimento do número de projetos de geração eólica, em operação, com fins de comercialização, aumentando de 273 projetos em 2015, para 510 em 2018 e nos últimos 2 anos para 690. Tal crescimento impacta de forma bastante efetiva na capacidade de geração do país, melhorando o crescimento da matriz eólica representado por um patamar de 4,78% da capacidade instalada em 2015, evoluindo para 7,9% em 2018 e crescendo para 9,87% em fevereiro de 2021 (ANEEL, 2021).

Por outro lado, existem projetos em elaboração relativos a uma geração outorgada de 4.282,16 MW extras que entrarão em operação nos próximos anos viabilizando um acréscimo de 36,48% da capacidade instalada atual (ANEEL, 2021).

A Região Nordeste do Brasil aparece liderando a capacidade de produção de energia a partir dos ventos no país. Dentre os seus estados o que mais produz energia eólica é o Rio

Grande do Norte com 135 parques, que integralizam capacidade instalada de 3.678,85 MW. Em seguida está a Bahia, com 93 parques e 2.410,04 MW de capacidade instalada. Em terceiro lugar vem o Ceará, com 74 parques e 1.935,76 MW de capacidade instalada.

Em quarto lugar no ranking nacional está o maior produtor da região sul, o Rio Grande do Sul. O estado tem 80 parques e 1.831,87 MW de capacidade instalada. Em quinto, voltando para a região nordeste, está o Piauí, com 52 parques e 1.443,10 MW instalados, e em sexto Pernambuco com 34 parques e 781,99 MW de capacidade instalada.

Tomando-se por base o ano de 2018, segundo a ABEOLICA a perspectiva é que até 2024 sejam adicionados apenas mais 1,45GW de capacidade eólica no país, este aumento seria decorrente de leilões de energia realizados em dezembro de 2017. Fato que prospecta que 18 milhões de residências sejam abastecidas. Este pequeno crescimento reduziu o ranking brasileiro para este tipo de geração passando de 2,02GW, 6ª posição na escala global em 2016 para a 7ª posição na escala global em 2017. Esta tendência de queda pode ser mais pronunciada nos próximos anos pois houve menos projetos durante o biênio 2019, 2020.

A Tabela 1 juntamente com as Figuras de 1 a 3 mostram, de forma codificada e em números, os patamares de geração das matrizes energéticas existentes que constroem a matriz energética brasileira. Sendo os valores percentuais presentes nas Figuras, referentes à Potência Fiscalizada, a Potência Outorgada igual à considerada no Ato de Outorga e a Potência Fiscalizada igual a considerada a partir da operação comercial da primeira unidade geradora.

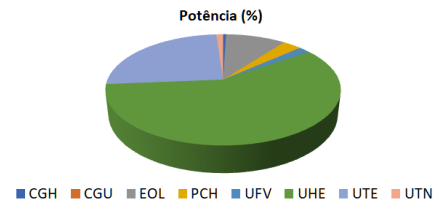
Tabela 1 – Codificação de legenda para as respectivas usinas e centrais energéticas

Legendas	
CGH	Central Geradora Hidrelétrica
CGU	Central Geradora Undi-elétrica
EOL	Central Geradora Eólica
PCH	Pequena Central Hidrelétrica
UFV	Central Geradora Solar Fotovoltaica
UHE	Usina Hidrelétrica
UTE	Usina Termelétrica
UTN	Usina Termonuclear

Fonte: Banco de Informações da Geração - BIG Aneel (ANEEL 2021).

Figura 1 – Empreendimentos de Geração em operação na Matriz Energética Brasileira.

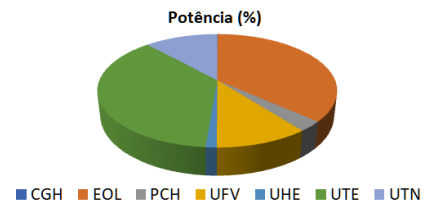
Empreendimentos em Operação				
Tipo	Quantidade	Potência Outorgada (Kw)	Potência Fiscalizada (kW)	%
CGH	739	827.479,59	827.479,59	0,47
CGU	1	50	50	0
EOL	690	17.411.958,86	17.247.262,86	9,87
PCH	424	5.493.802,57	5.451.181,57	3,11
UFV	4.055	3.292.423,25	3.288.423,25	1,87
UHE	219	102.990.428,00	103.026.876,00	58,36
UTE	3.071	44.457.852,89	42.954.895,39	25,19
UTN	2	1.990.000,00	1.990.000,00	1,13
Total	9.201,00	176.463.995,16	174.786.168,66	100



Fonte: Banco de Informações da Geração - BIG Aneel (ANEEL 2021).

Figura 2 – Empreendimentos de Geração em construção não iniciada na Matriz Energética.

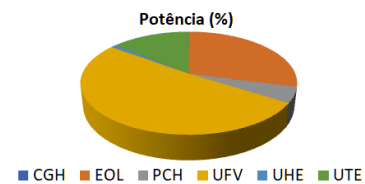
Empreendimentos em Construção			
Tipo	Quantidade	Potência Outorgada (Kw)	%
CGH	4	7.112,00	0,06
EOL	132	4.182.165,00	36,48
PCH	24	329.811,10	2,88
UFV	33	1.219.220,68	10,63
UHE	1	141.900,00	1,24
UTE	60	4.235.428,50	36,94
UTN	1	1.350.000,00	11,77
Total	255,00	11.465.637,28	100



Fonte: Banco de Informações da Geração - BIG Aneel (ANEEL 2021).

Figura 3 – Empreendimentos de Geração em construção na Matriz Energética Brasileira.

Empreendimentos com construção não iniciada			
Tipo	Quantidade	Potência Outorgada (Kw)	%
CGH	3	7.100,00	0,03
EOL	205	7.900.210,00	29,04
PCH	95	1.333.991,95	4,9
UFV	352	13.909.616,27	51,14
UHE	3	262.000,00	0,96
UTE	41	3.788.201,00	13,93
Total	699,00	27.201.119,22	100



Fonte: Banco de Informações da Geração - BIG Aneel (ANEEL 2021).

Os termos e siglas apresentadas na Tabela 1 e nas Figuras de 1 a 3 são explicados e definidos pelo Banco de Informações da Geração fornecido pela ANEEL.

De forma a permitir a regularidade no processo de balanço energético e acoplamento de unidades de geração eólica no sistema energético brasileiro objetivando mitigar possíveis impactos que alterem a estabilidade de funcionamento do sistema. É viável a discussão relativa à predição da geração em um horizonte determinado, que permita ao sistema escolher possíveis diretrizes de controle e capacidade de geração, condizentes com o valor demandado pela matriz energética em um dado momento.

Como a matriz eólica é peculiar quanto à variabilidade de sua grandeza base, para esta predição são indicados modelos previsores de geração de curto prazo. Investigando novas possibilidades, estão os Modelos previsores de Potência, sendo estes atrelados a parques ou máquinas eólicas. Em que o enfoque pode ser paramétrico como em RENANI, et. al., (2016) ou não paramétrico como em BAPTISTA, (2017).

Em se tratando de modelos não paramétricos destacam-se os modelos previsores que tem como tendência de ajuste, a curva de potência. Para o ajuste da curva de potência no modelo predictor são utilizadas ferramentas de inteligência artificial, bastante difundidas como: redes neurais artificiais (LUDERMIR, et. al. 2000), sistemas de inferência *fuzzy* (KOSKO, 1997), além de novas tendências de estudo como os modelos de *deep learning* (HINTON, et. al. 2006). Tendo essas ferramentas um papel essencial para agregar exatidão à resposta dos modelos previsores.

Esta tese propõe novas técnicas de desenvolvimento de modelos não paramétricos aproximadores de curvas de potência, empregando ferramentas de inteligência artificial, fortemente difundidas, cujo caráter aplicativo servirá para o controle de potência em aerogeradores. Para isso, foram utilizados como pontos de partida, estudos realizados em parques eólicos situados na região nordeste do Brasil, citados e avaliados por ALBUQUERQUE (2015), AQUINO, et. al. (2016) e AQUINO, et. al. (2017).

As principais contribuições dos estudos desenvolvidos nesta tese estão na proposição, criação e avaliação de novas formas e arquiteturas de modelos previsores e de curvas de potência utilizando ferramentas de IA que visem uma melhor exatidão tanto no que diz respeito a previsão de dados de vento quanto a conversão dos valores previstos de dados de vento em valores de potência através dos modelos de curvas de potência.

Uma destas criações, inova por propor um “Pré-treinamento” advindo das características provenientes da inferência do especialista seguido de uma transformação de um sistema *fuzzy* Mamdani já existente em um sistema *fuzzy* Sugeno, em que após a transformação a fase de aprendizagem é iniciada. Além do uso de redes iniciais de aprendizagem profunda para melhorar a exatidão de modelos de previsão de geração eólica através da simulação de curvas de potência mais exatas provenientes de bases reconstruídas.

1.1 Descrição do problema

A utilização de elementos de inteligência artificial como RNA e SIF, já é um assunto bastante difundido em previsão de geração eólica, pois existem vários trabalhos e publicações com esse enfoque. Porém, um problema notório é o erro relativo ao valor previsto de geração, que nestes sistemas apresenta um caráter de erro fixo considerável nas primeiras horas de previsão e aumenta para valores inaceitáveis nas horas finais previstas. Esse erro característico dependerá de cada sítio sob análise já que o erro da predição de geração eólica

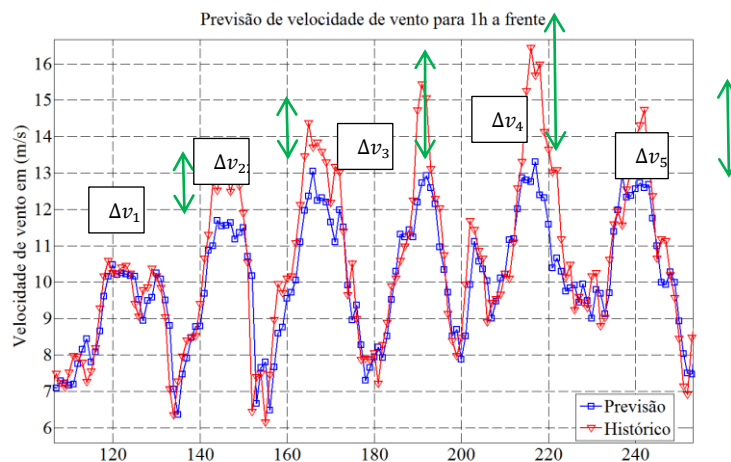
está fortemente ligado à característica (vento x potência), isto é, a curva de potência do parque em análise.

De forma a conter esta característica é necessária à procura por um tipo de solução que objetive diminuir este caráter de erro, além de trazer os erros nas últimas horas previstas para valores aceitáveis em termos de potência instalada. Assim o erro relativo à previsão diminuirá provocando uma melhora significativa no desempenho dos modelos de previsão de dados eólicos, velocidade e direção, e em sistemas previsores de geração.

Outro ponto importante é com relação à variabilidade da grandeza eólica abordada como em BITAR, E. Y., et. al., (2012) e WANG, D. et. al., (2015) de maneira bastante contundente em várias publicações, é a ocorrência da defasagem temporal entre as tendências do histórico da base de dados e os valores previstos, causando uma baixa no desempenho dos modelos previsores. Em alguns casos, esta defasagem é pequena, porém a faixa de variação de amplitude entre o valor previsto e o valor real do histórico pode sofrer discrepâncias em um determinado instante do horizonte, contribuindo para o baixo desempenho do previsor.

A Figura 4 mostra o que é descrito no parágrafo anterior no que diz respeito à discrepância na previsão de vento com 1h a frente, na qual se pode perceber que a tendência do histórico é mantida, porém as amplitudes previstas podem apresentar incoerências.

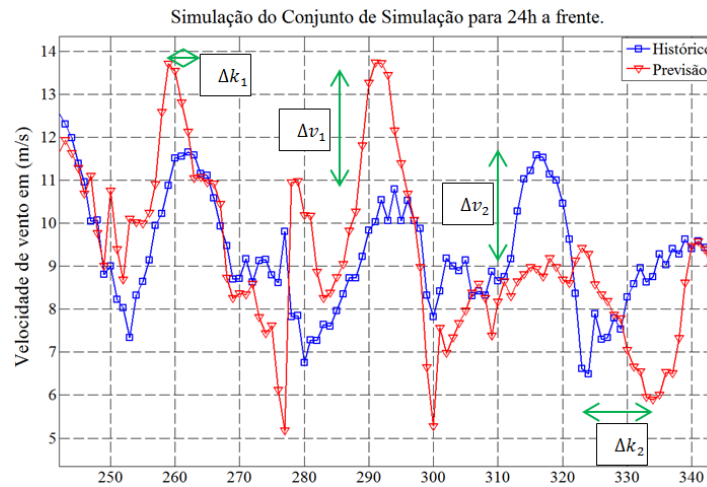
Figura 4 – Discrepância entre amplitude do valor previsto em relação ao histórico.



Fonte: O autor (2021).

A Figura 5 mostra a discrepância na previsão de vento com 24 h a frente, na qual se pode perceber que a tendência do histórico é deslocada por um termo Δk_i que pode variar de acordo com a posição do elemento a ser previsto no conjunto de simulação, além das amplitudes também apresentarem as incoerências comentadas anteriormente.

Figura 5 – Discrepância entre amplitude e posição no valor previsto em relação ao histórico.



Fonte: O Autor (2021).

Estas inconsistências podem afetar o controle de despacho eólico baseado em um horizonte, bem como o controle de operação de um aerogerador ou parque eólico, baseado num sistema de previsão de geração eólica. Outro ponto importante é que em um predictor, a incoerência gerada pelo erro da previsão poderá distorcer o funcionamento de um sistema de geração eólica que tenha em sua base de controle a curva de potência. Tal fato poderá acontecer em várias máquinas no caso de um parque ou em uma única máquina.

1.2 Justificativa

No que diz respeito à precisão de modelos preditores de geração que utilizem curvas de potência simuladas, muito já foi mostrado em publicações anteriores acerca do seu desempenho, porém com relação ao erro da previsão a diferença de desempenho entre os modelos não mostra avanços significativos. Com relação à construção, estes modelos podem ser de dois tipos específicos, os paramétricos e não paramétricos, tendo cada tipo a sua respectiva aplicação.

No primeiro caso, enfatizando os modelos paramétricos, o modelo fica refém de parâmetros balizadores que através de uma função característica, comumente proveniente de interpolação polinomial e seus respectivos contornos para o problema, devem convergir para um valor de previsão adequado. Nesta abordagem a falha está na função, pois como o parâmetro possui uma grande variabilidade, o mesmo tende a não expressar respostas condizentes nestas condições de oscilação.

No segundo caso, referenciando os modelos não paramétricos, o enfoque está em aprender a tendência da base de dados, que em uma nuvem de pontos define a tendência característica média, o que de certa forma é mais adequado do que no caso paramétrico, pois a tendência característica média leva em consideração também a variabilidade dos pontos que se apresentam distantes de uma região de concentração. Nesta abordagem a falha encontra-se na concentração de pontos em uma mesma área, pois a tendência irá priorizar regiões de maior concentração de dados, ou seja, dependendo da base, alguns pontos condizentes, porém mais distantes da característica média podem ser excluídos da análise e distorcer a característica de aprendizagem.

Outro ponto importante é que a maioria dos modelos, sejam os mesmos paramétricos ou não paramétricos, utilizam blocos de ferramentas de inteligência artificial sem fazer adequações em sua respectiva utilização. Assim o previsor ficará refém do desempenho da curva de potência simulada pelo bloco de IA.

Logo, surge a necessidade de criação de técnicas originais que busquem a diminuir este erro de previsão citado nas aplicações anteriores, além da busca pela melhora do desempenho da curva de potência simulada com o principal intuito de otimizar a previsão de geração eólica, já que esta fonte é a que mais tem crescido no Brasil nos últimos anos. Além de mostrar que modelos não paramétricos não estão estagnados em termos de heurísticas que tentem melhorar o seu desempenho nas previsões de geração.

1.3 Objetivos

De forma a apresentar os propósitos e contribuições descritas nesta tese, a seguir os objetivos são descritos de forma geral e específica, no que se refere a utilização, técnicas, e melhoramentos de modelagem a cerca das ferramentas de IA utilizadas e desenvolvidas neste trabalho. Sendo o principal enfoque: melhorar o desempenho dos modelos previsores desenvolvidos e apresentados nesta com enfoque no curto prazo.

1.3.1 OBJETIVOS GERAIS

Desenvolver técnicas originais de utilização de ferramentas de IA que melhorem o desempenho dos modelos previsores de geração eólica, inicialmente abordados em

ALBUQUERQUE, (2015). Objetivando mostrar que modelos previsores de geração eólica que utilizam ferramentas de IA comuns, (RNA e SIF) podem ter seus desempenhos melhorados significativamente, pela modificação na heurística e montagem destes modelos e na forma como os dados são inseridos no previsor.

1.3.2 OBJETIVOS ESPECÍFICOS

- Desenvolver abordagem original da curva de potência, utilizando sistemas de inferência *Fuzzy* com o intuito de melhorar o desempenho de modelos previsores de geração eólica que utilizem como base a curva de potência de parques eólicos.
- Melhorar o desempenho dos modelos previsores inicialmente abordados em ALBUQUERQUE, (2015).
- Avaliar os modelos previsores *fuzzy* com um critério de aprendizagem pré-estabelecido.
- Estudar os sistemas neurais de aprendizagem profunda (*deep learning*), utilizando seus conceitos principais para desenvolver novas heurísticas objetivando a criação de novos modelos de curva de potência que melhorem a eficiência de modelos previsores de potência;
- Entender as semelhanças e diferenças entre os sistemas de aprendizagem profunda e os sistemas neurais e de inferência *fuzzy* abordados anteriormente;
- Verificar o funcionamento dos modelos de curva de potência propostos com as novas técnicas em relação aos modelos de curvas iniciais baseados em redes neurais e em lógica *fuzzy*.
- Estudar técnicas de otimização para minimizar o erro entre a saída dos sistemas previsores *fuzzy* e respostas desejadas;
- Aplicar essa estratégia de minimização do erro aos sistemas previsores de vento e geração eólica, bem como, verificar seu comportamento e vantagens em relação a sistemas *fuzzy* mais simples.
- Propor a tecnologia com maior viabilidade de implantação, considerando as particularidades da região Nordeste do Brasil;
- Produzir a documentação em forma de artigos técnico-científicos utilizando ferramentas de inteligência artificial visando melhorar a exatidão da previsão de geração eólica, e serem submetidos a periódicos e congressos nacionais e/ou internacionais;

1.4 Trabalhos publicados

Publicação em Revista:

1. J. C. de Albuquerque, R. R. B. de Aquino, O. N. Neto, M. M. S. Lira, A. A. Ferreira and M. A. de Carvalho, "Power Curve Modelling for Wind Turbine Using Artificial Intelligence Tools and Pre-established Inference Criteria," in *Journal of Modern Power Systems and Clean Energy*, vol. 9, no. 3, pp. 526-533, May 2021, doi: 10.35833/MPCE.2019.000236.

Publicações em congressos:

2. DE AQUINO, R. R. B.; DE ALBUQUERQUE, JONATA C.; NETO, OTONI NOBREGA; LIRA, MILDE M. S.; CARVALHO, MANOEL A.; NETO, ALCIDES CODECEIRA; FERREIRA, AIDA A. Assessment of power curves in models of wind power forecasting. In: 2016 International Joint Conference on Neural Networks (IJCNN), 2016, Vancouver. 2016 International Joint Conference on Neural Networks (IJCNN). p. 3915.
3. DE AQUINO, R. R. B.; DE ALBUQUERQUE, JONATA C.; NETO, OTONI NOBREGA; LIRA, M. M. S.; FERREIRA, AIDA A.; CODECEIRA NETO, A. Influence of Deep Learning on Precision Improvement in Predictive Models of Wind Power Generation. In: International Conference on Computational Science and Computational Intelligence, 2017, Las Vegas. The 2017 International Conference on Computational Science and Computational Intelligence, 2017.

1.5 Estrutura do Trabalho

Este trabalho está organizado em oito seções. Nesta primeira seção descrevem-se as linhas que serão seguidas para o desenvolvimento deste trabalho, com a justificativa, objetivos e importância.

Na seção 2, apresenta-se uma revisão bibliográfica baseada nos artigos e trabalhos utilizados como base ferramental de pesquisa no desenvolvimento dos modelos e como ponto de partida para as técnicas e procedimentos propostos.

Na seção 3, tem-se uma breve análise funcional de uma curva de potência num modelo eólico.

Na seção 4, apresentam-se os tipos de ferramentas de inteligência artificial abordadas neste trabalho.

Na seção 5, apresenta-se uma revisão do estado da arte sobre a previsão de geração eólica.

Na seção 6, apresentam-se os modelos de curvas de potência utilizados neste trabalho bem como o modelo previsor que utilizará os modelos de curva de potência propostos.

Na seção 7, apresentam-se os resultados gerais obtidos com a aplicação dos modelos.

Na seção 8, apresentam-se as conclusões e propostas para trabalhos futuros.

2 REVISÃO BIBLIOGRÁFICA

Nesta seção é realizada uma revisão bibliográfica baseada nos artigos e trabalhos utilizados como base ferramental de pesquisa no desenvolvimento dos modelos e como ponto de partida para as técnicas e procedimentos propostos.

Em BOULARD, H. et. al. (2000), é feita uma reimpressão eletrônica com pequenas extensões e adaptações de um artigo intitulado “auto-associação por perceptrons multicamadas e decomposição de valor singular por H. Boulard e Y. Kamp e inicialmente publicado em *cibernética biológica*, vol. 59. pp. 291-294.

O artigo mostra que, para associação automática, as não-linearidades das unidades ocultas são inúteis e que os valores dos parâmetros ótimos podem ser derivados diretamente por técnicas puramente lineares baseadas na decomposição de valores singulares e na baixa matriz de aproximação, em espírito ao bem-estar, conhecida transformada de Karhunen-Loeve.

Essa abordagem aparece como uma alternativa eficiente ao algoritmo *backpropagation* comumente usados para treinar os *perceptrons* multicamadas. Além disso, dá uma interpretação clara da regra dos diferentes parâmetros.

Em JANGAMSHETTI, S. H. et. al. (2001), apresenta-se um método que adequa geradores eólicos a um local utilizando curvas de potência normalizadas e seus fatores de capacidade. A correspondência entre os locais é baseada na busca de parâmetros ótimos de velocidade a partir da curva de desempenho da turbina, que são obtidas pelas curvas normalizadas, aumentando a produção de energia com um maior fator de capacidade.

As velocidades são parametrizadas usando a média cúbica e a função de densidade de probabilidade de Weibull. Uma expressão para a potência normalizada é equacionada em função da velocidade nominal normalizada e do fator de capacidade. Um novo parâmetro de classificação, o índice de desempenho de turbinas eólicas, é definido combinando otimamente as turbinas com o local.

Ainda neste trabalho, são apresentados dois estudos ilustrativos que utilizam as curvas normalizadas para a identificação de parâmetros ótimos dos geradores de turbinas eólicas para um sítio, tais curvas, se usadas nas fases de planejamento e desenvolvimento das centrais eólicas, servirão como ferramenta útil para fazer uma escolha criteriosa de um gerador de turbina eólica que produza maior energia com maior fator de capacidade.

Em NORGAARD, P. et. al. (2005), apresenta-se uma nova formulação para o problema de correspondência de turbinas com base nas características da velocidade do vento

em qualquer local, nos parâmetros da curva de potência de qualquer turbina eólica, bem como no tamanho da turbina e altura da torre. A velocidade do vento em qualquer local é caracterizada pela função de distribuição Weibull de dois parâmetros e o valor do coeficiente de atrito do solo (a).

A curva de Potência é caracterizada pelas velocidades de corte (*Cut-in*) e (*Cut-out*), e a potência nominal. O novo Índice de Correspondência de Turbinas (TSMI) é derivado com base em uma formulação genérica para o Fator de Capacidade (FC), que inclui o efeito da altura da torre da turbina (h). O uso do FC como base para a correspondência entre turbinas e locais produz resultados que são tendenciosos em direção a torres mais altas, sem considerações para os custos associados.

A ideia principal do novo método é separar a dinâmica da saída de energia de uma turbina eólica em uma parte determinística e estocástica, correspondente ao comportamento real da turbina e influências externas, como a turbulência do vento, respectivamente.

Os resultados provam que uma abordagem dinâmica permite captar a dinâmica de conversão real de uma turbina e obter resultados mais precisos para a curva de potência, independentemente das influências específicas do local.

Em KHALFALLAHA, M. G. et. al, (2007a), um programa de medição abrangente foi experimentado através de turbinas Nordtank de 300kW. As máquinas foram erguidas na estação de teste para moinhos de vento no Centro de Tecnologia de Energia Eólica de Hurghada, onde, todos os sistemas experimentais da instalação foram descritos. No artigo descreve-se a variabilidade anual da velocidade do vento em Hurghada no período (1973-2001). A velocidade e a direção do vento foram medidas na latitude utilizando mastros de medição com, 10, 24,5 e 31 m, com estações de medição sendo operadas desde 1986. Os dados de vento foram gravados por instrumentos com uma taxa de amostragem de 2 Hz sendo os mesmos subsequentemente armazenados como valores médios consecutivos horários.

Assim, várias estatísticas mensais foram acumuladas. Além disso, se discute o efeito da variabilidade do vento na curva de potência das turbinas eólicas em Hurghada. Os resultados mostraram que, se a velocidade média anual do vento varia em $\pm 10\%$ em torno de um valor médio a longo prazo, a variabilidade natural correspondente da energia eólica disponível é de cerca de $\pm 25\%$.

Em KHALFALLAHA, M. G. et. al, (2007b), foram desenvolvidos métodos teóricos e experimentais para melhorar a potência em turbinas eólicas. Investigou-se a influência da mudança da velocidade de rotação do rotor de uma turbina de 300kW com controle de ‘*stall*’.

Tal controle foi feito através da alteração do programa de configuração para o funcionamento da turbina eólica.

Foi verificado nos experimentos que a variação no desempenho aerodinâmico do rotor da turbina ocorre devido à alteração do ângulo de inclinação. O ângulo da ponta definido pelo fabricante para a turbina de 300 kW é $(-1,8^\circ)$. Neste estudo, o ângulo foi alterado para dois valores $(-2,8^\circ)$ e $(-0,8^\circ)$. Para melhorar o desempenho da pá de uma turbina eólica operando com baixa velocidade de vento; podem ser utilizados os geradores de vórtice. Estes geradores foram colocados sobre a superfície de baixa pressão da pá ou ponta da pá. Neste estudo, dois tipos de geradores de vórtice, retângulo e formas triangulares foram investigados.

Em ŠIMIC, Z.et. al., (2007), são apresentados resultados da modelagem de um sistema construído a partir de fontes fotovoltaicas, eólicas e a diesel. Dados reais foram usados em relação à radiação solar global prevista, velocidade do vento e a carga. A simulação foi realizada com a ferramenta de otimização *micropower* HOMER.

O objetivo principal foi como estimar a influência da curva de potência no custo e na quantidade da energia produzida a partir do vento disponível. Sendo avaliadas neste estudo cinco diferentes turbinas eólicas por suas curvas de potência.

A configuração do sistema é alterada apenas em relação ao número ótimo de alguns componentes (i.e., potência fotovoltaica, bateria e capacidade do carregador / inversor). Os resultados mostraram a influência da curva de potência da turbina eólica no custo de energia e na configuração ótima do sistema.

Simula-se o desempenho do controlador através de um fator de controlabilidade. Os resultados mostram a efetividade do controlador baseado na curva de potência.

Em ZAMANI, M.H. et. al., (2007), é proposta uma curva em 3D nomeada curva controladora que representa a quantidade de potência elétrica gerada pelo controlador levando-se em conta as características do vento e as reações instantâneas do controlador à variação da velocidade do vento. Para tal, um método, chamado EPF de representação dinâmica do vento é proposto.

Em BENGIO, Y. et. al. (2007), estudou-se um algoritmo de aprendizagem não supervisionada introduzido por Hinton empiricamente, explorando variantes para entender melhor seu sucesso e estendê-lo a casos em que as entradas são contínuas ou onde a estrutura da distribuição de entrada não é suficientemente reveladora sobre a variável a ser prevista em uma tarefa supervisionada. Os experimentos confirmaram a hipótese de que a estratégia de treinamento não supervisionada em camadas ajuda a otimização, ao inicializar pesos em uma

região próxima a um mínimo local, dando origem a representações distribuídas internas que são abstrações de alto nível da entrada, trazendo melhor generalização

Em ÜSTÜNTAS, T. et al (2008), são desenvolvidos modelos em Lógica Fuzzy que solucionam a existência de dados nas regiões de (*Cut-in*) e (*Cut-out*). Este fato é extremamente prejudicial já que a previsão nas saídas dos modelos é bastante afetada pelas distorções de dados presentes nestas áreas da curva da turbina.

Para resolver este problema, é proposto um algoritmo de agrupamento central de classes que escolhe na curva da turbina, clusters que não são afetados por estas regiões, livrando assim a previsão destas imperfeições. Assim verifica-se que para uma curva de potência ser bem representada são necessários de 4 a 5 clusters.

Em SUHUA, L et. al., (2008), são proposto um método de agrupamento *Fuzzy* e a teoria de similaridade para classificar períodos diferentes em diferentes categorias de tempo. Escolhe-se um valor de saída fixo para representar a saída de energia eólica na categoria respectiva.

A velocidade do vento e a produção eólica são descritas pela distribuição da velocidade. Por fim uma curva típica de potência da turbina eólica é obtida e os resultados fornecem uma base teórica para a simulação da produção e análise de previsão do sistema integrado de energia eólica.

Em GOTTSCHALL, J. et. al., (2008), é introduzida uma abordagem dinâmica para a determinação de curvas de potência para turbinas eólicas, sendo esta comparada com dois métodos comuns, sendo um deles, a norma IEC 61400-12-1.

Em KUSIAK, A. et. al., (2009), é apresentada uma abordagem orientada a dados para a análise do desempenho de turbinas eólicas. O desempenho da turbina é capturado com uma curva de potência. As curvas de potência são construídas usando dados históricos de turbinas eólicas, três modelos de curvas de potência são desenvolvidos, dos quais, um baseado no método dos mínimos quadrados, outro baseado na abordagem de probabilidade máxima; o terceiro, é não paramétrico e é construído com o k-vizinho mais próximo utilizando o algoritmo (k-NN). Dois modelos, um baseado em mínimos quadrados (paramétrico) e um modelo não paramétrico, são utilizados para o monitoramento on-line da curva de potência e seu desempenho é analisado.

Em ALBADI, M. H. et. al., (2009), formula-se um novo fator de capacidade para aerogeradores, usando as características da velocidade do vento e os parâmetros da curva de potência da turbina. Os resultados mostraram uma maior precisão e simplicidade da nova formulação adequando de uma melhor forma a turbina a um projeto.

Em JIN, T. et. al., (2010), estima-se a incerteza de potência quando um aerogerador opera entre a velocidade de (*Cut-in*) e a velocidade nominal do vento. Um modelo probabilístico dinâmico da potência de saída é proposto, considerando que a potência segue uma distribuição normal, com média variável e desvio padrão constante.

Com base nessa condição, a média e a variância da produção eólica total são derivadas pela incorporação da velocidade do vento e da dinâmica de potência. Os resultados mostraram-se satisfatórios e ilustram a performance significativa do novo modelo proposto.

Em AKDAG, S. A. et. al., (2010), o modelo de distribuição *Weibull* foi utilizado para determinar a produção de energia de seis turbinas eólicas comerciais que variam de 335 a 1000 kW. As curvas de potência das turbinas eólicas foram modeladas com sete modelos de curva de potência e os resultados da produção de energia foram comparados com os resultados da Análise de Vento e do Programa de Aplicação (WASP).

Para comparar estes modelos, utilizaram-se dados médios horários de vento medidos a 50 m de altura na Turquia. Três modelos estimam a produção de energia menor para todos os casos e um modelo estima a produção de energia mais alta para todos os casos.

Em VINCENT, P. et. al. (2010), utiliza-se uma estratégia original para a construção de redes profundas, com base no empilhamento de camadas de decodificadores de distúrbios denominados *denoising autoencoders* que são treinados localmente para desativar versões corrompidas de suas entradas. O algoritmo resultante é uma variação direta no empilhamento de *autoencoders* comuns.

O modelo profundo de *autoencoders* do artigo é considerado como uma referência em problemas de classificação por produzir erros de classificação significativamente menores, preenchendo desta forma o modelo com redes DBN com o intuito de melhorar o desempenho alcançado, sendo este fato alcançado em vários casos.

Sendo verificado que características de nível mais elevado, profundo, aprendidas desta forma puramente não supervisionada também ajudam a melhorar o desempenho dos classificadores SVM subsequentes.

Os resultados mostraram que ao contrário dos *autoencoders* comuns, os *denoising autoencoders* são capazes de aprender detectores de bordas semelhantes ao *Gabor-like* a partir de amostras naturais e detectores de traços maiores a partir de imagens digitalizadas. Este trabalho estabelece claramente o valor de usar um critério de *denoising* como um objetivo não supervisionado tratável para guiar a aprendizagem de representações úteis de alto nível.

Em ULUYOL, et. al., (2011), apresenta-se uma análise de desempenho que utiliza a potência medida e a curva de potência para calcular uma potência residual. Como a curva de

potência não é específica do local, o resíduo é mascarado por ele e por outros fatores externos, como a degradação do desempenho dos componentes. Delinearam-se regimes operacionais e desenvolveram-se indicadores estatísticos de condição para adaptar o desempenho da turbina e isolar os componentes com defeito. No estudo foram inclusas turbinas eólicas que não dispunham da curva de potência do fabricante, tendo essas máquinas curvas de potência empíricas. Foram utilizados dados de sistema (SCADA) de duas turbinas eólicas de propriedade de diferentes operadores.

Em THAPAR, V. et. al., (2011), apresenta-se um estudo comparativo de vários métodos matemáticos, referenciados a três turbinas eólicas comercialmente disponíveis. Verificou-se que os métodos baseados em equações fundamentais de potência eólica, não replicam corretamente o comportamento das turbinas atuais.

Modelos baseados na curva de potência, embora simples de usar, dependem da precisão desejada; entretanto, respondem satisfatoriamente para velocidades médias de vento anuais mais altas.

Verificou-se que métodos matemáticos em que a curva de potência real de uma turbina é usada para desenvolver equações características, mostraram resultados precisos para turbinas eólicas com curva de potência suave; já para as turbinas que não possuem curva de potência suave, as curvas são modeladas pelo método de mínimos quadráticos, sendo o ajuste feito por esse método e por interpolação de (*splines cúbicos*), curvas suaves resultantes do processo de interpolação.

Em F. TRIVELLATO et., al., (2012), é feita uma avaliação sobre a curva de potência ideal de vários tipos de turbinas eólicas (velocidade fixa, passiva, eixo horizontal). A curva de potência ideal refere-se a condições ideais, i.e., o vento é estável, laminar, espacialmente uniforme e não perturbado pela turbina, sem erro de guinada, e a potência de saída pretendida é instável.

A curva de potência ideal tem duas aplicações principais: a previsão da energia eólica e a extensão da curva de potência para locais com diferentes níveis de turbulência a partir do local de teste primitivo.

Uma turbina eólica de 12 kW de velocidade fixa foi monitorada durante dois anos em um local turbulento, o trem de força da turbina não teve um efeito de filtração apreciável sobre as flutuações do vento. A curva de potência ideal foi analiticamente derivada da expansão de Taylor (cuja convergência foi reforçada pela transformação de *Shank*) e por uma suposição analítica precisa do coeficiente de potência ideal. A presente solução analítica é fácil de manusear e comparada com sucesso à curva baseada pela IEC.

A turbulência aumenta ou diminui a potência de saída para velocidades menores ou maiores que a velocidade no ponto de inflexão da curva de potência. As previsões de energia pela curva ideal ficaram dentro do erro experimental inerente.

Em SIMON GILL et. al., (2012), propõe-se o uso de dados operacionais para estimar funções de distribuição de probabilidade bivariada representando a curva de potência de turbinas existentes, para que desvios de comportamento sejam detectados. Como a dependência entre velocidade e potência é complexa, nenhuma distribuição parametrizada clássica pode aproximar este fato, sendo propostas as cópulas empíricas;

Esta teoria permite que a forma das distribuições marginais da velocidade e da potência do vento seja expressa separadamente das informações sobre a dependência entre elas. Faz-se então uma discussão sobre a utilidade na monitoração de condições de turbinas eólicas, particularmente no reconhecimento precoce de falhas incipientes, tais como degradação da pá, erro de guinada e de passo.

Em MARVUGLIA, A. Et. al., (2012), propõe-se um modelo não paramétrico de estado estacionário baseado na estimativa da curva de potência do parque eólico. Tal modelo tem como base o vento e sua exploração nas condições normais de funcionamento sendo seu objetivo detectar anomalias de funcionamento no parque eólico.

Utilizaram-se três modelos de aprendizagem de máquina para a curva de potência: Uma rede neural auto-supervisionada chamada GMR (*Generalized Mapping Regressor*), uma rede *feed-forward* MLP e uma Rede de regressão geral (GRNN).

O GMR é uma rede incremental auto-supervisionada que pode aproximar cada função ou relação multidimensional que apresentar qualquer tipo de descontinuidade; As MLPs são as mais utilizadas, modelos de rede neural de última geração e as GRNNs pertencem à família de redes neurais de *kernel*.

Este modelo não paramétrico da curva de potência pode ser usado como um perfil de referência para o monitoramento on-line do processo de geração de energia, e para as previsões de potência. Os resultados obtidos mostraram que a abordagem não paramétrica proporciona desempenho justo, desde que seja realizado um pré-processamento adequado dos dados de entrada.

Em COLAK, I. et. Al. (2012), é apresentada uma revisão incoerente em curto, médio e longo prazo. Os estudos disponíveis na literatura têm sido avaliados e criticados devido às suas precisões e deficiências. Os resultados mostram que sistemas adaptativos *Neuro-Fuzzy*, redes MLP obtêm os melhores resultados em termos de previsão eólica. Em alguns casos, o erro de estimativa da produção de energia pode chegar a 80%.

Em KUSIAK, A. et. al., (2013), são apresentadas três curvas operacionais diferentes com o intuito de monitorar o desempenho de um parque eólico. Sendo as mesmas: a curva de potência, a curva do rotor e a curva de inclinação da pá. Utilizou-se cinco anos de dados históricos sendo montado um banco de dados para construir as curvas de referência de energia eólica, cujas variáveis presentes foram: a velocidade do rotor e o ângulo de inclinação da pá, com a velocidade do vento como variável de entrada.

Uma abordagem de detecção de atributos multivariados baseada em *means clustering* e *Mahalanobis distance* é aplicada a estes dados produzindo um conjunto de dados para modelar turbinas. *Kurtosis* e *skewness* de dados bivariados são usados como métricas para avaliar o desempenho das turbinas eólicas. O monitoramento de desempenho das turbinas eólicas é realizado com o gráfico de controle da *Hotelling*.

Em Z. O. OLAOFE et. al., (2013), são apresentadas duas técnicas de análise para energia eólica: o uso da técnica direta, em que as potências de saída são estimadas pela curva de potência da turbina e o uso de técnicas baseadas em estatística onde as Potências de saída são estimadas com base na curva de potência do local desenvolvido.

A avaliação dos recursos eólicos foi realizada no Darling Wind Farm que está situado entre Darling city, uma pequena cidade em uma área agrícola na região da costa oeste do Cabo Ocidental, a cerca de 75 km (46,6 milhas) da Cidade do Cabo (África do Sul) e a cidade costeira de Yzerfontein (África do Sul) e está operacional com planos de expansão, sendo conduzida utilizando dados meteorológicos de séries temporais com amostragem de 5 min coletados a uma altura de 10 m durante um período de 24 Meses.

Devido à não-linearidade da velocidade do vento e sua potência correspondente no local, na modelagem dos recursos eólicos, a curva de potência do local foi utilizada para estimar a energia das turbinas eólicas e as alterações nas condições climáticas.

Três turbinas eólicas de classificação de 1,3 MW, 1,3 MW e 1,0 MW foram selecionadas para a geração de energia com base no vento medido a 50, 60 e 70 m de altura, respectivamente. As saídas de energia a 50 m de altura na turbina de 1,3 MW foram comparados com as saídas de potência a 60 m para determinar a altura padrão para utilidade da geração de energia em escala reduzida neste local.

Uma geração de energia adicional de 190,71 MWh estava disponível com a mesma turbina a uma altura de 60 m. Além disso, foram feitas comparações entre o uso da turbina e a curva de potência do local para análise de energia eólica nas alturas consideradas. Os resultados mostraram que a análise das Potências de saída das turbinas com base na curva de

potência do local é uma técnica precisa para a análise de energia eólica em comparação com a curva de potência da turbina.

Em LYDIA, M. et. al., (2013), desenvolvem-se modelos paramétricos e não paramétricos de curvas de potência em turbinas eólicas. Tais modelos paramétricos utilizam de quatro a cinco expressões de parâmetros logísticos definidos, utilizando algoritmos genéticos (GA), programação evolucionária (EP), otimização de enxame de partículas (PSO) e evolução diferencial (DE).

Os modelos não paramétricos têm evoluído usando algoritmos como redes neurais, Lógica difusa, *c-means clustering*, e mineração de dados. A modelagem da curva de potência da turbina eólica é feita usando cinco conjuntos de dados; um é um conjunto, estatisticamente gerado e os outros são conjuntos de dados em tempo real. Os resultados obtidos foram comparados com as métricas de desempenho adequado e o melhor método para a modelagem da curva de potência foi obtido.

Em M. SCHLECHTINGEN et. al., (2013), são comparadas quatro abordagens de mineração de dados para monitoramento de curvas de potência de turbinas eólicas. Esta supervisão pode ser aplicada para avaliar a geração de saída da máquina e detectar desvios causadores de prejuízos financeiros. Assim foram criados modelos baseados em lógica *Fuzzy* do centro de cluster, em redes neurais e em modelos vizinhos mais próximos, e seus desempenhos comparados com a literatura.

Tais modelos de centro de cluster foram adaptativos sendo estes sistemas de inferência *Neuro-Fuzzy* e seu desempenho comparado com os outros modelos, usando os mesmos dados. A literatura muitas vezes negligencia a influência da temperatura ambiente na direção do vento, podendo esta influência ser com relação à potência de até 20%. As abordagens propostas na literatura e os modelos ANFIS são comparados usando apenas a velocidade do vento e dois insumos adicionais. A comparação baseia-se no MAE, no MSE, no MAPE e no desvio padrão, utilizando dados provenientes de três turbinas de passo regulado com uma potência de 2 MW. A capacidade de destacar os desvios de desempenho é investigada pelo uso de medidas reais. A comparação mostra a diminuição das taxas de erro e dos modelos ANFIS quando se consideram as duas entradas adicionais e a capacidade de detectar falhas anteriormente.

Em CARRILLO, C. et. al. (2013) é realizada uma revisão das equações que comumente representam as curvas de potência de turbinas eólicas de velocidade variável (VSWTs): curva de potência polinomial, curva de potência exponencial, curva de potência cúbica e curva de potência cúbica aproximada. Estes tipos de curvas foram comparados com

as curvas do fabricante usando os coeficientes de determinação, como indicadores de aptidão, e usando a estimativa da produção de energia.

Foram utilizados dados de 200 aerogeradores comerciais, variando de 225 a 7500kw. Os resultados da análise mostram que as aproximações exponencial e cúbica dão os maiores valores de R^2 e o menor erro na estimação de energia. Com a curva de potência cúbica aproximada obtêm-se valores bastante altos de R^2 e baixa estimativa de energia de erro, o que torna este tipo de aproximação muito interessante devido à sua simplicidade. Já as curvas de potência polinomiais apresentam os piores resultados, devido à sua sensibilidade aos dados fornecidos pelo fabricante.

Em LYDIA, M. et. al., (2014), apresenta-se um panorama exaustivo sobre a necessidade de modelagem de curvas de potência de turbinas eólicas e as diferentes metodologias empregadas. Faz-se também uma análise e uma avaliação criticamente detalhada das técnicas de modelagem paramétricas e não paramétricas, além de apresentar o enfoque em outras áreas de pesquisa.

Em SARIKAYA, R. et. al., (2014), um algoritmo chamado de divergência contrastiva é proposto para auxiliar uma *Deep Belief Network* (DBN) em seu processo de aprendizagem, sendo a sua principal tarefa um problema de compreensão da linguagem.

O algoritmo *Contrastive Divergency* (CD) permite que a DBN aprenda um modelo generativo multicamada a partir de dados não rotulados e os recursos descobertos por este modelo são usados para inicializar uma rede neural *feed-forward* que é ajustada com o algoritmo *backpropagation*. Para tal, compara-se uma rede neural inicializada por DBN com três algoritmos de classificação de texto amplamente utilizados: máquinas de vetores de suporte (SVM), *boosting* e entropia máxima (*MaxEnt*). O modelo baseado em DBN simples fornece uma exatidão de classificação de roteamento de chamadas que é igual ao melhor dos outros modelos tomados como base no artigo. Porém, o artigo mostra que o uso de dados adicionais não rotulados para o pré-treinamento da DBN e a combinação de recursos aprendidos baseados na DBN com os recursos originais fornece ganhos significativos em relação às redes SVM, que, por sua vez, tiveram desempenho melhor que o entropia máxima (*MaxEnt*) e o *Boosting*.

Em KHODAYAR, M. et. al., (2015), propõe-se uma rede baseada em *stacked autoencoder* (SAE) para previsões em curto prazo e ultracurto-prazo de velocidade do vento. Um fato notório é que este foi o primeiro trabalho que aplicou o *Deep Learning* em previsão da velocidade do vento de um local de vento. Além disso, uma camada de regressão é

aplicada no topo deste modelo, a fim de lidar com fatores de incerteza existentes nos dados de velocidade do vento.

Os resultados mostraram uma melhora significativa do desempenho quando comparada a outros métodos de RNA que aplicaram arquiteturas superficiais e a SAE tradicional.

Em YAN, J. et. al., (2017), propõe-se um método WPF (*Wind Prediction Forward*) de várias escalas, estabelecendo uma rede de mapeamento *multi-to-multi* (*m2m*) e o uso de *stacked denoising autoencoders* (SDAE). O horizonte temporal previsto é de 24 a 72 horas. Em primeiro lugar, os *multi-NWPs* em uma região são corrigidos com base no SDAE para gerar melhores insumos para o seguinte WPF regional; em segundo lugar, vários SDAEs com diversos parâmetros de modelo e recursos de entrada são integrados ao conjunto SDAE para prever a energia eólica gerada a partir de vários parques eólicos em uma região.

Dois conjuntos de dados são utilizados neste estudo de caso para validar o método proposto. Os resultados mostram que o mapeamento *m2m* proposto e SDAE são capazes de capturar as correlações reais do vento em vários locais, e supera as outras contrapartes em termos de correção *multi-NWPs*, bem como o WPF para a região e o parque eólico individual em causa. Além disso, o conjunto SDAE funciona melhor do que qualquer outro modelo regional individual do WPF.

3 CURVA DE POTÊNCIA

Curvas de Potência de turbinas eólicas mostram a relação física existente entre a potência elétrica gerada pelo aerogerador e a velocidade do vento incidente na altura do centro do rotor do mesmo, esta relação é própria e característica de um parque ou turbina eólica. Sendo esta tendência apresentada para cada trecho de existência da curva, segundo o *International Energy Association* - I.E.A., a curva de potência é a identidade funcional de uma máquina eólica, sendo definida como um certificado de desempenho garantido pelo fabricante.

Assim, este órgão desenvolveu recomendações que têm sido continuamente melhoradas e adaptadas pelo *International Electrotechnical Commission* - I.E.C., que para normatizar tais indagações sobre o tema, criou a norma IEC61400-12, atualmente aceita como referência mundial na definição, aferição e medição de uma curva de potência.

Uma curva de potência típica é modelada por três características básicas, são elas: a velocidade (*Cut-in*), a região “Efetiva” de velocidades de vento e a velocidade (*Cut-out*). Estes aspectos são definidos a seguir:

A velocidade (*Cut-in*): é a velocidade mínima de vento, na altura do cone (*Hub*) do rotor, em regime permanente e sem turbulência, que permite o início de geração de energia elétrica por parte do aerogerador.

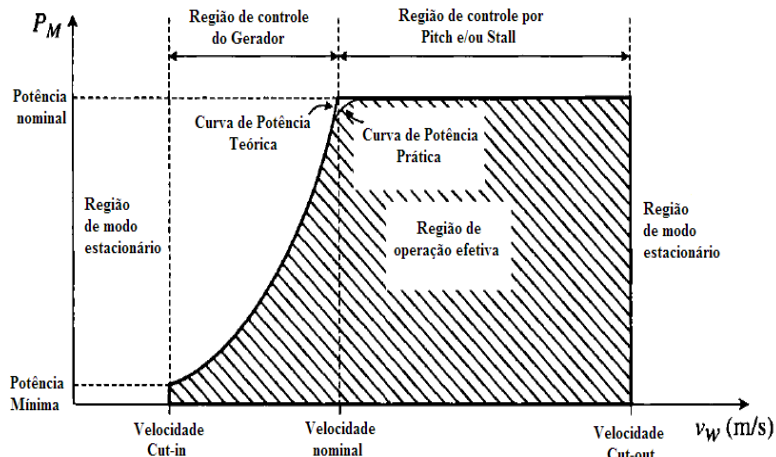
A região efetiva: é a faixa média crescente de velocidades de vento, na altura do cone do rotor, em regime permanente e sem turbulência, que permite a geração de energia elétrica. Esta região se estende da velocidade (*Cut-in*) até a velocidade (*Cut-out*).

Já a velocidade (*Cut-out*): é a velocidade máxima de vento, na altura do centro do rotor, em regime permanente e sem turbulência, que ainda permite a máquina operar antes de ser desligada. Assim, por se tratar de uma velocidade elevada, testa o limite físico-estrutural do aerogerador, sendo responsável por questões de segurança, pelo corte operacional da máquina eólica.

Estas três peculiaridades delimitam a curva de potência de qualquer turbina, além de estabelecer de forma efetiva qual a faixa de operação da mesma, definindo desta forma, os valores de operação mínimos e máximos de velocidades médias e de potência médias.

A Figura 6 mostra o diagrama esquemático de uma curva de potência.

Figura 6 – Demonstração qualitativa de uma curva de potência.



Fonte: O Autor (2015)

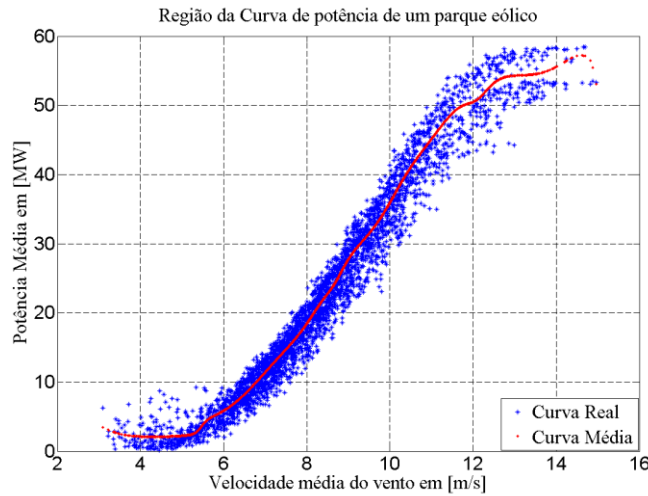
Segundo (STIEBLER, M. 2008, p. 23) mais importante do que a capacidade energética de um sistema de geração eólica, é a sua curva de potência. Pois, o levantamento desta, associado ao conhecimento da velocidade média e de distribuições estatísticas, como a de *Rayleigh*, fornecem informações indispensáveis para a previsão do rendimento anual de energia.

Outro enfoque bastante difundido da curva de potência é o uso da mesma para controle num parque eólico, tal caso é bastante diferente do abordado anteriormente, dizendo respeito ao limite de construção de uma máquina, pois é feita uma medição por um determinado período específico e calcula-se uma média das variáveis agregadas ao vento no período. Tais valores médios devem ser valores condizentes quanto à análise (local) ou (global) de controle incidente em um único aerogerador ou em vários aerogeradores no parque.

No final desta análise, tem-se uma região de pontos que comportam de forma singular a característica de acordo com o controle estabelecido a característica (velocidade x Potência) para um ou todos os aerogeradores do parque.

Na parte central da Figura 7 tem-se uma curva de potência média atestando à faixa média de velocidade atrelada a faixa média de potência para todos os aerogeradores sendo esta característica denominada curva de potência média do parque.

Figura 7 – Característica velocidade x Potência para um parque eólico.



Fonte: O Autor (2015)

A energia do vento disponível que atravessa o rotor de uma turbina eólica pode ser obtida a partir da Equação (3.1):

$$P_W(v) = \frac{1}{2} \rho A v^3 \quad (3.1)$$

Onde $P_W(v)$ é a potência em W associada a uma velocidade v do vento em m/s, A é a área do rotor da turbina em m^2 e ρ é a densidade mássica do ar, padrão que precisa ser ajustado a cada local para a correta produção de eletricidade, (típico. 1225 kg / m^3 IEC61400-12-1. (2005)). Tal potência se relaciona com a potência gerada por uma turbina de vento por meio do coeficiente de potência, sendo este resultado responsável pela potência elétrica extraída nos geradores eólicos, conforme mostrado pela Equação (3.2).

$$P_e = \eta C_p P_W \quad (3.2)$$

Em que, C_p é o coeficiente adimensional de potência que representa a quantidade teórica de força mecânica que pode ser extraída pelo rotor da turbina, porém não é eficiência, pois é medido em local diferente, e (η) é a eficiência total da máquina ACKERMANN, T. et. al. (2000). O coeficiente (C_p) de potência é uma função que relaciona na pá de uma turbina eólica, a velocidade do vento na ponta da pá (λ) e o ângulo de inclinação da pá (θ) ACKERMANN, T. (2012).

A energia mecânica teórica máxima que pode ser extraída por turbinas eólicas, i.e., o valor máximo do rendimento que o aerogerador pode atingir, é um valor adimensional aproximadamente igual a 0,5926 e é conhecido como o limite de Betz (Betz A. 1920).

4 FERRAMENTAS DE INTELIGÊNCIA ARTIFICIAL

Esta seção apresenta e expõe conceitos básicos sobre IA, ferramentas de IA, modelos de redes neurais artificiais, sistemas de inferência *fuzzy*, redes, modelagens e técnicas de aprendizagem profunda (*Deep Learning*) propostas ou ligadas a modelagem utilizada nesta tese. Além de modelos de redes com recorrência de aprendizagem, sendo este conceito bem como as novas redes profundas, um dos pontos principais responsáveis pelas inovações nos modelos de curto prazo com melhoria significativa na exatidão da previsão desenvolvidos neste trabalho.

4.1 Redes Neurais Artificiais

Uma Rede Neural Artificial (RNA) é um conjunto de neurônios artificiais que procura reproduzir a estrutura e o funcionamento do cérebro humano em um ambiente técnico. As redes neurais apresentam características que as tornam especialmente atrativas, que são: paralelismo, processamento de informação incerta, capacidade de aprendizagem, robustez e tolerância a falhas. Em outras palavras as redes neurais apresentam a capacidade de adaptar ou aprender, depois generalizar, agrupar ou organizar dados e por fim minimizar os erros atrelados às respostas.

As RNAs são sistemas paralelos distribuídos compostos por unidades de processamento simples, chamados de neurônios artificiais, que calculam determinadas funções matemáticas (normalmente não lineares). Tais unidades são dispostas em uma ou mais camadas e interligadas por um grande número de conexões. Na maioria dos modelos estas conexões possuem pesos associados, os quais armazenam o conhecimento representado no modelo e servem para ponderar a entrada recebida por cada neurônio da rede (LUDERMIR, T. B. et. al. 2000).

Hoje, as redes neurais servem de base para o sistema de reconhecimento óptico de caracteres empregados em digitalizadores, previsões meteorológicas, detectores de bombas e até mesmo em previsões para o mercado financeiro e apresentam-se como uma abordagem alternativa aos métodos estatísticos de previsão temporal.

Por se tratar de uma proposta de otimização de ferramentas IA já discutidas anteriormente e desenvolvimento de novas técnicas com outras ferramentas de IA, toda a

metodologia dos conceitos de redes neurais artificiais do tipo *Multi layer perceptron* podem ser encontradas na Dissertação: ALBUQUERQUE (2015).

4.2 Sistemas de Inferência Fuzzy

Sistemas *Fuzzy* ou nebulosos são sistemas que utilizam funções de pertinência sintonizadas por faixas que obedecem a uma listagem de regras delimitadas pelo conhecimento do especialista. Tal conhecimento poderá estar ligado a uma ou a várias características de um problema ou sistema, que poderão determinar parâmetros para as regras bem como uma dependência ou independência das variáveis de entrada com relação as variáveis de saída.

4.2.1 SISTEMA DE INFERÊNCIA FUZZY MAMDANI

O modelo de inferência *Mamdani* foi um dos primeiros sistemas constituído utilizando a teoria de conjuntos *fuzzy*, sendo proposto em 1975 por Ebrahim Mamdani (REZENDE, S. O., 2005). A regra de semântica tradicionalmente utilizada para o processamento de inferências com o modelo de *Mamdani* é chamada de inferência de (Máximos e Mínimos). Utilizando as operações de união e de interseção entre conjuntos da mesma forma de Zadeh, (REZENDE, S. O., 2005).

As regras de produção em um modelo de *Mamdani* possuem relações *fuzzy* tanto em seus antecedentes como em seus consequentes.

De forma semelhante, por se tratar de uma proposta de otimização das ferramentas IA já discutidas anteriormente e desenvolvimento de novas técnicas com outras ferramentas de IA, toda a metodologia dos conceitos de Sistemas de inferência *fuzzy* do tipo *Mamdani* podem ser encontradas na Dissertação: DE. ALBUQUERQUE, J.C. (2015), precursora deste projeto de tese.

4.2.2 SISTEMAS DE INFERÊNCIA FUZZY SUGENO

A aparição da inferência *fuzzy TSK* - (*Takagi-Sugeno-Kang*) se deu em meados de 1985, SUGENO, M. (1985), o mesmo é um tipo de inferência similar à inferência *Mamdani* em muitos aspectos, pois, as duas etapas iniciais da inferência, que são a fuzzificação das entradas e a aplicação do operador *fuzzy* são os mesmos do sistema *Mamdani*. A grande diferença está na saída, pois num sistema *TSK* a mesma poderá assumir um critério constante ou um caráter linear.

Assim considerando um sistema de duas entradas e uma saída o nível de saída Z assumirá a forma ditada pela Equação (4.1):

$$Z = ax + by + c \quad (4.1)$$

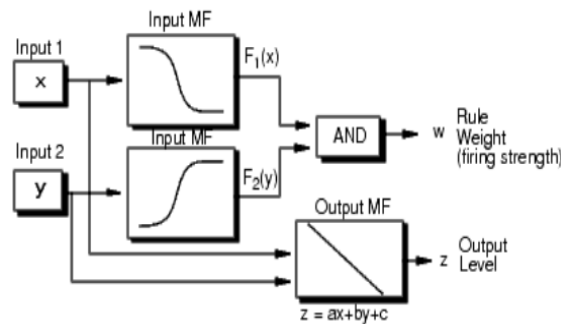
A inferência de cada entrada é controlada por uma função membro específica sendo o nível de resposta desta função um valor $F_i(x)$ que dependerá das funções de pertinência escolhidas relativas a cada entrada. De forma semelhante, os sistemas *fuzzy* sejam eles *Mamdani* ou *Sugeno* necessitam de uma lista de regras em sua parametrização objetivando uma ação de tendência específica, é o caso da variável w , no caso de um sistema com duas entradas, esta poderá assumir valores ditados conforme a Equação (4.2).

$$W_i = ([AND/OR] \text{ Method}(F_i(x), F_{i+1}(y))) \quad (4.2)$$

No sistema *fuzzy Sugeno* conforme ilustrado pela Figura 8 prevalece a heurística de dependência, papel da função (AND) ou independência papel da função (OR), que através da lista de regras fará a escolha de concordância de um nível de saída Z_i que poderá ser constante ou linear. A saída será calculada conforme a Equação (4.3) e considerará tanto o peso das regras quanto o nível de saída.

$$saída = \frac{\sum_{i=1}^N w_i Z_i}{\sum_{i=1}^N w_i} \quad (4.3)$$

Figura 8 – Inferência de regras em um sistema Fuzzy (Sugeno).



Fonte: Mathworks.com (2016).

Os Modelos *fuzzy Sugeno* tratados neste trabalho em termos práticos têm a mesma arquitetura dos *fuzzy Mamdani* diferindo apenas na sua forma de criação, pois o sistema Sugeno foi criado a partir de uma conversão de características dos modelos *Mamdani* anteriores.

Esta etapa pode ser considerada como um pré-treinamento, ou inferência inicial (ponto de partida), já que no processo de definição de aprendizagem o sistema *Sugeno* partiu com as características pré-estabelecidas pelo sistema *Mamdani*, parametrizado de acordo com a tendência de inferência entre a velocidade e/ou direção com a potência do parque.

Tal processo depende de cada parque e é denominado inferência do especialista.

Nesta tese é desenvolvida uma nova heurística que utiliza conceitos de aprendizagem profunda para o treinamento e parametrização dos modelos de curva de potência. Tal heurística consiste em um treinamento inicial advindo da inferência *fuzzy*, ou seja, um ponto de partida na parametrização dos modelos. Esta etapa é executada em modelos *fuzzy Mamdani*.

A segunda etapa consiste na transformação do modelo *Mamdani* em um modelo Sugeno com parâmetros semelhantes ao modelo anterior, a mesma é necessária, pois ela preparará o sistema para uma aprendizagem mais aprofundada. A terceira etapa consiste em utilizar o ANFIS como otimizador dos parâmetros do bloco *fuzzy Sugeno*, porém tendo um ponto de partida menos complexo do que o inicial.

Para o processo de aprendizagem secundária, utiliza-se o ANFIS estabelecendo através de redes neurais os melhores parâmetros do sistema *fuzzy*, como: o número de funções de pertinência, os tipos de funções de pertinência (Gaussianas), a melhor disposição das mesmas em cada intervalo específico, o raio de inferência e o número de clusters utilizados na aprendizagem.

Desta forma, pode-se considerar o ponto de partida do sistema com parametrização inicial como um conjunto menos entrópico do que o ponto inicial do sistema sem qualquer tipo de parâmetro definido (treinamento). Este enfoque mostrará uma significativa melhora na aprendizagem dos modelos de curvas de potência, podendo esta melhora ser verificada pelos índices avaliadores utilizados como MAE, NMAE e RMSE.

4.3 Adaptative Neuro Fuzzy Inference System (ANFIS)

Considerado um sistema neuro adaptativo, o ANFIS (*Adaptative Neuro Fuzzy Inference System*) é definido como uma rede neural artificial que tem por base além dos métodos de aprendizagem (característica neural), os métodos de inferência TSK, SUGENO, M. (1985). Foi desenvolvido como técnica no início dos anos 90, por JANG (1993), JANG & SUN (1995), Jang et al. (1997), possuindo a característica, de capturar os benefícios de ambos os métodos (Neural e *fuzzy*) em um único plano de fundo.

Seu sistema de inferência corresponde a um conjunto de regras IF-THEN *fuzzy* que possuem capacidade de aprendizado para aproximar funções não-lineares adequando a sintonia dos parâmetros das funções membro na entrada do sistema e na saída do mesmo. Esta característica serve para definir de forma menos custosa o tipo da arquitetura que compõe o sistema sobre análise.

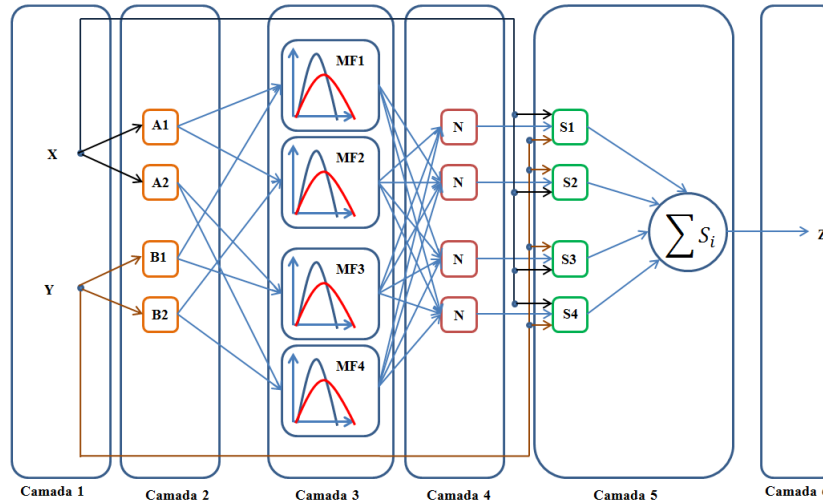
Outra diferença bastante notória é que nas RNAs os pesos armazenam o conhecimento enquanto no ANFIS este é indexado na lista de regras e no tipo de função membro utilizada.

Tais características neurais têm seus preceitos ligados fortemente a métodos de aprendizagem altamente difundidos em sistemas de IA como o método *backpropagation*, sendo esta, sua característica neural. Já sua capacidade de contorno em soluções e adequação ou ajuste de um conjunto específico, é herdada dos sistemas de inferência *fuzzy* TSK e suas funções de pertinência.

O ANFIS foi desenvolvido como técnica no início dos anos 90. JANG (1993), JANG & SUN (1995), portanto, é considerado um estimador universal, JANG, SUN, MIZUTANI (1997). Para a convergência de suas aproximações e definição de suas funções de pertinências são utilizadas apenas funções deriváveis (contínuas), sendo a família das gaussianas as mais utilizadas em algoritmos neuro adaptativos. Esta característica é possível graças ao uso combinado entre o algoritmo de retropropagação do erro (*backpropagation*) e o método dos mínimos quadráticos LSM (*Least-Square Methode*).

A Figura 9 apresenta a estrutura do ANFIS, no exemplo são utilizadas quatro entradas denominadas $A1$, $A2$, $B1$ e $B2$, em que para cada uma das quatro entradas são utilizadas duas funções membro também chamadas de função de pertinência, estando estas entradas interligadas através do ANFIS a quatro funções de saída.

Figura 9 – Estrutura de Aprendizado do ANFIS.



Fonte: O Autor (2021).

Para entender como se procede com a informação no ANFIS são mostradas a seguir as camadas inerentes ao processo de aquisição de aprendizagem:

Método de definição das camadas de aprendizagem para o ANFIS:

Camada 1:

A primeira camada é responsável pela recepção dos dados, que são lidos e anexados como informação, não ocorrendo assim processamento, portanto, as entradas não são consideradas como nodos da rede. Nesta camada são armazenados os parâmetros com o objetivo de definir as funções de pertinência, sendo este armazenamento representado por uma (*string*) variável linguística obtida pela Equação 4.4.

$$O_i^1 = \mu_{A_i}(x) \quad (4.4)$$

Em que x e y representam as entradas do nó i , A_i é o classificador definido pela função do nó e O_i^1 é a saída do nó i da primeira camada. Nesta camada as funções de pertinência que poderão ser escolhidas são do tipo triangular, trapezoide, sigmoide, forma de Z, Gaussiana e (*Bell*), formato de sino. Como exemplo, a Equação (4.5) define a função para os parâmetros da entrada x :

$$\mu_{A_i}(x) = \frac{1}{1 + \left[\left(\frac{x_i - c_i}{a_i} \right)^2 \right]^{b_i}} \quad (4.5)$$

Onde a_i , b_i e c_i são os parâmetros de ajuste, e normalmente são escolhidas as funções no formato de sino e gaussiana e sigmoide por apresentarem características não lineares, estabelecendo assim os conjuntos *Fuzzy*.

Camada 2:

Na segunda camada os dados são codificados, havendo uma ordenação entre os valores lidos e o posicionamento das funções de pertinência. Também é definido nesta camada, o grau de sintonia entre os conjuntos ou clusters do ANFIS e os respectivos valores das entradas.

Como cada conjunto é representado por um número de funções de pertinência específico, a função de ativação dos neurônios desta camada é uma combinação linear das respectivas funções de pertinência. Em que os ajustes de seus parâmetros são delineados por uma heurística de treinamento predefinida (algoritmo, métodos ou regras de convergência).

Outra característica é a definição dos nós de acordo com as regras adotadas, sendo definido para cada nó o (*threshold*), limiar ou grau de ativação conforme a regra que impera, ativando a função quando necessário. A Equação 4.6 mostra a equação do nó para a entrada x

$$w_i = \mu_{A_{i1}}(x_1) \cdot \mu_{A_{i2}}(x_2), i=1,2. \quad (4.6)$$

Camada 3:

A operação de interseção entre as funções de pertinência de entrada é realizada na terceira camada, de acordo com o operador escolhido. Os neurônios dessa camada possuem, portanto, funções de ativação fixas, não ajustáveis. Esta camada normaliza o grau de desempenho relativo as últimas regras consideradas para o somatório dos níveis de ativação dos nós, em que o valor calculado para o nível de ativação é dado pela Equação 4.7.

$$\bar{w}_i = \frac{w_i}{w_1 + w_2 + \dots + w_R}; i = 1,2 \quad (4.7)$$

A terceira camada é responsável por realizar o translado de forma ótima nas faixas de valores determinados pela camada dois, cabendo assim definir a posição exata das funções de pertinência bem como se haverá ou não intersecção entre duas ou mais funções numa mesma faixa. Uma diferença notória em relação à camada anterior é o valor fixo das faixas.

Camada 4:

A quarta camada é a camada responsável pela normalização dos valores ótimos definidos na terceira, sendo a informação anexada as quatro funções de pertinência da saída.

Esta camada calcula o produto entre o respectivo nível normalizado da camada 3 e uma função f_i linear das variáveis de entrada, conforme mostrado pela Equação 4.8.

$$O_i^4 = \overline{w}_i \cdot f_i = \overline{w}_i(p_i x + q_i y + r_i) \quad (4.8)$$

Onde: $\{p_i, q_i \text{ e } r_i\}$ é o conjunto de parâmetros do nó i

Na quarta camada da estrutura ilustrada na Figura 9, são realizadas as normalizações das saídas da camada anterior. Por realizar apenas essa operação, ela pode ser incorporada ao neurônio da última camada, ficando o sistema ANFIS com apenas cinco camadas, ou ainda se pode colocá-la como apenas um neurônio na última camada, permanecendo o sistema ANFIS com seis camadas. A função de ativação dos neurônios dessa camada é descrita na Equação 4.8.

Camada 5:

Na quinta camada as funções de ativação são do tipo *Sugeno*, sendo definidas como combinações lineares de ordem p das entradas, definindo desta forma os parâmetros consequentes de acordo com as regras criadas. A Equação (4.9) mostra a definição de uma função *Sugeno* de primeira ordem, tendo em seus parâmetros α , β e γ um contorno funcional que sintoniza entradas (x) com relação as suas respectivas saídas (y).

$$f(x, y) = \alpha x + \beta y + \gamma \quad (4.9)$$

Os parâmetros dessas funções também podem ser ajustados pelo método de otimização ou algoritmo de treinamento escolhido. A Equação 4.9 descreve uma função *Sugeno*, de primeira ordem, para a estrutura ilustrada na Figura 9 com parâmetros α , β e γ . Em muitos casos, o termo independente gama não é utilizado, reduzindo o número de parâmetros a serem ajustados.

Nesta camada ocorre a concatenação dos consequentes através de uma função somatório e posterior produção de uma saída *Fuzzy*, sendo responsabilidade de um neurônio fixo a produção da saída. Este cálculo é feito através de uma média ponderada do sinal de saída da camada de saída O conforme mostrado pela Equação 4.10.

$$O_i^5 = \sum_i \overline{w}_i \cdot f_i = \frac{\sum_i w_i \cdot f_i}{\sum_i w_i} \quad (4.10)$$

Camada 6:

Por fim, a camada 6 realiza o somatório das saídas da camada 5, apresentando então a saída do ANFIS. Por fim a função de ativação do neurônio da saída é definida conforme a Equação 4.11, definindo assim a ativação ou não do nodo da camada 6.

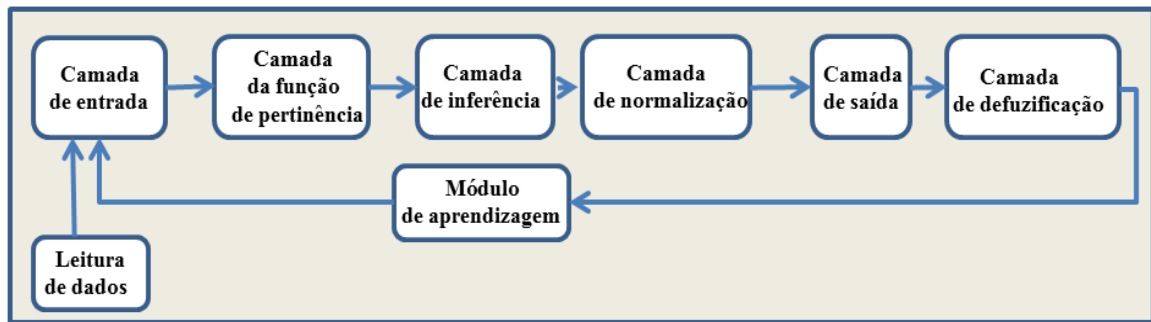
$$f_i(w) = \frac{w_i}{\sum_{j=1}^m w_j} \quad (4.11)$$

Onde: w_i é a saída do neurônio da camada anterior.

Dessa forma, em um ANFIS, o aprendizado consiste no ajuste dos parâmetros das funções de ativação dos neurônios da segunda e quinta camadas, ou seja, consiste em otimizar os parâmetros das funções de entrada e de saída de um sistema nebuloso *Takagi - Sugeno*.

A Figura 10 ilustra o conceito em diagrama de fluxo do ANFIS com todas as etapas do processo.

Figura 10 – Fluxograma conceitual do ANFIS



Fonte: AFONSO MARQUES, A. C., (2014)

4.4 Deep Learning

A aprendizagem profunda ou *Deep Learning* é uma das mais recentes abordagens no domínio das redes neurais artificiais, desde sua primeira aparição em (Hinton et. al., 2006), quando os modelos de aprendizagem profunda obtiveram o estado-da-arte em alguns problemas como a classificação e reconhecimento de padrões. Desta forma é considerada uma das vertentes da aprendizagem de máquinas (*Machine Learning*) responsáveis pela modelagem de sistemas com grau elevado de abstração, *i.e.*, sistemas com vários níveis ou características correlacionadas com uma ou várias grandezas, cujo objetivo do problema é a sua modelagem.

Neste tipo de aprendizagem também se utilizam vários tipos de camadas, porém diferentemente do *perceptron* a fonte das informações em cada camada é proveniente de uma característica específica, ou seja, uma camada poderá receber informações que outra camada não recebe.

Em (HINTON, G. Et. Al., 2006), mostraram que os resultados de uma rede neural profunda poderiam ser significativamente melhorados quando a mesma era pré-treinada com um algoritmo de aprendizagem não supervisionada, atuando em uma camada após outra, a partir da primeira. Este trabalho iniciou a área hoje conhecida como *Deep Learning*.

No início das pesquisas desta vertente, de um modo geral, os algoritmos tinham as seguintes características:

- 1) Aprendizado não supervisionado de representações para pré treinar cada uma das camadas;
- 2) Treinamento não supervisionado, aplicado a uma camada por vez, sendo a representação aprendida, de cada nível de entrada para a camada seguinte;
- 3) Utilização do treinamento supervisionado para ajuste fino de todas as camadas pré-formadas, bem como uma ou mais camadas adicionais dedicadas à produção de previsões.

Em termos de possibilidades de aplicações, a aprendizagem profunda dispõe das seguintes opções em caráter acadêmico e/ou industrial:

- a) No reconhecimento da fala e processamento de sinais, SEIDE , F. et. al. (2011);
- b) No reconhecimento de objetos, CIRESAN, D. C. et. Al.(2010);
- c) No processamento natural de linguagem, COLLOBERT, R. et. Al.(2011);
- d) Na transferência de aprendizagem, BENGIO, Y. (2012).
- e) Na previsão de séries temporais, FILDES, R. et. Al. (2008).

No entanto, tais modelos têm sido pouco utilizados na previsão de séries temporais.

4.4.1 REDES USADAS EM ARQUITETURAS DE APRENDIZAGEM PROFUNDA

No que se refere a aplicações de aprendizagem profunda, várias foram as redes neurais desenvolvidas dependendo logicamente, do tipo de problema, do método de aprendizado de máquina baseado em representações confiáveis de dados, bem como a viabilidade temporal do processo de aprendizagem, algoritmo e a exatidão relativos a um processo bem-sucedido.

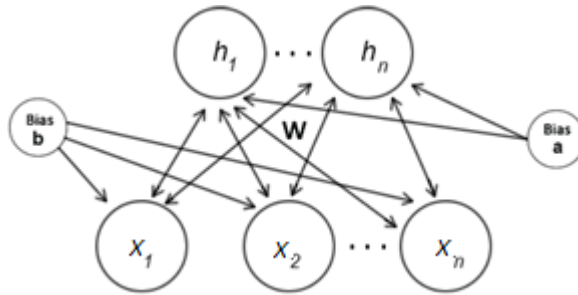
As principais redes, isto é, mais utilizadas na concepção de aprendizagem profunda são de 3 tipos:

- Máquina Restrita de Boltzman
- Deep Belief Networks
- Auto encoders

4.4.1.1 Máquina Restrita de Boltzman

Uma RBM (*Restricted Boltzman Machine*), é representada na Figura 11, são modelos gráficos não diretos sem ligações entre nós na mesma camada. Eles são compostos de unidades de entrada (ou visíveis) x_j e unidades neurais ocultas h_i , permitindo o cálculo de $P(h|x)$ e $P(x|h)$.

Figura 11 – Ilustração de uma Restricted Boltzman Machine.



Fonte: Adaptado de Rafael Hrasco et. al., (2015).

Uma RBM normalmente é treinada por um algoritmo de aprendizagem que treina cada camada a cada vez de uma forma gananciosa, para construção de representações mais abstratas dos dados originais através do $P(h^k|x)$.

Hinton chamou este algoritmo de *Contrastive Divergence (CD)*, HINTON, G. (2002).

O método de *Contrastive Divergence* é um método de análise probabilística que analisa o valor do logaritmo do valor de saída de uma RBM que é uma probabilidade e utiliza esse valor num critério de aproximação que toma como parâmetro duas fases uma chamada fase de divergência positiva e uma chamada de fase de divergência negativa.

Uma *Boltzmann Machine* (HINTON, G. E. et. al., 1985); (HINTON & SEJNOWSKI, 1986) é um modelo probabilístico da distribuição conjunta entre as unidades visíveis x , marginalizando sobre os valores das unidades escondidas h , conforme mostrado pela Equação (4.12).

$$P(x) = \sum_h P(x, h) \quad (4.12)$$

Em que a distribuição conjunta entre unidades escondidas e visíveis é associada a uma função quadrática de energia, mostrada pela Equação 4.13.

$$\varepsilon(x, h) = -b'x - c'h - h'Wx - x'Ux - h'Vh \quad (4.13)$$

Com:

$$P(x, h) = \frac{e^{-\varepsilon(x, h)}}{Z} \quad (4.14)$$

$$Z = \sum_{x,h} e^{-\varepsilon(x,h)} \quad (4.15)$$

Em que Z é uma constante de normalização (chamada função de partição) e (b, c, W, U, V) são parâmetros do modelo. b_j é chamado de bias do neurônio visível x_j ; c_i é o bias do neurônio visível h_i e as matrizes W, U e V representam os termos de interação entre os neurônios. Note-se que para U e V diferentes de 0, significa que existem interações entre unidades pertencentes à mesma camada (*layer* escondido ou camada visível).

A probabilidade atribuída para o neurônio da camada visível pode ser calculada pela soma de todas as conexões possíveis entre os neurônios da camada oculta. (aplica-se o mesmo processo para os neurônios da camada oculta).

A probabilidade atribuída para o neurônio da camada oculta pode ser calculada pela soma de todas as conexões possíveis entre os neurônios da camada visível.

$$P(x) = \frac{1}{Z} \sum_h e^{-\varepsilon(x,h)} \quad (4.16)$$

$$P(h) = \frac{1}{Z} \sum_x e^{-\varepsilon(x,h)} \quad (4.17)$$

Observação: Desde que a RBM tenha a forma de um gráfico bipartido, sem conexões entre os neurônios da entrada, as funções de ativação dos neurônios da camada oculta são mutuamente independentes dadas às funções de ativação da camada visível e de forma recíproca, as funções de ativação dos neurônios da camada oculta são mutuamente independentes dadas às funções de ativação da camada visível. Isto é, para ' m ' neurônios visíveis e para ' n ' unidades ocultas a probabilidade condicional de uma configuração das unidades visíveis x , dada uma configuração de h neurônios da camada oculta, é:

$$P(x|h) = \prod_{i=1}^m P(x_i|h) \quad (4.18)$$

Similarmente, a probabilidade condicional de ' h ' dado ' x ' é dada por:

$$P(h|x) = \prod_{j=1}^n P(h_j|x) \quad (4.19)$$

Calculando a função ativação de probabilidade:

Próximo, a função de ativação de probabilidade para os neurônios da camada oculta pode ser dada como:

$$P(h_j = 1|x) = \sigma(B_j + \sum_{i=1}^m W_{ij}x_i); \text{ com: } B_j + \sum_{i=1}^m W_{ij}x_i = \text{Rede}(h_j) \quad (4.20)$$

$$\text{Assim: } \sigma(x) = \frac{1}{1+e^{-x}} e [bias + \sum(W_{ij}x_i)] \Leftrightarrow [bias + x]$$

Similarmente, a função de ativação de probabilidade para os neurônios da camada visível pode ser dada como:

$$P(x_i = 1|h) = \sigma(A_i + \sum_{j=1}^n W_{ij}h_j); \text{ com: } A_i + \sum_{j=1}^n W_{ij}h_j = \text{Rede}(x_i) \quad (4.21)$$

$$\text{Assim: } \sigma(y) = \frac{1}{1+e^{-y}} e [bias + \sum(W_{ij}y_i)] \Leftrightarrow [bias + y]$$

Marginalmente distante de h o nível da energia produz a energia chamada energia livre:

$$\mathcal{F}(x) = -\log \sum_h e^{-\mathcal{E}(x,h)} \quad (4.22)$$

Pode-se descrever a probabilidade logarítmica de acordo com:

$$\log P(x) = \log \sum_h e^{-\mathcal{E}(x,h)} - \log \sum_{\tilde{x}, \tilde{h}} e^{-\mathcal{E}(\tilde{x}, \tilde{h})} = -\mathcal{F}(x) - \log \sum_{\tilde{x}} e^{-\mathcal{F}(\tilde{x})} \quad (4.23)$$

Diferenciando a expressão acima, a inclinação do log-probabilidade no que diz respeito a alguns dos parâmetros do modelo θ pode ser escrito como se segue:

$$\begin{aligned} \frac{\partial \log P(x)}{\partial \theta} &= - \sum_h \frac{e^{-\mathcal{E}(x,h)} \frac{\partial \mathcal{E}(x,h)}{\partial \theta}}{e^{-\mathcal{E}(x,h)}} + \sum_{\tilde{x}, \tilde{h}} \frac{e^{-\mathcal{E}(\tilde{x}, \tilde{h})} \frac{\partial \mathcal{E}(\tilde{x}, \tilde{h})}{\partial \theta}}{e^{-\mathcal{E}(\tilde{x}, \tilde{h})}} \frac{\partial \log P(x)}{\partial \theta} \Rightarrow \\ &\Rightarrow \frac{\partial \log P(x)}{\partial \theta} = \sum_h P(h|x) \frac{\partial \mathcal{E}(x,h)}{\partial \theta} + \sum_{\tilde{x}, \tilde{h}} P(\tilde{x}|\tilde{h}) \frac{\partial \mathcal{E}(\tilde{x}, \tilde{h})}{\partial \theta} \end{aligned} \quad (4.24)$$

Computando $\frac{\partial \mathcal{E}(x,h)}{\partial \theta}$ diretamente. Portanto, se a amostragem a partir do modelo foi possível, pode-se obter um gradiente estocástico para uso na formação do modelo, como se segue. Duas amostras são necessárias: $(h|x)$ de $P(h|x)$ para o primeiro mandato, que é chamada a **fase positiva**, e um par $(\tilde{x}|\tilde{h})$ de $P(\tilde{x}|\tilde{h})$ em que é chamada a **fase negativa**.

4.4.1.2 Deep Belief Networks

Deep Belief Networks (DBN) são redes neurais que seguem um modelo gerador probabilístico, foram introduzidas pela primeira vez em G. HINTON, et. Al., (2006). Modelos geradores proporcionam uma distribuição probabilística dos dados e etiquetas, proporcionando a estimação de $P(\text{Observação} | \text{Etiqueta})$ e $P(\text{Etiqueta} | \text{Observação})$. Modelos discriminativos, como a maioria das redes neurais convencionais, só fornecem $P(\text{Etiqueta} | \text{Observação})$.

Uma DBN, como é ilustrada pela Figura 12 visa mitigar alguns dos problemas de formação de redes neurais convencionais: a necessidade de maior número de dados marcados no conjunto de treinamento; técnicas de treinamento lento e de aprendizagem inadequada para os parâmetros de seleção que levam a mínimos locais.

Além disso, baseadas em redes de crenças sigmóides, redes cujas funções de aprendizagem são sigmóides logísticas, SBNs, uma rede neural multicamadas geradora

Neste caso, $P(x)$ é intratável na prática, exceto em modelos muito pequenos. A camada superior é então fatorada pelo produtório mostrado pela Equação (4.27).

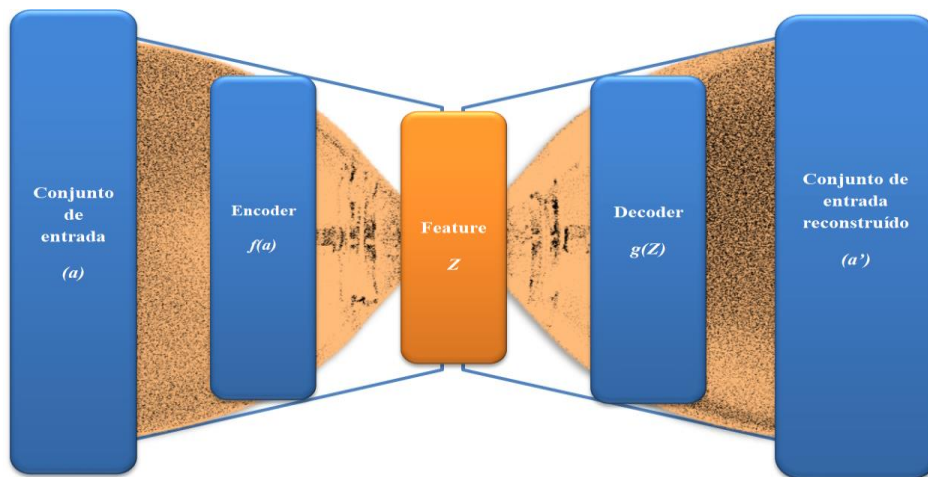
$$P(h^1) = \prod_i P(h_i^1). \quad (4.27)$$

4.4.1.3 Autoencoders

Autoencoders são redes neurais também conhecidas como “*redes diabolô*” BOURLARD, H., et. al., (1988), em que o diabolô é uma estrutura de meio cálice simétrica na direção horizontal formado por dois troncos de cone unidos por uma base mais fina, são denominadas de encriptadores e sua formação é mais simplória do que treinar uma RBM. Os *Autoencoders* têm sido utilizados como blocos de construção de arquiteturas de Deep Learning, em que cada nível é treinado separadamente, desta forma de acordo com a complexidade do problema estas redes podem constituir cadeias de encriptadores resultando em um *Stacked Autoencoder*, BENGIO, Y. et al., (2007) e RANZATO, M.A. et al., (2006). Porém quando não são constituídas em cadeia são consideradas redes de aprendizagem profunda de caráter inicial.

A Figura 14 ilustra a estrutura do diabolô que forma um *Autoencoder*.

Figura 14 – Estrutura diabolô componente da arquitetura de um Autoencoder.



Fonte: O Autor (2021).

Um *Autoencoder* (AE) é treinado para codificar uma entrada a para alguma representação $f(a)$, de modo que ela (entrada) possa ser reconstruída a partir desta representação. A função de decodificação, $g(Z)$, define a reconstrução da rede, e geralmente

é um vector de números obtidos através de uma função que poderá assumir características de uma função sigmoide logística, de saturação linear e característica de linearidade pura.

Para codificar as informações do conjunto de entrada e definir a *feature*, função (Z) do conjunto e posteriormente através desta *feature* decodificar os dados, reconstruindo assim a entrada, o AE realiza dois tipos de transformações ou operações lineares.

Em que, tais operadores lineares são a função extratora de características (*encoder*) que mapeia o conjunto de aprendizagem buscando uma representação otimizada do mesmo, conforme definido pela Equação 4.28. E a função de reconstrutora (*decoder*), que mapeia a representação produzida por $f(a)$ retornando para o espaço original, com $h(a)$ sendo a função identidade, e D e K respectivamente as dimensões dos domínios de codificação e de decodificação, conforme definido pela Equação 4.29.

$$\text{Função encoder: } f: \mathbb{R}^D \rightarrow \mathbb{R}^K; f(a) = s(W * a + b) = Z \quad (4.28)$$

$$\text{Função decoder: } f: \mathbb{R}^K \rightarrow \mathbb{R}^D; \begin{cases} g(Z) = s(W' * Z + b') = a' \\ h(a) = g(f(a)) = a' \end{cases} \quad (4.29)$$

A expectativa é que Z seja uma representação de dados distribuídos que capte o conjunto dos principais fatores das variações da entrada denominado de “*feature*”.

Geralmente a aprendizagem de um AE, a capacidade do mesmo de fixar a forma estrutural de uma base de dados, é feita de maneira não supervisionada. Em que a “*feature*” em questão preserva o conjunto de características concisas referentes aos dados da base sobre análise de reconstrução. Desta forma essa “*feature*” resultante do processo de aprendizagem do AE, pode ser utilizada como inferência num processo de aprendizagem supervisionada.

Quando se deseja aumentar a profundidade da arquitetura formam-se cascatas de *autoencoders*, sendo estas cascatas denominadas de *Stacked autoencoders* (SAE), em que o seu algoritmo de treinamento é descrito a seguir:

Algoritmo para treinamento de um SAE:

- (i) Treinar a primeira camada como um *autoencoder*, a fim de minimizar o erro de reconstrução da entrada original, de forma não supervisionada.
- (ii) As saídas dos neurônios das camadas ocultas são usadas como insumos para outra camada, também treinadas como um *Stacked autoencoder*. Estes dois passos não requerem dados rotulados.
- (iii) Itera-se a etapa dois para inicializar o número desejado de camadas adicionais.
- (iv) Utilize a saída da última camada oculta como entrada para uma camada supervisionada e inicialize aleatoriamente seus parâmetros ou use-os de forma supervisionada.

- (v) Realizar o ajuste fino de todos os parâmetros desta arquitetura profunda através de um critério supervisionado.

A abordagem sensata do pré-treinamento das camadas de uma rede profunda funciona através da formação de uma camada por vez. Os *autoencoders* podem ser "empilhados" de uma forma a conter camadas sábias detentoras de pré-treinamento. Um SAE é uma RNA que consiste em múltiplas camadas de *autoencoders* esparsos em que as saídas de cada camada são ligadas às entradas da camada sucessiva. Formalmente, considere um autoencoder empilhado com n camadas.

Usando a notação para um autoencoder, Suponha que $W^{(k,1)}, W^{(k,2)}, b^{(k,1)}, b^{(k,2)}$ denotem os parâmetros $W^{(1)}, W^{(2)}, b^{(1)}, b^{(2)}$ para o k -ésimo *autoencoder*. Em seguida, o passo de codificação para o *autoencoder* empilhado é dado por executar o passo de codificação de cada camada, do modo a seguir:

$$a^{(l)} = f(z^{(l)}) \quad (4.30)$$

$$z^{(l+1)} = W^{(l,1)}a^{(l)} + b^{(l,1)} \quad (4.31)$$

O passo de decodificação é dada por colocar em execução a pilha de decodificação de cada *autoencoder* em ordem inversa:

$$a^{(n+l)} = f(z^{(n+l)}) \quad (4.32)$$

$$z^{(n+l+1)} = W^{(n-l,2)}a^{(n+l)} + b^{(n-l,2)} \quad (4.33)$$

A informação de interesse está contida dentro de um $a^{(n)}$, que é a ativação da camada mais profunda de unidades escondidas. Este vector nos dá uma representação da entrada em termos das características de ordem superior.

As características do *sacked autoencoder* podem ser usadas para problemas de classificação por alimentação de um $a^{(n)}$ para um classificador *SoftMax*. Uma boa maneira de obter bons parâmetros para um *stacked autoencoder* é usar o treinamento sábio por camada de forma gananciosa. Para fazer isso, em primeiro lugar forma-se a primeira camada sobre a entrada com dados brutos para obter parâmetros $W^{(1,1)}, W^{(1,2)}, b^{(1,1)}, b^{(1,2)}$.

Utiliza-se a primeira camada para transformar a entrada bruta em um vector que consiste das ativações das unidades escondidas, procede-se o treinamento à segunda camada incidindo este vector na mesma para obter parâmetros de $W^{(2,1)}, W^{(2,2)}, b^{(2,1)}, b^{(2,2)}$. Repete-

se esse passo para as camadas subsequentes, utilizando a saída de cada camada como entrada para a camada posterior.

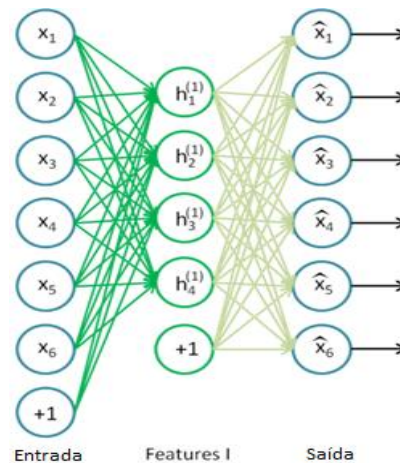
Este método treina os parâmetros de cada camada individual, enquanto congela parâmetros para o resto do modelo. Para produzir melhores resultados, após esta fase de formação ser concluída, é utilizado um ajuste fino por meio da retropropagação com o objetivo de melhorar os resultados para o ajuste dos parâmetros de todas as camadas, sendo esses alterados ao mesmo tempo.

Se o objetivo for apenas ajustes finos para efeitos de classificação, a prática comum é, descartar as camadas de "decodificação" do *autoencoder* empilhado, e vincular à última camada oculta um $a^{(n)}$ para o classificador *softmax*. Os gradientes da função (*softmax*) e o erro de classificação serão então retropropagados nas camadas de codificação.

As Figuras 15, 16, 17 e 18 mostram as etapas de treinamento de um SAE com duas camadas ocultas.

Primeiro, treina-se um autoencoder esparso sobre as matérias-primas $x^{(k)}$ para aprender características primárias $h^{(1)(k)}$ sobre a entrada bruta.

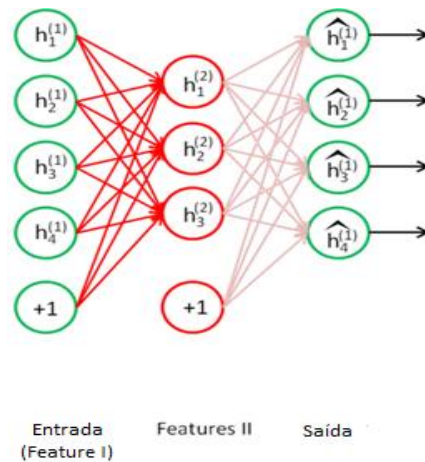
Figura 15 – Ilustração da primeira camada do stacked autoencoder.



Fonte: O Autor (2021).

Em seguida na Figura 16, alimenta-se a entrada bruta para este *autoencoder* escasso treinado, para a obtenção do recurso das Ativações primárias $h^{(1)(k)}$ para cada uma das entradas $x^{(k)}$. A partir daí usam-se esses recursos primários como a "entrada bruta" para outro *autoencoder* escasso subsequente aprender características secundárias $h^{(2)(k)}$ sobre esses recursos primários.

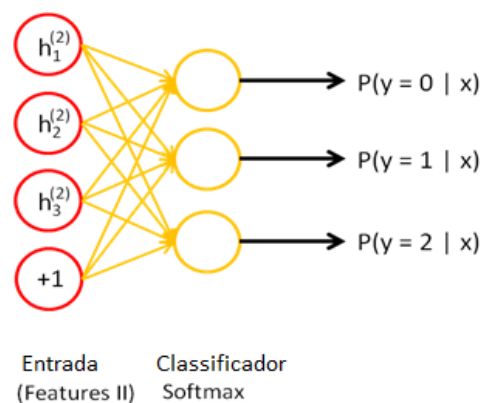
Figura 16 – Ilustração da segunda camada do stacked autoencoder.



Fonte: O Autor (2021).

Após isso, conforme ilustrado na Figura 17, alimenta-se os recursos primários para o segundo *autoencoder* escasso para obter os recursos secundários de ativação $h^{(2)(k)}$ para cada um dos recursos primários $h^{(1)(k)}$ (que correspondem às características primárias da correspondente entrada de dados $x^{(k)}$). Tratam-se então essas características secundárias como "dados brutos" a um classificador *softmax*, treinando-o para mapear rótulos com dígitos de características secundárias.

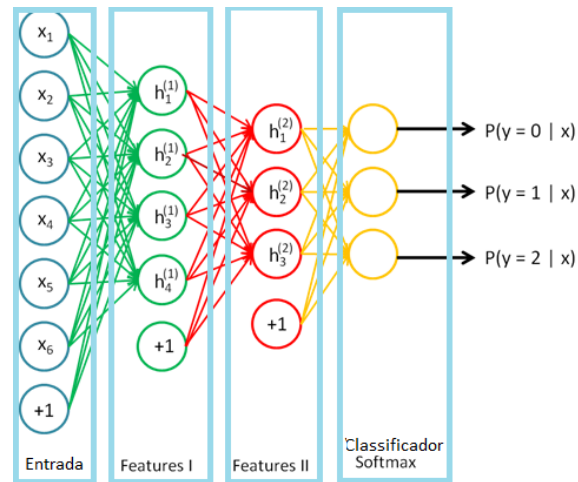
Figura 17 – Ilustração da terceira camada do stacked autoencoder.



Fonte: O Autor (2021).

Finalmente, como mostrado na Figura 18, combinam-se todas as três camadas para formar um *stacked autoencoder* com duas camadas ocultas e uma camada final classificadora com o (*softmax*), capaz de classificar base de dados inserida conforme desejado.

Figura 18 – Ilustração stacked autoencoder resultante do processo de aprendizagem.



Fonte: O Autor (2021).

4.4.2 REDES NEURAIS RECORRENTES

A aprendizagem humana não funciona a partir do “zero” de um determinado conhecimento, pois a cada segundo ao ler, por exemplo, uma informação o cérebro humano entende cada palavra com base em sua compreensão adquirida no conhecimento de palavras que foram lidas anteriormente. Desta forma, após adquirir uma informação o cérebro não descarta tudo que lhe fora apresentado e não começa a pensar do zero novamente. Assim, os pensamentos (informações) como no exemplo da leitura abordado neste parágrafo têm uma aprendizagem com uma fixação variável, esta função numa rede é chamada de *persistência significativa*.

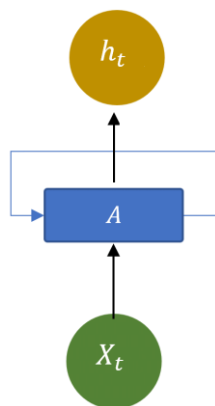
As redes neurais tradicionais do tipo MLP (Multi Layer Perceptron) não conseguem estabelecer bem este conhecimento anterior (Persistência significativa), e quando se tenta realizar esta inclusão de informações anteriores em seus treinamentos, os resultados não são satisfatórios, apresentando grandes erros na reconstrução de padrões, sendo estes erros cometidos em modulo e fase em relação a um instante da série em específico. Por exemplo, imagine que se queira classificar que tipo de evento está acontecendo em todos os momentos de uma série temporal.

Não fica claro como uma rede neural MLP poderia usar seu raciocínio sobre eventos anteriores na série temporal para informar os eventos posteriores com um grau de precisão adequado, mesmo utilizando para atenuar este fato um estudo prévio de correlação para a escolha das melhores entradas que serão utilizadas no conjunto de aprendizagem da rede

MLP. Criando assim um problema notório para um sistema previsor baseado nestas redes, que queira mostrar sua eficiência em todo o horizonte abordado, conforme foi mostrado anteriormente pelas Figuras 4 e 5.

Redes neurais recorrentes abordam esse tipo de problema, possuindo loops em suas arquiteturas, permitindo desta forma que as informações persistam ao longo do treinamento das mesmas.

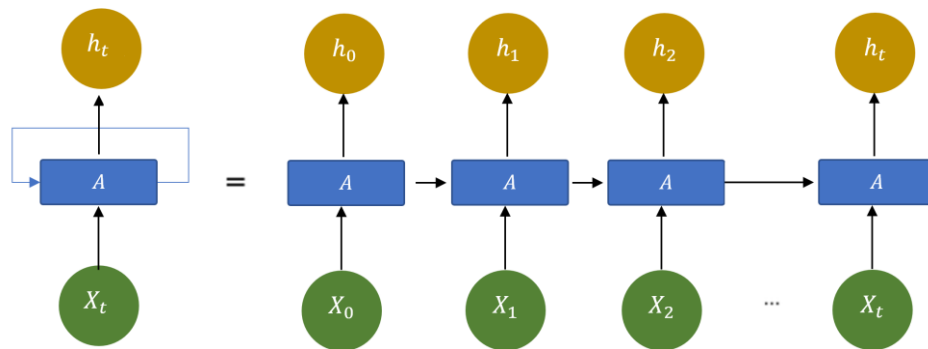
Figura 19 – Diagrama de uma rede neural recorrente.



Fonte: O Autor (2021).

No diagrama mostrado pela Figura 19, um pedaço da rede neural, A, analisa alguma entrada x_t e gera um valor h_t . Um loop permite que as informações sejam passadas de uma etapa da rede para a próxima. Desta forma uma rede neural recorrente pode ser vista como várias cópias da mesma rede, cada uma passando uma mensagem para um sucessor, conforme mostrado pela Figura 20. Essa natureza em cadeia revela que redes neurais recorrentes estão intimamente relacionadas a sequências, séries temporais e listas, sendo este tipo de arquitetura a natural a ser usada para dados dispostos desta forma.

Figura 20 – Significado de uma RNN (Rede Neural Recorrente)



Fonte: O Autor (2021).

Nos últimos anos, houve um sucesso incrível aplicando RNNs a uma variedade de problemas como: reconhecimento de fala, modelagem de idiomas, tradução de idiomas, legendas de imagens e etc. Essencial para esses sucessos é o uso de “LSTMs”, um tipo muito especial de rede neural recorrente que funciona, para muitas tarefas, de forma mais eficaz do que a versão padrão. Quase todos os resultados de referência baseados em redes neurais recorrentes são alcançados com este tipo de rede.

4.4.2.1 O problema das dependências de longo prazo

Uma das utilidades das RNNs é a ideia de que as mesmas podem conectar informações anteriores à tarefa atual, um exemplo disto é o uso de quadros de vídeo (Frames) anteriores que podem informar o entendimento do quadro atual ou futuro. Sua utilidade é notória, porém a sua eficácia na passagem da informação dependerá de algumas condições:

1) Às vezes, necessita-se apenas olhar para informações recentes para executar a tarefa atual. Por exemplo, considere um modelo de linguagem tentando prever a próxima palavra com base nas palavras anteriores. Se o objetivo é prever a última palavra em "as nuvens estão no céu", não é necessário um contexto maior, é bastante óbvio que a próxima palavra será céu. Nesses casos, onde a diferença entre as informações relevantes e o local necessário é pequena, as RNNs podem aprender a usar as informações anteriores.

Mas há também os casos em que se necessita de um maior contexto.

2) Considere tentar prever a última palavra no texto "Eu cresci na França ... falo francês fluentemente". Informações recentes sugerem que a próxima palavra provavelmente é o nome de um idioma, mas se o objetivo é restringir qual o idioma, precisa-se do contexto da França, mais adiante. É totalmente possível a lacuna entre as informações relevantes e o ponto

em que é necessário tornar-se muito grande. Desta forma, à medida que essa lacuna aumenta, as RNNs tornam-se mais incapazes de aprender a conectar tais informações.

Em tese, as RNNs são absolutamente capazes de lidar com essas “dependências de longo prazo”, O que na implementação pode ser extremamente difícil. Um exemplo discutido por dois grandes pesquisadores foi o seguinte:

Um ser humano poderia escolher cuidadosamente parâmetros para resolver problemas com altos níveis de discretização, i.e., que possuam grandes lacunas temporais em suas bases. Porém, Infelizmente, na prática, as RNNs não parecem ser capazes de preencher a tendência dessas grandes lacunas de maneira aceitável. Este problema foi estudado minuciosamente por HOCHREITER (1991) e BENGIO, et al. (1994), que encontraram algumas razões fundamentais que explicaram tal dificuldade.

Tais estudos resultaram em novas redes recorrentes que não apresentaram esta característica, estas redes foram denominadas LSTM (*Long Short Term Memory*).

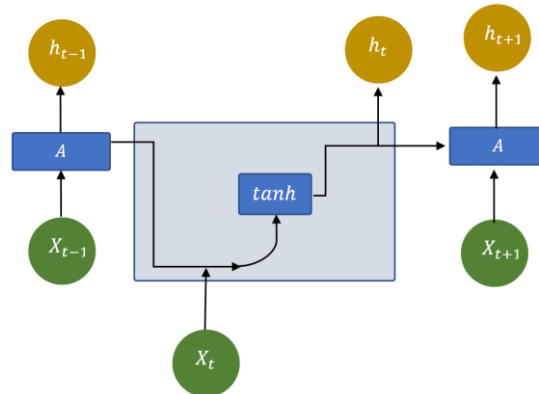
4.4.3 REDES LSTM (LONG SHORT-TERM MEMORY)

As redes de memória de longo prazo - geralmente chamadas de “LSTM” - são um tipo especial de RNN, capaz de aprender dependências de longo prazo. Elas foram introduzidas por HOCHREITER e SCHMIDHUBER (1997) e foram refinadas e popularizadas por muitos pesquisadores em diversos trabalhos. Este modelo de rede trabalha tremendamente bem em uma grande variedade de problemas sendo amplamente utilizado.

As redes LSTMs são projetadas explicitamente para evitar o problema de dependência em longo prazo. Seu comportamento padrão é o de lembrar informações por longos períodos de tempo.

Por se tratar de um tipo de rede neural recorrente, tem a forma de uma cadeia de módulos repetidos de rede neural. Nas RNNs padrão, este módulo de repetição terá uma estrutura muito simples, com uma única camada *tanh* conforme mostra a Figura 21 a seguir.

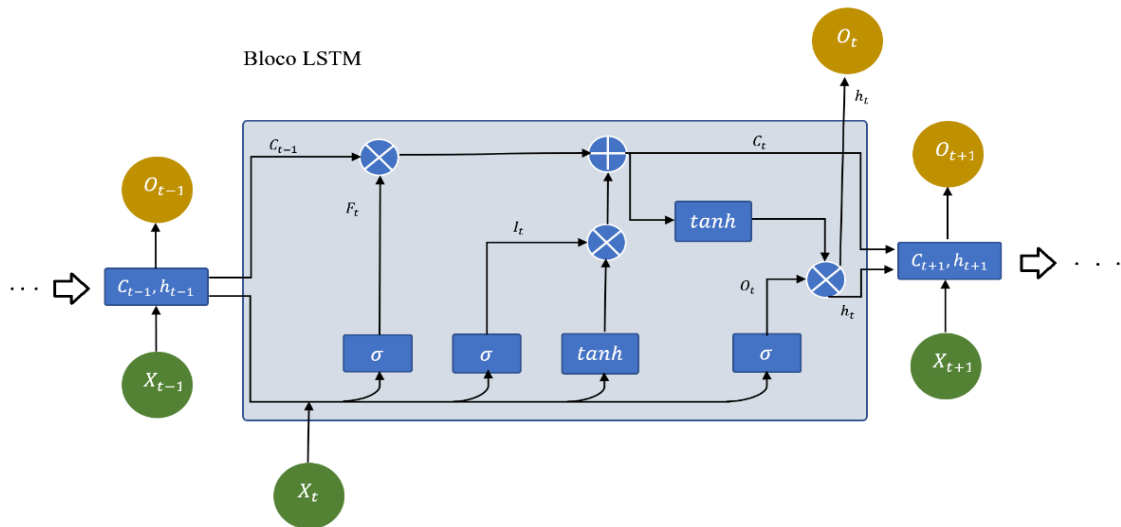
Figura 21 – Módulo de repetição padrão de uma RNN



Fonte: O Autor (2021).

Já em uma rede do tipo LSTM também há presença desta estrutura numa configuração em cadeia, mas o módulo de repetição tem uma estrutura diferente. Em vez de ter uma única camada de rede neural, há quatro, interagindo de uma maneira muito especial, conforme ilustra a Figura 22 a seguir:

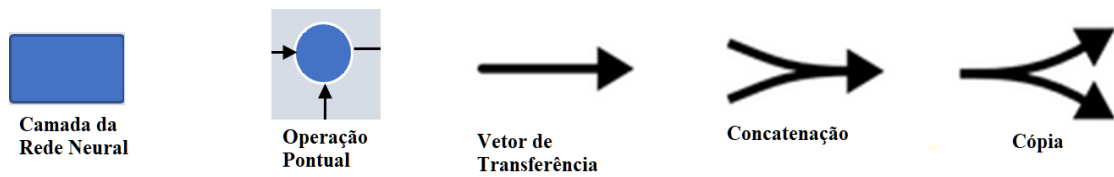
Figura 22 – Módulo de repetição padrão de uma LSTM com 4 camadas



Fonte: O Autor (2021).

Em que, analisando-se o módulo de repetição tomando por base as Figuras 21 e 25, podem-se notar algumas simbologias importantes às quais é necessário enfatizar algumas definições sobre as mesmas.

Figura 23 – Simbologias importantes do modelo de uma LSTM.

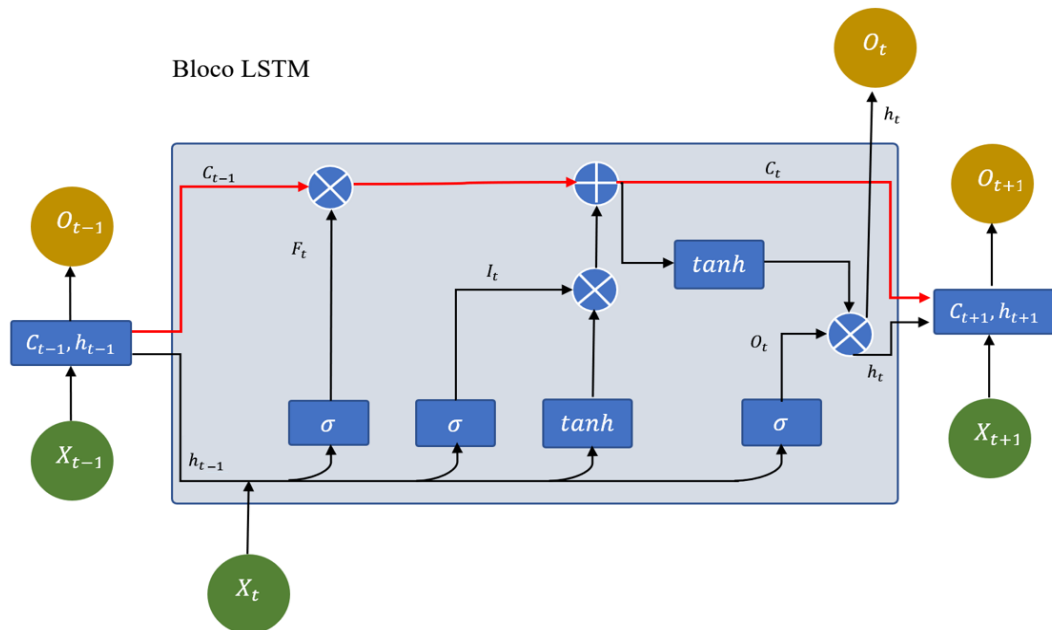


Fonte: O Autor (2021).

Em que, cada linha do modelo carrega um vetor inteiro, da saída de um nó, às entradas de outros nós. Os círculos cor-de-rosa representam operações pontuais, como adição de vetores, enquanto as caixas amarelas são as camadas de apreensão da rede neural. As linhas mescladas denotam concatenação, enquanto uma bifurcação indica o conteúdo sendo copiado e as cópias indo para locais diferentes.

A ideia central por trás do modelo das LSTMs está na linha horizontal que percorre a parte superior do diagrama, esta linha representa o estado da célula, um caminho transportador. Ele percorre toda a cadeia, com apenas algumas interações lineares menores, sendo responsável pela facilidade transporte de dados possibilitando que as informações fluam com suas formas inalteradas, tal linha pode ser vista de acordo com a Figura 24 a seguir:

Figura 24 – Caminho de transporte de informação em uma LSTM.



Fonte: O Autor (2021).

Outro ponto importante do aprendizado de uma LSTM é que nem toda informação é guardada pela célula apenas informações características que contribuam para a manutenção da longa confiabilidade temporal da informação. Para esta função de aproveitamento e triagem da informação que circula pela célula, a rede possui estruturas denominadas (Gates), compostas de uma camada de rede neural sigmoide e uma operação de multiplicação por pontos, que de maneira opcional, permitem a passagem da informação pelo caminho de transporte da informação conforme pode ser visto pela Figura 25, estas estruturas funcionam como portões cuja funcionalidade é a de remover ou adicionar informações ao estado da célula, sendo esta tarefa, cuidadosamente regulada pelo modelo da LSTM.

Esses *gates* são aplicados nos valores das entradas ou saídas da rede, como um vetor de “máscara”, com valores entre 0 e 1, que indicam o quanto de informação deve ser propagada. É bem similar ao *Dropout* 2 em sua forma de funcionamento, mas com essa “máscara” aprendida pela rede para reaproveitar informações da série, ao invés de ser definida aleatoriamente durante o treinamento.

4.4.3.1 Os Principais tipos de Gates utilizados em um bloco LSTM

A seguir são apresentados os três principais tipos de portas de controle de informação ou Gates de uma LSTM, além de alguns aspectos importantes quanto a sua finalidade e qual o impacto da equação funcional na atualização dos parâmetros da aprendizagem.

A rede LSTM utiliza três tipos de portões diferentes sendo os mesmos o *forget gate*, o *input gate* e o *output gate*.

O *forget gate*: é responsável por definir o que deve ser ignorado das entradas da etapa atual. Em que a sua função de transferência F_t é mostrada pela Equação 4.34.

$$F_t = \sigma(W_F * [h_{t-1}, x_t] + b_F) \quad (4.34)$$

O *input gate*: é responsável por definir o que deve ser propagado da etapa anterior para a etapa posterior tendo o caráter básico de um filtro, utilizando desta forma uma informação passada para ditar uma informação futura. Em que a sua respectiva função de transferência é mostrada pela Equação 4.35.

$$I_t = \sigma(W_i * [h_{t-1}, x_t] + b_i) \quad (4.35)$$

O output gate: é responsável por definir o que deve ser passado da etapa atual para a próxima, tendo o caráter de um filtro com uma memória mais curta, utilizando desta forma uma informação presente para ditar uma informação futura. Em que a sua respectiva função de transferência é mostrada pela Equação 4.36.

$$O_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \quad (4.36)$$

Mantendo o curso da informação na passagem pelos **gates** mencionados anteriormente os **gates** F_t e I_t são aplicados de forma a combinar $\tilde{c}$ e c_{t-1} para chegar na saída C_t , enquanto o **output gate** O_t é aplicado na saída da rede h_t . Em que as expressões para $\tilde{C}$, C_t e h_t são mostradas pelas respectivas Equações 4.37. a 4.39.

$$\tilde{C}_t = \tanh(W_c * [h_{t-1}, x_t] + b_c) \quad (4.37)$$

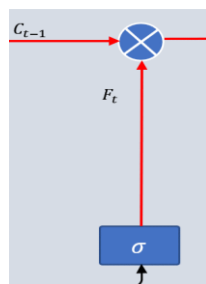
$$C_t = F_t * C_{t-1} + I_t + \tilde{C}_t \quad (4.38)$$

$$h_t = O_t * \tanh(C_t) \quad (4.39)$$

Dessa forma, a rede neural pode aprender com informações de etapas muito anteriores que podem influenciar na predição da etapa atual da série – parecido com a ideia dos métodos autorregressivos para séries temporais.

A camada sigmoide que compõe o portão gera números entre zero e um, descrevendo quanto de cada componente deve ser liberado. Um valor (zero) significa "não deixar passar nada", enquanto o valor (um) significa “deixar passar tudo”. Uma LSTM possui três dessas estruturas em sua concepção com a função de proteger e controlar o estado ótimo para o transporte da informação.

Figura 25 – Estrutura do portão (Gate) de permissão da passagem de informação de uma LSTM.



Fonte: O Autor (2021).

4.4.3.2 Princípio de funcionamento de uma LSTM

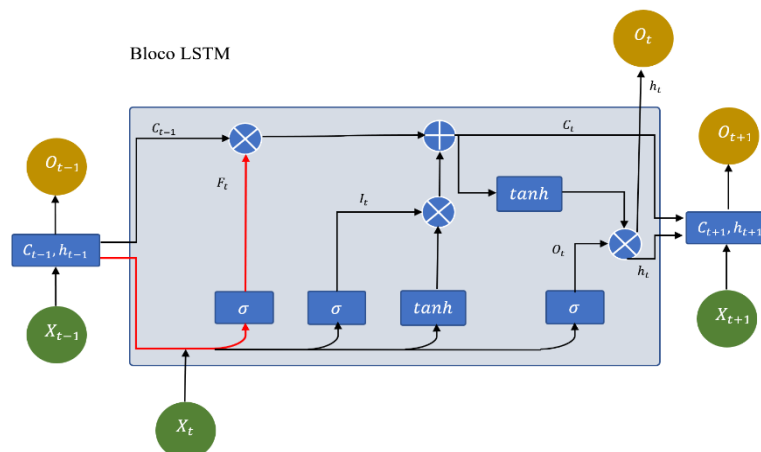
O primeiro passo no modelo de uma rede LSTM é decidir quais as informações devem ser descartadas do estado da célula. Essa decisão é tomada por uma camada sigmóide chamada "camada do portão do esquecimento". Ela olha para h_{t-1} e x_t , e gera um número entre (0) e (1) para cada número no estado da célula C_{t-1} . Um número (1) representa "mantenha isso completamente" enquanto um número (0) representa "se livrar completamente disso".

Um uso bastante relevante deste tipo de rede seria quando houvesse a necessidade de previsão de uma próxima característica atrelada a uma variável em uma série temporal tendo esta característica forte relação com termos anteriores da série sobre estudo. Neste tipo de ênfase de utilização a heurística determina o estado celular da LSTM podendo incluir outras características relevantes da variável que compõe a série. Este conceito é bastante importante e significativo pois esta característica não é bem fundamentada no se diz respeito ao cérebro humano, pois geralmente o cognitivo humano ao se deparar com um novo assunto, característica ou imagem tende a esquecer tudo que foi abordado em passos anteriores.

O passo seguinte é decidir quais as novas informações que devem ser armazenadas no estado da célula. Para isso, inicialmente uma camada *sigmoide* chamada "camada do portão de entrada" decide quais valores serão atualizados. Em seguida, uma camada *tanh* cria um vetor de novos valores candidatos, C_t , que pode ser adicionado ao estado. Na próxima etapa, esses dois (camada e vetor), serão combinados para criar uma atualização para o estado.

Esta camada é representada pela Figura 26 mostrada a seguir:

Figura 26 – Estrutura do portão (Gate) de permissão da passagem de informação de uma LSTM.

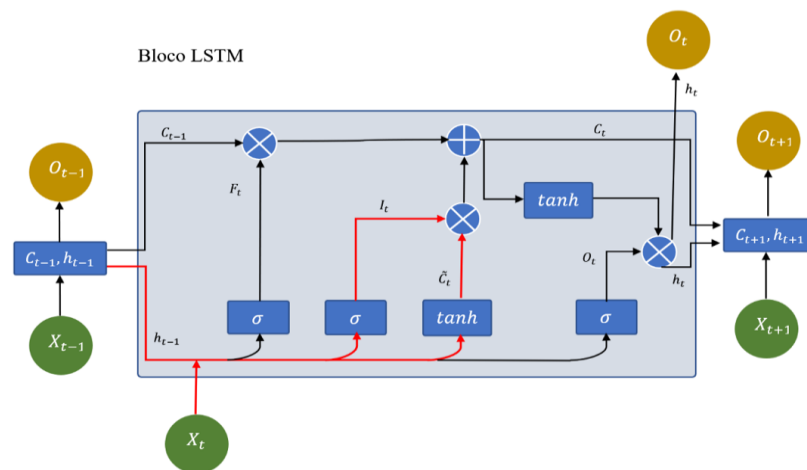


Fonte: O Autor (2021).

Para o caso de um modelo de previsão de uma variável qualquer, a característica futura da variável é tendenciada pela característica passada, porém a mesma poderá servir como mantenedora de estado passado ou alteradora do estado futuro, sendo esta percepção designada de estado celular controlada respectivamente por uma função (Sigmoide logística, Tangente hiperbólica, ...) atrelada ao *gate* da camada.

O próximo passo é adicionar uma nova informação ao estado presente da célula, com o objetivo de excluir o estado antigo que a rede está esquecendo, conforme é ilustrado pela Figura 27. a seguir:

Figura 27 – Adição de informação ao estado presente da Célula em uma LSTM.

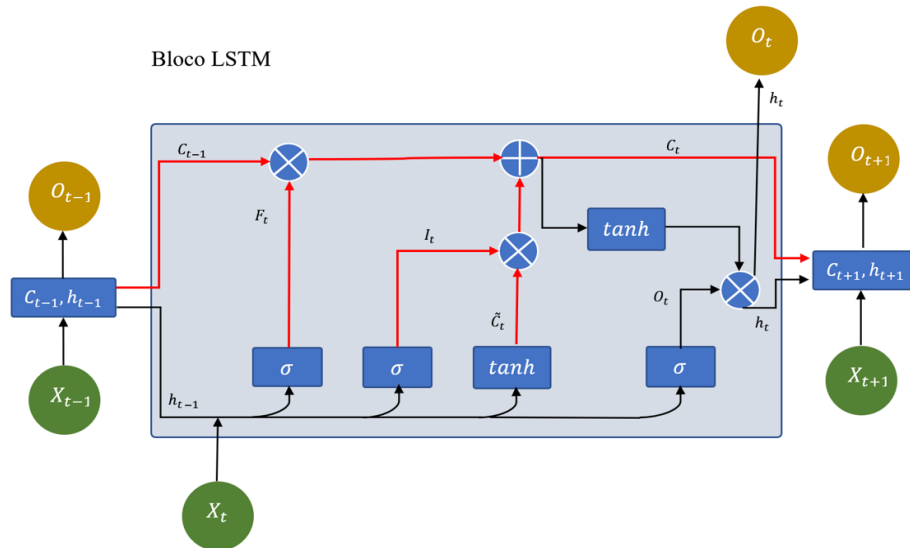


Fonte: O Autor (2021).

Para atualizar o antigo estado da célula, C_{t-1} , para o novo estado da célula, C_t , neste momento a heurística da LSTM definiu as etapas anteriores já decidindo desta forma o que fazer com a informação em trânsito. Durante o trânsito da informação, o estado antigo (C_{t-1}) é multiplicado por $F(t)$, definindo um novo estado antigo que não leva em conta as informações passadas, isto é, faz o novo estado da célula esquecer o que fora definido anteriormente. Então se adiciona o termo $(i(t) * \tilde{C}_t)$, ao estado novo estado antigo $(C_{t-1}) * F(t)$. Estes são os novos valores candidatos da série temporal, ditados pela LSTM, valores estes dimensionados quando o modelo da célula decide atualizar cada novo valor de estado.

Este passo de ação no funcionamento da LSTM pode ser mostrado pela Figura 28 a seguir:

Figura 28 – Passo de atualização da informação em uma LSTM.

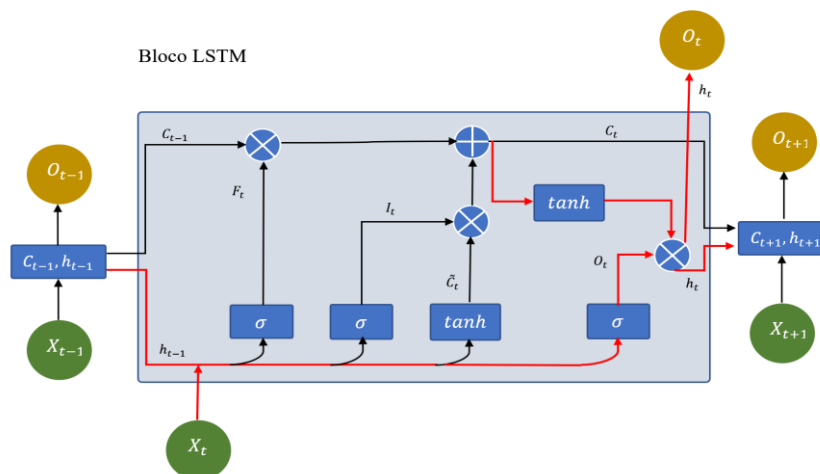


Fonte: O Autor (2021).

Por fim, decide-se o a saída que será produzida pelo modelo. Essa saída será baseada no estado da célula, mas será uma versão filtrada. Para isto primeiramente a informação é transformada pela camada sigmoide que decide quais partes do estado da célula deve-se produzir. Em seguida, é feita a inserção do estado da célula em $\tanh$ (com o objetivo de levar os valores das respectivas saídas numa faixa compreendida entre -1 e 1) e assim multiplicam-se estas saídas pelas saídas da porta sigmoide, de forma que apenas são produzidas as saídas em que há a decisão no transporte da informação.

Esta parte é ilustrada pela Figura 29, que é mostrada a seguir:

Figura 29 – Processo de Decisão na construção da saída em uma LSTM



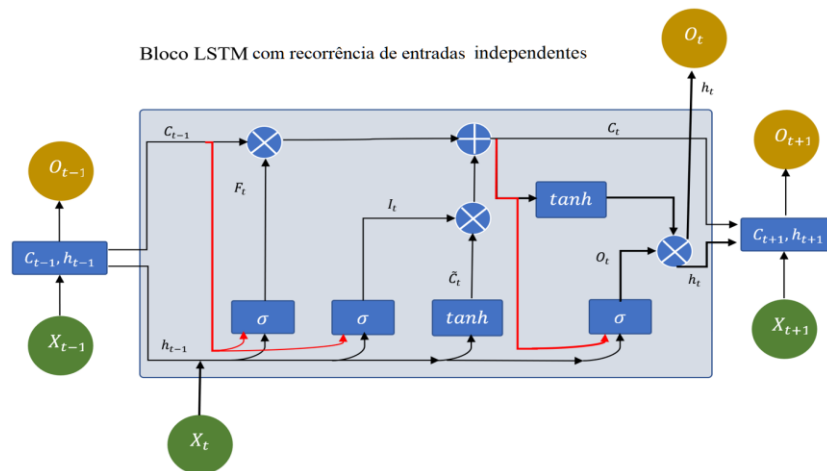
Fonte: O Autor (2021).

4.4.3.3 Principais variações no funcionamento dos Modelos de uma LSTM

Nesta Tese utilizou-se o modelo padrão de LSTM para formulação de modelos previsores de dados de ventos, porém nem todas as heurísticas sobre as redes LSTM são iguais aos tipos tratados nesta tese. De fato, vários trabalhos envolvendo estas redes, principalmente no que diz respeito ao tratamento de bases de imagens, utilizam uma versão ligeiramente modificada sendo desta forma, diferentes da implementação padrão.

Uma variante popular de LSTM, introduzida por GERS & SCHMIDHUBER (2000) e posteriormente explorada em GERS & SCHMIDHUBER (2002), está adicionando "conexões de olho mágico". Isso significa que este tipo de rede deixa as camadas de *gate* visualizarem o estado da célula.

Figura 30 – Diagrama relativo ao processo de recorrência independente por entrada, inerente ao processo de decisão na construção da saída.



Fonte: O Autor (2021).

De forma a ilustrar o comportamento das funções que compõe o bloco LSTM com recorrência independente por entradas as Equações 4.40 a 4.42 demonstram as mudanças associadas as funções F_t ; I_t e O_t respectivamente.

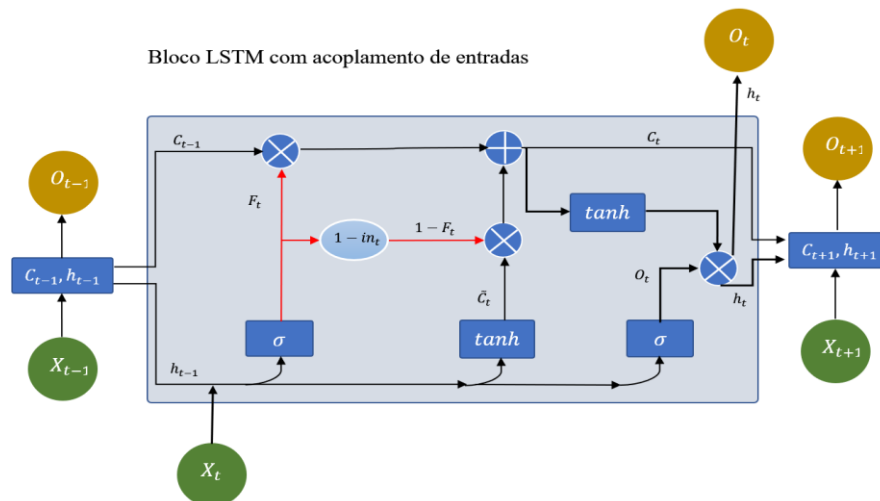
$$F_t = \sigma(W_F * [C_{t-1}, h_{t-1}, x_t] + b_F) \quad (4.40)$$

$$I_t = \sigma(W_i * [C_{t-1}, h_{t-1}, x_t] + b_i) \quad (4.41)$$

$$O_t = \sigma(W_o[C_t, h_{t-1}, x_t] + b_o) \quad (4.42)$$

O diagrama acima adiciona brechas para todos os *gates*, porém muitos dados fornecem brechas e outros não. Outra variação é usar os *gates* de esquecimento e de entradas acopladas em vez de decidir separadamente o que esquecer e para o qual devem-se adicionar novas informações, logo a decisão na célula é tomada por este conjunto (*Gate* (esquecimento) + *Gate* (acoplamento de entradas)). Assim a informação só é de fato esquecida quando se insere algo em seu lugar, sendo esta inserção com valores novos se e somente se, desprezar-se a informação mais antiga.

Figura 31 – Diagrama relativo ao processo de recorrência em uma LSTM GRU com acoplamento de entradas, inerente ao processo de decisão na construção da saída.



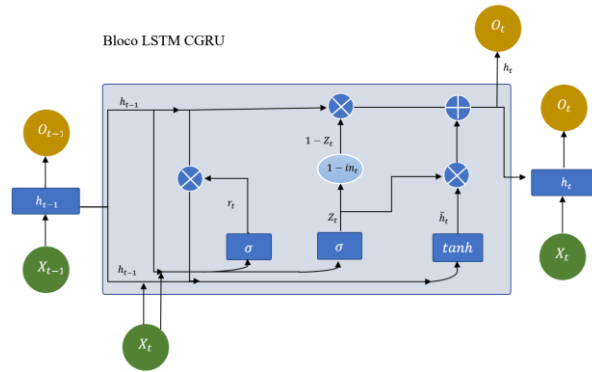
Fonte: O Autor (2021).

Com esse tipo de configuração de acoplamento entre as informações de entrada a função C_t sofre uma pequena alteração que é mostrada pela Equação 4.43 a seguir.

$$C_t = F_t * C_{t-1} + (1 - F_t) * \tilde{C}_t \quad (4.43)$$

Uma variação ligeiramente mais complexa de LSTM é a Unidade Recorrente Fechada, ou GRU (*Gated Recurrent Unit*), introduzida por CHO et al. (2014). Combina os *gates* de esquecimento e de entrada em um único "*gate* de atualização". Ele também mescla o estado da célula e o estado oculto e faz outras alterações. O modelo resultante é mais simples que os modelos de LSTM padrões e tem se tornado cada vez mais popular entre os pesquisadores.

Figura 32 – Diagrama relativo ao processo de recorrência em uma LSTM CGRU com acoplamento de entradas, inerente ao processo de decisão na construção da saída.



Fonte: O Autor (2021).

Assim novas alterações nas funções que descrevem o comportamento do bloco LSTM GRU surgem descrevendo novas composições, expressas pelas funções Z_t , r_t , $\tilde{h}_t$, e alterando a função h_t , estas alterações são mostradas respectivamente pelas Equações 4.44 a 4.47 a seguir.

$$Z_t = \sigma * (W_z * [h_{t-1}, x_t]) \quad (4.44)$$

$$r_t = \sigma * (W_r * [h_{t-1}, x_t]) \quad (4.45)$$

$$\tilde{h}_t = \tanh(W * [r_t * h_{t-1}, x_t]) \quad (4.46)$$

$$h_t = (1 - Z_t) * r_t * h_{t-1} + Z_t * \tilde{h}_t \quad (4.47)$$

Estas são apenas algumas das variantes LSTM mais notáveis. Existem muitos outros, como os RNNs fechadas por YAO, et al. (2015) e há também uma abordagem completamente diferente para lidar com dependências de longo prazo, como as *Clockwork* RNNs de KOUTNIK, et al. (2014).

O treinamento do bloco LSTM pode ser feito utilizando algoritmos de aprendizagem como o ADAM (*Adaptive moment estimation*), SGDM (*Stochastic gradient descent with momentum*) e RMSProp (*Root Mean Square Propagation*). Nesta Tese o algoritmo utilizado foi o ADAM, devido aos melhores resultados obtidos depois de sucessivos testes, cuja rotina de aprendizagem definida é mostrada abaixo:

Rotina de Aprendizagem das LSTM nos Parques 01 e 02

- (i) Escolhe-se o banco de dados, podendo ser do parque 01 ou parque 02;
- (ii) São definidas as dimensões do vetor de entrada $[v(t)]$ ou $[v(t), d(t)]$;
- (iii) É definida a arquitetura de Rede. Como a arquitetura MLP anterior tinha duas camadas ocultas, foram definidas 2 camadas profundas;
- (iv) Definem-se pesos iniciais aleatórios para as camadas profundas;
- (v) Define-se o intervalo de busca para definição da melhor quantidade de neurônios nas camadas profundas para ambos os parques;
 Busca Camada profunda 1- (38:1:110)
 Busca Camada profunda 2- (38:1:110)
Obs: os valores limites do intervalo de procura foram encontrados após sucessivos testes em que se verificou se a mudança alcançava um impacto significativo na diminuição do erro.
- (vi) Define-se o número de épocas, sendo este, igual a 1000,
- (vii) Utiliza-se o algoritmo de aprendizagem ADAM, proposto em KINGMA, D., P. (2015), para definir a melhor arquitetura e o melhor desempenho dos modelos de LSTM propostos, com a finalidade de previsão mais precisa do banco de dados do parque escolhido;
- (viii) Comparam-se os MSEs dos conjuntos de aprendizagem, (treinamento – 80%, validação – 10% e teste – 10%), com os modelos definidos e escolhe-se então o melhor modelo de arquitetura da LSTM relativo à base de dados.

5 REVISÃO DO ESTADO DA ARTE SOBRE A PREVISÃO DE GERAÇÃO EÓLICA

As descrições decorrentes neste capítulo dizem respeito às inovações propostas em termos de melhoria da previsão de curto prazo nas últimas décadas. Este capítulo é baseado em relatórios de (GIEBEL et al, 2011), tendo como base o quadro tecnológico de projetos europeus do segmento *SafeWind* e *ANEMOS.plus*.

As pesquisas no campo da previsão de energia eólica iniciaram no final dos anos 80, devendo-se este fato ao aumento do uso da energia eólica nas matrizes energéticas mundiais. Tal fato tornou a previsão uma ferramenta crucial para diversos operadores de sistemas de transmissão em todo mundo.

Para este enfoque um dos documentos mais relevantes é o gerado pelo *Argonne National Laboratory* (MONTEIRO et. al., 2009), este relatório introduz o conceito de previsões numéricas meteorológicas (NWP), detalhando a visão dos modelos de mercados comerciais de energia aglutinando os conceitos de NWP aos mercados de energia, especialmente nos Estados Unidos.

Em ERNST et. al. (2007) mostram-se alguns casos de aplicação dos NWPs para o uso internacional, e a conclusão desta aplicação foi que na combinação de modelos e previsões para regiões maiores e horizontes mais curtos os índices de erro médio foram reduzidos.

LANGE e FOCKEN (2005) enfatizam os desenvolvimentos na Alemanha, PINSON et. al. (2007a) abordaram o caráter probabilístico da previsão concluindo que as melhorias resultantes estavam ligadas a características específicas como regimes climáticos, previsões conjuntas e aspectos temporais relacionados à incerteza da previsão.

COSTA et al. (2007), fizeram uma revisão sobre os últimos 30 anos de aplicação da previsão de curto prazo em energia eólica, concluindo com uma lista de tópicos não resolvidos ou até inexplorados. Desta forma propuseram pesquisas adicionais sobre a estimação de parâmetros adaptativos, além de novas abordagens em terrenos complexos, como por exemplo: uma distribuição de ventos mais precisa e computacionalmente viável, sendo estes modelos, (Modelos de Fechamento para Ferramentas em Microescala).

Outra aplicação internacional em curto prazo foi o estudo encomendado pela Canadian Wind Energy Association, este trabalho realizado por Garrad Hassan (hoje parte do Germanischer Lloyd Group), faz uma varredura dos estudos em curto prazo, incluindo informações sobre modelos de previsão, sua avaliação, previsão de experiências no mundo inteiro e um resumo detalhado dos modelos previsores.

5.1 Horizontes de Previsão

Sobre o horizonte de abordagem, no curto prazo, está o fato da variabilidade do vento, tal comportamento acontece independente de escalas de análise temporal. Entretanto duas abordagens são mais relevantes: Uma para o controle da turbina em si (de milissegundos para segundos) e a outra para acoplamento da turbina à rede elétrica sendo esta última dependente das constantes de tempo da rede do sistema elétrico, podendo variar de (minutos a semanas).

Como horizontes de aplicações relevantes podem-se citar os seguintes exemplos:

- a) A Otimização do escalonamento de usinas convencionais por funções como despacho econômico, etc. Os horizontes de previsão podem variar entre 3 a 10 h, dependendo do tamanho do sistema e do tipo de unidades convencionais incluídas (*ou seja, para sistemas que incluem apenas unidades convencionais rápidas, como geradores a diesel ou turbinas a gás, o horizonte pode ser inferior a 3 h*).

Apenas algumas aplicações on-line totalmente integradas desse tipo são atendidas hoje normalmente, esses sistemas são usados para sistemas de energia menores ou isolados, como sistemas insulares, embora a otimização para sistemas maiores, como a Irlanda, esteja sendo avaliada.

- b) A Otimização do valor da energia produzida no mercado e em grandes sistemas elétricos. Tais previsões são exigidas por diferentes tipos de usuários finais (utilitários, operadores de mercado, TSOs, DSOs, ESPs, IPPs, comerciantes de energia, etc.) e para diferentes funções, tais como compromisso de unidade, despacho econômico, avaliação dinâmica de segurança, gerenciamento de congestionamento, participação no mercado de eletricidade, etc.

De acordo com, (MORENO et. al., 2003) o horizonte de previsão pode ir até 48 h a frente. Como relataram (STILL e GRAINGER, 2001): *horizontes mais curtos também podem ser considerados para manutenção, quando é importante que a tripulação possa retornar com segurança das turbinas offshore à noite*.

O Operador do Sistema de Distribuição do Noroeste da Alemanha, (EWE) (FOCKEN et. al., 2009) está integrando as previsões de vento nas rotinas de manutenção de transformadores para avaliar o carregamento de linha dos fluxos de eletricidade remanescentes.

- c) Alocação de reservas com base na alimentação eólica esperada. Esta alternativa visa a segurança do sistema sendo feita na Irlanda de acordo com (BARRY, 2006) e em Portugal de acordo com (BESSA et. al., 2012).

- d) Além disso, escalas de tempo ainda mais longas seriam interessantes para o planejamento de manutenção de grandes componentes de usinas de energia, turbinas eólicas ou linhas de transmissão. No entanto, de acordo com (MORENO et al., 2003), a precisão das previsões meteorológicas diminui fortemente olhando para 5 e 7 dias de antecedência .

5.2 Modelos Típicos Utilizados

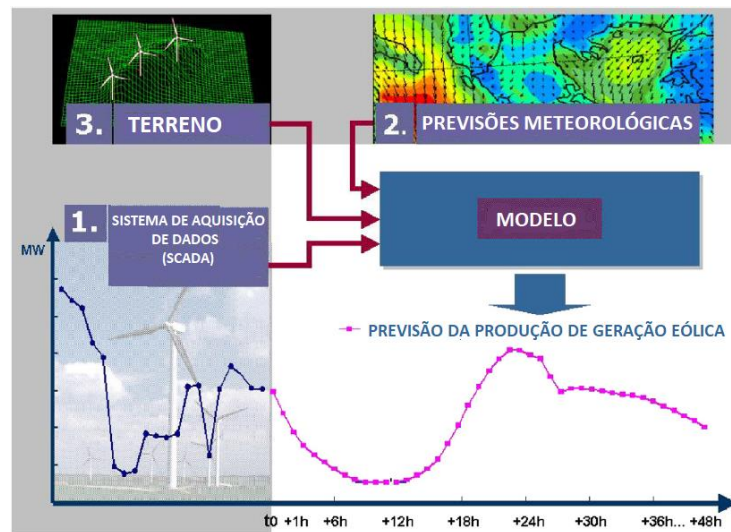
Geralmente, a classificação dos modelos está ligada a presença ou a falta de um modelo NWP. Onde a grande questão é se esforço e despesas compensam a inclusão do modelo NWP. Em termos de comparação modelos NWP superam as abordagens de modelos de séries temporais após cerca de 3 à 6h de tempo de antecipação.

Assim, concessionárias utilizam esta abordagem na maioria dos seus modelos operacionais, sendo estes uma combinação de modelos de fluxo atmosféricos e modelos estatísticos. Desta forma os modelos físicos averiguam o maior tempo possível para alcançar a melhor estimativa da velocidade do vento, antes de usar os modelos estatísticos de saída que tem o objetivo de minimizar o erro da previsão.

Os modelos estatísticos em sua forma pura tentam encontrar as relações entre uma riqueza de variáveis explicativas, incluindo os resultados do NWP, e dados de potência medidos on-line, geralmente empregando técnicas recursivas. Frequentemente, modelos de caixa-preta como mínimos quadrados recursivos avançados ou redes neurais artificiais (RNAs) são usados.

Os modelos estatísticos mais bem-sucedidos empregam, na verdade, modelos de caixa cinza, em que algum conhecimento das propriedades da energia eólica é usado para ajustar os modelos ao domínio específico. Alguns dos modelos estatísticos podem ser expressos analiticamente, outros (como as RNAs) não podem. Os modelos estatísticos podem ser usados em qualquer estágio da modelagem e, na maioria das vezes, combinam várias etapas em uma. A cadeia típica de modelos para modelos de previsão de energia eólica é ilustrada na Figura 33.

Figura 33 – As várias abordagens de previsão podem ser classificadas de acordo com o tipo de entrada.



Fonte: Adaptado de Kariniotakis, Renewable Energy Forecast (2017).

A resolução da escala de um modelo NWP pode resolver problemas com um valor um pouco elevado, segundo (PIELKE e KENNEDY, 1980); (YOUNG e PIELKE, 1983) e (SKAMAROCK, 2004), mesmo para uma resolução de 2 km, apenas recursos na ordem de 10 km são realmente levados em consideração. Isso significa que os modelos de microescala, exceto em casos de terreno muito simples, devem ser capazes de melhorar as previsões do NWP, embora essa questão permaneça um tópico de pesquisa.

O processo de *downscaling* produz uma velocidade e direção do vento para a localização da turbina. Baixas velocidades são convertidas para altas velocidades referentes a altura do cubo do aerogerador. Este vento é então convertido em potência através uma curva de potência.

A abordagem mais fácil é a da curva de potência do fabricante, embora pesquisas tenham mostrado vantagem em estimar a curva de potência a partir da velocidade e direção do vento prevista e da potência medida. A maioria dos modelos estatísticos atuais sai dessa etapa e faz uma previsão direta da produção de energia para turbinas simples ou parques eólicos inteiros. Dependendo do horizonte de previsão e disponibilidade, os dados de potência medidos podem ser usados como entrada adicional.

Na maioria dos casos, os dados reais melhoram os erros residuais em uma Abordagem estatística *Model Output Statistics* - (MOS). Se os dados on-line estiverem disponíveis, um modelo recursivo de autocalibração é altamente vantajoso. Isso faz parte da abordagem estatística. Esta abordagem pode ter a forma de um modelo estatístico explícito empregado com métodos estatísticos avançados autorregressivos, ou como uma caixa preta tipo RNA.

No entanto, às vezes apenas dados *off-line* estão disponíveis, com os quais o modelo pode ser calibrado em retrospectiva. Nos últimos anos, vários operadores de sistemas exigiram que os dados on-line dos parques eólicos fossem usados especificamente em suas ferramentas de previsão *on-line*.

5.3 Modelos de séries temporais

Se o horizonte de previsão não for muito longo, a velocidade e a potência do vento podem ser previstas usando apenas métodos de análise de séries temporais, sem recorrer a previsões meteorológicas reais. Modelos de séries temporais diretas são modelos que usam valores observados recentes de vento e outras variáveis para prever a velocidade futura do vento.

BOSSANYI (1985) usou um Filtro de Kalman (KF) com os últimos seis valores como entrada e obteve uma melhoria de até 10% no erro RMS sobre persistência para dados médios de 1 min para a previsão do próximo passo de tempo. Esta melhora diminuiu para médias mais longas e desapareceu completamente para as médias de 1 hora.

FELLOWS e HILL (1990) usaram previsões de duas horas à frente de velocidades de vento de 10 minutos em um modelo da rede elétrica das Ilhas Shetland. Sua abordagem foi usar a previsão otimizada e iterativa do método Box-Jenkins a partir de dados retificados, que então foram submetidos à suavização da média móvel central.

NOGARET et al., 1994, relataram que para o controle de um sistema insular de tamanho médio, a previsão simples com o melhor desempenho é uma média dos últimos dois ou três valores, ou seja, 20 e 30 min. TANTAREANU (1992) descobriu que modelos de média móvel autorregressiva (ARMA) podem executar até 30% melhor que a persistência para 3e10 passos à frente em médias de 4 segundos de dados amostrados de 2,5 Hz

MADSEN (1995) e NIELSEN e MADSEN (1996) descobriram que a modelagem de dois estágios (conversão de previsões de velocidade do vento em energia eólica, em que a estrutura de correlação em medições de potência é desconsiderada) é geralmente inferior a modelos que levam em conta a correlação de potência.

TORRES et al. (2005) usam um modelo ARMA para prever velocidades médias horárias do vento para cinco locais em Navarra. Eles usaram parâmetros específicos de site e mês para o modelo ARMA. O modelo ARMA geralmente superou a persistência para a previsão de 1 h, e sempre foi melhor no RMSE e no MAE para horizontes mais altos até 10h à

frente. Os dois sítios têm um RMSE ligeiramente superior em geral, mas ainda estão no mesmo intervalo que os outros sítios. Em geral, as melhorias de 2% e 5% para a previsão de 1 h correspondem a uma melhoria de 12% e 20% para a previsão de 10 h.

PINSON et al. (2008c) descobriram que a energia eólica e especialmente a variabilidade de energia eólica de grandes parques eólicos offshore (Horns Rev e Nysted) ocorrem em certos regimes e, portanto, testaram “abordagens de mudança de regime baseadas em critérios observáveis (baseadas na recente produção de energia eólica) ou sequências de regime não observáveis (isto é, uma cadeia de Markov oculta)” para uma previsão em uma etapa de dados de potência de 1 min, 5 min e 10 min. Em seu paper fica demonstrado que os modelos propostos MSAR superam os modelos base em performance cujo índice RMSE varia em redução de erro entre 19% a 32% dependendo do parque eólico e da resolução temporal a ser considerada.

KUSIAK et al. (2009b) explorou a viabilidade de cinco algoritmos de data mining para velocidade do vento e energia eólica. Previsão de 1 passo a 3 passos à frente. Dois dos cinco algoritmos tiveram um desempenho particularmente bom. O algoritmo de regressão máquina de vetores de suporte SVM fornece previsões precisas de energia eólica e velocidade do vento em intervalos de 10 min até 1h à frente, enquanto o algoritmo perceptron multicamada é preciso na previsão de geração em intervalos de uma hora até 4h.

Vários modelos para o curto prazo foram propostos com base na aprendizagem de máquina como: técnicas de aprendizado incluindo RNAs, redes neurais difusas e outras. Essas abordagens dão mais flexibilidade para considerar variáveis explicativas como entrada. As melhorias relatadas em relação à persistência são frequentemente comparáveis aos modelos clássicos de séries temporais, como AR, ARMA, etc.

No entanto, esses modelos exigem um esforço especial para lidar com questões de dimensionalidade. Um esforço considerável é necessário para selecionar apropriadamente a entrada mais relevante entre o conjunto de variáveis disponíveis e para selecionar os parâmetros estruturais dos modelos (ou seja, número de neurônios, camadas, etc.) de modo a evitar sobreajuste (*overfitting*).

Em DUTTON et al. (1999) um modelo autorregressivo linear e um modelo baseado em lógica *fuzzy* para os casos de Creta e Shetland são propostos. Eles encontraram pequenas melhorias em relação à persistência para um horizonte de previsão de 2 h, mas até 20% na melhoria do erro RMS para um horizonte de 8 horas. No entanto, para horizontes mais longos, a faixa de confiança de 95% continha a maioria dos valores prováveis de velocidade do vento

e, portanto, uma abordagem baseada em meteorologia foi considerada mais promissora nessa escala de tempo.

KARINIOTAKIS propôs diferentes tipos de modelos de previsão baseados em RNA e lógica *fuzzy* em KARINIOTAKIS (1996) e KARINIOTAKIS et al. (1996). A ênfase é dada na seleção dos parâmetros estruturais dos modelos através de várias abordagens, incluindo a validação cruzada. Em KARINIOTAKIS et al. (1999) vários métodos são testados com dados de ilhas na Grécia. Estes incluíram modelos lineares adaptativos, modelos de lógica difusa adaptativa e modelos baseados em *wavelets*.

Modelos adaptativos baseados em lógica *Fuzzy* foram instalados para operação online no quadro do projeto Joule II CARE (JOR3-CT96-0119).

5.4 Modelos de séries temporais para previsão de prazos muito curtos

Se o horizonte de previsão não for muito longo, a velocidade e a potência do vento podem ser modeladas usando métodos de análise de séries temporais, sem recorrer a previsões meteorológicas reais. Modelos de séries temporais diretas são modelos que usam valores observados recentes de vento e outras variáveis para prever a velocidade futura do vento.

Para horizontes de muito curto prazo, as escalas de tempo relevantes são influenciadas por:

- a) A mecânica da turbina eólica: tipicamente o gerador, a caixa de engrenagens, o mecanismo de guinada e, acima de tudo, a regulação do passo (da lâmina). As escalas de tempo envolvidas estão na ordem da turbulência, isto é, segundos. O objetivo é o controle ativo das turbinas eólicas. O vento nessas escalas de tempo é inerentemente não-estacionário, conforme mostrado em (VINCENT et al., 2010b), e pode ser melhor previsto com um lidar com o vento e um simples esquema de advecção do campo de alguns segundos à frente do rotor.
- b) O tipo do sistema de energia no qual as turbinas eólicas estão integradas. É típico que os sistemas insulares menores consistam em geradores a diesel com escalas de tempo bastante curtas.

A abordagem típica é usar técnicas de análise de séries temporais, incluindo métodos baseados em mineração de dados e técnicas de inteligência artificial, como redes neurais.

5.5 Modelagem meteorológica para energia eólica

Os dados de previsão numérica do tempo, *Numerical Weather Prediction* (NWP) são a forma mais familiar de dados de um modelo meteorológico. Os modelos de computador NWP processam as observações meteorológicas atuais para prever a condição meteorológica futura. A saída é baseada em observações meteorológicas atuais, que são assimiladas na estrutura do modelo e usadas para produzir previsões de temperatura, precipitação e centenas de outros elementos meteorológicos dos oceanos ao topo da atmosfera.

O principal erro na previsão final vem da entrada meteorológica. Por exemplo, SÁNCHEZ et al. (2002) mostraram que o instrumento estatístico espanhol SIPREOLICO executado com a entrada de velocidade do vento no local tem um grau de explicação muito maior do que as previsões do HIRLAM (Modelo de área limitada de alta resolução, um modelo NWP desenvolvido pela parceria dos institutos da Dinamarca, França, Noruega, Finlândia, Espanha e Irlanda.).

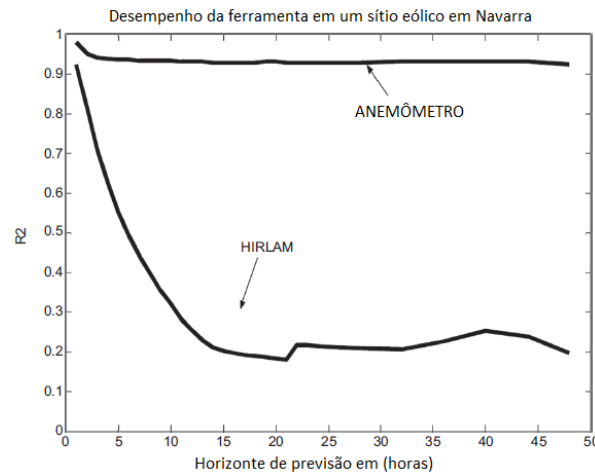
MÖHRLÉN (2001) analisou a resolução necessária para uma aplicação bem-sucedida da previsão do NWP. Em um estudo com o modelo dinamarquês HIRLAM para um sítio eólico na Irlanda, ela aponta as razões pelas quais os modelos NWP estão fornecendo uma precisão inadequada das velocidades do vento na superfície.

Outro fator é que a resolução topográfica não é boa o suficiente para explicar, por exemplo, efeitos de túnel de vento em vales. Previsões precisas exigem modelos de alta resolução cobrindo uma grande área. No entanto, executar ambos é numericamente muito caro. Para melhorar o estado das coisas, ela calculou a Potência diretamente dentro do modelo NWP. Esse fato trouxe a vantagem de que “*propriedades físicas importantes como aspereza de direção, densidade e estratificação da camada limite atmosférica pode ser usada nos cálculos*”.

O erro principal na previsão final vem da grandeza meteorológica. Por exemplo, SÁNCHEZ et al. (2002) mostraram que o instrumento estatístico espanhol SIPREOLICO executado com a entrada de velocidade do vento no local tem um grau de explicação muito maior do que as previsões do HIRLAM. Isto significa que, dada uma velocidade de vento representativa, SIPREOLICO pode prever a potência de forma confiável, já a entrada de velocidade do vento do modelo NWP contribui significativamente para a diminuição da precisão. Portanto, é lógico tentar melhorar a entrada do NWP para melhoria significativa na precisão da previsão como mostra a Figura 5.2.

Isto significa que, dada uma velocidade representativa do vento, o SIPREOLICO pode prever a potência muito bem. Desta forma a entrada de velocidade do vento do modelo NWP diminui significativamente a precisão. Portanto, a tentativa lógica foi melhorar a entrada do NWP para obter melhorias significativas na precisão da previsão, conforme ilustra a Figura 34:

Figura 34 – O erro vem das previsões meteorológicas numéricas.



Fonte: Adaptado de Kariniotakis, Renewable Energy Forecast (2017).

A Figura 34 mostra a diferença no grau de explicação do vento entre o SIPREOLICO executado com entrada HIRLAM (*de uma antiga versão do espanhol HIRLAM*) e o SIPREOLICO executado com entrada de velocidade do vento no local tendo como base de aquisição um anemômetro.

5.6 Modelos de previsão de curto prazo com NWP

Esta seção trata de modelos usando dados do NWP como entrada que inclui vários modelos comerciais. Provavelmente, o modelo mais antigo foi desenvolvido por McCarthy (1998) para a Área de Recursos Eólicos da Califórnia Central. Foi executado nos verões de 1985 e 87 em uma calculadora programável HP 41CX, usando observações meteorológicas e observações aéreas superiores locais. O programa foi construído em torno de um estudo climatológico do local e teve um horizonte de previsão de 24 h. Ele prevê velocidades médias diárias de vento com melhor habilidade do que a persistência ou a climatologia sozinha.

Em 1990, TROEN e LANDBERG (1990) e LANDBERG (1994) desenvolveram um modelo de previsão de curto prazo, agora conhecido como PREDIKTOR, baseado no

raciocínio físico semelhante à metodologia desenvolvida para o atlas Eólico Europeu (TROEN e PETERSEN, 1998). A ideia é usar a velocidade do vento e a direção de um NWP, transformar esse vento para o local do parque eólico, usar a curva de potência e, finalmente, com isso modificar a eficiência do parque.

Observe que o módulo de melhoria estatística MOS pode ser aplicado antes da transformação para o vento local, antes da transformação para alimentação, ou no final da cadeia de modelos para operar na própria energia. Uma combinação de tudo isso também é possível.

O WPPT (*Wind Power Prediction Turbine*) foi desenvolvido pelo Instituto de Informática e Modelagem Matemática (IMM) da Universidade Técnica da Dinamarca. O WPPT tem funcionado operacionalmente na parte ocidental da Dinamarca desde 1994 e na parte oriental desde 1999.

O modelo usa a estimativa adaptativa do método recursivo dos mínimos quadrados (RLS) nos parâmetros de modelos paramétricos condicionais para encontrar a melhor conexão entre as velocidades previstas do vento do NWP para o local e a potência medida para cada horizonte de previsão. Uma parte central deste sistema são os modelos estatísticos para a previsão de curto prazo da produção de energia eólica em parques eólicos ou áreas. Para aplicações *on-line*, é vantajoso permitir que as estimativas de função sejam modificadas à medida que os dados se tornam disponíveis.

Além disso, como o sistema pode mudar lentamente com o tempo, as observações devem ser reduzidas à medida que se tornam mais velhas. Por este motivo, é aplicado um método de estimação adaptativo e recursivo no tempo. Dependendo dos dados disponíveis, o sistema de modelagem WPPT emprega uma hierarquia de modelagem altamente flexível para calcular as previsões da energia eólica disponível de turbinas eólicas em uma região. Para uma região maior, isso geralmente é feito separando a região em várias subáreas.

As previsões de energia eólica são então calculadas para cada subárea e, em seguida, resumidas para obter uma previsão para a região total. As previsões para a região total são calculadas para um número de parques eólicos de referência usando medições *on-line* da produção de energia, bem como NWP como entrada.

Na Universidade de Oldenburg (BEYER et al., 1999) desenvolveram, uma abordagem bastante semelhante aos sistemas de previsão da época, semelhante aos sistemas como o PREDIKTOR, e o HIRLAM. Eles o chamaram de PREVIENTO (FOCKEN et al., 2001a). Já MÖNNICH (2000) mostrou uma boa visão geral sobre os parâmetros e modelos que

influenciam o resultado de um o sistema de previsão de curto prazo. Ele descobriu que o modelo mais importante dos vários submodelos usados é o de estabilidade atmosférica.

O uso de MOS (*Model Output Statistics*), um meio para remover o erro residual, foi considerado muito útil. No entanto, como o modelo NWP foi alterado com frequência, o uso de uma técnica recursiva foi recomendado. Uma grande influência foi encontrada em relação à curva de potência. A curva teórica de potência dada pelo fabricante e a curva de potência encontrada a partir dos dados pode ser bastante diferente.

Na verdade, até mesmo a curva de potência estimada a partir de dados de diferentes anos pode mostrar fortes diferenças. Este último pode ser devido a uma revisão completa da turbina. A maior influência no erro foi considerada proveniente do próprio modelo NWP.

5.7 Modelagem da velocidade do vento versus energia eólica

As ferramentas de previsão de energia eólica que consideram a modelagem “física” podem integrar módulos específicos para prever inicialmente a velocidade do vento e convertê-la em energia através de um modelo de curva de potência mais ou menos avançada do parque eólico (isto é, integração de modelagem de efeitos de ativação). É provável que a cadeia de modelos se beneficie de formas de pós-processamento estatístico, como o sistema MOS.

Alguns modelos envolvem modelagem explícita da curva de potência do parque eólico. Tais modelos podem receber uma previsão de velocidade do vento e convertê-la em previsões de Geração eólica. A modelagem da curva de potência a partir da velocidade do vento foi feita por CABEZON et al. (2004).

Eles usaram cinco métodos baseados em ferramentas estatísticas (modelos lineares com métodos de *binning*, uma técnica de pré-processamento que objetiva uma forma de agrupar um número de valores contínuos maiores ou menores para um número menor de "compartimentos", e um modelo de lógica *fuzzy*) e encontraram uma melhoria, sendo que quanto mais precisos os modelos foram, mais efeitos levaram em consideração.

COLLINS et al. (2009) apontaram que, para grandes parques eólicos (Potência instalada acima de 100 MW), os efeitos locais variam em todo o parque, portanto uma simples aplicação melhorada de uma curva de potência de fabricantes não é boa o suficiente. Um modelo de curva de potência de parque eólico avançado que leva em consideração a densidade do ar, o campo de fluxo heterogêneo e o efeito esteira e com as medições locais

sendo alimentado com velocidades de vento entre mastros reduziu o MAE da potência para o modelo de alimentação de 7,5% para 1,5% em um parque no Reino Unido.

Alimentado com as previsões reais do NWP com horizonte de 24h, o erro do dia seguinte foi reduzido em 1,2%. Para um parque nos Estados Unidos, os números foram de 11,6%, 4,6% e 0,9% de melhoria, respectivamente.

KUSIAK et al. (2009a) usaram as abordagens de mineração de dados de MÖHRLÉN (2001) com entrada de NWP para prever até 12h e até 84h adiante. Como entrada, eles usaram dois modelos NWP dos EUA, o modelo *Rapid Update Cycle* (RUC) e o *North American Mesoscale* (NAM), com os 16 pontos de dados ao redor do parque eólico. Ambos os modelos tinham velocidade e direção do vento em vários níveis, além de densidade do ar e diferença de temperatura potencial.

5.8 Os projetos ANEMOS e outros grandes projetos empresariais de atividades P & D

O projeto ANEMOS (“Desenvolvimento de um Sistema de Previsão de Recursos Eólicos de Próxima Geração para a Integração em Grande Escala de Parques Eólicos “Onshore” e “Offshore”, foi um projeto de P & D de 4 anos iniciado em outubro de 2002. Foi financiado pela Comissão Europeia sob o quinto programa-quadro (ENK5-CT-2002-00665). Vinte e dois parceiros de sete países participaram, incluindo institutos de pesquisa, universidades, empresas industriais, empresas de serviços públicos, TSOs e agências de energia.

Este projeto reuniu pela primeira vez os principais atores europeus no campo, sob um projeto de pesquisa ambicioso que permitiu avanços significativos em vários tópicos, incluindo modelos probabilísticos, modelagem física, pré-padronização de sistemas de previsão etc. Uma ação de coordenação chamada POW'WOW (Previsão de Ondas), Wakes and Offshore Wind, ver (powwow.risoe.dk) foi o próximo passo para os parceiros principais, juntamente com especialistas em ondas e vigílias. Funcionou de 2006 a 2009, financiado também pelo Comissão [Contract No 019898 (SES6)].

Como principal acompanhamento, foram desenvolvidos dois projetos, *ANEMOS.plus* e *SafeWind*, ambos apoiados pela Comissão Europeia. O primeiro teve um forte aspecto demonstrativo, concentrando-se no benefício máximo para os usuários finais através de ferramentas para programação de usinas e armazenamento, para melhorar o comércio nos mercados, para integrar vento ao curto prazo, gerenciar sistemas insulares e outros aplicativos.

O objetivo era desenvolver e demonstrar, em ambientes reais, ferramentas estocásticas de tomada de decisão e otimização capazes de levar em conta as incertezas relacionadas à previsão de energias renováveis.

Mais estudos de pesquisa foram feitos no último, *SafeWind* (Nele, a ênfase principal era em eventos extremos, sejam os mesmos meteorológicos, característicos do sistema elétrico ou de extremidade financeira. Pinson e o *SafeWind Consortium* (2010) compilaram um catálogo de eventos extremos, resultado de uma pesquisa entre os participantes do projeto. Mais importante ainda, um evento extremo era muito dependente do usuário. Um meteorologista responderia “um desvio significativo da climatologia”, um ator no sistema de energia responderia que precisa ter consequências consideráveis, seja financeiramente ou para a segurança do sistema, e um operador de parques eólicos está interessado em eventos de pequena escala que poderiam ser perigosos para as turbinas. Como novo parceiro importante existe o ECMWF, para melhorar o uso de produtos de conjunto para energia eólica.

Os três projetos de R & D acima mencionados produziram, no período 2002 a 2012, mais de 250 publicações em periódicos revisados por pares e conferências internacionais e trouxe à tona várias ferramentas usadas em práticas de negócios.

Um grande projecto nacional de R & D na Europa no domínio das energias renováveis previsão nos últimos anos tem sido o projeto EWELINE na Alemanha (PROJEKTE-EWELINE). O objetivo geral do projeto é melhorar a energia eólica e Previsões de energia fotovoltaica combinando melhores modelos de previsão de potência e otimizando previsões do tempo.

Os objetivos técnicos incluem, entre outros, a otimização das previsões meteorológicas para a melhoria da energia eólica e das previsões de PV. A assimilação de novos tipos de dados nos modelos NWP, por exemplo, dados de produção de energia, é considerada. Com relação às previsões probabilísticas, o foco é na geração de conjuntos e calibração de conjuntos.

Outro projeto importante foi o Projeto de Melhoria da Previsão do Vento (WFIP) no EUA (2011e13). Seu objetivo era melhorar a precisão do vento de curto prazo (0 a 6 h) previsões de energia para a indústria de energia eólica. A WFIP foi patrocinada pelos EUA Departamento de Energia (DOE), com parceiros que incluíam a Administração Nacional Oceânica e Atmosférica (NOAA), empresas de previsão privada, laboratórios nacionais do DOE, operadores de redes e universidades. A WFIP empregou dois caminhos para melhorar as previsões de energia eólica: primeiro, através da coleta de observações especiais a serem

assimiladas pelos modelos de previsão e, segundo, pela atualização dos modelos e conjuntos de previsão do NWP (JAMES WILCZAC et al., 2015).

Em termos de redes, o COST Action WIRE ES1002 (Weather Intelligence for Renewable Energies) reuniu atores de 27 países (2010e14, <http://www.wire1002.ch>). Esta rede desenvolveu uma série de atividades, incluindo uma competição internacional para previsão de energia eólica e solar. Um grande grupo de especialistas internacionais em previsão de curto prazo reunidos no *IEA Wind Task 36* (ver ieawindforecasting.dk) apresentou recentemente uma lista de tópicos de pesquisa para os próximos anos (GIEBEL et al., 2016). Entre outras coisas, recomenda-se a utilização de métodos de ciclo de atualização rápida, uma interação modelada melhor entre parques eólicos, melhorias no modelo de NWP e assimilação de dados, por exemplo, de energia de turbinas eólicas.

5.9 Conclusões

A previsão de curto prazo consiste em várias etapas. Para um horizonte de previsão de mais de 6 h em frente, começa com um modelo NWP. Outras etapas são o *downscaling* dos resultados do modelo NWP para o local, a conversão da velocidade do vento local em energia, e o *upscaling* do único vento leva a energia a toda uma região. Em todas essas frentes, melhorias foram feitas desde os primeiros modelos. Os números típicos de precisão são um RMSE de cerca de 9% e 14% da capacidade instalada de energia eólica para um horizonte de 24 horas. Quando sobe para a previsão regional ou nacional, o desempenho é da ordem de 3% e 5% da capacidade instalada.

O principal erro em um modelo de previsão de curto prazo origina-se do modelo NWP. Uma estratégia atual para superar essa fonte de erro e fornecer uma estimativa da incerteza de uma previsão específica é usar conjuntos de modelos, seja usando vários modelos de NWP ou usando condições iniciais diferentes dentro deles.

Espera-se que a tecnologia de informação e comunicação desempenhe um papel importante na integração de previsões de energia eólica na infra-estrutura de mercado. Outro aspecto da “mercantilização” da previsão de curto prazo é a integração de ferramentas de tomada de decisão para os usuários finais, como o módulo de otimização de agendamento ou o módulo de negociação desenvolvido e integrado pelo *ANEMOS.plus*.

6 MODELOS PREVISORES DE VELOCIDADE DO VENTO E VELOCIDADE E DIREÇÃO DO VENTO, MODELOS DE CURVAS DE POTÊNCIA E MODELO PREVISOR DE GERAÇÃO EÓLICA PROPOSTOS

A fim de esclarecer características, conceitos e definições da heurística a respeito dos modelos previsores de velocidade de vento, velocidade e direção de vento e dos modelos previsores de potência propostos nesta tese, a seção 6 procura explicar e comentar os principais conceitos relativos ao desenvolvimento dos modelos previsores deste trabalho.

6.1 Modelos previsores de velocidade do vento e modelos previsores de velocidade e direção do vento

O modelo previsor de velocidades e velocidades e direções do vento é responsável por alimentar os modelos da curva de potência tratados neste trabalho, o mesmo poderá tomar duas heurísticas distintas sendo uma com redes do tipo MLP e outra com redes profundas do tipo LSTM e cada um dos tipos foi configurado em duas arquiteturas distintas de acordo com o modelo da curva de potência escolhido, tais arquiteturas dependerão do número de variáveis de análise, podendo estas ser (velocidade média [m/s]) ou (velocidade média [m/s] e direção média[°]).

6.1.1 HEURÍSTICA MULTI LAYER PERCEPTRON (MLP)

Esta heurística utilizou parâmetros descritos, estudados e implementados em ALBUQUERQUE, J.C. (2015) e em AQUINO R. R. B. et. al. (2016), as mesmas são nomeadas RNA1 e RNA2, sendo definidas conforme a seguir:

6.1.1.1 Previsor RNA1 e seus parâmetros

A rede neural MLP1 foi nomeada de RNA1 possuindo vinte e sete variáveis de entrada ($[h(t+1)] + [Mês(t+1) (seno)] + [Mês(t+1) (cosseno)] + [vinte e quatro séries horárias de velocidades médias do vento reais]$) e vinte e quatro variáveis de saída (séries horárias de velocidades médias do vento previstas).

Como o objetivo inicial desse modelo foi uma previsão horária de velocidades médias de vento, a RNA1 ficou responsável por esta predição, tendo a tarefa de alimentar a CP-RNAM1A do modelo M1a com suas previsões.

Para tal, a mesma consistiu em uma rede neural de previsão de velocidades médias de vento com horizonte de 24h, ou seja, o passo de previsão variou da hora um até a hora vinte e quatro. Assim, para um valor de entrada $v(t)$ a RNA1 irá prever valores conforme descrito pela Expressão 6.1.

$$Previsões = \{\hat{v}(t+1), \hat{v}(t+2), \dots, \hat{v}(t+24)\} \quad (6.1)$$

No que diz respeito à preparação da RNA1, o tamanho dos conjuntos utilizados dentro do padrão montado, seguiu o que fora descrito no conjunto de aprendizagem. Para garantir uma igualdade de influências entre todas as variáveis no treinamento da rede, foi necessária a normalização dos conjuntos de entrada e saída. Para entender tal fato, é preciso lembrar que as funções de ativação em um neurônio artificial são definidas em faixas limitadas. Assim, normalizando os conjuntos, estes valores de entrada e saída são definidos nestas funções em seus extremos segundo GOUVEIA, H. T. V. (2011).

Com base no que foi dito acima, neste trabalho utilizou-se a função de normalização conforme descrito pela Expressão 6.2, com valores definidos no intervalo amostral fechado de $[0,1]$.

$$x_{Norm} = \frac{x(t) - x_{Min}}{x_{Max} - x_{Min}} \quad (6.2)$$

Quanto à escolha da RNA1 e sua respectiva parametrização, escolheu-se uma rede do tipo MLP (*Multi layer perceptron*), que utiliza retro propagação do erro (*Feed -forward backpropagation error*), com a seguinte arquitetura: $[27 - 14:34 -24]$, conforme descrito abaixo:

- 27 entradas
- [14 até 34 neurônios] na camada oculta.
- [24 neurônios] na saída

Sobre a escolha dessa arquitetura, foi adotado o procedimento definido em AQUINO, R. R. B. et.al (2009), em que, o número de nodos da camada oculta, é discriminado através de um processo de tentativas, nas quais, busca-se um melhor desempenho por parte da rede. Em que tal procedimento é discriminado a seguir:

- 1) Primeiramente, foi criada uma rede com quatorze neurônios na camada oculta;
- 2) Foram atribuídos pesos aleatórios para todas as conexões da rede, e realizou-se o treinamento da rede através do algoritmo RP (*Resilient Propagation*) e foi calculado o MSE para o conjunto de treinamento, validação e teste;
- 3) O passo anterior foi repetido dez vezes. Após a décima repetição foi calculada a média do MSE para cada conjunto (treinamento, validação e teste);
- 4) Outro neurônio foi adicionado à camada oculta, e os passos 2, 3 e 4 foram repetidos, até a rede possuir trinta e quatro nodos na camada oculta;

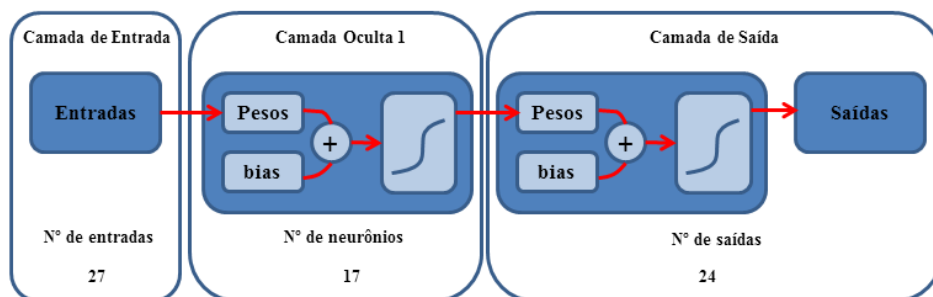
Assim, foi escolhida a arquitetura que apresentou menor MSE para o conjunto de validação.

Sobre o método de treinamento que adequou os valores das entradas às suas respectivas saídas, validando assim a aprendizagem por parte da RNA1, utilizou-se o algoritmo “*Resilient Propagation*”, conhecido comumente no simulador como “*Train RP*”. Enfatizando que, o mesmo só foi adotado em preferência ao outro método, “*Train LM*”, que utiliza o algoritmo de Levenberg Marquardt, e é tecnicamente mais preciso que o seu anterior, por questões de esforço computacional.

Tal fato pôde ser notado nas tentativas de implementação das simulações relativas ao treinamento da RNA1, em que as mesmas resultaram em sucessivos erros com a seguinte mensagem, “*Out of Memory*”. Logo, concluiu-se que, seria necessário mais memória para tal nível de precisão, e assim tomou-se a escolha de trocar o algoritmo de treinamento.

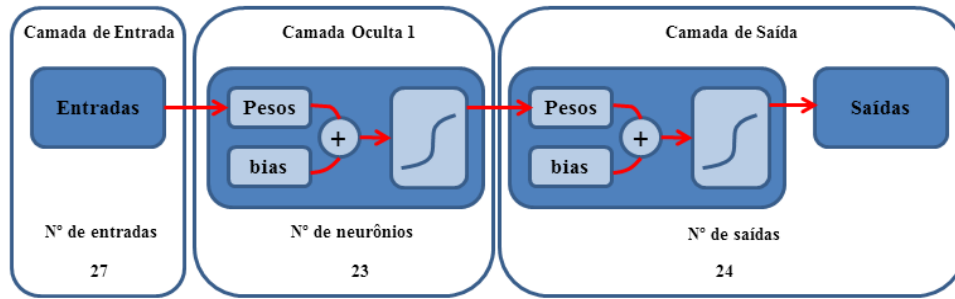
Em que, as melhores RNA1, denominadas “RNA1_BEST”, obtida nas inicializações que simularam os treinamentos, nas bases dos Parques 01 e 02 são mostradas nas Figuras 35 e 36 respectivamente:

Figura 35 – Arquitetura da melhor RNA1 para o Parque 01 escolhida com a execução dos treinamentos.



Fonte: O Autor (2015).

Figura 36 – Arquitetura da melhor RNA1 para o Parque 02 escolhida com a execução dos treinamentos.

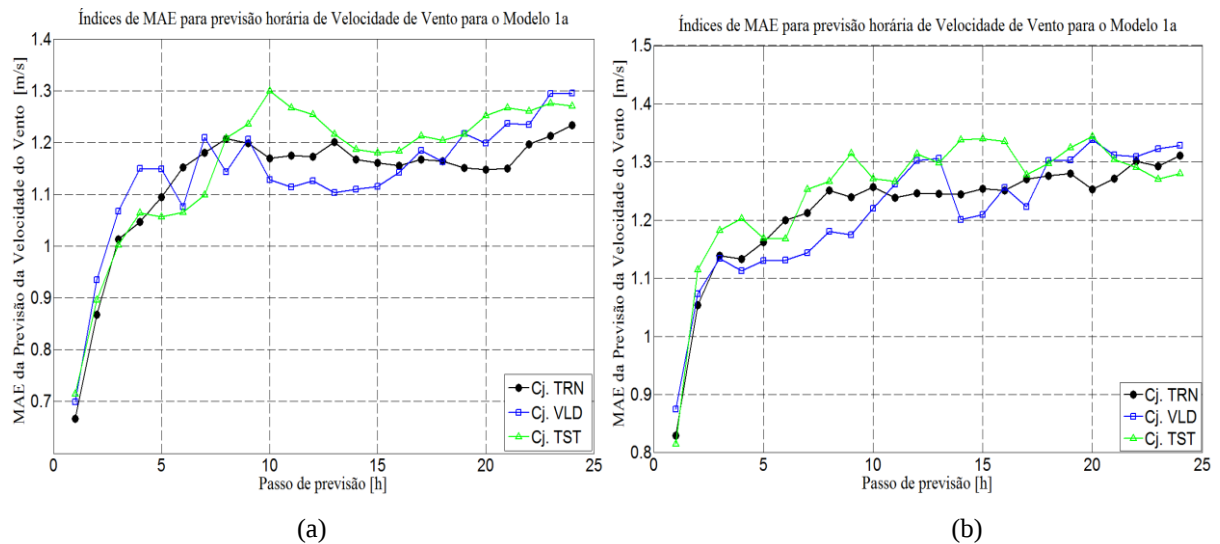


Fonte: O Autor (2015).

Das Figuras 35 e 36, pode-se notar que, para a melhor inicialização, o número de neurônios definidos pela camada oculta foi de 17 para o Parque 01 e 23 para o Parque 02, descrevendo assim a arquitetura da RNA1 após os treinamentos.

Outros resultados importantes relativos a estes treinamentos são descritos pela Figura 37 que descreve o desempenho do erro médio absoluto cometido para ambos os parques nas previsões das respectivas RNA1 dos Parques 01 e 02. Em que os mesmos são mostrados a seguir.

Figura 37 – Comportamento gráfico do MAE referente às simulações para o período de aprendizagem da RNA1 (a) aplicado ao Parque 01 e (b) aplicado ao Parque 02.



Fonte: O Autor (2015).

6.1.1.2 Previsor RNA2 e seus parâmetros

A rede neural MLP2 foi nomeada de RNA2 possuindo cinquenta e um conjuntos de variáveis de entrada, sendo estas ([h(t+1)]+[Mês(t+1)(seno)]+[Mês(t+1)(cosseno)]+[vinte e quatro séries horárias de velocidades médias do vento reais] +[vinte e quatro séries horárias de direções médias do vento reais]) e 48 variáveis de saída, sendo estas ([vinte e quatro séries horárias de velocidades médias do vento previstas] +[vinte e quatro séries horárias de direções médias do vento previstas]).

A mesma consistiu-se em uma rede neural de previsão de velocidades e direções médias de vento com horizonte de 24h, ou seja, o passo de previsão variou da hora um até a hora 24. Assim para valores de entrada $v(t)$ e $d(t)$ a RNA2 irá prever valores de velocidades e direções médias, conforme descrito pela Expressão 6.4:

$$\text{Previsões} = \{\hat{v}(t+1), \hat{v}(t+2), \dots, \hat{v}(t+24), \hat{d}(t+1), \hat{d}(t+2), \dots, \hat{d}(t+24)\} \quad (6.4)$$

De forma semelhante ao definido anteriormente no período de aprendizagem tamanho dos conjuntos, adotou-se a mesma divisão estabelecida para os subconjuntos deste período sendo eles os conjuntos de treinamento, validação e teste.

Da mesma forma que foi feito na RNA1, para garantir uma igualdade de influências entre todas as variáveis no treinamento da RNA2, foi necessária a normalização dos conjuntos de entrada e saída. Sendo utilizado o mesmo critério de normalização que fora descrito anteriormente na parametrização da RNA1.

Quanto à escolha da RNA2 e sua parametrização, escolheu-se uma rede do tipo MLP (*Multi layer perceptron*), que utiliza retro propagação do erro (*Feed -forward backpropagation error*), com a seguinte arquitetura: [51 – 14:40 -48], conforme descrito abaixo:

- 51 entradas
- [38 até 58 neurônios] na camada oculta.
- [48 neurônios] na saída

Sobre a escolha dessa arquitetura, também foi adotado o procedimento definido em AQUINO, R. R. B. et al (2009), em que, o número de nodos da camada oculta, é discriminado através de um processo de tentativas, nas quais, busca-se um melhor desempenho por parte da rede. Em que tal procedimento é discriminado a seguir:

- 1) Primeiramente, foi criada uma rede com quatorze neurônios na camada oculta;

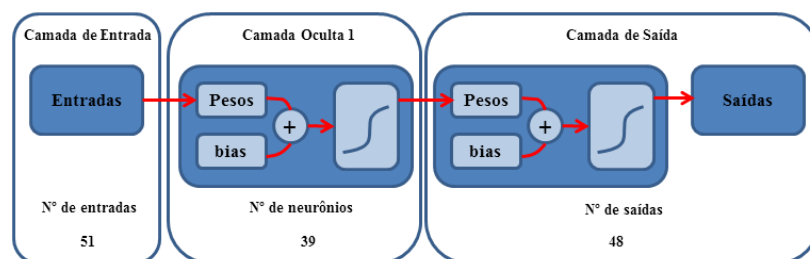
- 2) Foram atribuídos pesos aleatórios para todas as conexões da rede, e realizou-se o treinamento da rede através do algoritmo RP (*Resilient Propagation*) e foi calculado o MSE para o conjunto de treinamento, validação e teste;
- 3) O passo anterior foi repetido dez vezes. Após a décima repetição foi calculada a média do MSE para cada conjunto (treinamento, validação e teste);
- 4) Outro neurônio foi adicionado à camada oculta, e os passos 2, 3 e 4 foram repetidos, até a rede possuir quarenta e oito nodos na camada oculta;

Sobre o método de treinamento que adequou os valores das entradas às suas respectivas saídas, validando assim a aprendizagem por parte da RNA2, utilizou-se o algoritmo “*Resilient Propagation*”, representado no simulador como “*Train RP*”. Enfatizando que o mesmo só foi adotado em preferência ao “*Train LM*”, que utiliza o algoritmo de Levenberg Marquardt, tecnicamente mais preciso, por questões de esforço computacional.

Tal fato, assim como ocorrido na RNA1 também pôde ser notado nas tentativas de implementação da simulação relativa ao treinamento da RNA2, em que as mesmas resultaram em sucessivos erros com a seguinte mensagem “*Out of Memory*”. Logo, concluiu-se que seria necessária mais memória para tal nível de precisão, e assim tomou-se a escolha de trocar o algoritmo de treinamento.

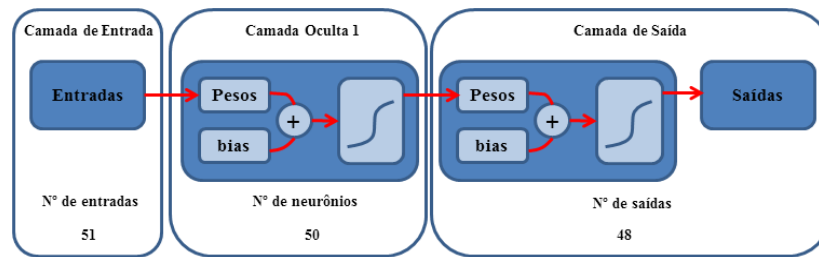
As melhores RNA2, denominadas “RNA2_BEST”, obtidas nas inicializações que simularam os treinamentos, nas bases dos Parques 01 e 02 são mostradas pelas Figuras 38 e 39 respectivamente:

Figura 38 – Arquitetura da melhor RNA2 para o Parque 01 escolhida com a implementação dos treinamentos.



Fonte: O Autor (2015).

Figura 39 – Arquitetura da melhor RNA2 para o Parque 02 escolhida com a implementação dos treinamentos.

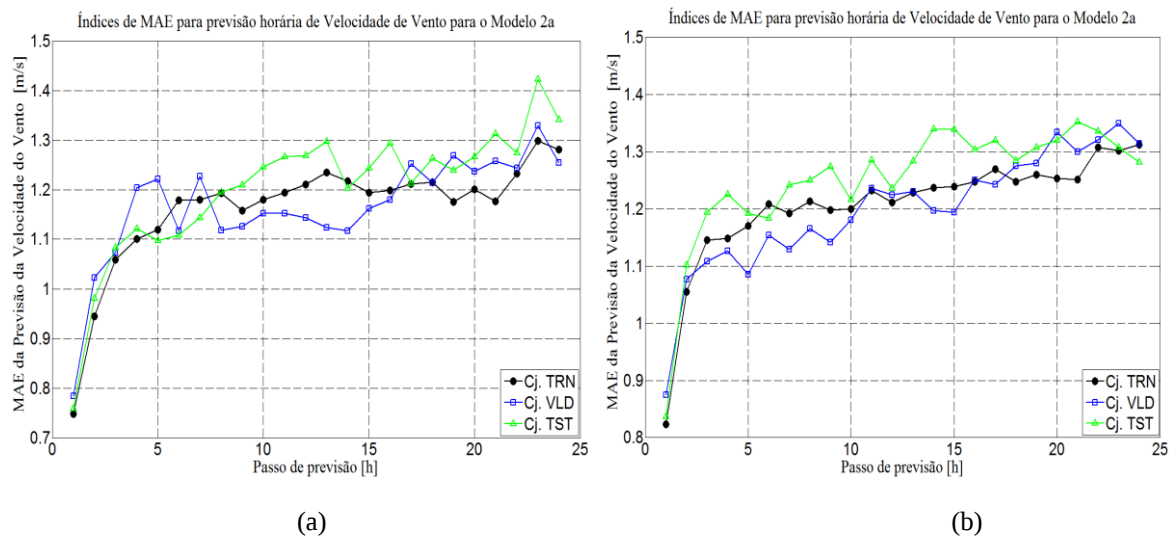


Fonte: O Autor (2015).

Das figuras acima, pode-se notar que, para a melhor inicialização, o número de neurônios definidos pela camada oculta foi de 39 para o Parque 01 e 50 para o Parque 02, descrevendo assim a arquitetura da RNA2 após os treinamentos.

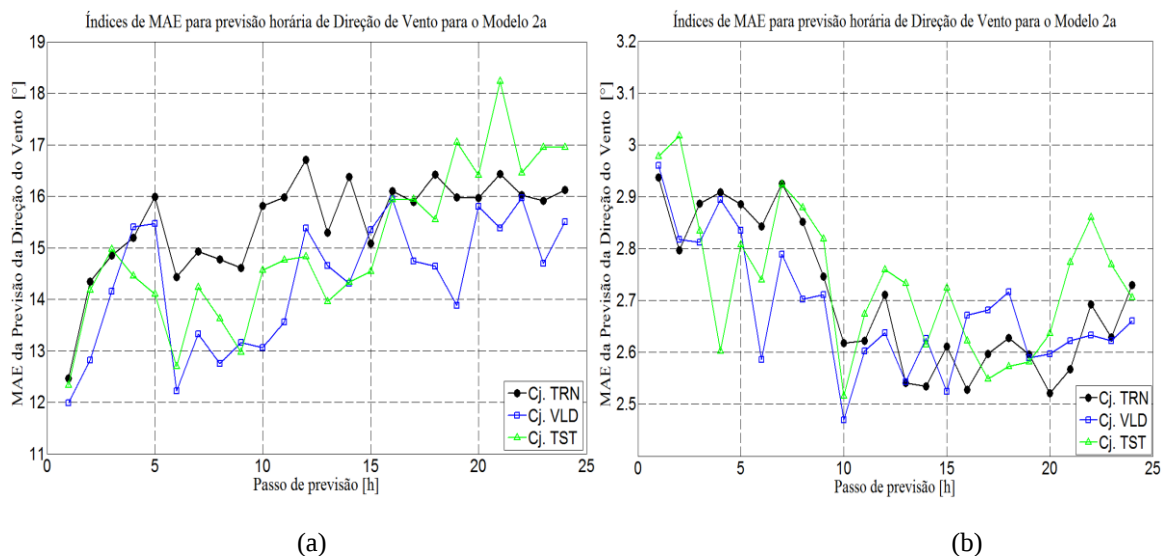
Os resultados do treinamento são mostrados pelas Figuras 40 e 41 que descreve o desempenho do erro médio absoluto realizado em ambos os parques nas previsões da respectiva RNA2.

Figura 40 – Comportamento gráfico do MAE da velocidade média referente às simulações para o período de aprendizagem da RNA2 aplicado em (a) para Parque 01 e em (b) para o Parque 02.



Fonte: O Autor (2015).

Figura 41 – Comportamento gráfico do MAE da direção média referente às simulações para o período de aprendizagem da RNA2 aplicado em (a) para Parque 01 e em (b) para o Parque 02.



Fonte: O Autor (2015)

Na Figura 40 pode-se ver que o comportamento respectivo nos Parques em questão no período de aprendizagem não possui muita diferença com relação à previsão de velocidades. Com respeito à previsão de direções o comportamento é bastante distinto verificando-se um crescimento do MAE à medida que o passo de previsão cresce no Parque 01 e um decréscimo do MAE para o mesmo horizonte no parque 02.

Analisando a Figura 41 é possível perceber a diferença no comportamento do MAE, pois no Parque 01 o erro aumenta ao longo do horizonte de previsão, enquanto no Parque 02 o erro diminui. Outro fator perceptível é a diferença no patamar do erro sendo 12° na 1ª hora e subindo até um patamar de 18° no fim do horizonte. Já o Parque 02 o erro é da ordem de 2,9° na 1ª hora e decai para 2,7° no final do horizonte.

Para simulação das RNAs foi utilizado o Software MATLAB versão R2018a, e cujo micro computador utilizado possui a seguinte configuração: cpu Intel (R) Core (TM) I7-3770, 3,4GHz, com uma memória total instalada de 8GB de RAM.

6.1.2 HEURÍSTICA DEEP LEARNING LONG SHORT-TERM MEMORY (LSTM)

Os previsores LSTM1 e LSTM2 foram concebidos sendo a base do parque sobre estudo a mesma adotada no treinamento e análises das rede neurais MLP RNA1 e RNA2, a arquitetura desta nova heurística dispõe de camadas profundas com suas próprias funções de ativação e conexões específicas.

Desta forma, após uma criação de padrões semelhante as redes MLP para o seu treinamento, os dados foram passados em formas de padrões para compor o vetor alimentador do conjunto de aprendizagem da rede e assim proceder com a sua aprendizagem. No caso respectivo específico do previsor LSTM1, foram utilizados vinte e sete conjuntos de variáveis como padrões de entrada ($[h(t+1)] + [Mês(t+1)(seno)] + [Mês(t+1)(cosseno)]$, [vinte e quatro séries horárias de velocidades médias do vento reais]) e [vinte e quatro variáveis de saída (séries horárias de velocidades médias do vento previstas)].

Já no caso do previsor LSTM2 foram utilizados cinquenta e um conjuntos de variáveis de entrada, sendo estas ($[h(t+1)] + [Mês(t+1)(seno)] + [Mês(t+1)(cosseno)]$ [vinte e quatro séries horárias de velocidades médias do vento reais] +[vinte e quatro séries horárias de direções médias do vento reais]) e 48 variáveis de saída, sendo estas ([vinte e quatro séries horárias de velocidades médias do vento previstas] +[vinte e quatro séries horárias de direções médias do vento previstas]).

De forma semelhante aos *autoencoders* que buscam uma função que retrate a base sobre estudo (*feature*), as LSTM também buscam tal função, sendo esta *feature* estabelecida por camadas profundas, assim cada camada profunda terá a sua *feature* específica. Outro detalhe importante é o tipo de algoritmo de treinamento utilizado no processo de aprendizagem da rede, para o caso desta tese foram testados 3 tipos de algoritmo sendo eles: SGDM (Gradiente descendente estocástico com momentum), ADAM (Estimativa de momentum adaptativo) e o RMSPROP (Propagação quadrática Média), os quais são descritos a seguir.

6.1.2.1 Algoritmos de aprendizagem

Algoritmos de aprendizagem são formas de organizar logicamente através de passos condicionais iterativos a aplicação ótima de um ou vários métodos de otimização ao treinamento das ferramentas de IA. Seu fundamento principal é estabelecer o aprendizado da modelagem através da exploração e análise de conjuntos de dados, estabelecendo funções objetivo e encontrando padrões de comportamento desses dados.

A respeito dos algoritmos e métodos de otimização são abordados nesta tese com a função de otimizar o aprendizado das LSTM os algoritmos SGDM, ADAM e RMSPROP. Os mesmos são abordados a seguir:

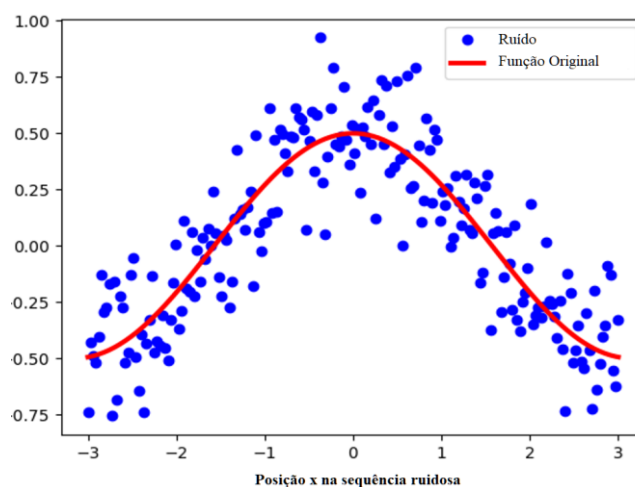
a) Algoritmo SGDM

O SGDM é o método que ajuda a acelerar os vetores de gradientes nas direções corretas, levando a uma convergência mais rápida. É um dos algoritmos de otimização mais populares e muitos modelos de última geração são treinados para usá-lo. Para entender o mesmo são necessárias algumas explicações específicas conforme se segue:

1) Médias exponencialmente ponderadas

As médias exponencialmente ponderadas lidam com sequências numéricas. Suponha que exista alguma sequência S que é ruidosa (variante). Como por exemplo, a função cosseno adicionada de um ruído gaussiano conforme ilustrada abaixo pela Figura 42:

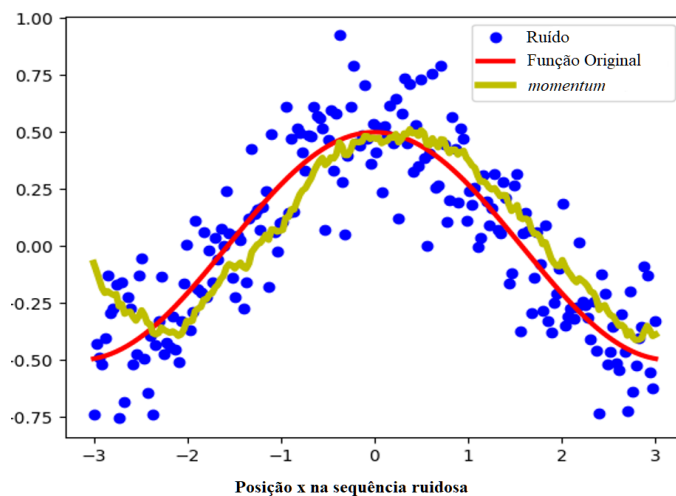
Figura 42 – Comportamento gráfico de uma sequência ruidosa.



Fonte: O Autor (2021).

Observando que, embora esses pontos pareçam muito próximos uns dos outros, nenhum deles compartilha a coordenada x , isto é, existe um número x único para cada ponto. Esse é o número que define o índice de cada ponto em nossa sequência S . O que se deseja fazer com esses dados é, em vez de usá-los, deseja-se uma média móvel (*momentum*) que concentre ordenadamente os dados, aproximando-os da função original. Conforme mostra a Figura 43 a seguir através da média exponencial ponderada.

Figura 43 – Comportamento gráfico da tendência de média móvel (momentum) em uma sequência ruidosa.



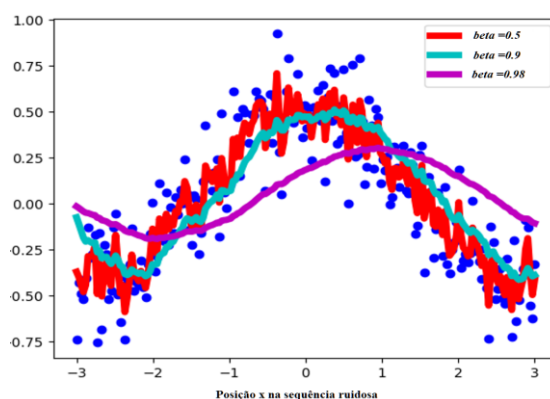
Fonte: O Autor (2021).

Conforme mostrado pela Figura 43 um resultado muito razoável. Em vez de ter dados com muito ruído, obtém-se uma linha muito mais suave, que está mais próxima da função original do que os dados que havia anteriormente. As médias exponencialmente ponderadas definem uma nova sequência V com a seguinte Equação 6.5:

$$V_t = \beta V_{t-1} + (1 - \beta)S_t; \text{ com } \beta \in [0,1] \quad (6.5)$$

Essa sequência V_t é aquela plotada em amarelo conforme pode-se notar na Figura 43 acima. Beta é outro hiperparâmetro que leva valores de 0 a 1. Pela Figura 6.10 abaixo o valor utilizado para $\beta = 0,9$. É um bom valor e mais frequentemente usado em SGDM. Intuitivamente, pode-se pensar em beta como o cálculo para uma média aproximada dos últimos $1 / (1 - \beta)$ pontos da sequência. Assim pode-se acompanhar pela Figura 44, como a escolha de beta afeta a nova sequência V_t .

Figura 44 – Variação comportamental da tendência do momentum para três diferentes valores do hiperparâmetro beta.



Fonte: O Autor (2021).

Notoriamente, para valores menores de beta, a nova sequência acaba tendo uma flutuação maior, devido ao cálculo da média de um número menor de exemplos e, portanto, aproximando-se assim dos dados ruidosos. Para valores maiores de beta, $\beta = 0,98$, tem-se uma curva mais atenuada e deslocada para a direita, devido ao fato de se executar uma média sobre um número maior de exemplos (cerca de 50 para $\beta = 0,98$). Assim para $\beta = 0,9$ tem-se um bom equilíbrio entre esses dois extremos.

2) O momentum gradual de descendência:

Existem basicamente apenas duas maneiras diferentes de escrever e definir o momentum gradual de descendência de um gradiente. A primeira maneira é a determinação do momentum que é uma média móvel dos gradientes, que em seguida é utilizada para atualizar o peso da rede conforme descreve a Equação 6.6 a seguir:

$$V_t = \beta V_{t-1} + (1 - \beta) \nabla_W L(W, X, y); \quad \text{com } W = W - \alpha V_t \quad (6.6)$$

Em que, L é a função de perda, o símbolo triangular é o gradiente com peso e α é taxa de aprendizagem. A Segunda maneira é a forma mais popular de escrever regras de atualização de momentum, é menos intuitiva e apenas omite o termo $(1 - \beta)$ conforme demonstra a Equação 6.7 a seguir:

$$V_t = \beta V_{t-1} + \alpha \nabla_W L(W, X, y); \quad \text{com } W = W - V_t \quad (6.7)$$

Esta forma é praticamente idêntica Equação 6.6, a única diferença é que se necessita dimensionar a taxa de aprendizado por um fator $(1 - \beta)$.

b) ADAM (Adam)

Em resumo, o algoritmo de otimização intitulado Adam é uma extensão do algoritmo do gradiente descendente estocástico que recentemente teve uma adoção mais ampla para aplicativos de aprendizado profundo em visão computacional e processamento de linguagem natural. Por ser uma extensão é amplamente utilizado em substituição ao seu algoritmo de origem, devido a maior eficiência computacional com relação a atualização iterativa dos pesos da rede com base nos dados de treinamento.

Foi proposto por Diederik Kingma (OpenAI) e Jimmy Ba (Toronto University) em seu artigo publicado no ICLR (*International Conference On Learning Representations*) de 2015 Com o Título “Adam: Um Método para Otimização Estocástica”. O interessante é que apesar

de parecer uma sigla a nomenclatura Adam é derivada da estimativa de momento adaptativa e não faz referência a sigla ADAM.

No artigo os autores listam os benefícios do uso do Adam em problemas de otimização não convexos, como: a sua fácil implementação, a eficiência computacional, a baixa requisição de memória, a invariância no redimensionamento diagonal dos gradientes, a sua boa adequação para problemas que são grandes em termos de dados e / ou parâmetros, sua boa adequação para implementações com funções objetivo não estacionárias, sua boa utilização para problemas com gradientes muito variantes / ou esparsos.

Por fim seus hiperparâmetros têm interpretação intuitiva e normalmente requerem poucos ajustes.

A forma de trabalho do Adam é diferente do SGDM que mantém uma única taxa de aprendizado (alfa) para todas as atualizações de peso sendo esta taxa de aprendizado fixa para todo o treinamento.

No Adam, a taxa de aprendizado é mantida para cada peso de rede (parâmetro) e separadamente adaptada conforme o aprendizado se desenvolve. O método então de forma adaptativa calcula as taxas de aprendizagem de maneira individual para diferentes parâmetros a partir de estimativas de primeiro e segundo momentos dos gradientes.

Segundo seus criadores, o Adam une as vantagens de duas outras extensões do SGDM: o algoritmo de gradiente adaptativo (*AdaGrad*) que mantém uma taxa de aprendizagem por parâmetro que melhora o desempenho em problemas com gradientes esparsos (por exemplo, linguagem natural e problemas de visão computacional). E o *Root Mean Square Propagation* (RMSProp), que também mantém taxas de aprendizagem por parâmetro que são adaptadas com base na média das magnitudes recentes dos gradientes para o peso (por exemplo, a rapidez com que está mudando). Isso significa que o algoritmo funciona bem, isto é, percebe os benefícios do AdaGrad e do RMSProp, em problemas on-line e não estacionários (por exemplo, variantes).

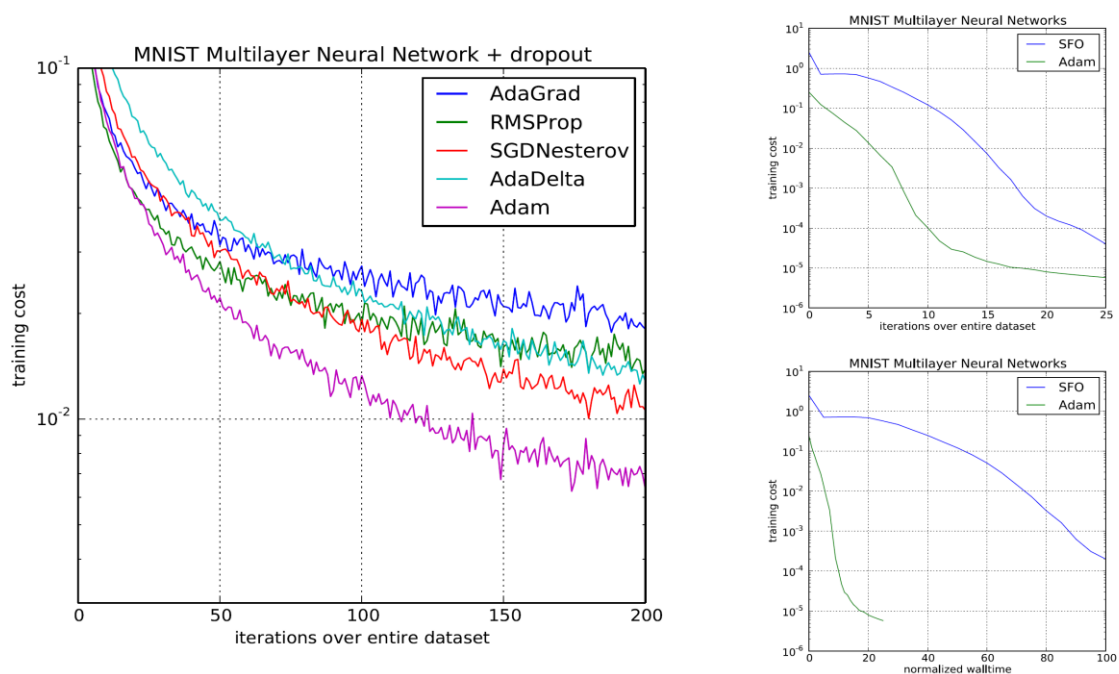
Em vez de adaptar as taxas de aprendizagem dos parâmetros com base na média do primeiro momento (a média) como no RMSProp, Adam também faz uso da média dos segundos momentos dos gradientes desta forma considerando que a variância não é centralizada). Especificamente o algoritmo calcula uma média móvel exponencial do gradiente e do gradiente quadrado, e os parâmetros β_1 e β_2 controlam as taxas de decaimento dessas médias móveis.

O valor inicial das médias móveis e os valores de β_1 e β_2 próximos de 1,0 (recomendado) resultam em uma tendência de estimativas de momento para zero. Esse viés é

superado calculando primeiro as estimativas tendenciosas antes de calcular as estimativas corrigidas de viés.

No artigo, foi demonstrado empiricamente que a convergência atende às expectativas da análise teórica. Adam foi aplicado ao algoritmo de regressão logística nos conjuntos de dados de reconhecimento de dígitos MNIST e análise de sentimento IMDB, um algoritmo Multilayer Perceptron no conjunto de dados MNIST e Redes Neurais Convolucionais no conjunto de dados de reconhecimento de imagem CIFAR-10. Eles concluem:

Figura 45 – Comportamento gráfico do MAE da direção média referente às simulações para o período de aprendizagem da RNA2 aplicado em (a) para Parque 01 e em (b) para o Parque 02.



Fonte: Adaptado de Kingma D., Ba J. (2015).

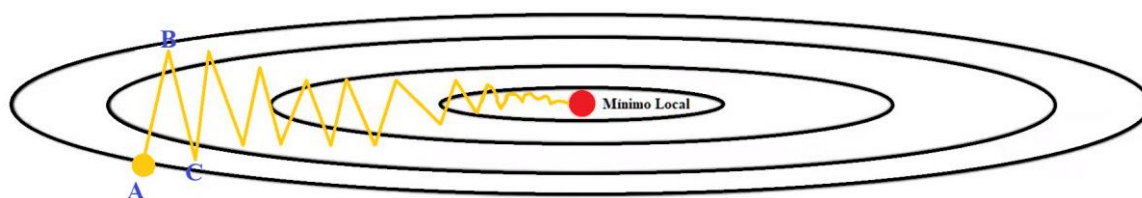
Usando grandes modelos e conjuntos de dados, foi demonstrado que Adam pode resolver problemas práticos de aprendizado profundo com eficiência, Kingma D., Ba J. (2015).

c) RMSPROP

Root mean square propagation ou RMSprop utiliza o mesmo conceito da média exponencialmente ponderada dos gradientes como gradiente descendente com momentum, mas a diferença está na atualização dos parâmetros.

Considere um exemplo onde tenta-se otimizar uma função de custo que tem contornos como abaixo e o ponto vermelho denota a posição do ótimo local (mínimo).

Figura 46 – Comportamento gráfico do movimento de convergência do gradiente em um processo iterativo de otimização.

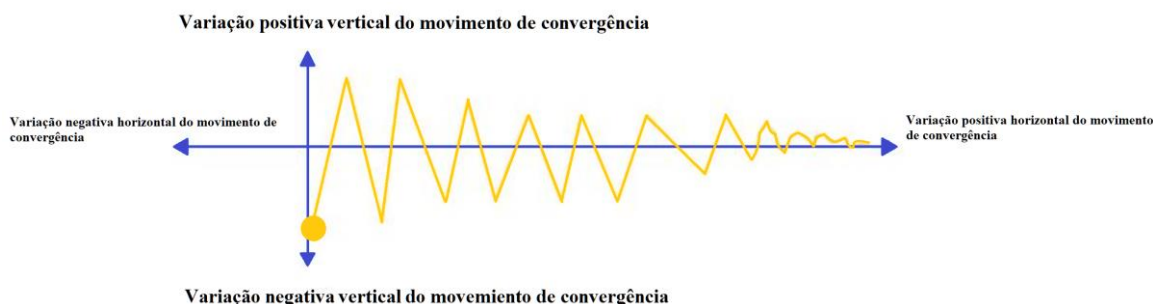


Fonte: Adaptado de Deep Learning Specialization by Andrew Ng (2018)

Inicia-se a descida do gradiente a partir do ponto 'A' e após uma iteração de descida do gradiente pode-se chegar no ponto 'B', o outro lado da elipse. Em seguida, outra etapa de descida gradiente pode chegar no ponto 'C'. Com cada iteração de gradiente descendente, move-se em direção ao mínimo ótimo local com oscilações para cima e para baixo. Se utiliza-se uma taxa de aprendizado maior, a oscilação vertical terá magnitude maior. Portanto, essa oscilação vertical retarda o gradiente descendente e impossibilita o uso de uma taxa de aprendizado muito maior.

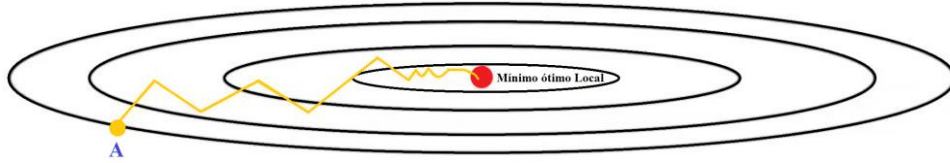
O valor 'bias' é responsável pelas oscilações verticais, enquanto o valor do 'peso' define o movimento na direção horizontal. Se a atualização para "polarização" diminuir, então podem-se amortecer as oscilações verticais e se atualizarem os "pesos" com valores mais altos, então ainda pode-se mover rapidamente em direção ao ótimo local.

Figura 47 – Comportamento planar variacional do movimento de convergência do gradiente em um processo iterativo de otimização.



Fonte: Adaptado de Deep learning Specialization by Andrew Ng (2018).

Figura 48 – Comportamento variacional otimizado do movimento de convergência do gradiente em um processo iterativo com atualização dos pesos regulada pelo algoritmo backpropagation.



Fonte: Adaptado de Deep Learning Specialization by Andrew Ng (2018).

Durante a retropropagação, usa-se dW e db para atualizar nossos parâmetros W e b da seguinte forma:

$$W = W - \text{taxa de aprendizagem} * dW \quad (6.8)$$

$$b = b - \text{taxa de aprendizagem} * db \quad (6.9)$$

No RMSprop, em vez de usar dW e db independentemente para cada época, toman-se as médias exponencialmente ponderadas do quadrado de dW e db .

$$S_{dW} = \beta * S_{dW} + (1 - \beta) * dW^2 \quad (6.10)$$

$$S_{db} = \beta * S_{db} + (1 - \beta) * db^2 \quad (6.11)$$

Em que, beta ' β ' é outro hiperparâmetro e assume valores de 0 a 1. Ele define o peso entre a média dos valores anteriores e o quadrado do valor atual para calcular a nova média ponderada.

Após o cálculo exponencial das médias ponderadas, atualizam-se os parâmetros.

$$W = W - \text{taxa de aprendizagem} * \frac{dW}{\text{Sqrt}(S_{dW})} \quad (6.12)$$

$$b = b - \text{taxa de aprendizagem} * \frac{db}{\text{Sqrt}(S_{db})} \quad (6.13)$$

S_{dW} é relativamente pequeno, então S_{dW} é o resultado da divisão de dW por um número relativamente grande, enquanto S_{db} é relativamente grande, então S_{db} é o resultado da divisão de db por um número relativamente pequeno, o motivo para isto está na maior desaceleração das atualizações em uma dimensão vertical.

6.1.2.2 Arquitetura dos Previsores LSTM

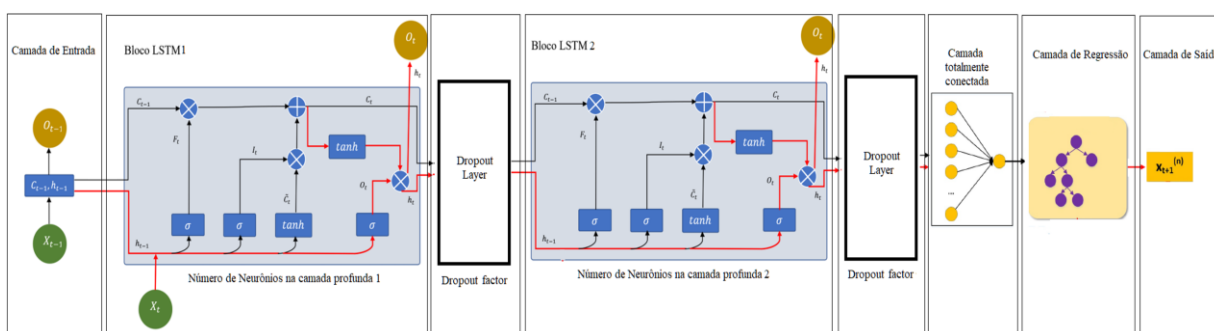
Para a definição da arquitetura utilizaram-se os respectivos algoritmos de aprendizagem descritos anteriormente, em que o algoritmo que obteve o melhor desempenho foi o Adam definindo assim o formato da arquitetura das redes LSTM. Levou-se em consideração o mesmo número de épocas utilizados nas MLPs definidas anteriormente e com a mesma taxa de aprendizagem. Além disso também se preservou a mesma separação de

dados para os conjuntos de aprendizagem (Treinamento, validação e teste) como também o respectivo conjunto de simulação.

As LSTM após sucessivos testes apresentaram resultados condizentes com arquitetura profunda a partir da primeira camada de profundidade, isto é, segunda camada da arquitetura da rede. As redes do tipo LSTM desta tese tiveram em sua composição 6 camadas sendo uma de entrada, 2 camadas profundas, uma camada totalmente conectada, uma camada de regressão e uma camada de saída.

Para cada camada profunda os parâmetros são os já enfatizados anteriormente sobre o funcionamento de uma LSTM, como a função $\tanh$ e a função σ além de um bloco de otimização chamado *Dropout Layer*, que por sua vez necessita de um parâmetro de ação o *Dropout factor*. A arquitetura de composição da rede LSTM previsora de velocidade de vento é mostrada a seguir pela Figura 49.

Figura 49 – Arquitetura genérica das LSTMs criadas para os Parques sobre estudo.



Fonte: O Autor (2021).

6.1.2.2.1 Previsor LSTM1 e seus parâmetro

A rede neural LSTM1 consiste em uma rede do tipo LSTM com os procedimentos para a sua obtenção semelhantes aos utilizados para a obtenção do modelo da RNA1, no que diz respeito a composição do seu vetor de entrada em que conforme explicado na teoria da heurística das LSTM possui os dados relativos a velocidade de vento e dados sazonais como o mês e as horas na estrutura do seu padrão, estando esta tarefa fortemente conectada as características relativas da série temporal definida para a composição dos padrões do modelo.

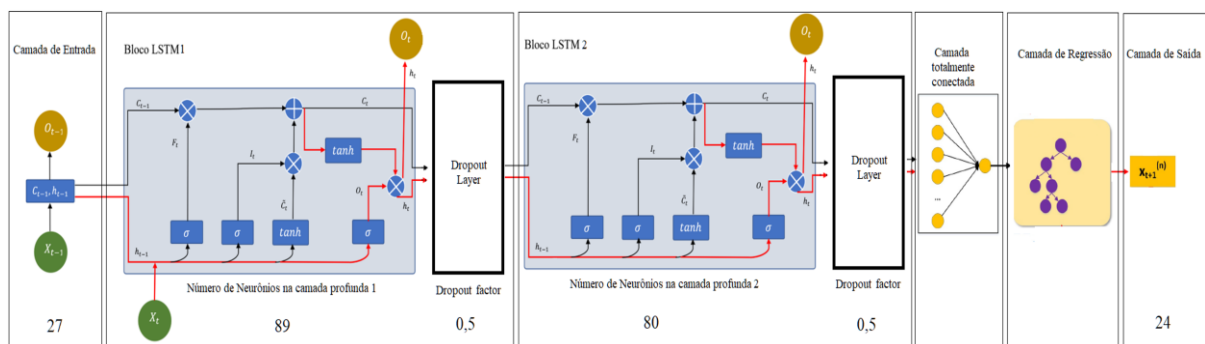
Assim a composição dos padrões de entrada da LSTM1 possuiu vinte e sete variáveis de entrada ($[[h(t+1)] + [Mês(t+1)(seno)] + [Mês(t+1)(cosseno)]$), [vinte e quatro séries

horárias de velocidades médias do vento reais]) e vinte e quatro variáveis de saída (séries horárias de velocidades médias do vento previstas).

Para tal, a mesma consistiu de uma rede neural de previsão de velocidades médias de vento com horizonte de 24h, ou seja, o passo de previsão variou da hora um até a hora vinte e quatro. Assim, para um valor de entrada $v(t)$ a LSTM1 irá prever valores conforme descrito pela Expressão 4.27.

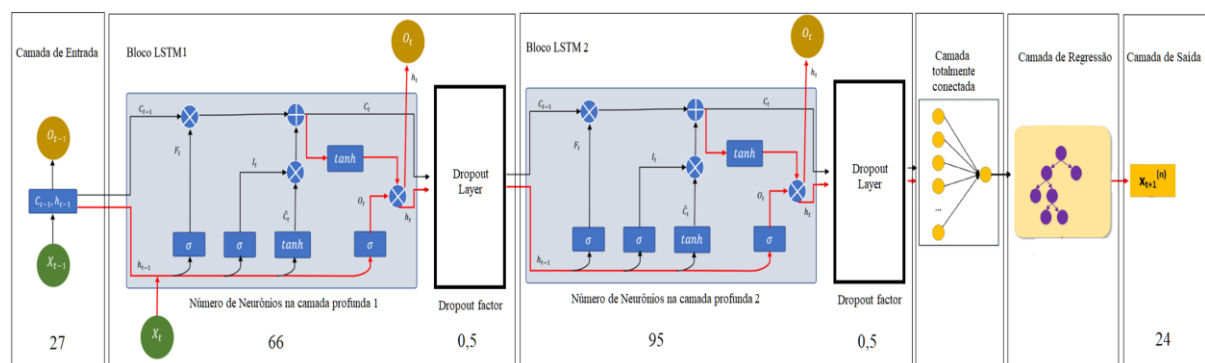
Em que, as melhores LSTM1, denominadas “LSTM1_BEST”, obtidas nas inicializações que simularam os treinamentos, nas bases dos Parques 01 e 02 são mostradas nas Figuras 50 e 51 respectivamente, com seus respectivos parâmetros utilizados:

Figura 50 – Arquitetura da melhor LSTM1 para o Parque 01 escolhida com a execução dos treinamentos.



Fonte: O Autor (2021).

Figura 51 – Arquitetura da melhor LSTM1 para o Parque 02 escolhida com a execução dos treinamentos.

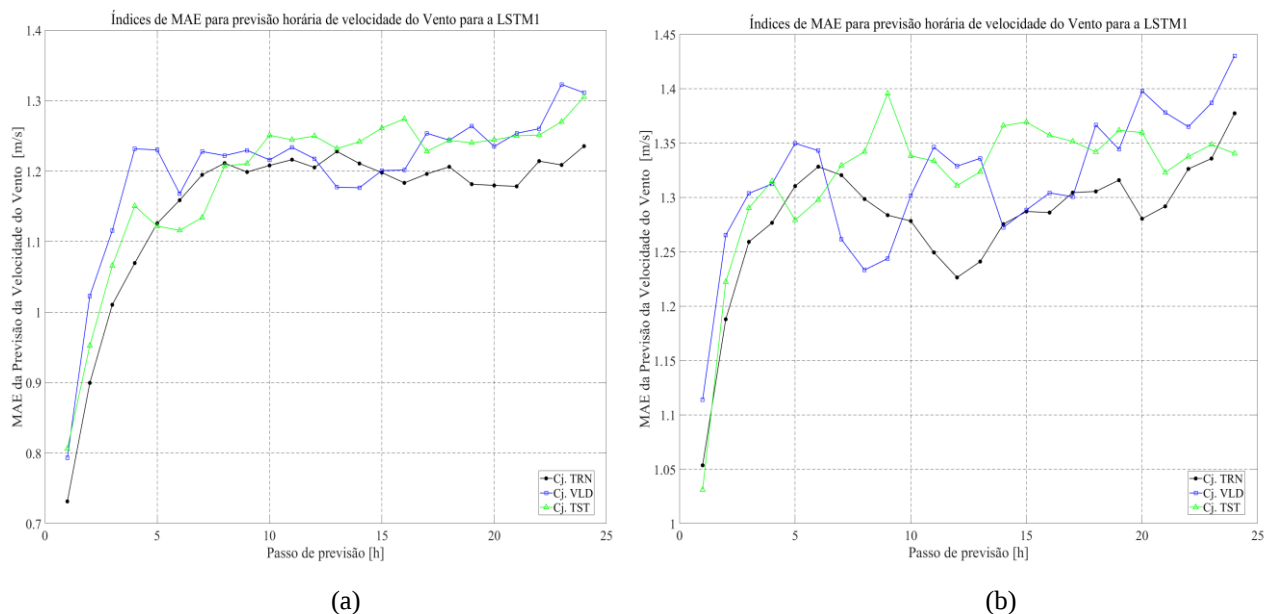


Fonte: O Autor (2021).

Das Figuras 50 e 51, pode-se notar que, para a melhor inicialização, o número de neurônios definidos pela camada profunda 1 foi de 89 e pela camada profunda 2 foi de 80 para o Parque 01 e respectivamente 66 e 95 para o Parque 02, descrevendo assim a arquitetura da LSTM1 após os treinamentos.

Outros resultados importantes relativos a estes treinamentos são descritos pela Figura 52 que descreve o desempenho do erro médio absoluto cometido para ambos os parques nas previsões das respectivas LSTM1 dos Parques 01 e 02. Onde os mesmos são mostrados a seguir.

Figura 52 – Comportamento gráfico do MAE referente às simulações para o período de aprendizagem da LSTM1 (a) aplicado ao Parque 01 e (b) aplicado ao Parque 02.



Fonte: O Autor (2021).

6.1.2.2.2 Previsor LSTM2 e seus parâmetros

De forma semelhante ao que foi implementado no previsor LSTM1 tomando por base a RNA1 o LSTM2 também tomou por base a RNA2 possuindo desta maneira da mesma forma que seu predecessor cinquenta e um conjuntos de variáveis de entrada, sendo estas $([h(t+1)] + [Mês(t+1)(seno)] + [Mês(t+1)(cosseno)] + [vinte e quatro séries horárias de velocidades médias do vento reais] + [vinte e quatro séries horárias de direções médias do vento reais])$ e 48 variáveis de saída, sendo estas $([vinte e quatro séries horárias de velocidades médias do vento previstas] + [vinte e quatro séries horárias de direções médias do vento previstas])$.

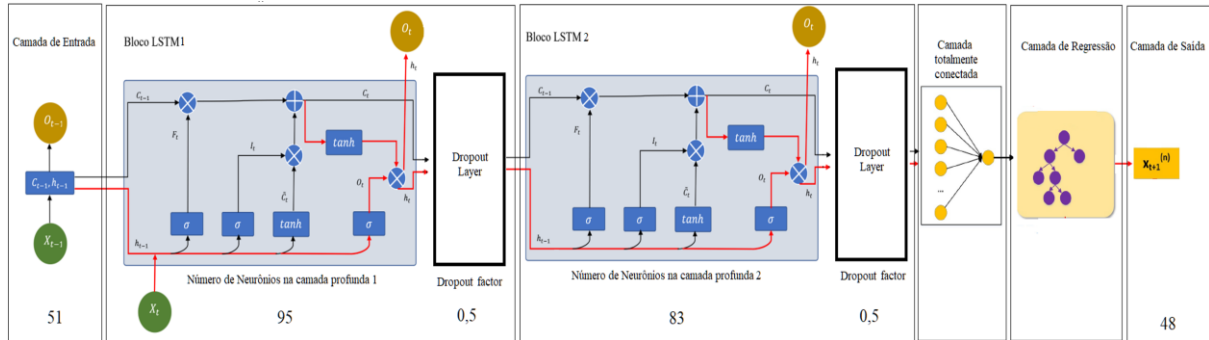
Para o caso dos novos padrões que alimentam o modelo da LSTM2, a composição do mesmo é dado da seguinte forma $X_t = [v_t, d_t]$, logo em termos de arquitetura não há mudanças, na assimilação das entradas por parte da rede, porém novos valores para as

variáveis paramétricas da LSTM serão atribuídos. Caso das variáveis i_t ; σ ; W_i ; h_{t-1} ; b_i ; C_t ; $\tilde{C}_t$; W_C ; b_C ; W_F ; b_F e F_t , anteriormente definidas conforme Equações 4.34 a 4.39.

A mesma consistiu de uma rede neural de previsão de velocidades e direções médias de vento com horizonte de 24h, ou seja, o passo de previsão variou da hora um até a hora 24. Assim para valores de entrada $v(t)$ e $d(t)$ a LSTM2 irá prever valores de velocidades e direções médias, conforme descrito pela Expressão 4.29.

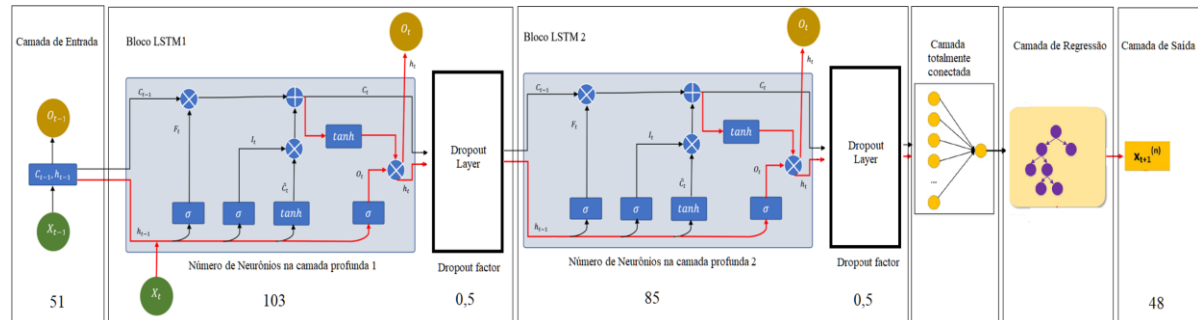
As melhores LSTM2, denominadas “LSTM2_BEST”, obtidas nas inicializações que simularam os treinamentos, nas bases dos Parques 01 e 02 são mostradas pelas Figuras 53 e 54 respectivamente:

Figura 53 – Arquitetura da melhor LSTM2 para o Parque 01 escolhida com a implementação dos treinamentos.



Fonte: O Autor (2021).

Figura 54 – Arquitetura da melhor LSTM2 para o Parque 02 escolhida com a implementação dos treinamentos.

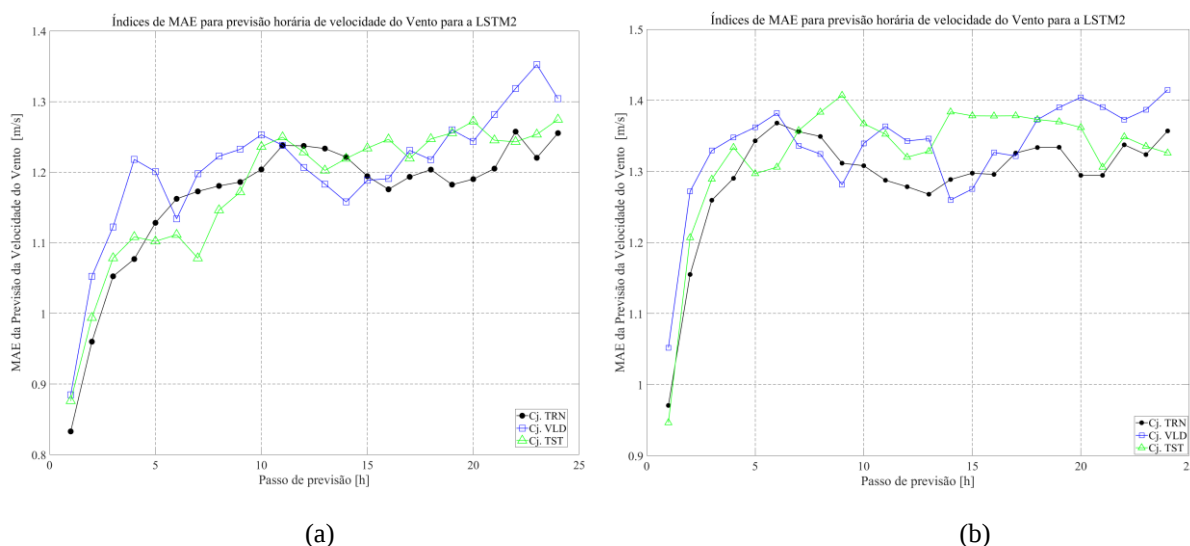


Fonte: O Autor (2021).

Das Figuras 53 e 54, pode-se notar que, para a melhor inicialização, o número de neurônios definidos pela camada profunda 1 foi de 95 e pela camada profunda 2 foi de 83 para o Parque 01 e respectivamente 103 e 85 para o Parque 02, descrevendo assim a arquitetura da LSTM2 após os treinamentos.

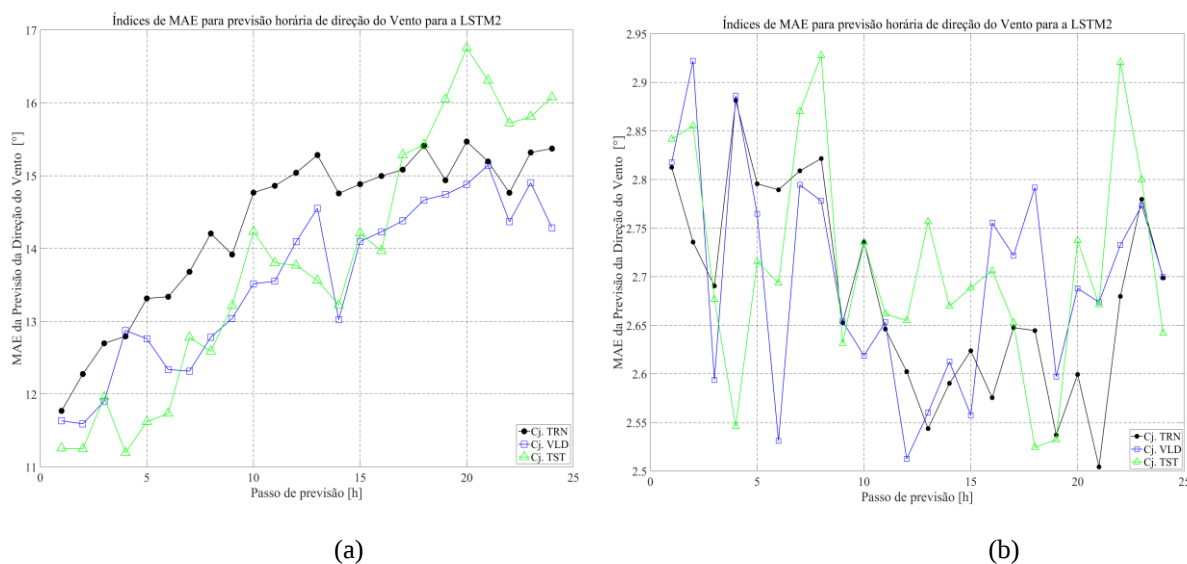
Os resultados do treinamento são mostrados pelas Figuras 55 e 56 que descrevem o desempenho do erro médio absoluto realizado em ambos os parques nas previsões da respectiva LSTM2.

Figura 55 – Comportamento gráfico do MAE da velocidade média referente às simulações para o período de aprendizagem da LSTM2 aplicado em (a) para Parque 01 e em (b) para o Parque 02.



Fonte: O Autor (2021).

Figura 56 – Comportamento gráfico do MAE da direção média referente às simulações para o período de aprendizagem da LSTM2 aplicado em (a) para Parque 01 e em (b) para o Parque 02.



Fonte: O Autor (2021).

Na Figura 55 pode-se ver que o comportamento respectivo nos Parques em questão no período de aprendizagem não possui muita diferença com relação à previsão de velocidades. Com respeito à previsão de direções o comportamento é bastante distinto verificando-se um

crescimento do MAE à medida que o passo de previsão cresce no Parque 01 e um decréscimo do MAE para o mesmo horizonte no Parque 02.

Analisando a Figura 56 é possível perceber a diferença no comportamento do MAE, pois no Parque 01 o erro aumenta ao longo do horizonte de previsão, enquanto no Parque 02 o erro diminui. Outro fator perceptível é a diferença no patamar do erro sendo 12° na 1ª hora e subindo até um patamar de 18° no fim do horizonte. Já o Parque 02 o erro é da ordem de 2,9° na 1ª hora e decai para 2,7° no final do horizonte.

6.1.3 HEURÍSTICA MULTI LAYER PERCEPTRON ADEQUADA POR AUTOENCODER (AE-MLP)

Esta heurística utilizou dois tipos de configurações de autoencoders objetivando desta forma entender a influência que a precisão das variáveis teria sobre o modelo predictor utilizado. Esta influência é possível graças às modificações propostas pela reconstrução dos padrões da base de dados do parque sobre estudo, e desta forma ao consequente novo treinamento das RNAs que compõem os modelos previsores de velocidade e/ou direção até então utilizados anteriormente, caso dos modelos RNA1 e RNA2.

Para tal, foram definidas duas configurações topológicas de autoencoders, sendo as mesmas denominadas AET1 e AET2, ambas seguindo a lógica de construção que será mostrada a seguir.

Assim após a reconstrução foram criados outros cinco novos modelos previsores de velocidade e/ou direção, modelos estes “filtrados ou adequados” pela lógica de reconstrução do AE dependendo da configuração topológica definida. Em que a configuração AET1 cria os modelos previsores: RNA1 AET1 e LSTM1 AET1 e a configuração AET2 cria os modelos previsores: RNA1 AET2, RNA2 AET2, LSTM2 AET2. Vale salientar que tais modelos previsores serão explicados posteriormente.

Outra informação importante é que estas topologias também definiram novas aprendizagens para os modelos de curva de potência abordados anteriormente nesta tese, envolvendo as CPs RNA e CPs Fuzzy. Desta forma foram criados mais 3 modelos aproximadores de Curvas de potência, sendo os mesmos: o CP1RNA AET1, o CP1RNA AET2 e o CP2RNA AET2. De forma semelhante, vale salientar que tais modelos de curva de potência serão explicados posteriormente.

Nesta tese as duas configurações de *Autoencoders* possuíram 1 camada de entrada e uma camada de saída, sendo a diferença básica entre as topologias o número de séries de variáveis componentes da série padrão da topologia. No geral, a formação de um *Autoencoder* não é uma tarefa difícil, seu treinamento tem o objetivo de codificar uma entrada em um domínio (a) em uma saída definida em um contra domínio de representação codificado cuja função de codificação é denominada $g(a)$ ou “*Encoder*”, de maneira que posteriormente a entrada (a) possa ser reconstituída através da sua função de representação codificada.

A função de decodificação ou “*Decoder*”, $f(g(a))$ produz a reconstrução da rede, sendo geralmente um vetor de valores obtidos através de uma função de transferência, podendo esta ser: sigmoide logística “*logsig*”, linear pura “*pureln*” ou linear saturada “*satlin*”. O objetivo do treinamento do *Autoencoder* é a criação de uma função $g(a)$ cuja representação distribuída dos dados capture os fatores principais de suas variações.

Para o treinamento dos *Autoencoders* utilizaram-se os padrões de sequência denominados 1 e 2 mostrados pelas Equações 6.14 e 6.15, em que a sequência deverá ser reconstruída pelo AE, de forma independente por variável. Isto é necessário, pois os padrões definem a configuração dos dados de velocidade, direção e Potência utilizados no treinamento das novas redes previsoras de velocidade e / ou direção do tipo MLP.

A topologia do padrão 1, (TP=1), possui metade dos seus elementos compostos por valores médios de velocidade e metade composta por valores médios de potência. Por padronização o código da topologia adotado será composto pela seguinte nomenclatura: [AE]+[TP]: [Autoencoder] + [Topologia Padrão], sendo o código mostrado pela Expressão 6.14 a seguir.

$$\text{AET1: } [v_{input} P_{output}] \quad (6.14)$$

A topologia do padrão 2, (TP=2), possui 1/3 dos seus elementos compostos por valores médios de velocidade, 1/3 de dos seus elementos compostos por valores médios de direção e 1/3 dos seus elementos compostos por valores médios de Potência. Por padronização o código da topologia adotado será composto pela seguinte nomenclatura: [AE]+[TP]: [Autoencoder]+[Topologia Padrão], sendo o código mostrado pela Expressão 6.15 a seguir.

$$\text{AET2: } [v_{input} d_{input} P_{output}] \quad (6.15)$$

Cada variável é aproximada por uma função de codificação diferente, conforme mostrado pelas Expressões 6.16, 6.17 e 6.18:

$$v_{input} \rightarrow g_v(v_{input}) \quad (6.16)$$

$$d_{input} \rightarrow g_d(d_{input}) \quad (6.17)$$

$$P_{output} \rightarrow g_P(P_{output}) \quad (6.18)$$

Assim internamente o *Autoencoder* terá a tarefa de reconstruir as bases de acordo com a topologia adotada, melhorando as possíveis variações que ocorram nas mesmas, funcionando assim como um filtro que diminuirá as varrições e imperfeição do conjunto de dados original.

É importante salientar que a decisão quanto a definição da ordem da topologia, que diz respeito à posição das variáveis de velocidade, direção e potência dentro da série padrão de reconstrução, levou em consideração o erro no padrão de reconstrução do AET em questão, sendo as duas topologias que apresentaram menor imprecisão de construção as topologias AET1 e AET2.

A Figura 57 mostra a arquitetura da heurística de reconstrução proposta:

Figura 57 – Arquitetura da heurística de reconstrução proposta.



Fonte: O Autor (2021).

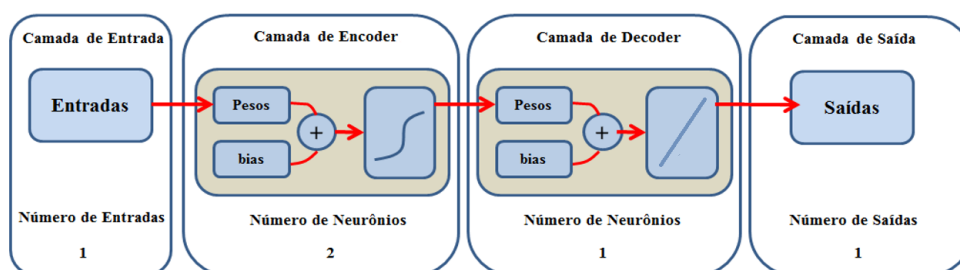
A seguir são mostrados a rotina para treinamento dos AE para os bancos de dados avaliados nesta tese e em seguida os parâmetros descriminados por parque de acordo com a topologia adotada.

Rotina para treinamento dos AE para os bancos de dados avaliados:

- (i) Escolhe-se o banco de dados a ser reconstruído, podendo ser do parque 01 ou parque 02.
- (ii) São definidas as Topologias, podendo ser adotada a AET1 ou a AET2.
- (iii) São definidas as funções de *encoder*, com 2 tipos, podendo ser (Sigmoide logística ou Linearmente saturada) e de *decoder*, com 3 tipos, podendo ser (Sigmoide logística, ou linearmente saturada ou puramente linear), de forma a estabelecer a criação de seis modelos de AE.
- (iv) Definem-se pesos iniciais aleatórios para as camadas de *encoder* e *decoder*
- (v) Define-se o número de épocas, sendo este, igual a 100.
- (vi) Utiliza-se o algoritmo de aprendizagem do método do gradiente conjugado em escala, proposto em Moller, M. F. (1993), para definir através do algoritmo backpropagation, a melhor arquitetura e o melhor desempenho dos modelos de AE propostos com a finalidade de reconstrução mais precisa do banco de dados do parque escolhido.
- (vii) Comparam-se os MSEs dos conjuntos de aprendizagem, (treinamento, validação e teste), respectivamente com percentuais definidos em 80%, 10%, 10% do conjunto de aprendizagem, entre os modelos definidos e escolhe-se então o melhor modelo de AE para reconstruir a base.

A Tabela 2 mostra os parâmetros escolhidos por parque de acordo com esta rotina de aprendizagem, para o melhor modelo de AE de acordo com a topologia adotada AET1 ou AET2. Já a Figura 58 apresenta a respectiva arquitetura das topologias escolhidas AET1 e AET2 dos Autoencoders para o Parque 01 e para o Parque 02.

Figura 58 – Arquitetura das Topologias AET1 e AET2 para os Parques 01 e 02 escolhidas durante o período de aprendizagem.



Fonte: O Autor (2021).

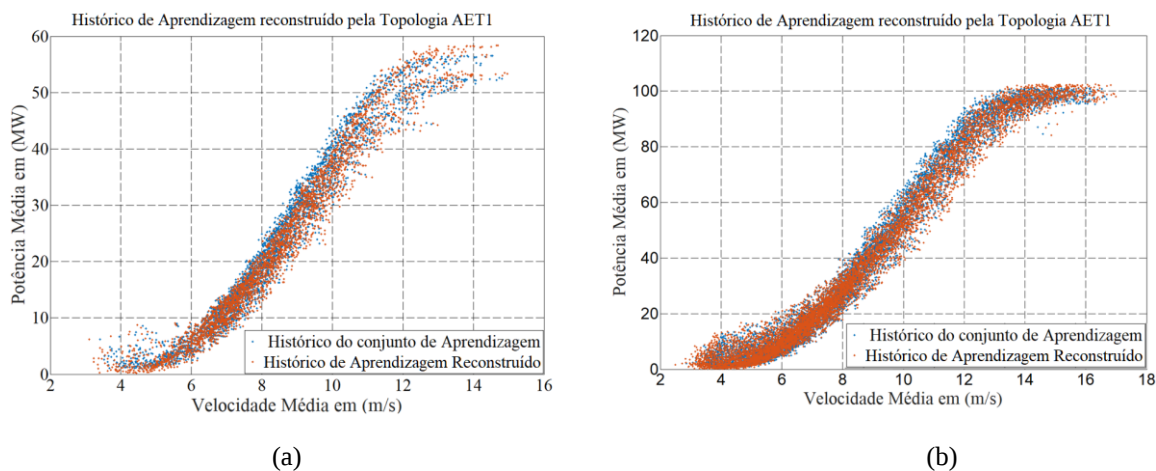
Tabela 2 – Parâmetros escolhidos por parque de acordo com esta rotina de aprendizagem, para o melhor modelo de AE de acordo com a topologia adotada AET1 ou AET2 nos Parques 01 e 02.

Parque 01						
Topologia	Função encoder	Função decoder	Nº de camadas	Nº de nodos na camada de encoder	Nº de nodos na camada de decoder	Mse da série de aprendizagem
AET1	logsig	logsig	2	2	1	0,2305
	logsig	satlin	2	2	1	0,2587
	logsig	purelin	2	2	1	0,1274
	satlin	logsig	2	2	1	0,3756
	satlin	satlin	2	3	1	0,5367
	satlin	purelin	2	2	1	0,3453
AET2	logsig	logsig	2	5	1	0,2964
	logsig	satlin	2	2	1	0,3027
	logsig	purelin	2	2	1	0,1924
	satlin	logsig	2	5	1	0,3128
	satlin	satlin	2	3	1	0,4263
	satlin	purelin	2	2	1	0,2981
Parque 02						
Topologia	Função encoder	Função decoder	Nº de camadas	Nº de nodos na camada de encoder	Nº de nodos na camada de decoder	Mse da série de aprendizagem
AET1	logsig	logsig	2	2	1	0,5302
	logsig	satlin	2	5	1	0,6583
	logsig	purelin	2	2	1	0,1969
	satlin	logsig	2	5	1	0,3754
	satlin	satlin	2	5	1	0,5362
	satlin	purelin	2	3	1	0,3457
AET2	logsig	logsig	2	2	1	0,2263
	logsig	satlin	2	2	1	0,2045
	logsig	purelin	2	2	1	0,1151
	satlin	logsig	2	3	1	0,6962
	satlin	satlin	2	2	1	0,8024
	satlin	purelin	2	2	1	0,3987

Fonte: O Autor (2021).

Como exemplo de boa margem de reconstrução dos dados A Figura 59 ilustra a comparação dos dados reconstruídos em relação ao histórico original de aprendizagem proposta pelo AET1, para os parques 01 e 02.

Figura 59 – Comparação entre histórico do conjunto de aprendizagem e a sua reconstrução proposta pela Topologia do AET1 para os parques 01 (a) e 02 em (b).



Fonte: O Autor (2021).

6.1.3.1 Previsor RNA1_AET1 e seus parâmetros

A rede neural MLP1_AET1 foi nomeada de RNA1_AET1 e consiste em uma rede do tipo MLP com os procedimentos para a sua obtenção sendo os mesmos que a RNA1 exceto no que diz respeito a sua base de dados. A base de dados utilizada no processo de aprendizagem desta rede foi reconstruída pelo autoencoder da topologia AET1. Desta forma os dados relativos ao seu conjunto de aprendizagem, que de forma diferente do caso anterior fora obtido pela reconstrução da base por intermédio do autoencoder da topologia AET1.

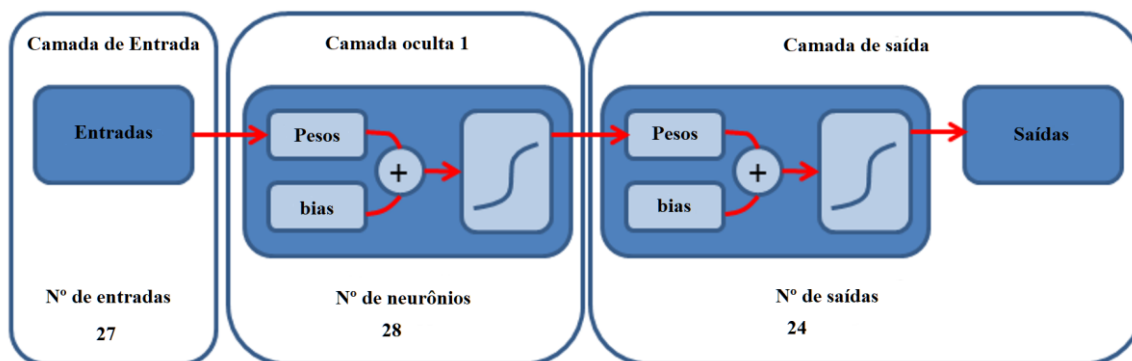
Assim o AET1 reconstruiu as bases relativas aos parques em questão redefinindo um novo conjunto de aprendizagem, com o mesmo número de padrões do conjunto base, e foi utilizado no treinamento do previsor RNA1_AET1.

De forma semelhante a RNA1_AET1 possuiu vinte e sete variáveis de entrada ($[h(t+1)] + [Mês(t+1)(seno)] + [Mês(t+1)(cosseno)] + [vinte e quatro séries horárias de velocidades médias do vento reais]$) e vinte e quatro variáveis de saída (séries horárias de velocidades médias do vento previstas).

Para tal, a mesma consistiu de uma rede neural de previsão de velocidades médias de vento com horizonte de 24h, ou seja, o passo de previsão variou da hora um até a hora vinte e quatro. Assim, para um valor de entrada $v(t)$ a RNA1_AET1 irá prever valores conforme descrito pela Expressão 4.27.

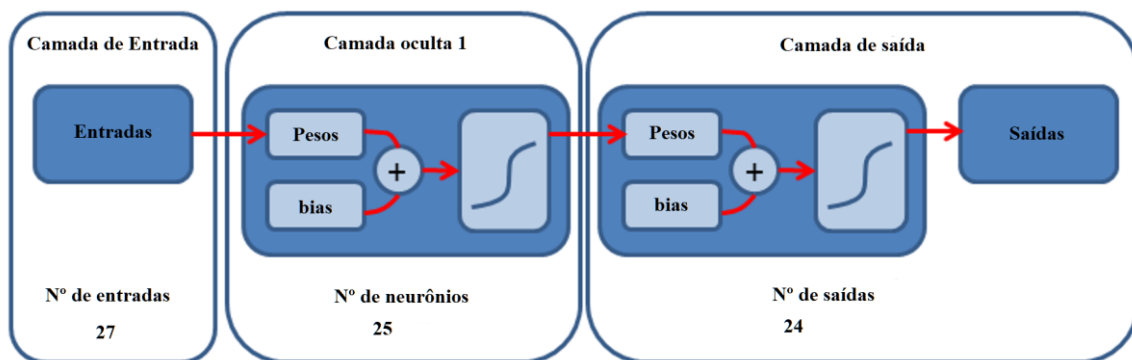
Em que, as melhores RNA1_AET1, denominadas “RNA1_AET1_BEST”, obtida nas inicializações que simularam os treinamentos, nas bases dos Parques 01 e 02 são mostradas nas Figuras 60 e 61 respectivamente:

Figura 60 – Arquitetura da melhor RNA1_AET1 para o Parque 01 escolhida com a execução dos treinamentos.



Fonte: O Autor (2021).

Figura 61 – Arquitetura da melhor RNA1_AET1 para o Parque 02 escolhida com a execução dos treinamentos.



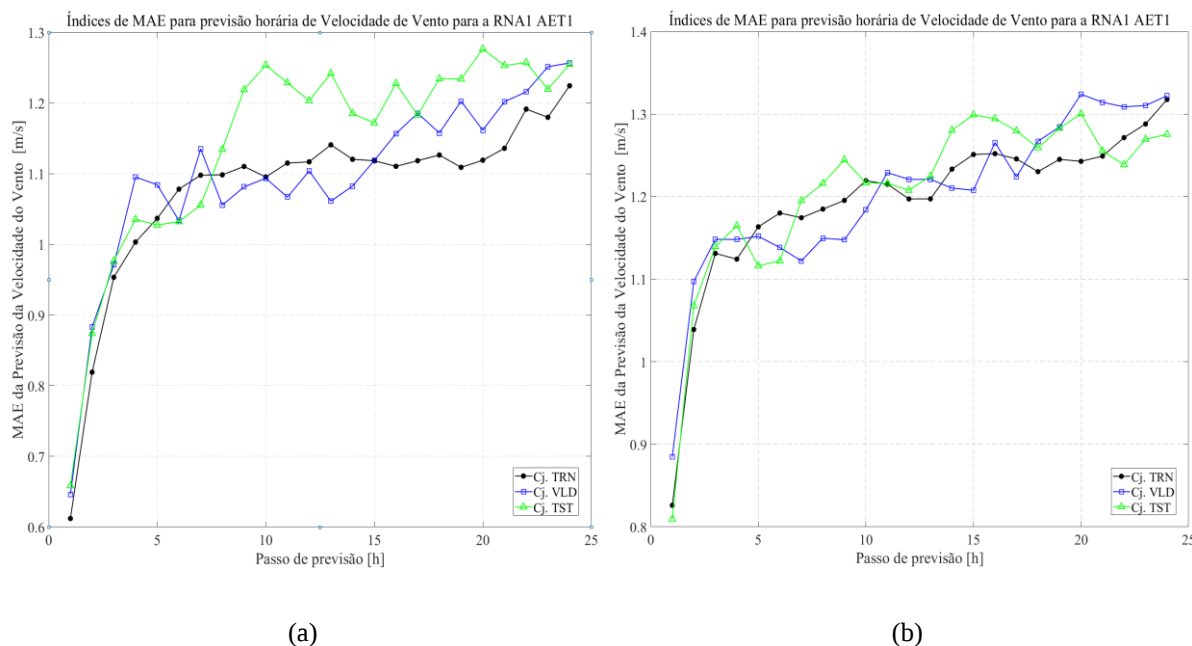
Fonte: O Autor (2021).

Das Figuras 60 e 61, pode-se notar que, para a melhor inicialização, o número de neurônios definidos pela camada oculta foi de 28 para o Parque 01 e 25 para o Parque 02, descrevendo assim a arquitetura da RNA1_AET1 após os treinamentos.

Outros resultados importantes relativos a estes treinamentos são descritos pela Figura 62 que descreve o desempenho do erro médio absoluto cometido para ambos os parques nas previsões das respectivas RNA1_AET1 dos Parques 01 e 02.

Onde os mesmos são mostrados a seguir.

Figura 62 – Comportamento gráfico do MAE referente às simulações para o período de aprendizagem da RNA1 (a) aplicado ao Parque 01 e (b) aplicado ao Parque 02.



Fonte: O Autor (2021).

6.1.3.2 Previsor RNA1 AET2 e seus parâmetros

A rede neural MLP1 AET2 foi nomeada de RNA1 AET2 e consiste em uma rede do tipo MLP com os procedimentos para a sua obtenção sendo os mesmos que a RNA1 exceto no que diz respeito a sua base de dados. A base de dados utilizada no processo de aprendizagem desta rede foi reconstruída pelo autoencoder da topologia AET2. É importante salientar que o fato de se criarem duas topologias de autoencoders AET1 e AET2 serviu para verificar o grau de influência dos parâmetros de velocidade e direção do vento no que diz respeito a precisão na reconstrução da série de dados executada pela topologia do autoencoder escolhida.

Desta forma os dados relativos ao seu conjunto de aprendizagem, diferentemente do caso anterior fora obtido pela reconstrução da base por intermédio do autoencoder da topologia AET2.

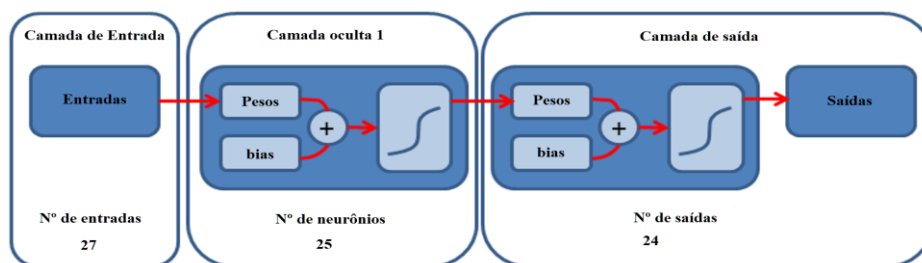
Assim o AET2 reconstruiu as bases relativas aos parques em questão redefinindo um novo conjunto de aprendizagem, com o mesmo número de padrões do conjunto base, e foi utilizado no treinamento do previsor RNA1 AET2.

De forma semelhante a RNA1 AET2 possuiu vinte e sete variáveis de entrada ($[h(t+1)] + [Mês(t+1)(seno)] + [Mês(t+1)(cosseno)] + [vinte e quatro séries horárias de velocidades$

médias do vento reais]) e vinte e quatro variáveis de saída (séries horárias de velocidades médias do vento previstas).

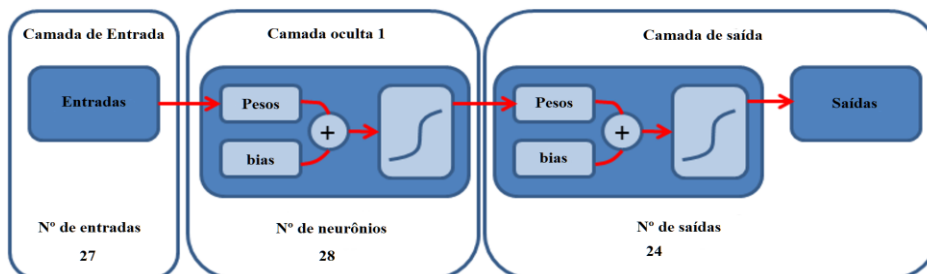
Para tal, a mesma consistiu de uma rede neural de previsão de velocidades médias de vento com horizonte de 24h, ou seja, o passo de previsão variou da hora um até a hora vinte e quatro. Assim, para um valor de entrada $v(t)$ a RNA1 AET2 irá prever valores conforme descrito pela Expressão 4.27. Em que, as melhores RNA1 AET2, denominadas “RNA1 AET2 *BEST*”, obtida nas inicializações que simularam os treinamentos, nas bases dos Parques 01 e 02 são mostradas nas Figuras 63 e 64 respectivamente:

Figura 63 – Arquitetura da melhor RNA1_AET2 para o Parque 01 escolhida com a execução dos treinamentos.



Fonte: O Autor (2021).

Figura 64 – Arquitetura da melhor RNA1_AET2 para o Parque 02 escolhida com a execução dos treinamentos.

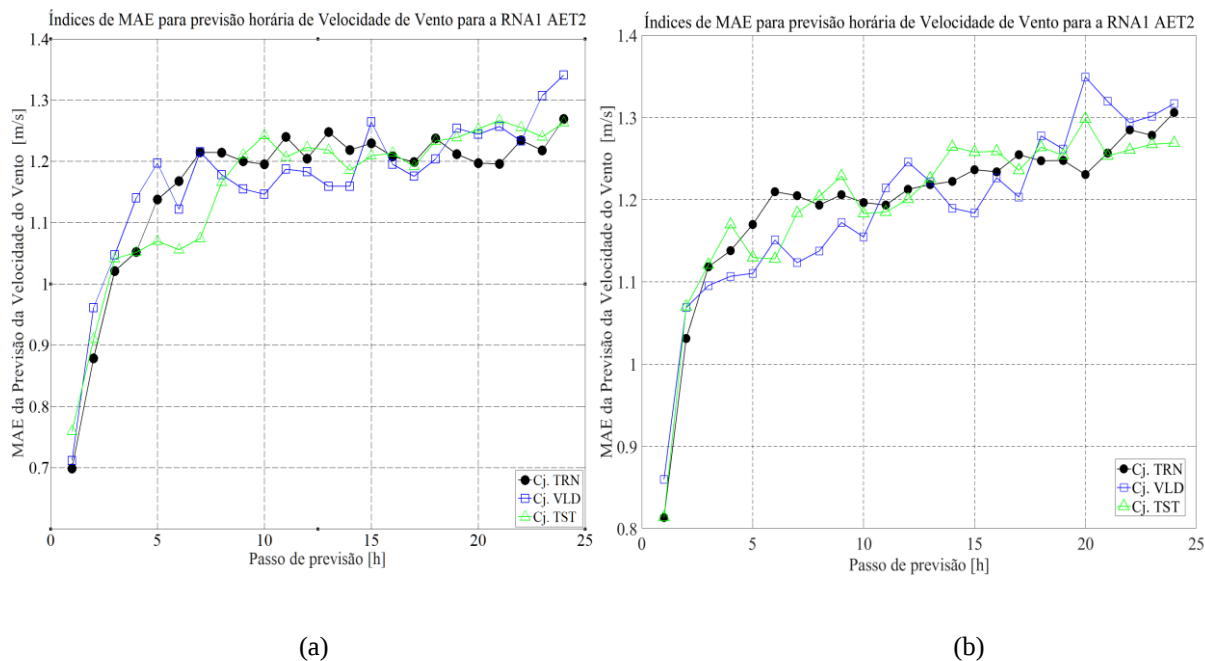


Fonte: O Autor (2021).

Das Figuras 63 e 64, pode-se notar que, para a melhor inicialização, o número de neurônios definidos pela camada oculta foi de 25 para o Parque 01 e 28 para o Parque 02, descrevendo assim a arquitetura da RNA1 AET2 após os treinamentos.

Outros resultados importantes relativos a estes treinamentos são descritos pela Figura 65 que descreve o desempenho do erro médio absoluto cometido para ambos os parques nas previsões das respectivas RNA1 AET2 dos Parques 01 e 02.

Figura 65 – Comportamento gráfico do MAE referente às simulações para o período de aprendizagem da RNA1_AET2 (a) aplicado ao Parque 01 e (b) aplicado ao Parque 02.



Fonte: O Autor (2021).

6.1.3.3 Previsor RNA2 AET2 e seus parâmetros

A rede neural MLP2 AET2 foi nomeada de RNA2 AET2 e consiste em uma rede do tipo MLP com os procedimentos para a sua obtenção sendo os mesmos que a RNA2 exceto no que diz respeito a sua base de dados. A base de dados utilizada no processo de aprendizagem desta rede foi reconstruída pelo autoencoder da topologia AET2. Desta forma os dados relativos ao seu conjunto de aprendizagem, que de forma diferente do caso anterior fora obtido pela reconstrução da base por intermédio do autoencoder da topologia AET2.

Assim o AET2 reconstruiu as bases relativas aos parques em questão redefinindo um novo conjunto de aprendizagem, com o mesmo número de padrões do conjunto base, e foi utilizado no treinamento do previsor RNA2 AET2.

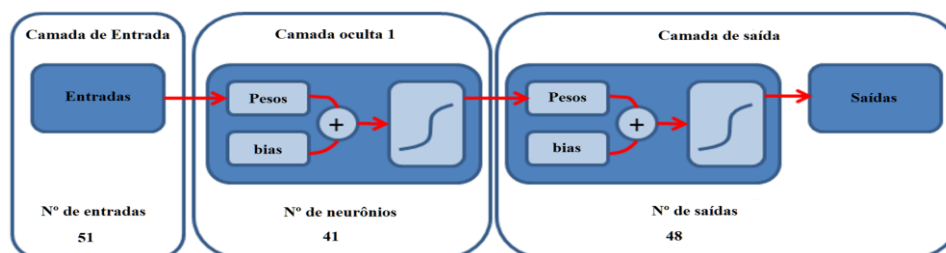
De forma semelhante a RNA2 AT2 possui cinquenta e um conjuntos de variáveis de entrada, sendo estas ([$h(t+1)$] + [$Mês(t+1)(seno)$] + [$Mês(t+1)(cosseno)$] + [vinte e quatro séries horárias de velocidades médias do vento reais] + [vinte e quatro séries horárias de direções médias do vento reais]) e 48 variáveis de saída, sendo estas ([vinte e quatro séries horárias de velocidades médias do vento previstas] + [vinte e quatro séries horárias de direções médias do vento previstas]).

A mesma consistiu de uma rede neural de previsão de velocidades e direções médias de vento com horizonte de 24h, ou seja, o passo de previsão variou da hora um até a hora 24.

Assim para valores de entrada $v(t)$ e $d(t)$ a RNA2_AET2 irá prever valores de velocidades e direções médias, conforme descrito pela Expressão 4.29.

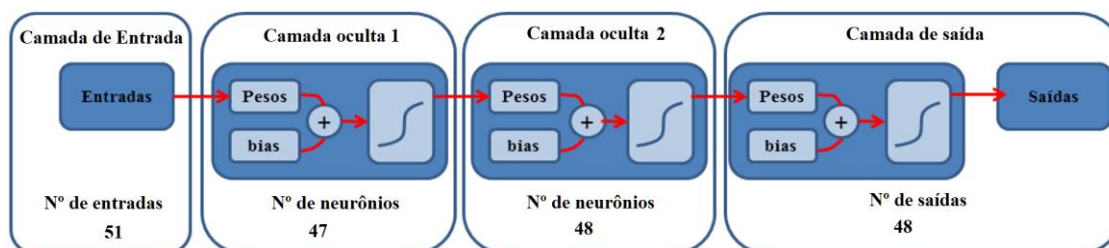
As melhores RNA2_AET2, denominadas “RNA2_AET2_BEST”, obtidas nas inicializações que simularam os treinamentos, nas bases dos Parques 01 e 02 são mostradas pelas Figuras 66 e 67 respectivamente:

Figura 66 – Arquitetura da melhor RNA2_AET2 para o Parque 01 escolhida com a implementação dos treinamentos.



Fonte: O Autor (2021).

Figura 67 – Arquitetura da melhor RNA2_AET2 para o Parque 02 escolhida com a implementação dos treinamentos.



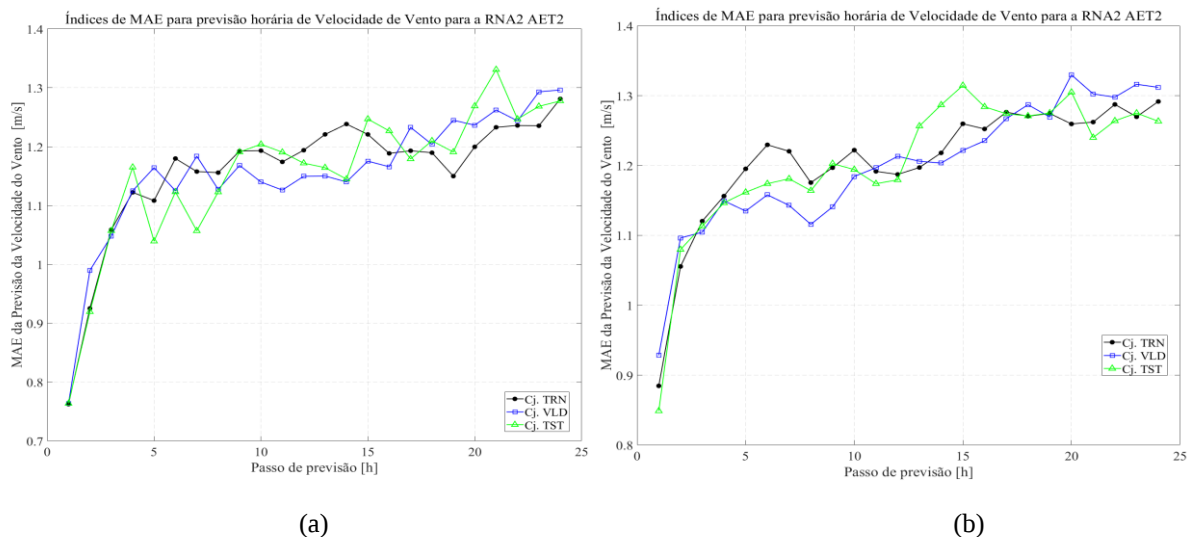
Fonte: O Autor (2021).

Das Figuras 66 e 67, pode-se notar que, para a melhor inicialização, o número de neurônios definidos pela camada oculta foi de 41 para o Parque 01, já para o Parque 02, pode-se notar a mudança na arquitetura da rede RNA2_AET2 tendo a mesma duas camadas ocultas com respectivamente 47 e 48 neurônios, descrevendo assim a arquitetura da RNA2_AET2 após os treinamentos. Assim como em ALBUQUERQUE, J.C. (2015) e em AQUINO R. R. B. et. al. (2016), foram testados dois tipos de arquiteturas uma com uma camada oculta e outra com duas, as quais tiveram seus respectivos acertos e ponderações. Desta forma, pode-se notar que a arquitetura de da Rede RNA2_AET2 para o Parque 01 mostrou-se mais precisa com uma camada oculta, e para o Parque 02 com duas.

Os resultados do treinamento são mostrados pelas Figuras 68 e 69 que descrevem o desempenho do erro médio absoluto realizado em ambos os parques nas previsões da

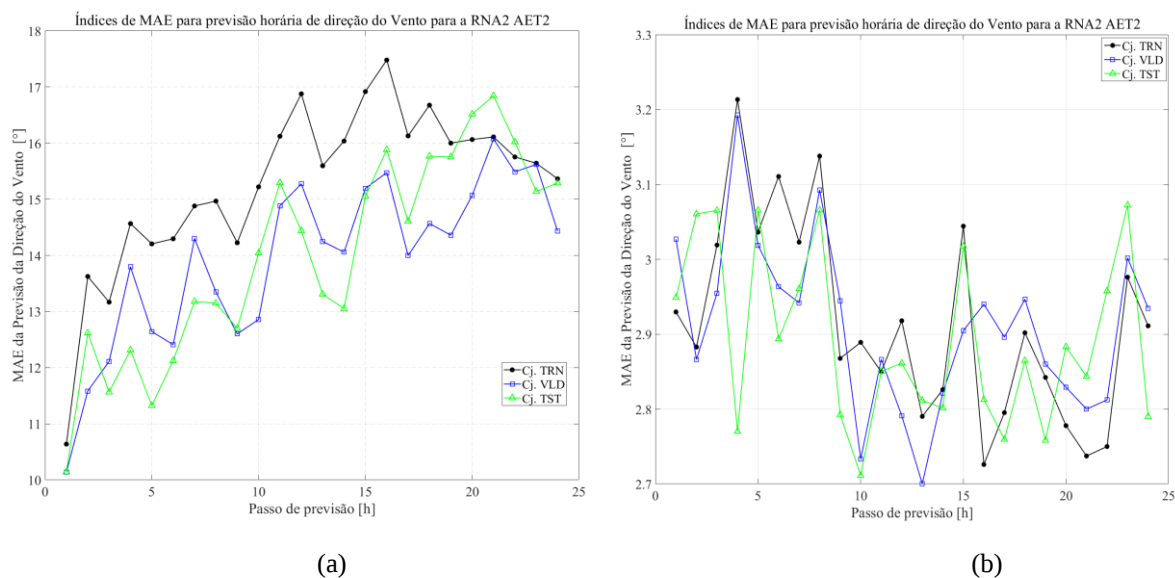
respectiva RNA2 AET2, sobre dados de velocidade de vento e direção do vento, respectivamente.

Figura 68 – Comportamento gráfico do MAE da velocidade média referente às simulações para o período de aprendizagem da RNA2 AET2 aplicado em (a) para Parque 01 e em (b) para o Parque 02.



Fonte: O Autor (2021).

Figura 69 – Comportamento gráfico do MAE da direção média referente às simulações para o período de aprendizagem da RNA2 AET2 aplicado em (a) para Parque 01 e em (b) para o Parque 02.



Fonte: O Autor (2021).

Na Figura 68 pode-se ver que o comportamento respectivo nos Parques em questão no período de aprendizagem não possui muita diferença com relação à previsão de velocidades. Com respeito à previsão de direções o comportamento é bastante distinto verificando-se um

crescimento do MAE à medida que o passo de previsão cresce no Parque 01 e um decréscimo do MAE para o mesmo horizonte no Parque 02.

Analisando a Figura 69 é possível perceber a diferença no comportamento do MAE, pois no Parque 01 o erro aumenta ao longo do horizonte de previsão, enquanto no Parque 02 o erro diminui. Outro fator perceptível é a diferença no patamar do erro sendo 10° na 1ª hora e subindo até um patamar de 17° no fim do horizonte. Já o Parque 02 o erro é da ordem de $2,9^\circ$ na 1ª hora e decai para $2,7^\circ$ no final do horizonte.

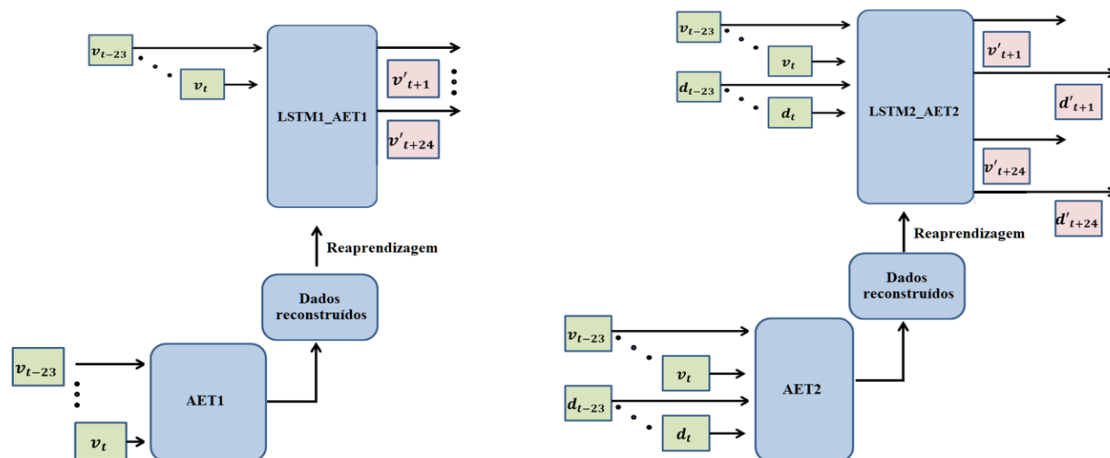
6.1.4 HEURÍSTICA DEEP LEARNING LONG SHORT-TERM MEMORY ADEQUADA POR AUTOENCODER (LSTM-AE)

Como enfatizado por Hochreiter e Schmidhuber (1997), é notório que as redes LSTM se comportam bem no que diz respeito a relevância de tendência a longo prazo, porém nesta tese propõem-se uma nova heurística com o autoencoder. A reconstrução da base de dados sobre análise proposta pelo AE, pode tornar a imprecisão da mesma mais ou menos relevante e desta forma fazer com que o desafio da LSTM seja mais significativo ainda.

Nesta tese foram criadas 2 arquiteturas de modelos previsores de velocidade e direção com a heurística das LSTM, semelhantes em arquitetura aos já comentados modelos de redes MLP puros ou adequados por AE. Porém a diferença básica está na nova heurística de predição implementada em que a característica principal é ter menor imprecisão no longo prazo em relação aos modelos anteriores.

Para o caso da heurística das LSTMs adequadas pelas topologias de autoencoder AET1 e AET2 a Figura 70 a seguir dispõe a modelagem do bloco em questão.

Figura 70 – Composição das arquiteturas das LSTM adequadas pelas topologias dos autoencoders para os Parques 01 e 02.



Fonte: O Autor (2021).

6.1.4.1 Previsor LSTM1_AET1 e seus parâmetros

O previsor LSTM1 AET1 foi concebido sendo a base do parque sobre estudo reconstruída pelo autoencoder da topologia AET1 de forma semelhante ao que fora feito com a rede LSTM1. A nova arquitetura desta nova heurística dispõe de camadas ocultas e camadas profundas com suas próprias funções de ativação e conexões específicas, semelhante ao que fora feito na LSTM1.

Desta forma, após a ação de reconstrução dos padrões da base sobre análise do AE da topologia AET1 os dados reconstruídos foram passados em formas de padrões para compor o vetor alimentador do conjunto de aprendizagem da rede e assim proceder com o seu treinamento. No caso específico do previsor LSTM1 AET1 foram utilizados como padrões de entrada [vinte e quatro séries horárias de velocidades médias do vento reais], assim como nas heurísticas anteriores que utilizaram apenas uma entrada reconstruída (velocidade de vento), caso do previsor RNA1 AET1.

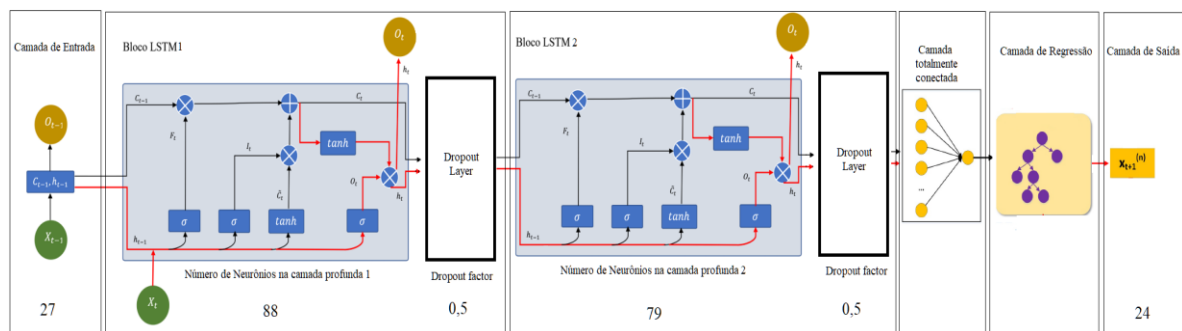
De forma semelhante aos *autoencoders* que buscam uma função que retrate a base sobre estudo (*feature*), as LSTM também buscam tal função, sendo esta *feature* estabelecida por camadas profundas, assim cada camada profunda terá a sua *feature* específica. Outro detalhe importante é o tipo de algoritmo de treinamento utilizado no processo de aprendizagem da rede, para o caso desta tese foram testados 3 tipos de algoritmo sendo eles: SGDM (Gradiente descendente estocástico com momentum), ADAM (Estimativa de momentum adaptativo) e o RMSPROP (Propagação quadrática Média).

Assim a composição dos padrões reconstruídos de entrada da LSTM1 AET1 possuiu as mesmas vinte e sete variáveis de entrada sendo $[h(t+1)] + [Mês(t+1)(seno)] + [Mês(t+1)(cosseno)] + [vinte e quatro séries horárias de velocidades médias do vento reais]$ e [vinte e quatro variáveis de saída (séries horárias de velocidades médias do vento previstas)].

Para tal, a mesma se consistiu em uma rede neural de previsão de velocidades médias de vento com horizonte de 24h, ou seja, o passo de previsão variou da hora um até a hora vinte e quatro. Assim, para um valor de entrada $v(t)$ a LSTM1_AET1 irá prever valores conforme descrito pela Expressão 4.27.

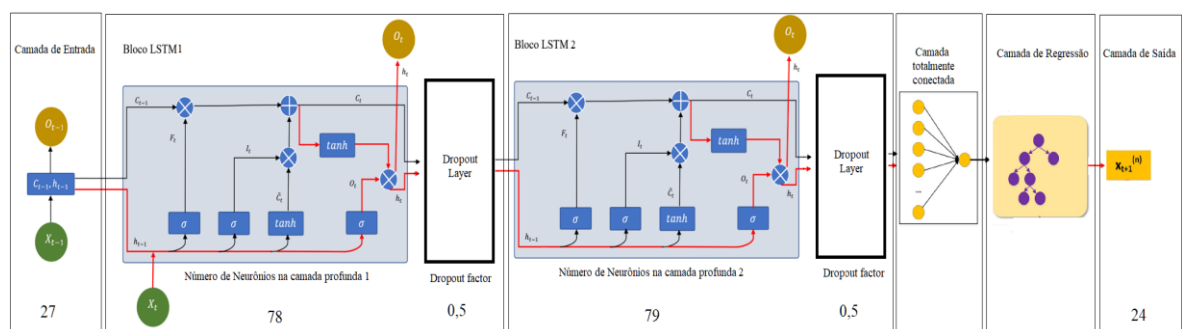
Em que, as melhores LSTM1 AET1, denominadas “LSTM1_AET1_BEST”, obtidas nas inicializações que simularam os treinamentos, nas bases dos Parques 01 e 02 são mostradas nas Figuras 71 e 72 respectivamente, com seus parâmetros utilizados:

Figura 71 – Arquitetura da melhor LSTM1 AET1 para o Parque 01 escolhida com a execução dos treinamentos.



Fonte: O Autor (2021).

Figura 72 – Arquitetura da melhor LSTM1 AET1 para o Parque 02 escolhida com a execução dos treinamentos.

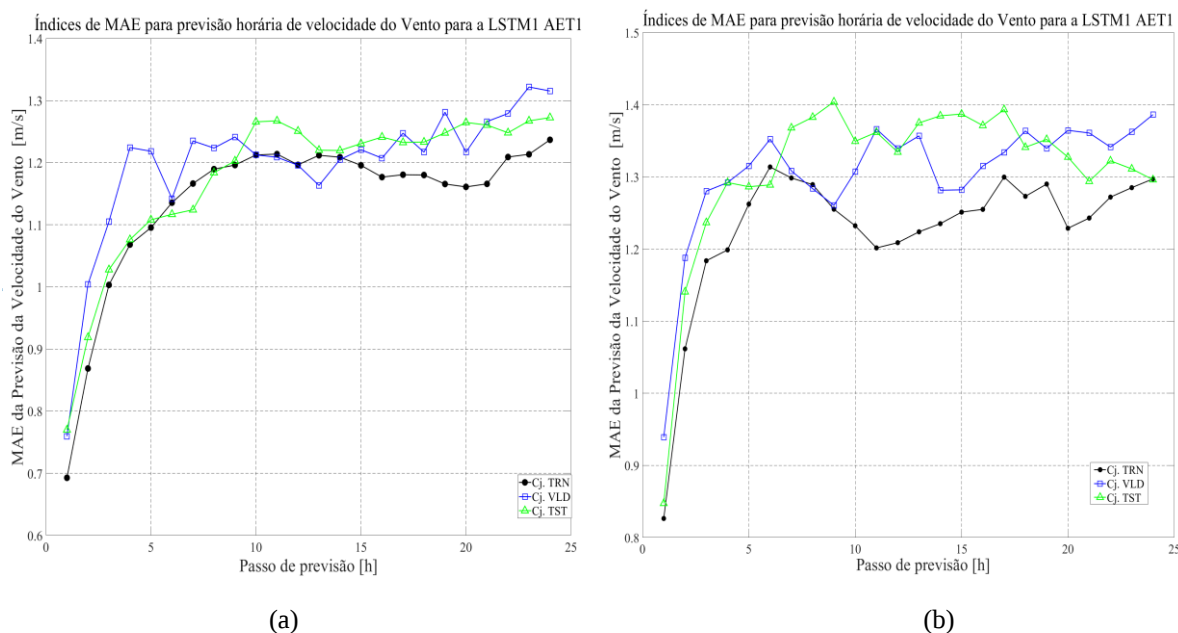


Fonte: O Autor (2021).

Das Figuras 71 e 72, pode-se notar que, para a melhor inicialização, o número de neurônios definidos pela camada profunda 1 foi de 89 e pela camada profunda 2 foi de 80 para o Parque 01 e respectivamente 66 e 95 para o Parque 02, descrevendo assim a arquitetura da LSTM1 após os treinamentos.

Outros resultados importantes relativos a estes treinamentos são descritos pela Figura 73 que descreve o desempenho do erro médio absoluto cometido para ambos os parques nas previsões das respectivas LSTM1 dos Parques 01 e 02. Em que os mesmos são mostrados a seguir.

Figura 73 – Comportamento gráfico do MAE referente às simulações para o período de aprendizagem da LSTM1 AET1 (a) aplicado ao Parque 01 e (b) aplicado ao Parque 02.



Fonte: O Autor (2021).

6.1.4.2 Previsor LSTM2_AET2 e seus parâmetros

O previsor LSTM2 AET2 foi concebido sendo a base do parque sobre estudo reconstruída pelo autoencoder da topologia AET2 de forma semelhante ao que fora feito com a rede LSTM2. A nova arquitetura desta nova heurística dispõe de camadas ocultas e camadas profundas com suas próprias funções de ativação e conexões específicas, semelhante ao que fora feito na LSTM2.

Desta forma, após a ação de reconstrução dos padrões da base sobre análise do AE da topologia AET2 os dados reconstruídos foram passados em formas de padrões para compor o vetor alimentador do conjunto de aprendizagem da rede e assim proceder com o seu treinamento. No caso específico do previsor LSTM2 AET2 foram utilizados como padrões de entrada [vinte e quatro séries horárias de velocidades médias do vento reais], assim como nas heurísticas anteriores que utilizaram apenas uma entrada reconstruída (velocidade de vento), caso do previsor RNA2 AET2.

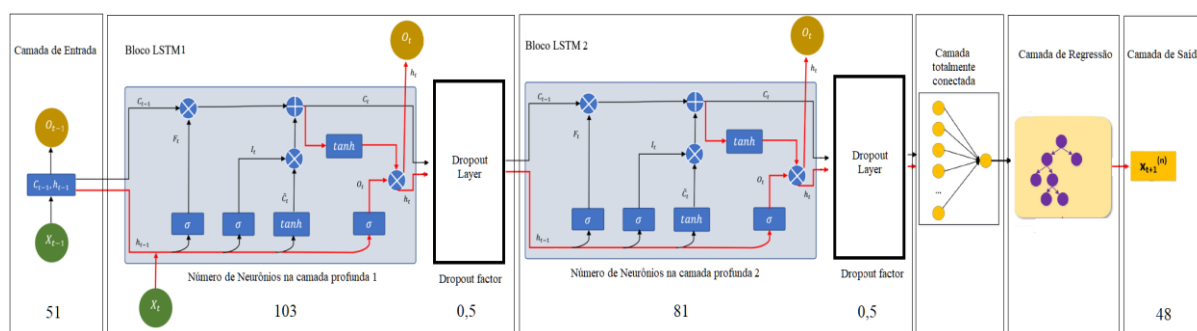
De forma semelhante aos *autoencoders* que buscam uma função que retrate a base sobre estudo (*feature*), as LSTM também buscam tal função, sendo esta *feature* estabelecida por camadas profundas, assim cada camada profunda terá a sua *feature* específica. Outro detalhe importante é o tipo de algoritmo de treinamento utilizado no processo de aprendizagem da rede, para o caso desta tese foram testados 3 tipos de algoritmo sendo eles: SGDM (Gradiente descendente estocástico com momentum), ADAM (Estimativa de momentum adaptativo) e o RMSPROP (Propagação quadrática Média).

Assim a composição dos padrões reconstruídos de entrada da LSTM1_AET1 possuiu as mesmas cinquenta e uma variáveis de entrada sendo $([h(t+1)] + [Mês(t+1)(seno)] + [Mês(t+1)(cosseno)] + [quarenta e oito séries horárias de velocidades médias do vento reais])$ e $[quarenta e oito variáveis de saída (séries horárias de velocidades médias do vento previstas)]$.

Para tal, a mesma se consistiu em uma rede neural de previsão de velocidades médias de vento com horizonte de 24h, ou seja, o passo de previsão variou da hora um até a hora vinte e quatro. Assim, para um valor de entrada $v(t)$ a LSTM2_AET2 irá prever valores conforme descrito pela Expressão 4.27.

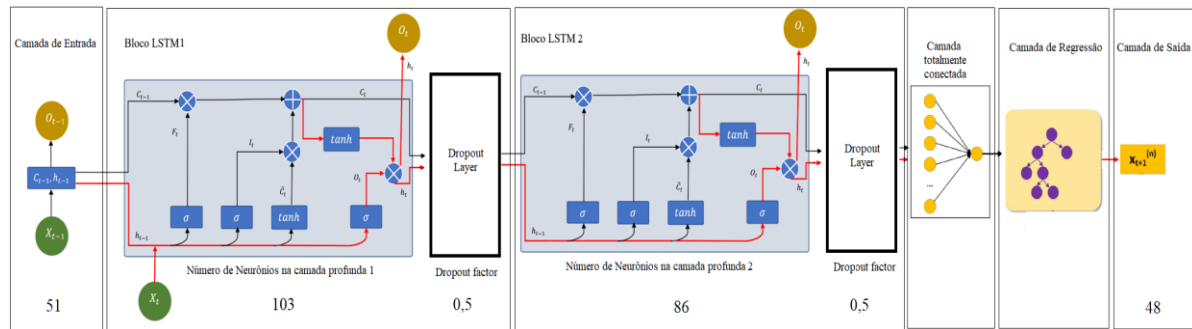
Em que, as melhores LSTM2_AET2, denominadas “LSTM2_AET2_BEST”, obtida nas inicializações que simularam os treinamentos, nas bases dos Parques 01 e 02 são mostradas nas Figuras 74 e 75 respectivamente, com seus respectivos parâmetros utilizados:

Figura 74 – Arquitetura da melhor LSTM2 AET2 para o Parque 01 escolhida com a execução dos treinamentos.



Fonte: O Autor (2021).

Figura 75 – Arquitetura da melhor LSTM2 AET2 para o Parque 02 escolhida com a execução dos treinamentos.



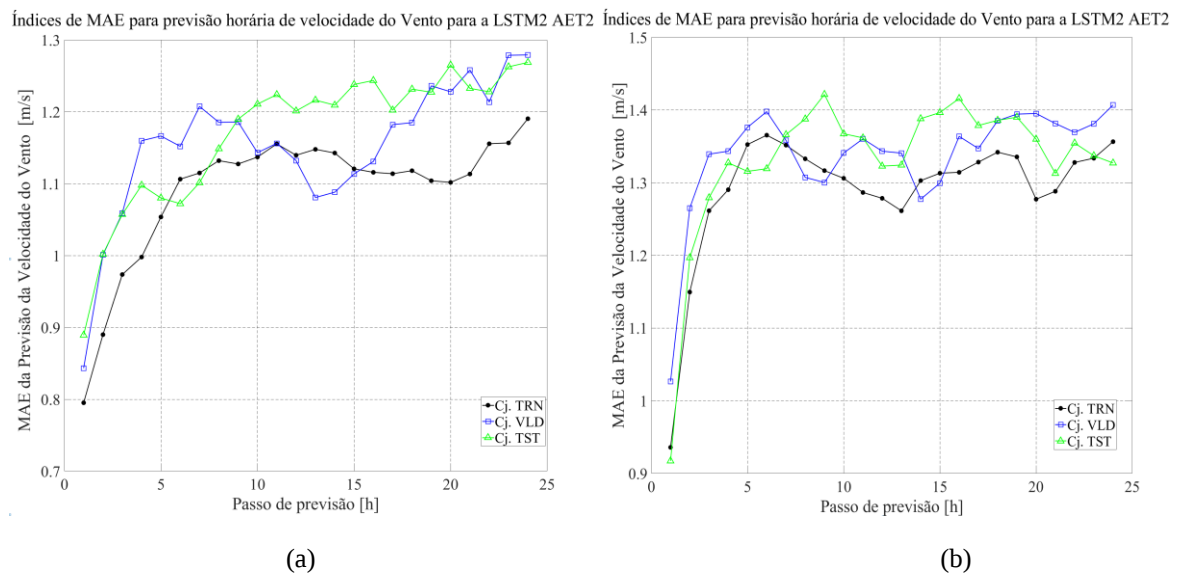
Fonte: O Autor (2021).

Das Figuras 74 e 75, pode-se notar que, para a melhor inicialização, o número de neurônios definidos pela camada profunda 1 foi de 103 e pela camada profunda 2 foi de 81 para o Parque 01 e respectivamente 103 e 86 para o Parque 02, descrevendo assim a arquitetura da LSTM2 AET2 após os treinamentos.

Outros resultados importantes relativos a estes treinamentos são descritos pelas Figuras 76 e 77 que descreve o desempenho do erro médio absoluto cometido para ambos os parques nas previsões das respectivas LSTM2 AET2 dos Parques 01 e 02.

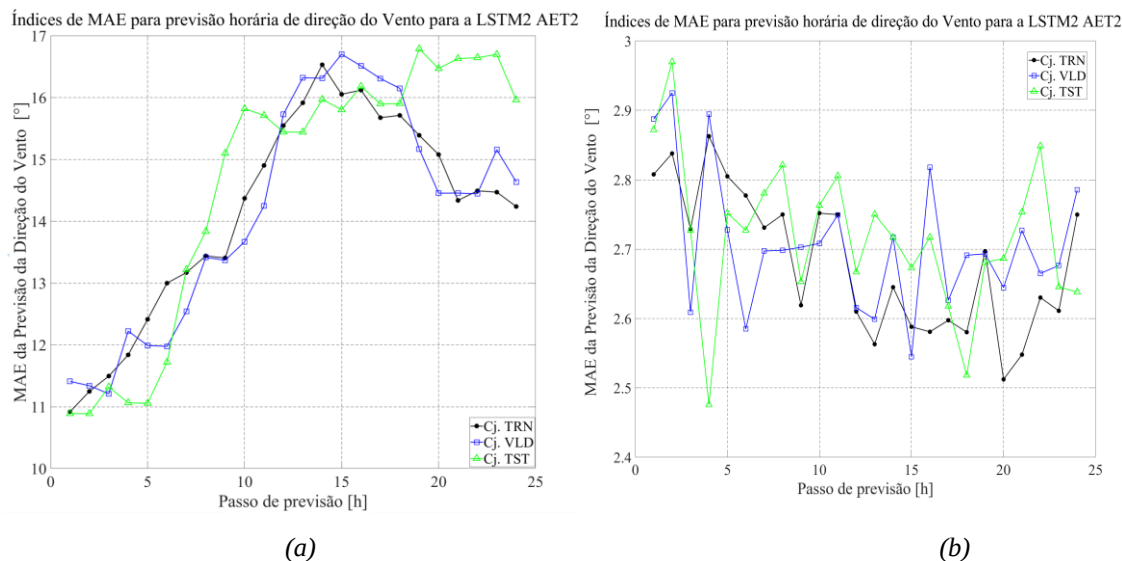
Em que os mesmos são mostrados a seguir.

Figura 76 – Comportamento gráfico do MAE referente às simulações para o período de aprendizagem da LSTM2_AET2 (a) aplicado ao Parque 01 e (b) aplicado ao Parque 02.



Fonte: O Autor (2021).

Figura 77 – Comportamento gráfico do MAE da direção média referente às simulações para o período de aprendizagem da LSTM2_AET2 aplicado em (a) para Parque 01 e em (b) para o Parque 02.



Fonte: O Autor (2021).

Na Figura 76 pode-se ver que o comportamento respectivo nos Parques em questão no período de aprendizagem não possui muita diferença com relação à previsão de velocidades. Com respeito à previsão de direções o comportamento é bastante distinto verificando-se um crescimento do MAE à medida que o passo de previsão cresce no Parque 01 e um decréscimo do MAE para o mesmo horizonte no Parque 02.

Analisando a Figura 77 é possível perceber a diferença no comportamento do MAE, pois no Parque 01 o erro aumenta ao longo do horizonte de previsão, enquanto no Parque 02 o erro diminui. Outro fator perceptível é a diferença no patamar do erro sendo 11° na 1ª hora e subindo até um patamar de 17° no fim do horizonte. Já o Parque 02 o erro é da ordem de 2,9° na 1ª hora, oscila até 2,5° e eleva-se para 2,8° no final do horizonte.

6.2 Metodologia utilizada para os Modelos de curvas de Potência Eólica (WTPCs) descritos nesta tese

Apesar de fornecer uma relação entre a velocidade do vento e Potência gerada considerando uma densidade de ar específica, a curva de potência fornecida pelos fabricantes, não leva em consideração o local de instalação e nem o desgaste da máquina eólica. Pois segundo KUSIAK, A. et. al. (2013), foi verificado que não há discrepâncias entre as curvas de

potência dos fabricantes de pequenas turbinas eólicas e os resultados de testes realizados em altas velocidades de vento.

Outro ponto importante é que a variabilidade da velocidade do vento faz com que a curva de potência do fabricante não seja considerada um modelo adequado para a estimativa da potência gerada, uma vez que a mesma ignora a tendência dinâmica do vento, LYDIA, M. et. al. (2013). Assim para levar em consideração esse caráter variável dos parâmetros envolvidos na estimativa da potência pela incidência de ventos, cria-se a necessidade, do desenvolvimento de modelos de curvas de potência eólicas, intitulados (WTPCs).

Foram desenvolvidos, doze modelos aproximadores de curvas de potência, dos quais nove foram utilizados efetivamente na construção dos previsores de geração desenvolvidos, sendo: dois utilizando RNAs, três utilizando RNAs adequadas por autoencoder, dois utilizando SIFs (*Mamdani*), dois utilizando SIFs (*Sugeno*) (Não utilizados), dois utilizando SIFs (*Sugeno*) com auxílio do ANFIS e um polinomial (Não utilizado), sendo a inovação proposta nos dois últimos modelos *Fuzzy* e na influência da reconstrução das topologias dos autoencoders nos 3 modelos das novas CPs RNA.

É importante salientar dois aspectos:

1) O modelo de Referência paramétrico polinomial da curva média do parque foi utilizado apenas como fator de comparação na análise do erro cometido na representação dos modelos de curvas de potência desenvolvidos. Desta forma o modelo polinomial foi a base de referência na análise do erro médio dos modelos de CP desenvolvidos e assim não foi utilizado como componente na arquitetura dos previsores de geração.

2) Os modelos Fuzzy Sugeno simples desta Tese por serem modelos intermediários no processo de construção do modelo das CPFuzzy Sugeno ANFIS não estarão relacionados na análise dos resultados no modelo predictor final de geração. Isto deve-se ao fato de os mesmos serem modelos que ainda não se encontram sintonizados, e desta forma a sua apresentação dificultaria a análise do processo, devido à quantidade de modelos de CP desenvolvidos. Porém, a sua criação foi importante em termos científicos para comprovar que a heurística MAMDANI pode ser representada com confiabilidade por uma transformação de um sistema MAMDANI em um sistema SUGENO.

A aplicação foi feita para dois parques situados na região nordeste do Brasil, que por critérios confidenciais serão chamados neste trabalho de Parque 01 e Parque 02.

6.3 Modelo de Referência - Curva de potência Média (Polinomial)

O modelo paramétrico da curva média é fortemente baseado na curva média do parque sobre estudo, assim para garantir uma boa aproximação para tal tendência construiu-se uma função polinomial que melhor adequasse seus parâmetros a estas características.

A curva média utilizada consiste num modelo paramétrico que determina uma função polinomial estabelecendo de forma aproximada a relação média do caráter (Velocidade x Potência) para um dado parque, utilizando para isso a base de dados do sítio respectivo.

Tal função polinomial é composta por trechos específicos, bastante conhecidos na CP, sendo esses trechos a região de (*Cut in*), A região de operação efetiva, a Região (*Cut out*). Nos modelos de curvas médias desse artigo, por se tratar de modelos de operação, a região considerada na criação foi a região de operação efetiva (Após a região de (*Cut-in*) e antes da Região de (*Cut-out*)).

Para tal foi utilizada a função *polyfit* do Matlab na geração de tais curvas médias, foram consideradas nas montagens das mesmas 4 regiões, sendo 3 cúbicas e uma constante descritas a seguir. Tais curvas tiveram a referente aproximação polinomial, devido a tendência esboçada pelo trecho sobre aproximação, dizendo respeito a cada parque sobre análise:

1) Região cúbica 1, 2) Região Cúbica 2, 3) Região Cúbica 3 e 4) Região constante. Como o próprio nome indica cada região tem uma característica de relação entre pontos especificada de acordo com um polinômio característico para cada parque em análise:

$$1) \text{ Região cúbica 1 } \Rightarrow P_{c1}(v) = \alpha_{11}v^3 + \alpha_{21}v^2 + \alpha_{31}v + c_1 \quad (6.19)$$

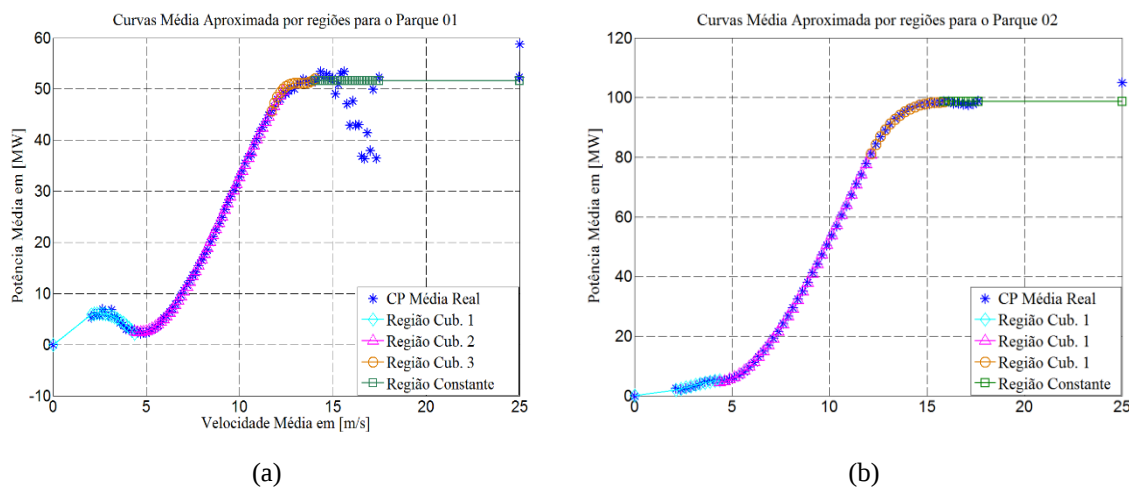
$$2) \text{ Região cúbica 2 } \Rightarrow P_{c2}(v) = \alpha_{12}v^3 + \alpha_{22}v^2 + \alpha_{32}v + c_2 \quad (6.20)$$

$$3) \text{ Região cúbica 3 } \Rightarrow P_{c3}(v) = \alpha_{13}v^3 + \alpha_{23}v^2 + \alpha_{33}v + c_3 \quad (6.21)$$

$$4) \text{ Região constante } \Rightarrow P_k(v) = c_4 \quad (6.22)$$

Na Figura 78 se observa a tendência aproximada da curva média para cada região de aproximação em cada parque respectivo.

Figura 78 – Aproximações polinomiais relativas às curvas de potências médias no (a) Parque 01 e no (2) Parque 02.



Fonte: O Autor (2020).

6.4 Modelos não paramétricos de Curvas de Potência utilizando RNAs

Neste trabalho foram utilizados duas arquiteturas de modelos aproximadores de CPs que utilizam RNAs em sua estrutura estando a diferença entre as mesmas, no número de entradas, no qual o objetivo foi verificar a influência no comportamento da CP para uma entrada diferente da velocidade do vento no caso (a direção do vento). A escolha desta variável se deu devido a sua boa correlação com a velocidade do vento.

Assim foram desenvolvidos 2 modelos, o primeiro CP1RNA, consiste em uma RNA com entradas de velocidades médias de vento e saídas de potência. O segundo CP2RNA consiste de uma RNA com entradas de velocidades médias e direções médias e saídas de potência.

Para a fase de aprendizagem dos modelos a RNAs, foram tomados dois bancos de dados referentes a parques eólicos do nordeste do brasileiro, que por questões de sigilo não serão divulgados neste trabalho. Para que fique clara a distinção entre os mesmos, neste trabalho serão denominados como parque 01 e parque 02, o mesmo é feito para a fase de simulação. Conforme descrito em ALBUQUERQUE (2015), os parques estão situados na região nordeste do Brasil, mais especificamente no estado do Ceará, e já se encontram em pleno funcionamento, tendo 28 máquinas no Parque 01 com a capacidade total instalada de 70,56MW e, 60 máquinas no Parque 02 com a capacidade total instalada de 126MW.

Os parques são característicos da região e, portanto, os modelos podem ser aplicados aos demais parques da região sem grandes perdas na precisão das previsões. Desta forma,

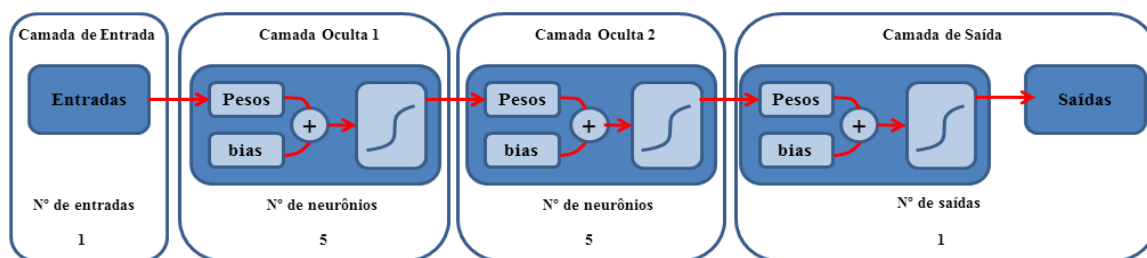
pode-se ter uma ideia da variação do desempenho dos modelos para parques na mesma região em termos de valores percentuais do erro de previsão da geração.

6.4.1 MODELO DA CP1RNA

O modelo da CP1RNA consiste em uma rede neural MLP com uma entrada sendo esta, a velocidade média do vento e uma saída sendo esta, a Potência média do vento, tais modelos de CPs a RNAs são os mesmos utilizados em DE. ALBUQUERQUE, J.C. (2015) e em AQUINO R. R. B. et. al. (2016) cujos parâmetros são mostrados nestas referências. Outra particularidade é que estes trabalhos utilizaram mais de uma camada oculta em suas arquiteturas, fato que agregou maior precisão nas aproximações, podendo este enfoque ser considerado como uma característica de aprofundamento da informação mesmo se tratando de redes do tipo MLP.

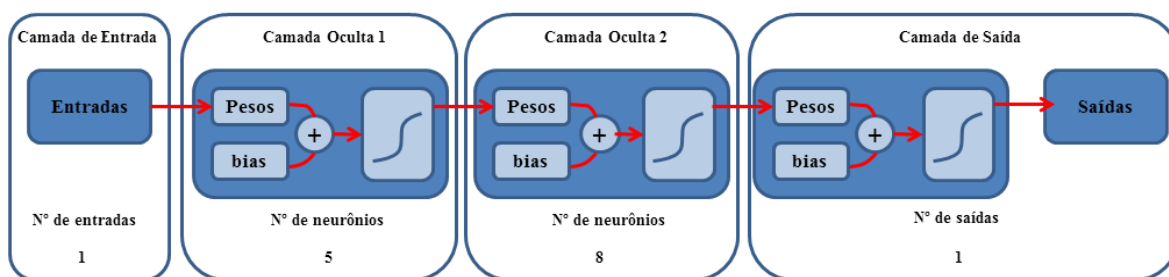
No processo de aprendizagem das mesmas foi utilizado o algoritmo de Levenberg Marquardt, outro fator importante são as arquiteturas destas CPs como dito anteriormente, possuindo mais de uma camada oculta. Tais arquiteturas referentes à CP1RNA são mostradas para ambos os parques 01 e 02 pelas Figuras 79 e 80.

Figura 79 – Arquitetura da CP1RNA para o Parque 01 escolhida no período de aprendizagem.



Fonte: DE ALBUQUERQUE, J.C. (2015).

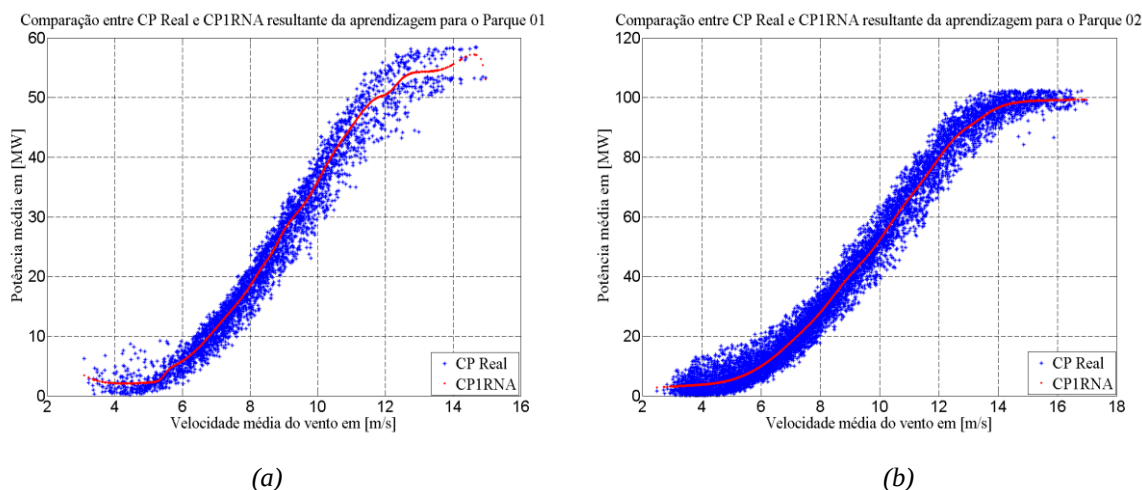
Figura 80 – Arquitetura da CP1RNA para o Parque 02 escolhida no período de aprendizagem.



Fonte: DE ALBUQUERQUE, J.C. (2015).

Outro ponto importante diz respeito às características de aproximação de tais redes para ambos os parques como mostrado a seguir pela Figura 81:

Figura 81 – Comparação entre CP Real e a curva CPRNA1 treinada para o (a) Parque 01e (b) para o Parque 02.



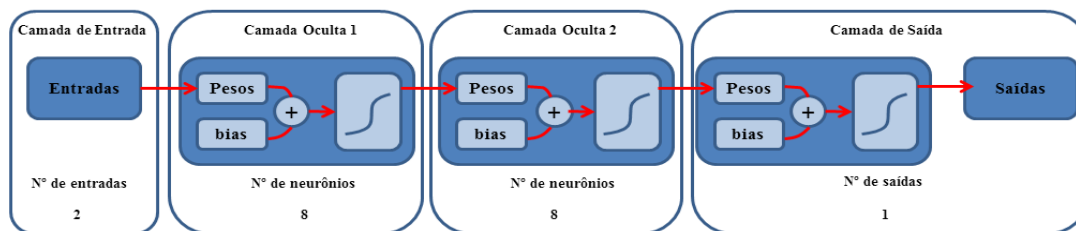
Fonte: DE ALBUQUERQUE, J.C. (2015).

6.4.2 MODELO DA CP2RNA:

O modelo CP2RNA consiste de uma rede neural MLP com duas entradas sendo estas a velocidade média do vento e a direção média do vento, e uma saída sendo esta a Potência média do vento, tais modelos de CPs a RNAs também são os mesmos utilizados em DE. ALBUQUERQUE, J.C. (2015) e em AQUINO R. R. B. et. al. (2016) cujos parâmetros são mostrados nestas referências. De forma semelhante ao utilizado no modelo anterior na aprendizagem das mesmas foi utilizado o algoritmo de *Levenberg Marquardt*, sendo utilizada como função de ativação em cada camada específica a tangente hiperbólica e as arquiteturas destas CPs também possuem mais de uma camada oculta.

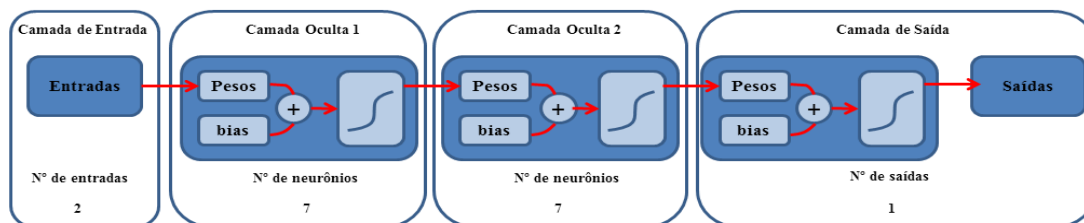
Tais arquiteturas referentes à CP2RNA são mostradas para ambos os Parques 01 e 02 pelas Figuras 82 e 83.

Figura 82 – Arquitetura da CP2RNA para o Parque 01 escolhida no período de aprendizagem.



Fonte: DE ALBUQUERQUE, J.C. (2015).

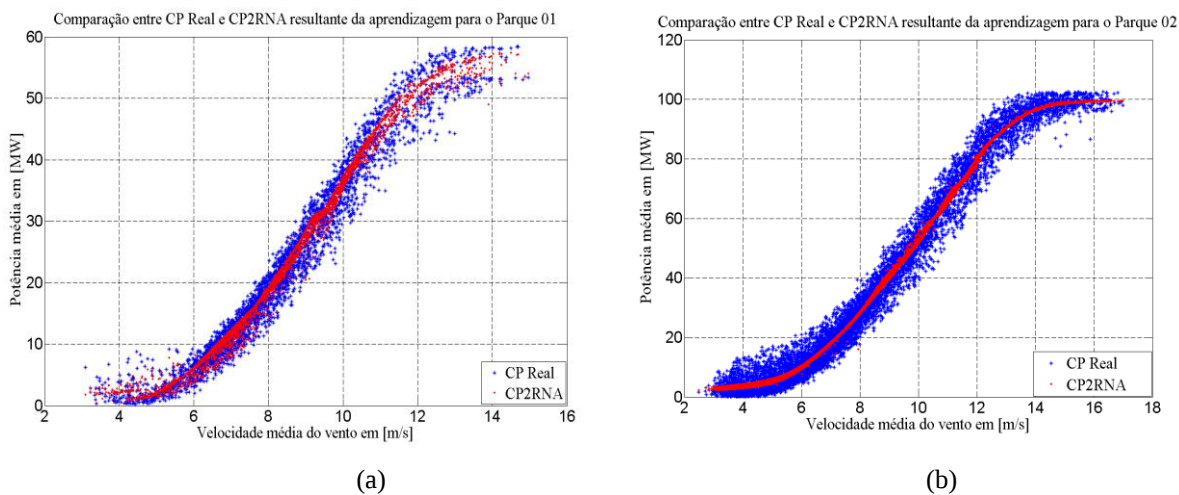
Figura 83 – Arquitetura da CP2RNA para o Parque 02 escolhida no período de aprendizagem.



Fonte: DE ALBUQUERQUE, J.C. (2015).

Em que de forma semelhante às características de aproximação de tais redes para ambos os parques como mostrado a seguir pela Figura 84:

Figura 84 – Comparação entre CP Real e a curva CP2RNA treinada para o (a) Parque 01e (b) para o Parque 02.



Fonte: DE ALBUQUERQUE, J.C. (2015).

6.5 Modelos não paramétricos de Curvas de Potência utilizando RNAs adequadas por Autoencoder

Modelos não paramétricos não possuem uma função contínua descrita por um polinômio ou função paramétrica e sim uma função iterativa objetiva discreta que depende de um determinado conjunto de dados, intervalo de análise e também de um padrão relacional característico entre os dados, advindo de um processo de aprendizagem.

A seguir são descritos os modelos de RNAs desenvolvidos para simular as curvas de potência dos Parques 01 e 02 sendo os dados utilizados, reconstruídos por duas topologias de autoencoders diferentes.

6.5.1 MODELO DA CP1-RNA RETREINADA COM DADOS FILTRADOS PELO AE TOPOLOGIA AET1

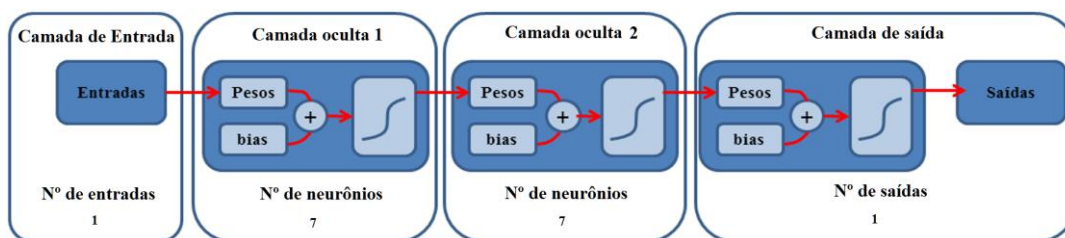
A rede neural MLPCP1RNA AET1 foi nomeada de CP1RNA AET1 e consiste em uma rede do tipo MLP com os procedimentos para a sua obtenção sendo os mesmos que a CP1RNA exceto no que diz respeito a sua base de dados. A base de dados utilizada no processo de aprendizagem desta rede foi reconstruída pelo autoencoder da topologia AET1. Desta forma os dados relativos ao seu conjunto de aprendizagem, que de forma diferente do caso anterior fora obtido pela reconstrução da base por intermédio do autoencoder da topologia AET1.

Assim o AET1 reconstruiu as bases relativas aos parques em questão redefinindo um novo conjunto de aprendizagem, com o mesmo número de padrões do conjunto base, e foi utilizado no treinamento do previsor CP1RNA AET1.

O modelo da CP1RNA AET1 consiste de uma rede neural MLP com uma entrada sendo esta a velocidade média do vento e uma saída sendo esta a Potência média do vento, tais modelos de CPs a RNAs são os mesmos utilizados em DE. ALBUQUERQUE, J.C. (2015) e em AQUINO R. R. B. et. al. (2016) cujos parâmetros são mostrados nestas referências. Outra particularidade é que estes trabalhos utilizaram mais de uma camada oculta em suas arquiteturas, fato que agregou maior precisão nas aproximações, podendo este enfoque ser considerado como uma característica de aprofundamento da informação mesmo se tratando de redes do tipo MLP.

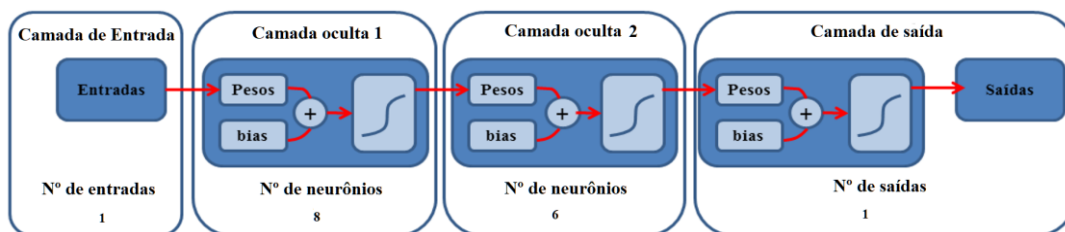
No processo de aprendizagem das mesmas foi utilizado o algoritmo de Levenberg Marquardt, outro fator importante são as arquiteturas destas CPs como dito anteriormente, possuindo mais de uma camada oculta. Tais arquiteturas referentes à CP1RNA_AET1 são mostradas para ambos os Parques 01 e 02 pelas Figuras 85 e 86.

Figura 85 – Arquitetura da CP1RNA AET1 para o Parque 01 escolhida no período de aprendizagem.



Fonte: O Autor (2021).

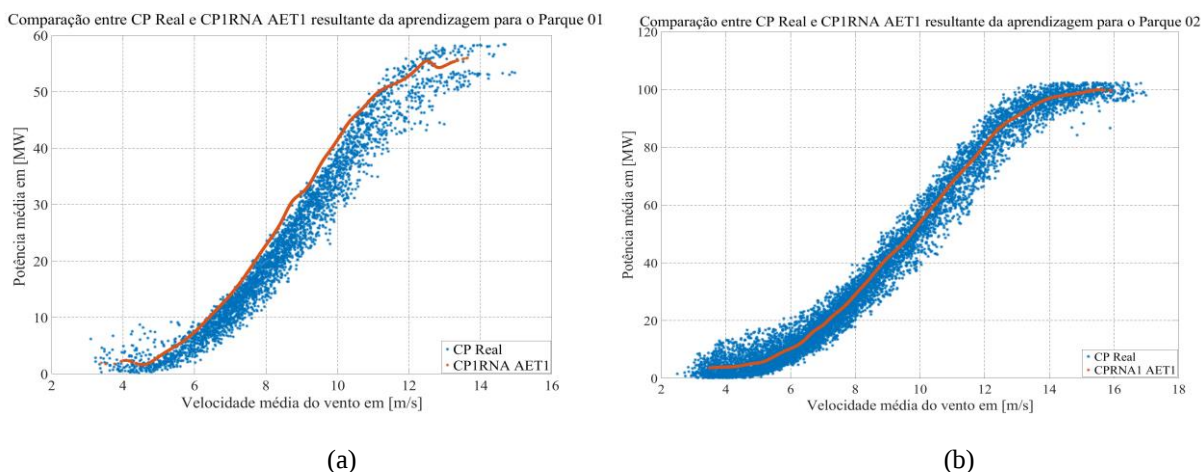
Figura 86 – Arquitetura da CP1RNA AET1 para o Parque 02 escolhida no período de aprendizagem.



Fonte: O Autor (2021).

Outro ponto importante diz respeito às características de aproximação de tais redes para ambos os parques como mostrado a seguir pela Figura 87:

Figura 87 – Comparação entre CP Real e a curva CPRNA1_AET1 treinada para o (a) Parque 01e (b) para o Parque 02.



Fonte: O Autor (2021).

6.5.2 MODELO DA CP1-RNA RETREINADA COM DADOS FILTRADOS PELO AE TOPOLOGIA AET2

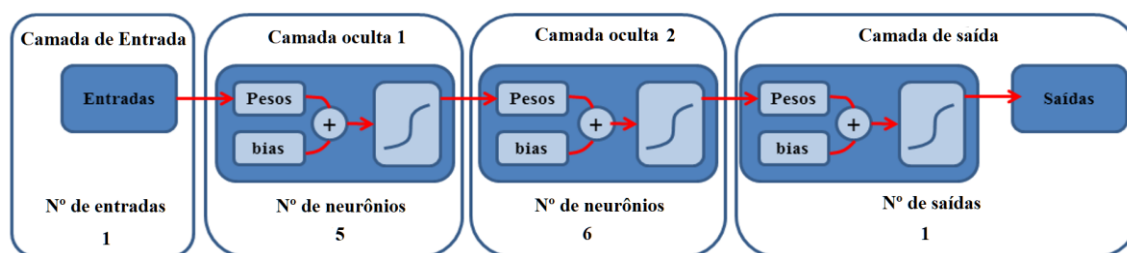
A rede neural MLPCP1RNA AET2 foi nomeada de CP1RNA AET2 e consiste em uma rede do tipo MLP com os procedimentos para a sua obtenção sendo os mesmos que a CP1RNA exceto no que diz respeito a sua base de dados. A base de dados utilizada no processo de aprendizagem desta rede foi reconstruída pelo autoencoder da topologia AET2. Desta forma os dados relativos ao seu conjunto de aprendizagem, que de forma diferente do caso anterior fora obtido pela reconstrução da base por intermédio do autoencoder da topologia AET2.

Assim o AET2 reconstruiu as bases relativas aos parques em questão redefinindo um novo conjunto de aprendizagem, com o mesmo número de padrões do conjunto base, e foi utilizado no treinamento do previsor CP1RNA AET2.

O modelo da CP1RNA AET2 consiste em uma rede neural MLP com uma entrada sendo esta a velocidade média do vento e uma saída sendo esta a Potência média do vento, tais modelos de CPs a RNAs são os mesmos utilizados em DE. ALBUQUERQUE, J.C. (2015) e em AQUINO R. R. B. et. al. (2016) cujos parâmetros são mostrados nestas referências. Outra particularidade é que estes trabalhos utilizaram mais de uma camada oculta em suas arquiteturas, fato que agregou maior precisão nas aproximações, podendo este enfoque ser considerado como uma característica de aprofundamento da informação mesmo se tratando de redes do tipo MLP.

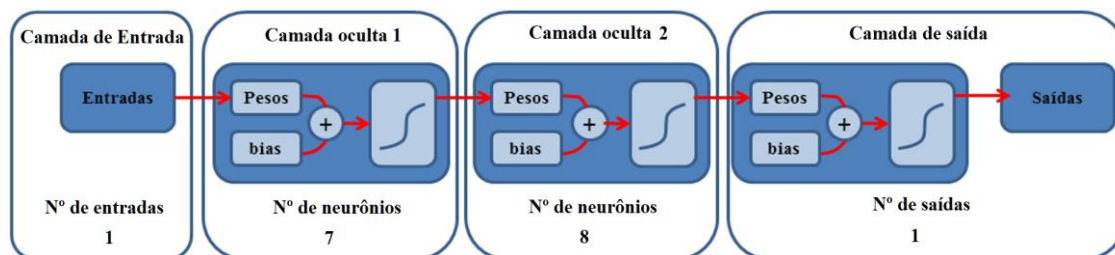
No processo de aprendizagem das mesmas foi utilizado o algoritmo de Levenberg Marquardt, outro fator importante são as arquiteturas destas CPs como dito anteriormente, possuindo mais de uma camada oculta. Tais arquiteturas referentes à CP1RNA AET2 são mostradas para ambos os parques 01 e 02 pelas Figuras 88 e 89.

Figura 88 – Arquitetura da CP1RNA AET2 para o Parque 01 escolhida no período de aprendizagem.



Fonte: O Autor (2021).

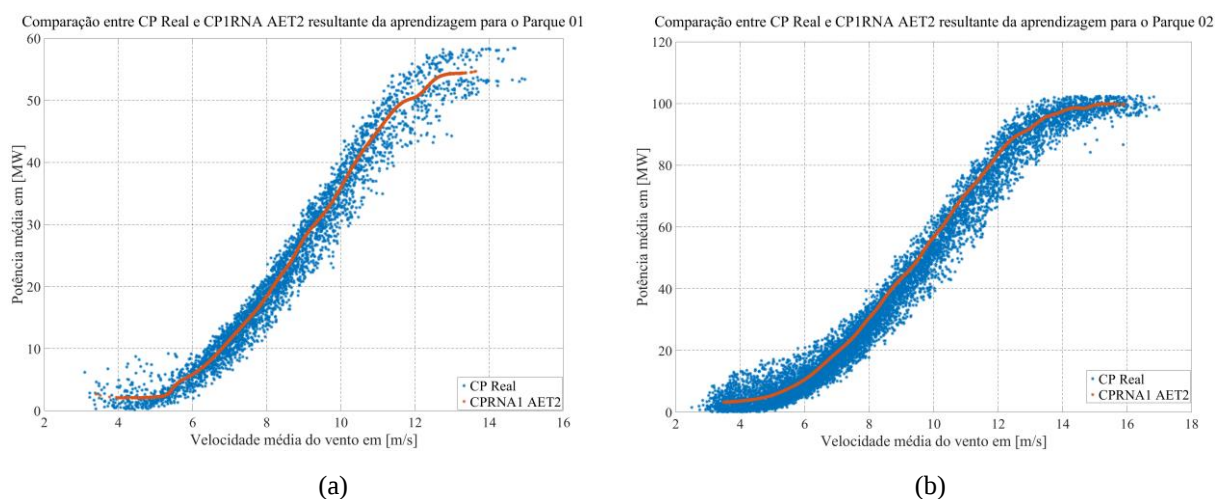
Figura 89 – Arquitetura da CP1RNA AET2 para o Parque 02 escolhida no período de aprendizagem.



Fonte: O Autor (2021).

Outro ponto importante diz respeito às características de aproximação de tais redes para ambos os parques como mostrado a seguir pela Figura 90:

Figura 90 – Comparação entre CP Real e a curva CPRNA1_AET2 treinada para o (a) Parque 01e (b) para o Parque 02.



Fonte: O Autor (2021).

6.5.3 MODELO DA CP2-RNA RETREINADA COM DADOS FILTRADOS PELO AE TOPOLOGIA AET2

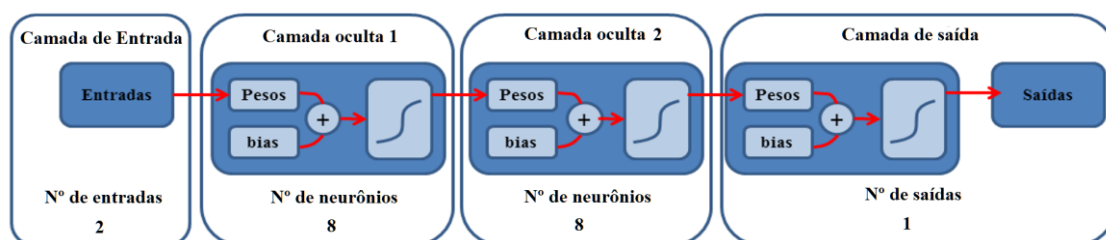
A rede neural do tipo MLP CP2RNA AET2 foi nomeada de CP2RNA AET2 e consiste em uma rede do tipo MLP com os procedimentos para a sua obtenção sendo os mesmos que a CP2RNA exceto no que diz respeito a sua base de dados. A base de dados utilizada no processo de aprendizagem desta rede foi reconstruída pelo autoencoder da topologia AET2. Desta forma os dados relativos ao seu conjunto de aprendizagem, que de forma diferente do caso anterior fora obtido pela reconstrução da base por intermédio do autoencoder da topologia AET2.

Assim o AET2 reconstruiu as bases relativas aos parques em questão redefinindo um novo conjunto de aprendizagem, com o mesmo número de padrões do conjunto base, e foi utilizado no treinamento do previsor CP1RNA AET2.

O modelo CP2RNA AET2 consiste de uma rede neural MLP com duas entradas sendo estas a velocidade média do vento e a direção média do vento, e uma saída sendo esta a Potência média do vento, tais modelos de CPs a RNAs também são os mesmos utilizados em DE. ALBUQUERQUE, J.C. (2015) e em AQUINO R. R. B. et. al. (2016) cujos parâmetros são mostrados nestas referências. De forma semelhante ao utilizado no modelo anterior na aprendizagem das mesmas foi utilizado o algoritmo de *Levenberg Marquardt*, sendo utilizada como função de ativação em cada camada específica a tangente hiperbólica e as arquiteturas destas CPs também possuem mais de uma camada oculta.

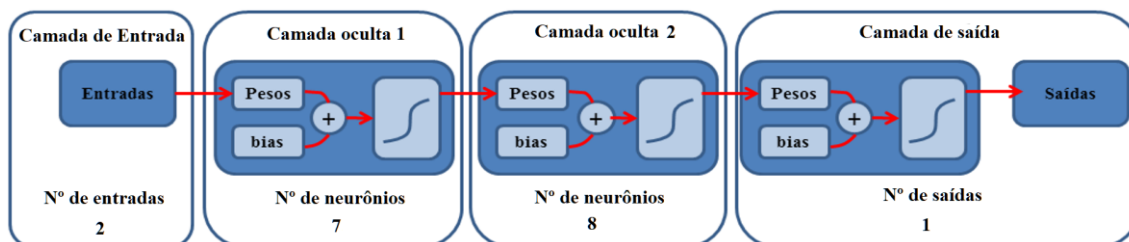
Tais arquiteturas referentes à CP2RNA AET2 são mostradas para ambos os parques 01 e 02 pelas Figuras 91 e 92.

Figura 91 – Arquitetura da CP2RNA AET2 para o Parque 01 escolhida no período de aprendizagem.



Fonte: O Autor (2021).

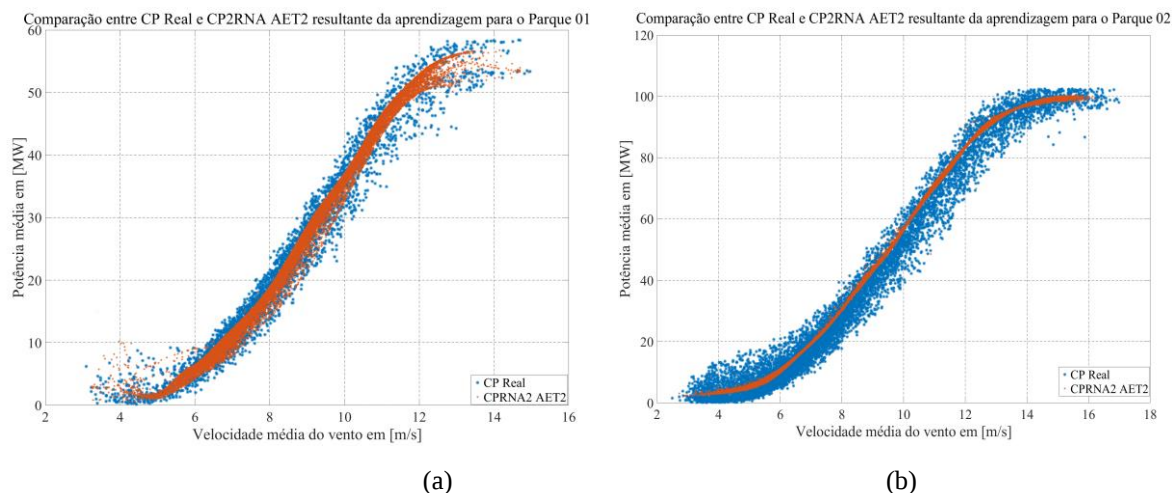
Figura 92 – Arquitetura da CP2RNA AET2 para o Parque 02 escolhida no período de aprendizagem.



Fonte: O Autor (2021).

De forma semelhante às características de aproximação de tais redes para ambos os parques como mostrado a seguir pela Figura 93:

Figura 93 – Comparação entre CP Real e a curva CP2RNA AET2 treinada para o (a) Parque 01e (b) para o Parque 02.



Fonte: O Autor (2021).

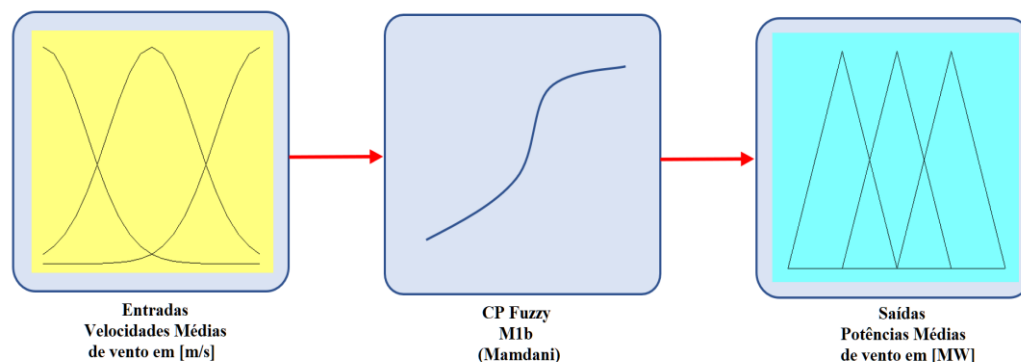
6.6 Modelos não paramétricos de Curvas de Potência utilizando Lógica Fuzzy

Foram utilizados dois tipos de sistemas *Fuzzy* distintos, um com a inferência *Mamdani* e outro com a inferência *Takagi-Sugeno-Kang*, sendo um, com o auxílio do ANFIS e um, com inferência manual. Onde de forma semelhante aos modelos com redes neurais foram utilizadas uma entrada no caso dos modelos CP1Fuzzy, CP1Fuzzy-Sugeno-ANFIS ou duas entradas no caso do modelo CP2Fuzzy e CP2Fuzzy-Sugeno-ANFIS.

6.6.1 MODELO DA CP1FUZZY (MAMDANI)

O modelo CP1Fuzzy consiste em um Sistema de Inferência *Fuzzy* (SIF) com uma entrada sendo esta, a velocidade média do vento e uma saída sendo esta, a Potência média do vento, tais modelos de CPs a SIFs são os mesmos utilizados em DE. ALBUQUERQUE, J.C. (2015) e em AQUINO R. R. B. et. al. (2016) cujos parâmetros também são mostrados nestas referências. Em que a arquitetura para a CP1Fuzzy é mostrada pela Figura 94, sendo a mesma heurística utilizada em ambos os Parques 01 e 02.

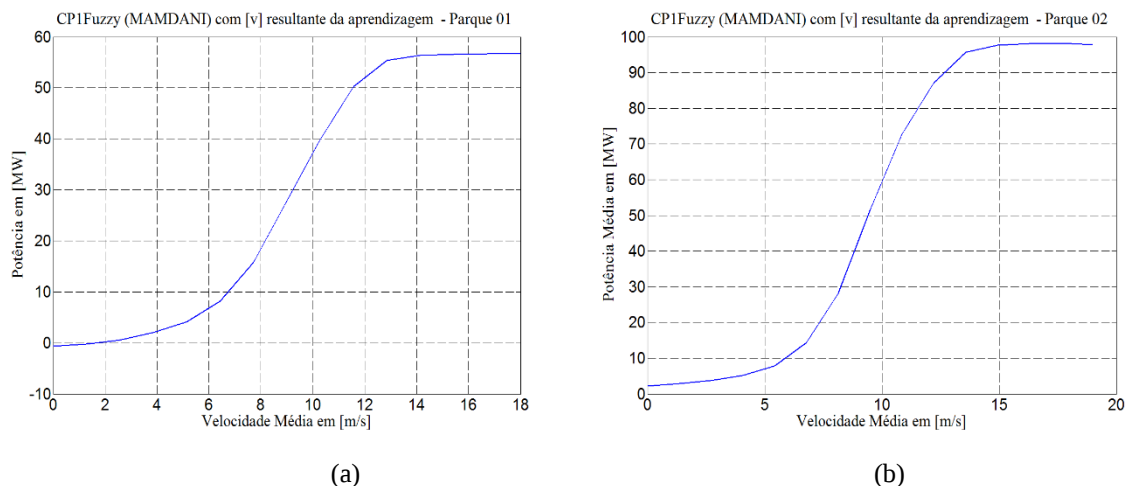
Figura 94 – Arquitetura do SIF Fuzzy, Mamdani, referente à CP1Fuzzy.



Fonte: DE ALBUQUERQUE, J.C. (2015).

Em que a Figura 95 mostra a curva de tendência resultante do processo de aprendizagem, considerando os limites manualmente impostos pelas funções de pertinência e suas relações.

Figura 95 – CP1Fuzzy resultante do processo de aprendizagem relativa a o (a) Parque 01 e ao (b) Parque 02.



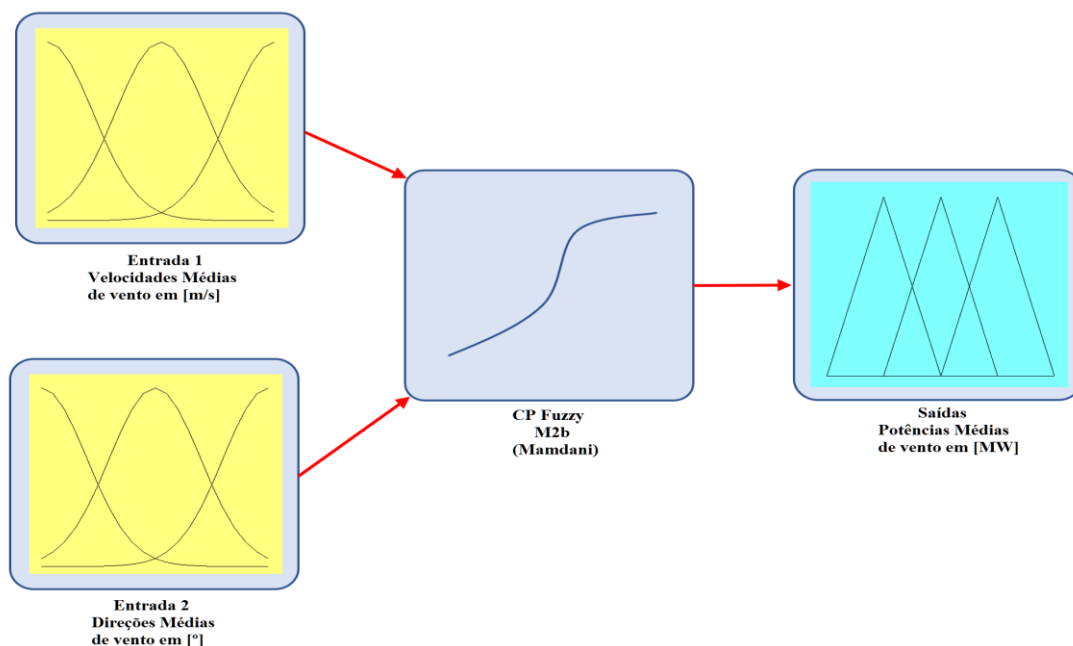
Fonte: O Autor (2015).

6.6.2 MODELO DA CP2FUZZY (MAMDANI)

O modelo CP2Fuzzy, mesmo definido, proposto e avaliado em DE. ALBUQUERQUE, J.C. (2015) e em AQUINO R. R. B. et. al. (2016), mostrado pela Figura 96, consiste de um Sistema de Inferência Fuzzy (SIF) com duas entradas sendo estas, a velocidade média do vento e a direção média do vento e uma saída sendo esta, a Potência média do vento. De forma semelhante ao caso com uma entrada, para a fase de aprendizagem

o modelo da CP2Fuzzy (*Mamdani*) foi construído de forma manual utilizando o método de tentativa e erro descriminando a posição adequada para a função de pertinência numa faixa estabelecida, comparando-se o valor obtido e o esperado.

Figura 96 – Arquitetura do SIF Fuzzy (Mamdani), referente a CP2Fuzzy.

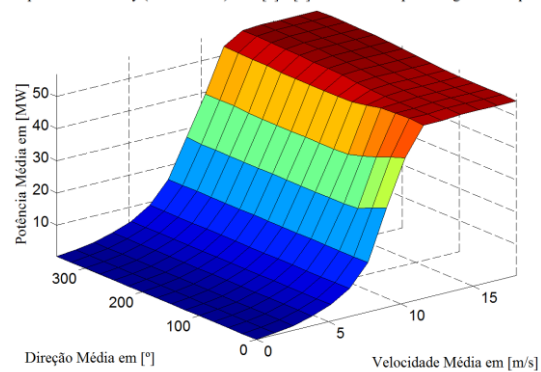


Fonte: DE ALBUQUERQUE, J.C. (2015).

Em que a Figura 97, mostra as superfícies de tendência resultantes do processo de aprendizagem, considerando os limites manualmente impostos pelas funções de pertinência e suas relações.

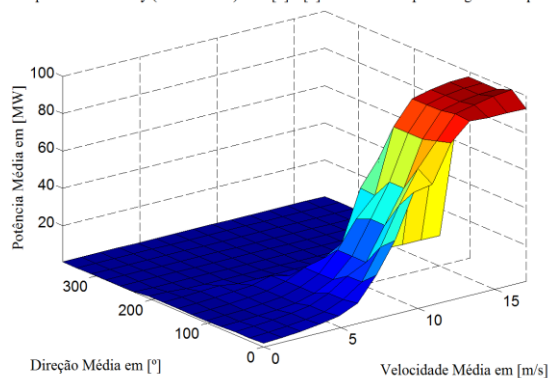
Figura 97 – CP2Fuzzy resultante do processo de aprendizagem relativa ao (a) Parque 01 e ao (b) Parque 02.

Superfície CP2Fuzzy (MAMDANI) com [v] e [d] resultante da aprendizagem - Parque 01



(a)

Superfície CP2Fuzzy (MAMDANI) com [v] e [d] resultante da aprendizagem - Parque 02



(b)

Fonte: O Autor (2020).

6.6.3 MODELO DA CP1FUZZY (SUGENO-ANFIS)

Considerando as diferenças relativas à criação dos modelos *Sugeno* em relação aos modelos *Mamdani*, o modelo CP1Fuzzy (Sugeno-ANFIS) tem a mesma entrada do CP1Fuzzy (*Mamdani*), a velocidade do vento, e uma saída sendo esta a Potência média do vento. A inferência utilizada é a inferência *Sugeno* onde a saída é discriminada por faixas e cada faixa obedece a uma função de inferência específica podendo esta função ser: constante ou linear.

Para fazer a conversão das características de um sistema *Mamdani* em sistema *Sugeno* utilizou-se a transformada ‘*mam2sug*’ do Matlab. Esta transformada converte um sistema *Mamdani* onde a inferência é dada pelo critério de máximos e mínimos por faixa, em um sistema *Sugeno*, cuja inferência (função de saída) é dada por funções relativas a cada faixa, podendo estas funções ser lineares ou constantes.

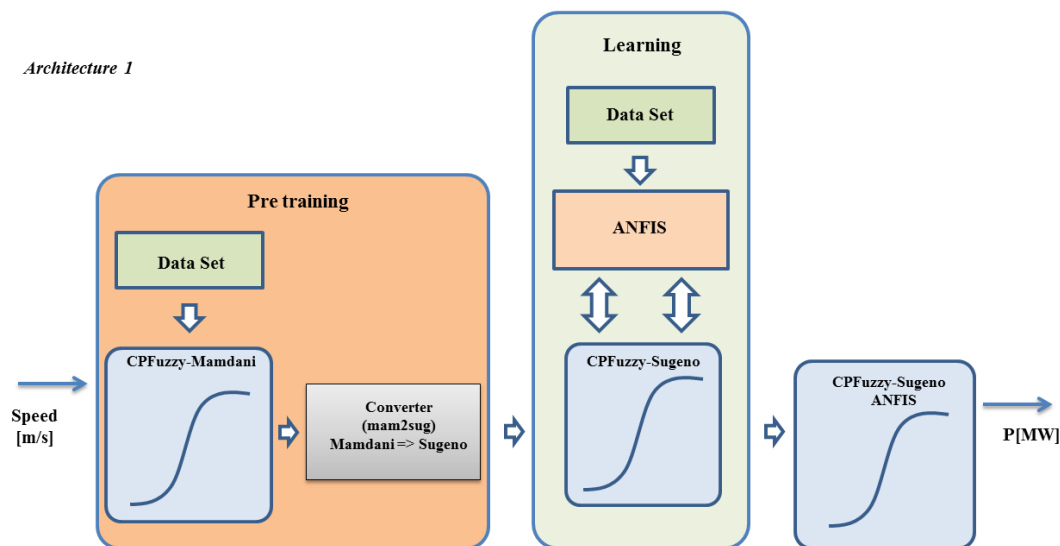
A principal diferença entre os dois modelos *Fuzzy* está no fato de haver um pré treinamento, resultante do uso da transformada *mam2sug*, após esta transformação o nível entrópico dos clusters das entradas tornasse efetivamente menor. Esta organização faz com que a procura pelos novos clusters na fase de aprendizagem seja mais bem definida e assim faz a aproximação da CPFuzzy (Sugeno ANFIS) se tornar mais verossímil em relação ao histórico de dados original.

A seguir é mostrada a rotina para criação das CPs *Fuzzy* (Sugeno-ANFIS):

- A) Criação do SIF(*Mamdani*):
- 1) Inicia-se no SIF o banco de dados de aprendizagem, carregando no mesmo, entradas e saídas.
 - 2) Escolhem-se através de testes sucessivos o número de intervalos das respectivas entradas e saídas.
 - 3) Escolhe-se o tipo de função membro inferente em cada intervalo, através da tendência correlacional entre as variáveis de entrada e saída.
 - 4) Ajustasse manualmente, por tentativas a melhor disposição das funções membro em cada intervalo, até que o valor da saída atinge um nível de aproximação satisfatório.
- B) Criação da Inferência pré-estabelecida (pré-treinamento):
- 5) Utiliza-se a transformada *Mam2Sug*, e converte-se o SIF(*Mamdani*) em um SIF(*Sugeno*).
 - 6) Inicia-se o Banco de dados de aprendizagem no ANFIS.
 - 7) Define-se no ANFIS por tentativas um raio amostral satisfatório.
 - 8) De posse do valor do raio amostral, escolhe-se a função *grid partition*, que fará o ajuste automático da melhor disposição das funções membro por intervalo. Sendo ajustado para isto o mesmo n° de épocas utilizado nos modelos anteriores.
 - 9) Cria-se o SIF(Sugeno-ANFIS).

Em que, a arquitetura do modelo da CP1Fuzzy (Sugeno-ANFIS) é mostrada abaixo pela Figura 98:

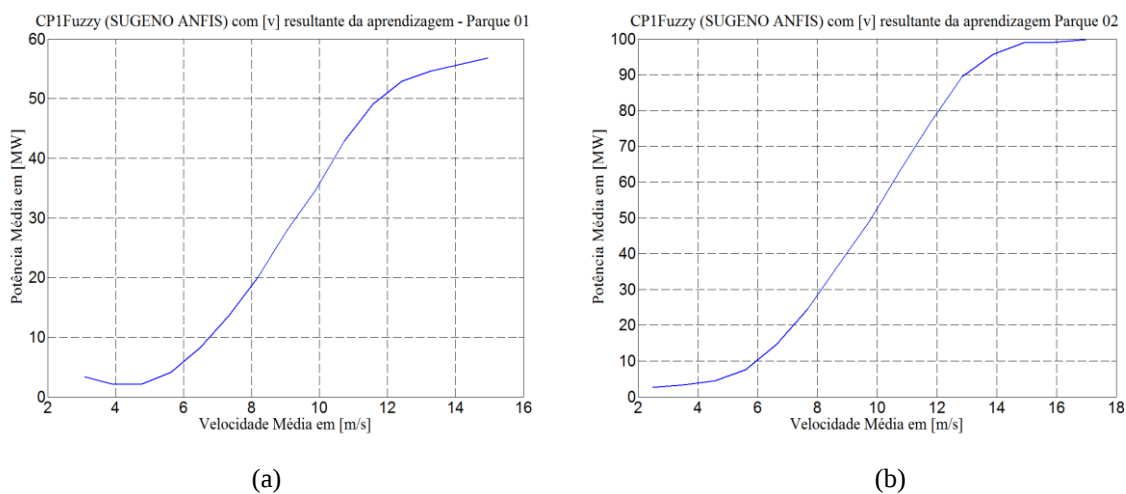
Figura 98 – CP1Fuzzy (Sugeno-ANFIS) resultante do processo de aprendizagem relativa aos Parques 01 e 02.



Fonte: O Autor (2020).

A Figura 99, mostra as curvas de tendência resultantes do processo de aprendizagem, considerando os limites impostos no processo de aprendizagem pelo ANFIS na melhor colocação das funções de pertinência e suas relações.

Figura 99 – CP1Fuzzy (Sugeno ANFIS) resultante do processo de aprendizagem relativa a o (a) Parque 01 e ao (b) Parque 02.



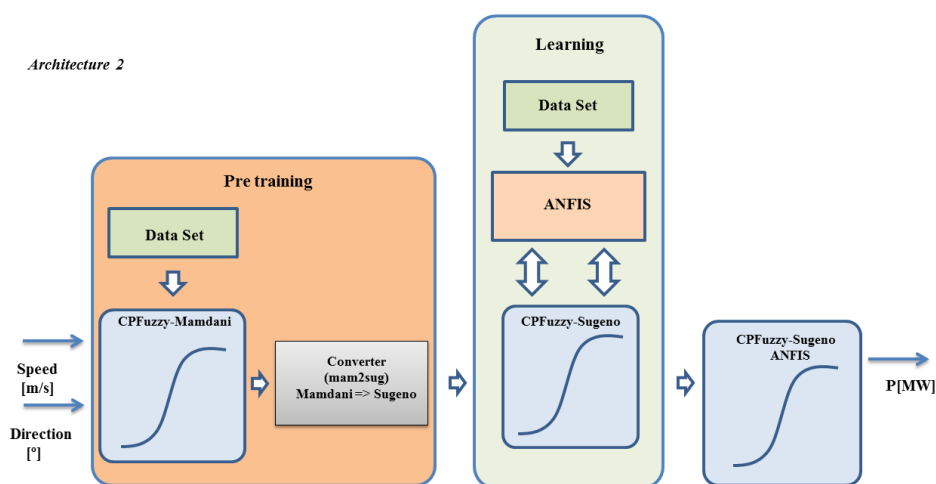
Fonte: O Autor (2020).

6.6.4 MODELO DA CP2FUZZY (SUGENO-ANFIS)

Semelhantemente ao que fora feito com a CP1Fuzzy (Sugeno-ANFIS), o modelo da CP2Fuzzy(Sugeno ANFIS) tem características semelhantes diferindo-se apenas na sua arquitetura, por possuir duas entradas, a velocidade média e a direção média do vento e uma saída a potência média do vento. A heurística e os métodos utilizados para a definição dos parâmetros principais deste bloco aproximador de potência, são semelhantes ao descrito no bloco *Fuzzy Sugeno* anterior.

Onde a arquitetura dos modelos das CP1Fuzzy (Sugeno-ANFIS) é mostrada abaixo pela Figura 100:

Figura 100 – CP2Fuzzy (Sugeno-ANFIS) resultante do processo de aprendizagem relativa a o (a) Parque 01 e ao (b) Parque 02.



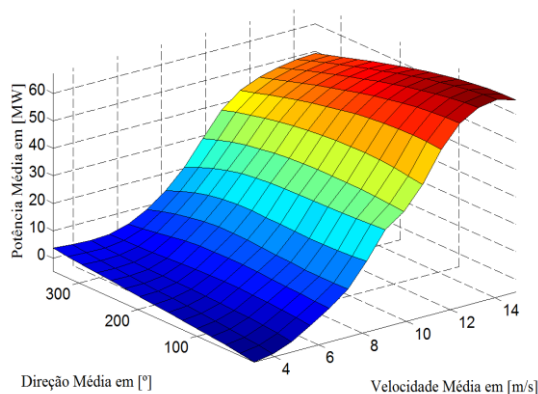
Fonte: O Autor (2020).

Em que a Figura 101, mostra as superfícies resultantes do processo de aprendizagem, considerando os limites impostos no processo de aprendizagem pelo ANFIS na melhor disposição das funções de pertinência e suas relações para ambos os parques.

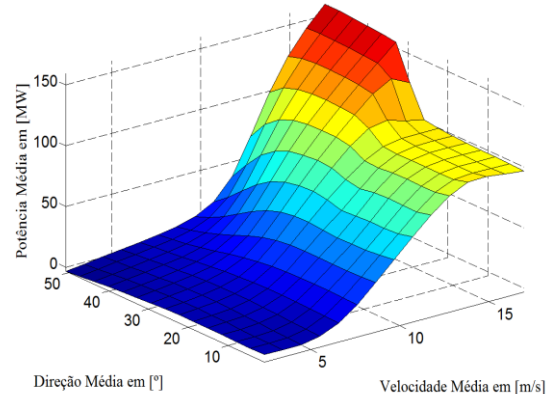
Figura 101 – CP2Fuzzy (Sugeno ANFIS) resultante do processo de aprendizagem relativa a o (a) Parque 01 e ao (b) Parque 02.

Superfície CP2Fuzzy (SUGENO ANFIS) com [v] e [d] resultante da aprendizagem - Parque 01

Superfície CP2Fuzzy (SUGENO ANFIS) com [v] e [d] resultante da aprendizagem Parque 02



(a)



(b)

Fonte: O Autor (2020).

6.7 Modelos Previsores de Potência

Nesta tese os modelos previsores de potência possuíam arquiteturas formadas por dois blocos principais intitulados bloco 1 e bloco 2, ligados em cascata, sendo o bloco 1 responsável pela previsão de variáveis relativas ao vento como a velocidade do vento ou conjunto velocidade de vento e direção do vento em um horizonte definido de 24h. E o bloco 2 ficou responsável por converter os valores previstos pelo bloco 1 em ambas as configurações de variáveis, isto é, tanto a conversão do padrão de velocidade em potência ou a conversão do padrão [velocidade direção] em potências previstas, através dos modelos de curvas de potência definidos anteriormente nesta tese.

Além disso, como a arquitetura dos modelos previsores de potência foram compostas pelos modelos previsores do bloco 1, estas potências previstas terão dependendo do previsor em questão, influência das heurísticas citadas anteriormente.

Assim foram definidos 32 modelos previsores de potência que tiveram a sua arquitetura definida conforme a Tabela 5, e também ilustrado pela Figura 102 de forma genérica, e pela Figura 106 de forma específica, ambas mais a diante nesta seção.

6.7.1 ESCOLHA DOS PREVISORES LSTM UTILIZADOS NA COMPOSIÇÃO DA ARQUITETURA DOS MODELOS PREVISORES DE POTÊNCIA

Nesta Tese foram desenvolvidos um conjunto de quatro modelos de previsão de velocidade de vento e previsão de velocidade e direção do vento, que como são advindos da mesma heurística das LSTM, faz-se necessária uma avaliação entre os mesmos para escolher quais farão parte da composição dos modelos previsores de potência.

Para tal avaliação foram utilizadas nas bases dos respectivos Parques 01 e 02 os conjuntos de simulação descritos em ALBUQUERQUE, J.C. (2015) e em AQUINO R. R. B. et. al. (2016), conforme mostrado a seguir:

Para o estudo de comparação utilizou-se na base de dados o período de simulação compreendido das:

(23h) 01/01/2012 - (9h) 10/11/2012 para o Parque 01

(23h) 01/04/2013 - (0h) 08/04/2014 para o Parque 02

Em que, com o objetivo de comparar o desempenho da previsão de velocidade de vento ou velocidade e direção de vento dos modelos propostos baseados em LSTM, as Tabelas 3 e 4 mostram o desempenho comparativo dos quatro Modelos desenvolvidos. Para isso as tabelas objetivaram mostrar o MAE horário alcançado ao longo do horizonte de previsão, bem como a média deste erro para as doze primeiras horas e para todo o horizonte, isto é, 24h.

Analisando os resultados em relação a escolha dos modelos previsores com LSTM desenvolvidos para os Parques 01 e 02 tem-se as seguintes percepções:

Na Tabela 3, estudando o Parque 01, a melhora nos resultados já é perceptível nas primeiras horas do horizonte de previsão de velocidades, e apenas apresentam resultados não satisfatórios em algumas horas do horizonte, caso das horas 7 a 9, 11 e 20. Porém, mesmo tendo desempenho não satisfatório nestas faixas do horizonte em termos médios o modelo LSTM1_AET1 apresenta médias do MAE de direção do vento nas 12 primeiras horas como para todo o horizonte menores do que o modelo LSTM1.

Desta forma a média do MAE tanto nas doze primeiras horas como para todo o horizonte é menor, para o modelo LSTM1_AET1 (Média MAE 12h=1,3106 m/s; Média MAE 24h =1,3513 m/s), em comparação ao modelo LSTM1 (Média MAE 12h=1,3185 m/s; Média MAE 24h =1,3768 m/s).

Ainda na Tabela 3, só que agora analisando o desempenho dos modelos com duas entradas, velocidades e direções, acontece algo semelhante ao caso da previsão de velocidades nos modelos com uma entrada, sendo a melhora nos resultados já perceptível nas primeiras horas, e apenas apresentando resultados não satisfatórios em algumas horas do horizonte caso das horas 8 a 13 e 15 a 20, considerando a previsão de velocidades, e as horas 5 a 7, 15 e 17 a 24 considerando a previsão de direções. Porém, mesmo tendo desempenho não satisfatório nestas faixas do horizonte em termos médios o modelo LSTM2_AET2 apresenta médias do MAE de velocidade do vento e direção do vento nas 12 primeiras horas como para todo o horizonte menores do que as médias de MAE apresentadas pelo modelo LSTM2.

Desta forma a média do MAE tanto nas doze primeiras horas como para todo o horizonte é menor, para o modelo LSTM2_AET2 (Média MAE 12h = [1,2674 m/s], [11,8427 °]; Média MAE 24h = [1,3224 m/s], [13,1565 °]), em comparação ao modelo LSTM2 (Média MAE 12h = [1,2816 m/s], [12,6968 °]; Média MAE 24h = [1,3224 m/s], [13,1680 °]).

Na Tabela 4, estudando o Parque 02, a melhora nos resultados já é perceptível nas primeiras horas do horizonte de previsão de velocidades, e apenas apresentam resultados não satisfatórios em algumas horas do horizonte, caso das horas 5, 7, 8 e 10 a 19. Porém mesmo tendo desempenho não satisfatório nestas faixas do horizonte em termos médios o modelo LSTM1_AET1 apresenta médias do MAE de direção do vento nas 12 primeiras horas como para todo o horizonte menores do que o modelo LSTM1.

Desta forma a média do MAE tanto nas doze primeiras horas como para todo o horizonte é menor, para o modelo LSTM1_AET1 (Média MAE 12h=1,2350 m/s; Média MAE 24h =1,2616 m/s), em comparação ao modelo LSTM1 (Média MAE 12h=1,2667 m/s; Média MAE 24h =1,2734 m/s).

Ainda na Tabela 4, analisando o desempenho dos modelos com duas entradas, velocidades e direções, verificando a previsão de velocidades de vento, acontece algo semelhante ao caso da previsão de velocidades nos modelos com uma entrada, sendo a melhora nos resultados já perceptível nas primeiras horas, e apenas apresentando resultados não satisfatórios em algumas horas do horizonte, caso das horas 8 a 13 e 15 a 20.

Para o caso da previsão de direções ocorre um desempenho insatisfatório nas primeiras horas do horizonte caso das horas 1, 2, 10, 13 a 16, 21, 23 e 24. Porém, mesmo tendo desempenho não satisfatório nestas faixas do horizonte, em termos médios, o modelo LSTM2_AET2 apresenta médias do MAE da previsão de velocidades do vento e direção do vento nas 12 primeiras horas, como para todo o horizonte, menores do que as médias de MAE apresentadas pelo modelo LSTM2

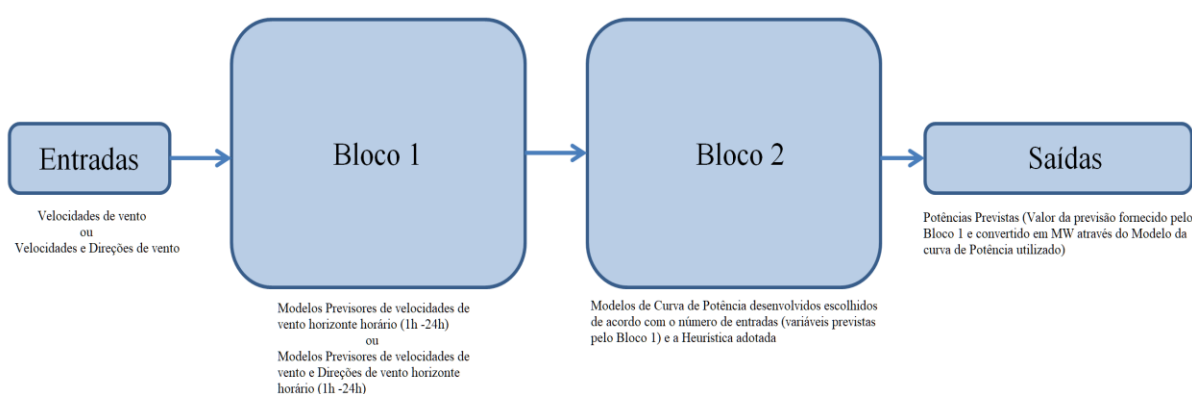
Desta maneira, a média do MAE tanto nas doze primeiras horas, como para todo o horizonte, é menor para o modelo LSTM2_AET2 (Média MAE 12h = [1,2617 m/s], [3,9356 °]; Média MAE 24h = [1,2705 m/s], [3,9847 °]), em comparação ao modelo LSTM2 (Média MAE 12h = [1,2710 m/s], [4,0092 °]; Média MAE 24h = [1,2718 m/s], [4,0037 °]).

Portanto, fica evidente que as topologias AET1 e AET2 causam impacto efetivo de melhora no desempenho dos modelos previsores de velocidade e direção. Este impacto de melhora pode ser notado seja no início do horizonte, quando de fato o previsor já se comporta bem nas primeiras horas, seja quando o mesmo se recupera não tendo um desempenho satisfatório nas primeiras horas mas, impedindo que o erro aumente ao longo do horizonte.

Esta característica em específico é um ponto levantado pelos criadores das LSTM, Hochreiter e Schmidhuber (1997), enfatizando que dentre os objetivos principais do modelo LSTM estão: a aprendizagem da dependência em longo prazo e a manutenção de informação por longos períodos.

Assim, tendo por base a análise comparativa demonstrada anteriormente, foram escolhidos os modelos LSTM1 AET1 e LSTM2 AET2 para compor a arquitetura dos modelos previsores de potência.

Figura 102 – Arquitetura Genérica de composição dos Modelos Previsores de Potência definida na Tabela 5 adotada para os Parques 01 e 02.



Fonte: O Autor (2021).

Tabela 3 – Desempenho das previsões no período de Simulação das redes (LSTM1, LSTM2, LSTM1AET1, LSTM2AET2) para o Parque 01.

MAE em [m/s]													
Horizonte	1h	2h	3h	4h	5h	6h	7h	8h	9h	10h	11h	12h	Média MAE 12h
LSTM1.	0,8056	1,0305	1,1624	1,2167	1,2905	1,3589	1,4314	1,5261	1,4867	1,5769	1,4607	1,4754	1,3185
LSTM1AET1	0,7682	0,9986	1,1252	1,1600	1,2769	1,3342	1,4543	1,5494	1,4917	1,5619	1,5474	1,4598	1,3106
Horizonte	13h	14h	15h	16h	17h	18h	19h	20h	21h	22h	23h	24h	Média MAE 24h
LSTM1.	1,4986	1,465	1,4554	1,4181	1,4041	1,4008	1,3982	1,3903	1,3996	1,4672	1,447	1,4764	1,3768
LSTM1AET1	1,4467	1,3708	1,3846	1,3932	1,3599	1,3460	1,3410	1,3909	1,3649	1,4482	1,4344	1,4218	1,3509
MAE em [m/s]													
Horizonte	1h	2h	3h	4h	5h	6h	7h	8h	9h	10h	11h	12h	Média MAE 12h
LSTM2.	0,8959	1,1582	1,2823	1,2981	1,2911	1,3184	1,3494	1,3556	1,3408	1,3354	1,4215	1,3320	1,2816
LSTM2AET2	0,8866	1,0282	1,1240	1,1896	1,2418	1,2888	1,3461	1,3915	1,4183	1,4275	1,4462	1,4202	1,2674
Horizonte	13h	14h	15h	16h	17h	18h	19h	20h	21h	22h	23h	24h	Média MAE 24h
LSTM2.	1,3527	1,4300	1,3578	1,3456	1,3411	1,3699	1,3647	1,3437	1,3645	1,3692	1,3869	1,3640	1,3237
LSTM2AET2	1,4114	1,4075	1,4187	1,3975	1,3729	1,3783	1,3786	1,3709	1,3609	1,3370	1,3425	1,3524	1,3224
MAE em [°]													
Horizonte	1h	2h	3h	4h	5h	6h	7h	8h	9h	10h	11h	12h	Média MAE 12h
LSTM2.	10,5052	10,3737	10,5562	10,9810	10,5303	11,2811	12,7914	15,5363	15,9224	15,1000	15,9558	12,8287	12,6968
LSTM2AET2	9,9522	10,0915	10,2203	10,4964	11,2248	12,2024	12,8336	13,2366	13,1129	13,0286	13,0935	12,6193	11,8427
Horizonte	13h	14h	15h	16h	17h	18h	19h	20h	21h	22h	23h	24h	Média MAE 24h
LSTM2.	13,5938	15,6360	12,9524	15,7847	13,2871	13,2759	12,9442	14,6895	13,1991	11,9083	13,4851	12,9133	13,1680
LSTM1AET2	12,4700	12,8511	13,1049	14,4470	15,0980	15,6515	15,6499	15,3671	15,0314	15,3302	14,8011	13,8418	13,1565

Fonte: O Autor (2021).

Tabela 4 – Desempenho das previsões no período de Simulação das redes (LSTM1, LSTM2, LSTM1AET1, LSTM2AET2) para o Parque 02.

MAE em [m/s]													
Horizonte	1h	2h	3h	4h	5h	6h	7h	8h	9h	10h	11h	12h	Média MAE 12h
LSTM1.	1,0503	1,1663	1,2472	1,3057	1,3323	1,3221	1,3078	1,3012	1,2908	1,2855	1,2953	1,2959	1,2667
LSTM1AET1	0,8281	1,0493	1,2083	1,2567	1,3329	1,3123	1,3335	1,3102	1,2731	1,2991	1,3134	1,3032	1,2350
Horizonte	13h	14h	15h	16h	17h	18h	19h	20h	21h	22h	23h	24h	Média MAE 24h
LSTM1.	1,3033	1,3171	1,3193	1,2859	1,2526	1,2381	1,2464	1,2698	1,2743	1,2665	1,2787	1,3102	1,2734
LSTM1AET1	1,3478	1,3391	1,3433	1,2915	1,2819	1,2498	1,2806	1,2678	1,2645	1,2506	1,2753	1,2669	1,2616
MAE em [m/s]													
Horizonte	1h	2h	3h	4h	5h	6h	7h	8h	9h	10h	11h	12h	Média MAE 12h
LSTM2.	0,8982	1,1023	1,2613	1,3453	1,3323	1,3244	1,3320	1,3585	1,3279	1,3426	1,3231	1,3036	1,2710
LSTM2AET2	0,8900	1,1158	1,2154	1,2766	1,3322	1,3340	1,3398	1,3226	1,3320	1,3311	1,3409	1,3104	1,2617
Horizonte	13h	14h	15h	16h	17h	18h	19h	20h	21h	22h	23h	24h	Média MAE 24h
LSTM2.	1,2884	1,3072	1,2909	1,2729	1,2880	1,2482	1,2466	1,2370	1,2811	1,2758	1,2807	1,2552	1,2718
LSTM2AET2	1,3090	1,3030	1,3091	1,3066	1,2771	1,2822	1,2892	1,2338	1,2389	1,2607	1,2640	1,2773	1,2705
MAE em [°]													
Horizonte	1h	2h	3h	4h	5h	6h	7h	8h	9h	10h	11h	12h	Média MAE 12h
LSTM2.	4,0391	3,9784	3,7865	4,1610	3,9044	3,8510	4,2669	4,4432	4,0622	4,0320	3,6905	3,8956	4,0092
LSTM2AET2	4,3394	4,4274	3,6909	3,9162	3,8711	3,7024	3,9560	3,9532	3,9718	4,1035	3,5011	3,7948	3,9356
Horizonte	13h	14h	15h	16h	17h	18h	19h	20h	21h	22h	23h	24h	Média MAE 24h
LSTM2.	3,7218	3,7997	3,9288	4,1121	4,2538	4,0066	3,9782	3,8718	4,0561	4,2376	4,0464	3,9652	4,0037
LSTM2AET2	3,8169	4,5694	3,9973	4,2465	4,0718	3,7049	3,9931	3,5838	4,1452	4,1621	4,0732	4,0409	3,9847

Fonte: O Autor (2021).

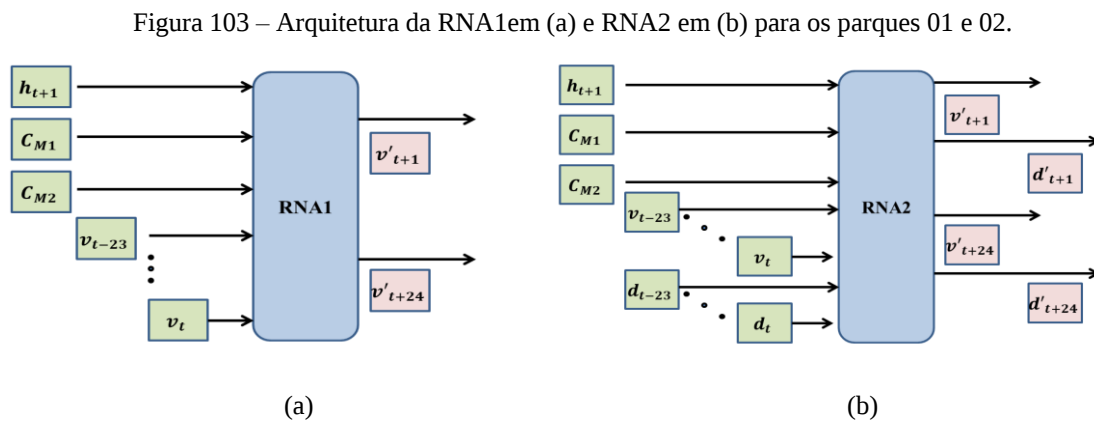
Tabela 5 – Composição de formação dos Modelos previsores de potência desenvolvidos.

Cód. Modelo Prev. Pot. (MW)	Prev. (v) ou Prev. (v, dir.) utilizado	Modelo de CP utilizado
M22	RNA2 AET2	CP2-RNA
M23	RNA2 AET2	CP2-RNA AET2
M24	RNA2 AET2	CP2-Fuzzy-Mandani
M25	RNA2 AET2	CP2-Fuzzy-Sugeno ANFIS
M2A	RNA2	CP2-RNA
M21	RNA2	CP2-RNA AET2
M2B	RNA2	CP2-Fuzzy-Mandani
M2C	RNA2	CP2-Fuzzy-Sugeno ANFIS
M32	RNA1 AET2	CP1-RNA
M33	RNA1 AET2	CP1-RNA AET1
M34	RNA1 AET2	CP1-RNA AET2
M35	RNA1 AET2	CP1-Fuzzy-Mandani
M36	RNA1 AET2	CP1-Fuzzy-Sugeno ANFIS
M1A	RNA1	CP1-RNA
M11	RNA1	CP1-RNA AET1
M20	RNA1	CP1-RNA AET2
M1B	RNA1	CP1-Fuzzy-Mandani
M1C	RNA1	CP1-Fuzzy-Sugeno ANFIS
M42	RNA1 AET1	CP1-RNA
M43	RNA1 AET1	CP1-RNA AET1
M44	RNA1 AET1	CP1-RNA AET2
M45	RNA1 AET1	CP1-Fuzzy-Mandani
M46	RNA1 AET1	CP1-Fuzzy-Sugeno ANFIS
M52	LSTM1 AET1	CP1-RNA
M53	LSTM1 AET1	CP1-RNA AET1
M54	LSTM1 AET1	CP1-RNA AET2
M55	LSTM1 AET1	CP1-Fuzzy-Mandani
M56	LSTM1 AET1	CP1-Fuzzy-Sugeno ANFIS
M62	LSTM2 AET2	CP2-RNA
M64	LSTM2 AET2	CP2-RNA AET2
M65	LSTM2 AET2	CP2-Fuzzy-Mandani
M66	LSTM2 AET2	CP2-Fuzzy-Sugeno ANFIS

Fonte: O Autor (2021).

Para entregar valores previstos para as curvas de potência aproximadas dependendo da arquitetura, o modelo predictor terá sua faixa de análise nas 24h seguintes. Assim poderá dependendo de sua configuração utilizar, seis modelos de redes neurais predictoras de acordo com as arquiteturas citadas anteriormente nesta tese, os parâmetros dessas redes são descritos em ALBUQUERQUE, J.C. (2015) e em AQUINO R. R. B. et. al. (2016) e também ao longo de suas definições neste trabalho, as mesmas são nomeadas RNA1 e RNA2, RNA1 AET1, RNA1 AET2 e RNA2 AET2, LSTM1 AET1 e LSTM2 AET2.

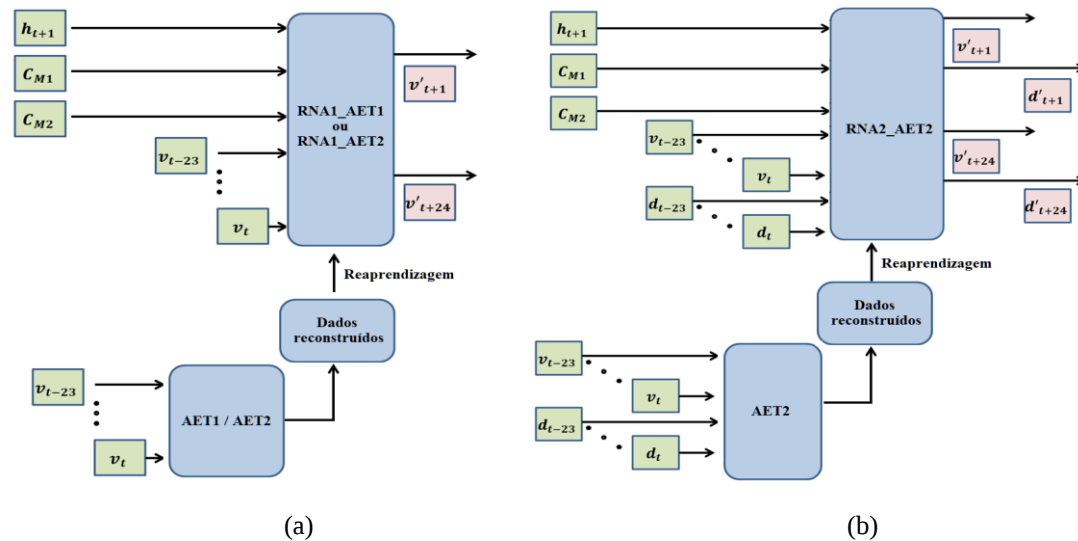
A Figura 103 mostra as arquiteturas das redes RNA1 em (a) e RNA2 em (b).



Fonte: O Autor (2021).

A Figura 104 mostra as arquiteturas das redes RNA1 AET1 e RNA1 AET2 em (a) e RNA2 AET2 em (b).

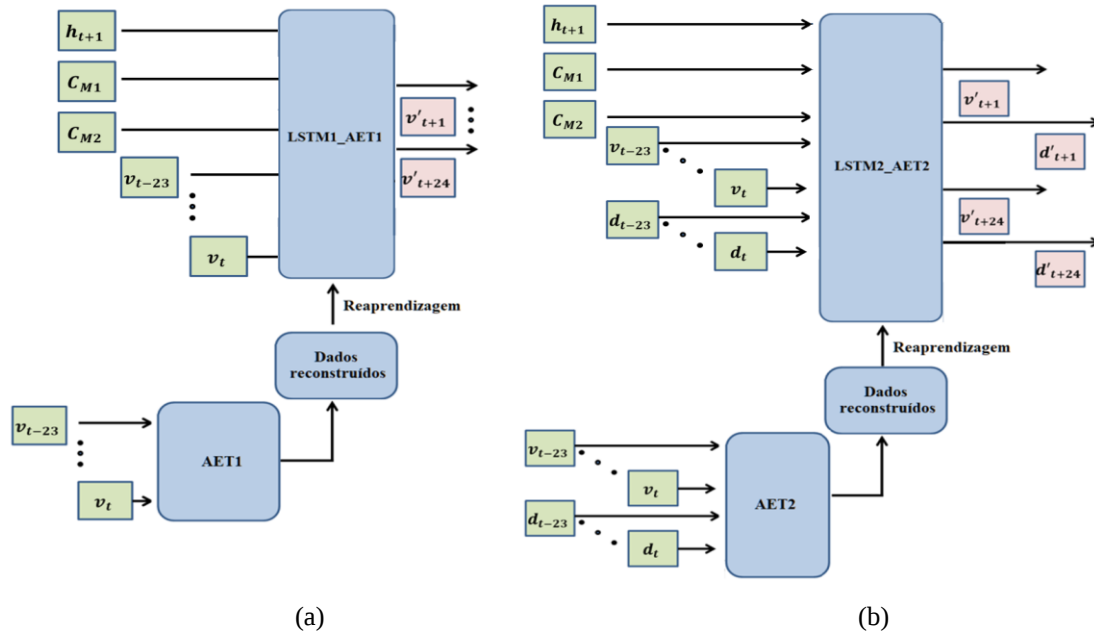
Figura 104 – Arquitetura da RNA1 AET1 ou RNA AET2 em (a) e RNA2 AET2 em (b) para os parques 01 e 02.



Fonte: O Autor (2021).

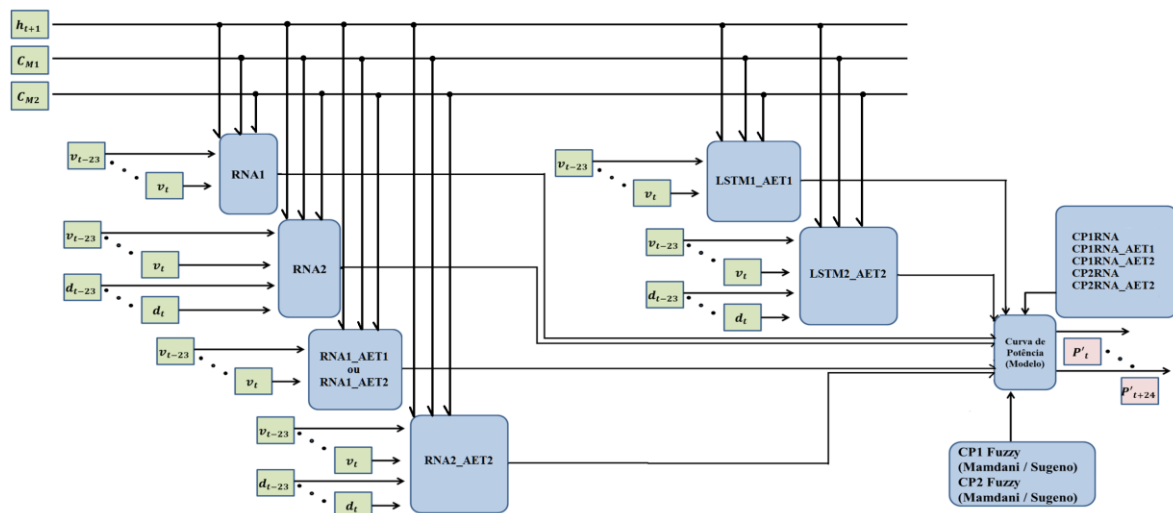
A Figura 105 mostra as arquiteturas das redes LSTM1 AET1 em (a) LSTM2 AET2 em (b), e a Figura 106, mostra o fluxograma genérico de ação na arquitetura do modelo Previsor geral de potências adotado.

Figura 105 – Arquitetura da LSTM1 AET1 (a) e LSTM2 AET2 em (b) para os parques 01 e 02.



Fonte: O Autor (2021).

Figura 106 – Alimentação de variáveis no fluxograma genérico de ação na arquitetura do modelo do Previsor geral de potências adotado para os parques 01 e 02.



Fonte: O Autor (2021).

7 ANÁLISE GERAL E RESULTADOS

Neste capítulo são mostrados os resultados e comparações obtidos na implementação dos modelos e novas heurísticas propostas por esta tese. As análises são realizadas em dois momentos em que no primeiro são analisados e comparados os resultados referentes aos modelos previsores relativos aos dados de vento, isto é, velocidade e direção. Já no segundo momento são analisados os resultados dos modelos de curva de potência propostos e também os respectivos modelos previsores de Potência desenvolvidos.

7.1 Discussão dos resultados

Para avaliação dos modelos previsores e curvas de potência desenvolvidos neste trabalho, foram utilizados como referência na aproximação dos mesmos, os índices de MAE, NMAE e RMSE, sendo estes levantados após sucessivas comparações entre os valores aproximados pelos modelos desenvolvidos e seus respectivos valores reais de acordo com o banco de dados do parque sobre estudo.

De forma a mostrar os índices avaliadores das aproximações dos modelos desta tese, as Equações (7.1), (7.2), (7.3) e (7.4) discriminam o cálculo utilizado dos Erros de aproximação, sendo os mesmos o MAE, o NMAE e o RMSE:

$$e_p(k) = P(k) - \bar{P}(k), \quad (7.1)$$

$$MAE_p(k) = média(|e_p(k)|), \quad (7.2)$$

$$NMAE_p(k) = \frac{1}{P_{(Máx-Instal)}} média(|e_p(k)|) \quad (7.3)$$

$$RMSE_p(k) = \sqrt{média|e_p^2(k)|} \quad (7.4)$$

Para mostrar avaliar o desempenho dos modelos avaliados de acordo com os índices respectivos já citados, a Equação (7.5) mostra o cálculo do ganho do modelo respectivo da curva de potência quando o mesmo é comparado com o modelo de referência, neste o modelo paramétrico polinomial da curva média.

$$G_{ref,CA}(k) = 100 \frac{CA_{ref} - CA(k)}{CA_{ref}(k)} \quad (7.5)$$

Onde, CA_{ref} é o critério de avaliação do modelo de referência e $CA(k)$ é o critério de avaliação do modelo proposto. De forma que o MAE, o NMAE e o RMSE podem ser usados como critérios de avaliação.

Como já enfatizado estudo de comparação aconteceu utilizando a base de dados no período de simulação compreendido das:

(23h) 01/01/2012 – (9h) 10/11/2012 para o parque 01

(23h) 01/04/2013 - (0h) 08/04/2014 para o parque 02

7.1.1 DISCUSSÃO DOS RESULTADOS RELATIVOS AOS PREVISORES DE DADOS DE VENTO

Conforme arguido anteriormente nesta tese, foram desenvolvidos mais cinco novos modelos previsores de velocidade do vento e velocidade e direção do vento, que em conjunto com os três modelos previsores de dados de vento definidos em ALBUQUERQUE (2015), complementam o banco de modelos para a análise em questão, totalizando oito modelos.

No caso dos modelos definidos em ALBUQUERQUE (2015), tem-se: a RNA1, RNA2 e o Modelo da Persistência. No caso dos modelos novos da tese tem-se: RNA1 AET1, RNA1 AET2, RNA2 AET2, LSTM1 AET1 e a LSTM2 AET2.

Para a melhora do entendimento e posterior análise de comparação foi criada a Tabela 6 que visa a simplificação por códigos relativos aos modelos previsores de vento, cuja letra escolhida para codificar tais modelos foi a letra (W) referência a palavra vento (*Wind*). Em que a mesma é apresentada a seguir:

Tabela 6 – Codificação dos Modelos previsores de dados de vento desenvolvidos.

Modelo	Cod.	Modelo	Cod.
RNA1	(W1)	RNA2 AET2	(W5)
RNA2	(W2)	LSTM1 AET1	(W6)
RNA1 AET1	(W3)	LSTM2 AET2	(W7)
RNA1 AET2	(W4)	Persistência	(W8)

Fonte: O Autor (2021).

De forma a demonstrar a avaliação comparativa entre os desempenhos dos já citados modelos previsores de dados de vento as Tabelas 7 a 10, ilustram os índices de MAE relativos a previsão dentro do horizonte de 24h para análise descrita ao período de simulação. Em seguida, as Figuras 107 a 110 mostram o desempenho do índice de MAE no mesmo horizonte comparativo destes modelos, para os parques sobre estudo.

Mais adiante as Tabelas 11 a 14 mostram o desempenho médio além dos ganhos percentuais envolvidos.

Tabela 7 – Desempenho no período de Simulação dos modelos previsores de dados de vento para as 12h iniciais do horizonte para o Parque 01.

MAE em [m/s] dos modelos previsores de dados de vento desenvolvidos para as 12h iniciais do horizonte para o Parque 01													
Horizonte	1h	2h	3h	4h	5h	6h	7h	8h	9h	10h	11h	12h	Média 12h
W1	0,7569	1,0016	1,1732	1,1923	1,2734	1,2997	1,3693	1,3879	1,4359	1,4161	1,3924	1,3879	1,2572
W2	0,7984	1,0449	1,1575	1,2225	1,2430	1,2939	1,3392	1,3296	1,3163	1,3514	1,3704	1,3783	1,2371
W3	0,7761	0,9905	1,1492	1,2433	1,2317	1,3104	1,4028	1,3748	1,4120	1,3915	1,4268	1,4301	1,2616
W4	0,8028	1,0171	1,2138	1,2378	1,2976	1,2965	1,4268	1,3835	1,3923	1,4801	1,4162	1,3931	1,2798
W5	0,8945	1,0373	1,1797	1,3016	1,4020	1,3930	1,5308	1,4336	1,5603	1,6863	1,4943	1,4387	1,3627
W6	0,7682	0,9986	1,1252	1,1600	1,2769	1,3342	1,4543	1,5494	1,4917	1,5619	1,5474	1,4598	1,3106
W7	0,8866	1,0282	1,1240	1,1896	1,2418	1,2888	1,3461	1,3915	1,4183	1,4275	1,4462	1,4202	1,2674
W8	0,7144	1,1208	1,4139	1,6524	1,8581	2,0438	2,2078	2,3388	2,427	2,4847	2,524	2,5388	1,9437
MAE em [°] dos modelos previsores de dados de vento desenvolvidos para as 12h iniciais do horizonte Parque 01													
Horizonte	1h	2h	3h	4h	5h	6h	7h	8h	9h	10h	11h	12h	Média 12h
W2	11,9779	13,3570	16,1980	15,2005	15,0883	12,5283	12,7632	12,9194	12,8972	13,8040	15,0162	14,1159	13,8222
W5	10,5954	12,5538	12,2933	14,3326	13,6574	14,4621	13,6690	15,1035	12,6236	14,5567	14,6625	15,2014	13,6426
W7	9,9522	10,0915	10,2203	10,4964	11,2248	12,2024	12,8336	13,2366	13,1129	13,0286	13,0935	12,6193	11,8427
W8	6,9712	12,4501	17,3000	21,6489	25,5776	29,0097	31,8689	34,2129	36,0089	37,3164	38,1023	38,3684	27,4029

Fonte: O Autor (2021).

Tabela 8 – Desempenho no período de Simulação dos modelos previsores de dados de vento para as 12h finais do horizonte para o Parque 01.

MAE em [m/s] dos modelos previsores de dados de vento desenvolvidos para as 12h finais do horizonte para o Parque 01													
Horizonte	13h	14h	15h	16h	17h	18h	19h	20h	21h	22h	23h	24h	Média 24h
W1	1,4165	1,4284	1,4249	1,4415	1,4012	1,4073	1,3685	1,3819	1,4154	1,3828	1,3988	1,4094	1,3318
W2	1,3618	1,3804	1,3878	1,3811	1,4293	1,4575	1,3504	1,3765	1,3727	1,4127	1,4710	1,4329	1,3191
W3	1,4038	1,4172	1,4220	1,4459	1,4490	1,4715	1,4631	1,4910	1,5085	1,5241	1,5280	1,5939	1,3691
W4	1,4354	1,3910	1,4672	1,4121	1,4377	1,4942	1,4259	1,4254	1,4020	1,4886	1,4321	1,3941	1,3568
W5	1,4318	1,4909	1,4880	1,3928	1,4762	1,4027	1,6116	1,4773	1,5054	1,5146	1,4577	1,3418	1,4143
W6	1,4467	1,3708	1,3846	1,3932	1,3599	1,3460	1,3410	1,3909	1,3649	1,4482	1,4344	1,4218	1,3513
W7	1,3090	1,3030	1,3091	1,3066	1,2771	1,2822	1,2892	1,2338	1,2389	1,2607	1,2640	1,2773	1,2733
W8	2,5221	2,4833	2,4285	2,3428	2,2415	2,1238	2,011	1,8993	1,7779	1,6694	1,5868	1,5603	1,9988
MAE em [°] dos modelos previsores de dados de vento desenvolvidos para as 12h finais do horizonte Parque 01													
Horizonte	13h	14h	15h	16h	17h	18h	19h	20h	21h	22h	23h	24h	Média 24h
W2	12,9085	13,7793	13,1237	14,3205	13,8271	14,7802	13,4886	14,7917	14,7216	13,7403	13,6298	13,3394	13,8465
W5	13,3491	13,7589	16,4577	18,1359	13,7922	14,7810	16,1287	15,5057	17,1599	15,1351	15,3437	14,9404	14,5083
W7	12,4700	12,8511	13,1049	14,4470	15,0980	15,6515	15,6499	15,3671	15,0314	15,3302	14,8011	13,8418	13,1565
W8	38,1690	37,4946	36,3086	34,5569	32,3162	29,5976	26,4669	22,9493	19,1897	15,4920	12,4705	11,2108	26,8774

Fonte: O Autor (2021).

Analisando os dados de MAE presentes nas Tabela 7 e 8, inerentes ao parque 01, podem-se notar em duas perspectivas médias, uma inicial e outra final, o comportamento relativo ao erro de previsão de dados de vento. Os dados mostram que para as 12h iniciais do horizonte o modelo W2 apresenta a menor média de erro **1,2371 m/s**. Para o horizonte completo o destaque em desempenho é do modelo W7 com MAE de **1,2733 m/s**.

Uma segunda análise já comentada a respeito em ALBUQUERQUE (2015), é sobre o erro de previsão em relação a primeira hora, em que para o Parque 01 devido a uma característica de distribuição muito homogênea quanto aos valores da base histórica de dados de velocidades, o modelo da Persistência W8 se destaca com o menor erro, **0,7144 m/s** contra a média **0,7569 m/s** apresentado pelo modelo W1. Porém este erro se modifica drasticamente nas horas posteriores, tendo o modelo W8 um péssimo desempenho médio ao longo do horizonte.

Desta forma, neste parque especificamente utilizou-se como comparação a segunda hora do horizonte, nesta, o modelo que se destaca é o modelo W3 com **0,9905 m/s** de MAE.

Em relação a previsão de direções, os dados mostram que para as 12h iniciais do horizonte, o modelo W7 apresenta a menor média de erro **11,8427 °**. Para o horizonte completo o destaque em desempenho também é do modelo W7 com MAE de **13,1565°**. Este fato atesta que o modelo W7 destaca-se em relação ao parque 01 como um modelo que se comporta bem tanto na previsão de velocidades e também no que diz respeito as direções.

Portanto após as análises relativas aos desempenhos apresentados nas Tabelas 7 e 8 fica comprovado que as novas heurísticas propostas apresentaram um desempenho satisfatório em relação a previsão de dados de vento. Especificamente a heurística AE, caso do modelo W3 (RNA1 AET1) e a heurística LSTM, caso do modelo W7 (LSTM2 AET2).

Outra percepção importante é que para o parque 01 assim como atestado pelos autores da heurística LSTM, os modelos que utilizaram estas heurísticas se destacaram com relação a resiliência de desempenho ao longo do horizonte sendo assim uma boa solução quanto a precisão da previsão quando se avalia o médio e longo prazo do horizonte. Caso exemplificado pelo modelo W7 tanto em questão de velocidades quanto em questão de direções.

Tabela 9 – Desempenho no período de Simulação dos modelos previsores de dados de vento para as 12h iniciais do horizonte para o Parque 02.

MAE em [m/s] dos modelos previsores de dados de vento desenvolvidos para as 12h iniciais do horizonte para o Parque 02													
Horizonte	1h	2h	3h	4h	5h	6h	7h	8h	9h	10h	11h	12h	Média 12h
W1	0,8086	1,0484	1,1523	1,2066	1,2260	1,2261	1,2245	1,2566	1,2528	1,2287	1,2309	1,2452	1,1756
W2	0,8324	1,0523	1,1334	1,1753	1,2022	1,2313	1,2458	1,2486	1,2341	1,2231	1,2624	1,2899	1,1776
W3	0,7912	0,9866	1,1020	1,1651	1,2026	1,2050	1,2135	1,2363	1,2170	1,2354	1,2166	1,2003	1,1476
W4	0,7971	1,0082	1,0722	1,1454	1,1806	1,2184	1,2353	1,2583	1,2725	1,2432	1,2332	1,2617	1,1605
W5	0,8479	1,0228	1,0881	1,1819	1,2296	1,2941	1,2715	1,2674	1,2963	1,2906	1,2502	1,2507	1,1909
W6	0,8281	1,0493	1,2083	1,2567	1,3329	1,3123	1,3335	1,3102	1,2731	1,2991	1,3134	1,3032	1,2350
W7	0,8900	1,1158	1,2154	1,2766	1,3322	1,3340	1,3398	1,3226	1,3320	1,3311	1,3409	1,3104	1,2617
W8	0,9230	1,5040	1,982	2,4164	2,7999	3,1590	3,4744	3,7519	3,9667	4,1064	4,169	4,1831	3,0363
MAE em [°] dos modelos previsores de dados de vento desenvolvidos para as 12h iniciais do horizonte Parque 02													
Horizonte	1h	2h	3h	4h	5h	6h	7h	8h	9h	10h	11h	12h	Média 12h
W2	4,1197	4,2371	4,1241	4,0822	4,3823	4,0155	4,4101	4,2540	4,1887	3,9888	4,0878	3,8926	3,9424
W5	4,7422	4,3338	4,2746	4,6507	4,5084	4,3951	4,5730	4,6628	4,5769	4,5066	4,7169	4,7403	4,5568
W7	4,3394	4,4274	3,6909	3,9162	3,8711	3,7024	3,9560	3,9532	3,9718	4,1035	3,5011	3,7948	3,9356
W8	3,1162	3,3207	3,4286	3,5069	3,5396	3,5836	3,5874	3,6858	3,6596	3,8394	3,7673	3,1162	2,9410

Fonte: O Autor (2021).

Tabela 10 – Desempenho no período de Simulação dos modelos previsores de dados de vento para as 12h finais do horizonte para o Parque 02.

MAE em [m/s] dos modelos previsores de dados de vento desenvolvidos para as 12h finais do horizonte para o Parque 02													
Horizonte	13h	14h	15h	16h	17h	18h	19h	20h	21h	22h	23h	24h	Média 24h
W1	1,2310	1,2350	1,2569	1,2234	1,1986	1,2117	1,2155	1,2301	1,2204	1,2300	1,2480	1,2710	1,2033
W2	1,2566	1,2738	1,2834	1,2413	1,2361	1,2295	1,2347	1,2445	1,2766	1,2750	1,2772	1,2732	1,2180
W3	1,1992	1,2198	1,2495	1,2390	1,2368	1,2168	1,2063	1,2373	1,2521	1,2439	1,2474	1,2247	1,1894
W4	1,2523	1,2706	1,2675	1,2255	1,1974	1,2130	1,2148	1,2361	1,2654	1,2951	1,2698	1,2528	1,2036
W5	1,2432	1,2764	1,2928	1,2972	1,2974	1,2742	1,2828	1,2924	1,2841	1,3062	1,2533	1,2594	1,2354
W6	1,3478	1,3391	1,3433	1,2915	1,2819	1,2498	1,2806	1,2678	1,2645	1,2506	1,2753	1,2669	1,2616
W7	1,3090	1,3030	1,3091	1,3066	1,2771	1,2822	1,2892	1,2338	1,2389	1,2607	1,2640	1,2773	1,2705
W8	4,1793	4,1143	4,0006	3,7883	3,5229	3,2147	2,8601	2,5056	2,1898	1,8867	1,6099	1,4293	2,9891
MAE em [°] dos modelos previsores de dados de vento desenvolvidos para as 12h finais do horizonte Parque 02													
Horizonte	13h	14h	15h	16h	17h	18h	19h	20h	21h	22h	23h	24h	Média 24h
W2	3,8407	3,7776	3,8481	3,9704	3,9758	3,9081	3,9618	3,9606	4,0502	3,9988	3,9886	4,0277	4,0455
W5	4,5599	4,4862	5,0462	4,5876	4,6177	4,6305	4,7765	4,4359	4,3048	4,4681	4,8629	4,6555	4,5880
W7	3,8169	4,5694	3,9973	4,2465	4,0718	3,7049	3,9931	3,5838	4,1452	4,1621	4,0732	4,0409	3,9847
W8	3,8152	3,8406	3,8247	3,7495	3,7991	3,7730	3,7218	3,7250	3,5689	3,5033	3,5314	3,454	3,3166

Fonte: O Autor (2021).

Analisando os dados de MAE presentes nas Tabela 9 e 10, inerentes ao parque 02, podem-se notar em duas perspectivas médias, uma inicial e outra final, o comportamento relativo ao erro de previsão de dados de vento. Os dados mostram que para as 12h iniciais do horizonte o modelo W3 apresenta a menor média de erro **1,1476 m/s**. Para o horizonte completo o destaque em desempenho é do modelo W3 com MAE de **1,1894 m/s**.

Uma segunda análise, é sobre o erro de previsão em relação a primeira hora, em que para o Parque 02, o modelo W3 se destaca com o menor erro, **0,7912 m/s** contra a média 0,8086 m/s apresentado pelo modelo W1.

No caso da análise relativa a segunda hora do horizonte, o modelo que se destaca também é o modelo W3 com **0,9866 m/s** de MAE.

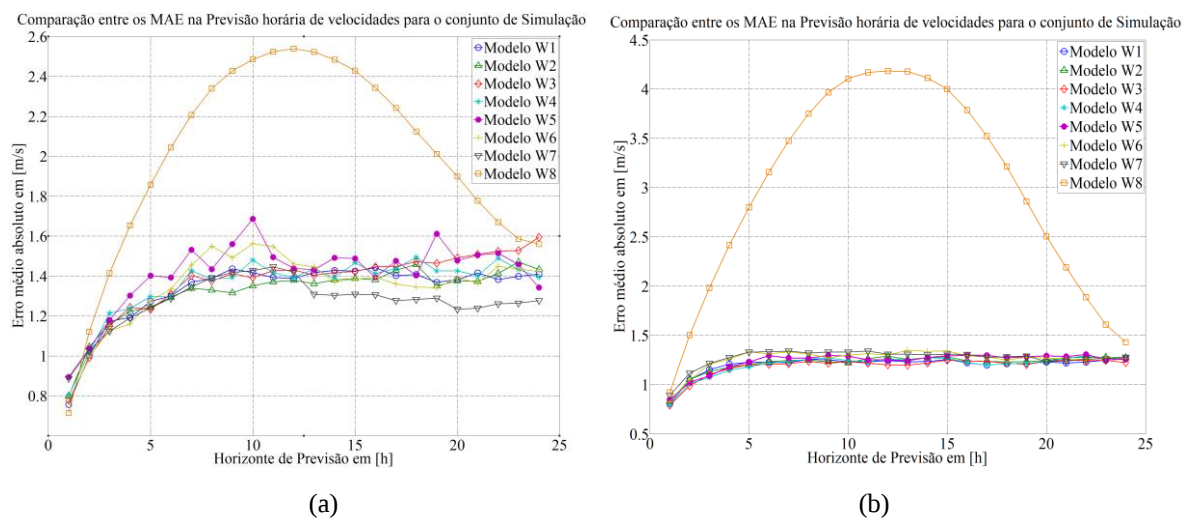
Em relação a previsão de direções, uma informação já verificada a respeito em ALBUQUERQUE (2015), é sobre o erro de previsão da direção em relação a primeira hora, em que para o Parque 02 devido a uma característica de distribuição muito homogênea quanto aos valores da base histórica de dados de direções, o modelo da Persistência W8 se destaca com o menor erro; **3,1162°** contra a ; **4,1197°** apresentado pelo modelo W2 os dados mostram que para as 12h iniciais do horizonte, o modelo W8 apresenta a menor média de erro 2,9410°.

Para o horizonte completo o destaque em desempenho também é do modelo W8 com MAE de **3,3166°** contra **3,9847°** do modelo W7. Este fato atesta que para o parque 02, apesar de um comportamento promissor nas previsões de velocidade, os modelos desenvolvidos devido a esta característica de distribuição homogênea na base histórica de dados de direções, não apresentaram um resultado satisfatório quanto a previsão de direções. Porém o modelo que mais se aproxima do W8 é o modelo W7 no caso de avaliação no horizonte completo.

Portanto após as análises relativas aos desempenhos apresentados nas Tabelas 9 e 10 fica comprovado que as novas heurísticas propostas apresentaram um desempenho satisfatório em relação a previsão de dados de velocidade de vento. Especificamente a heurística AE, caso do modelo W3 (RNA1 AET1) e a heurística LSTM, caso do modelo W7 (LSTM2 AET2).

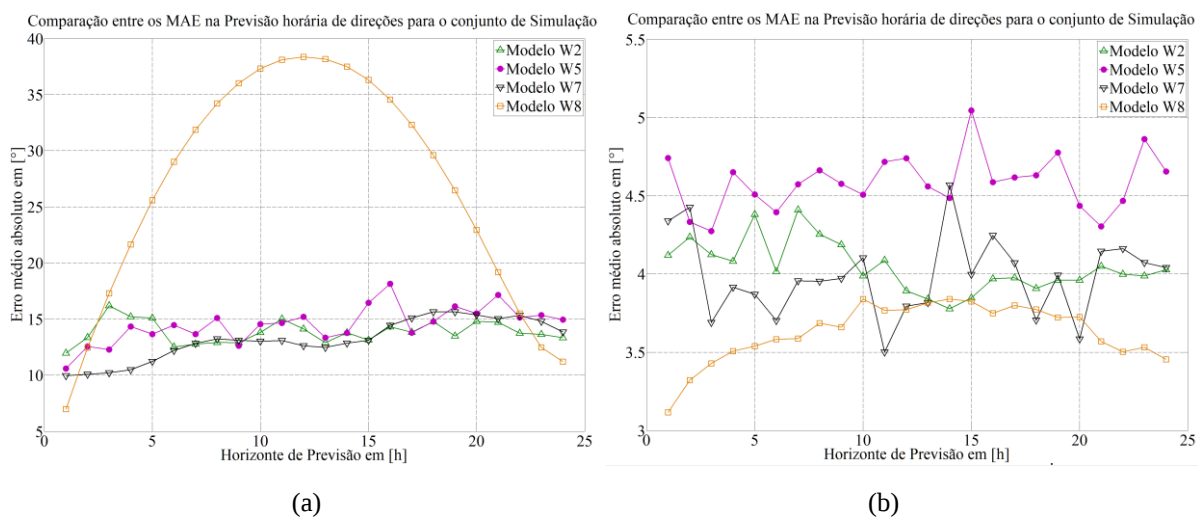
Outra percepção importante é que para o parque 02 assim como atestado pelos autores da heurística LSTM, os modelos desta heurística se destacaram com relação a resiliência de desempenho ao longo do horizonte, mesmo que esta resiliência não obtivesse desempenho satisfatório. Caso exemplificado pelo modelo W7 que apesar de não superar o modelo W8 é o que mais se aproxima, no caso da previsão de direções.

Figura 107 – Comparação do MAE de velocidade em (m/s) do Conjunto de Simulação para os modelos previsores de dados de vento com o modelo da Persistência atuantes (a) no Parque 01 e (b) no Parque 02.



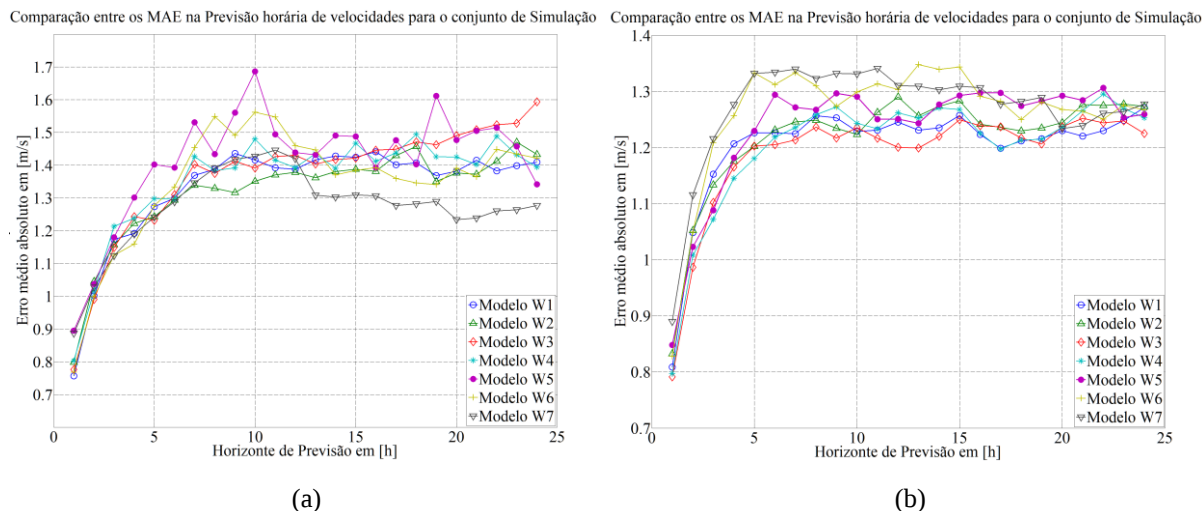
Fonte: O Autor (2021).

Figura 108 – Comparação do MAE de direção em (°) do Conjunto de Simulação para os modelos previsores de dados de vento com o modelo da Persistência atuantes (a) no Parque 01 e (b) no Parque 02.



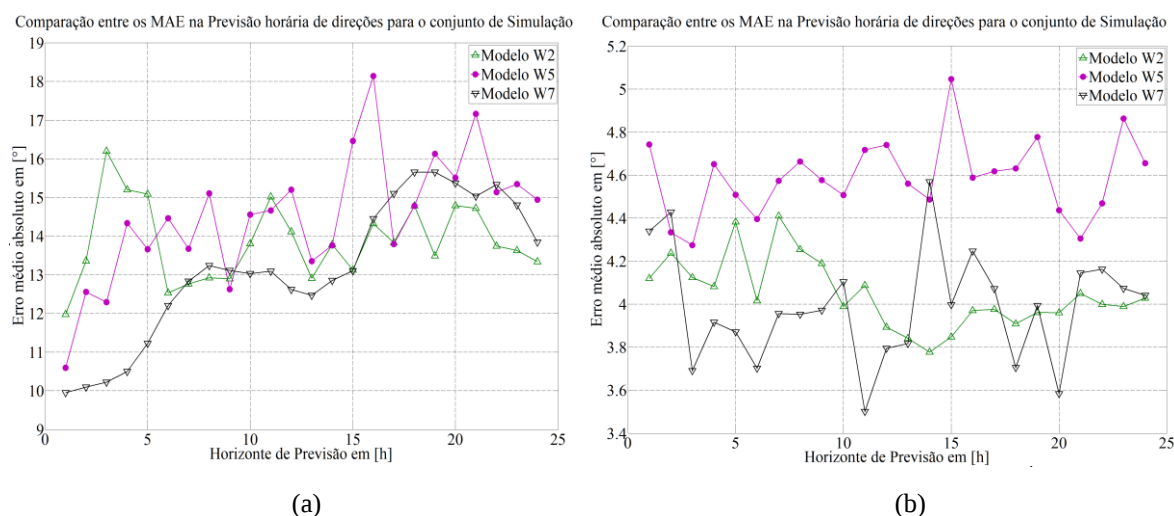
Fonte: O Autor (2021).

Figura 109 – Comparação do MAE de velocidade em (m/s) do Conjunto de Simulação para os modelos previsores de dados de vento sem o modelo da Persistência atuantes (a) no Parque 01 e (b) no Parque 02.



Fonte: O Autor (2021).

Figura 110 – Comparação do MAE de direção em ($^{\circ}$) do Conjunto de Simulação para os modelos previsores de dados de vento sem o modelo da Persistência atuantes (a) no Parque 01 e (b) no Parque 02.



Fonte: O Autor (2021).

Analisando a Figura 107, pode-se verificar o que fora atestado e explicado anteriormente sobre as comparações entre as previsões de velocidade de vento realizadas pelos modelos propostos (W1, W2, W3, W4, W5, W6 e W7) em relação ao modelo da persistência (W8), conforme mostrado pelas Tabelas 7 e 8.

Pelas análises realizadas nos resultados referentes aos previsores de velocidade e ou direção de vento, pode ser notado o melhor desempenho do modelo W7 em relação aos demais modelos, no caso do Parque 01 e destacando-se o modelo W3 no caso do parque 02. Já a Figura 108 evidencia o comportamento também explícito nas Tabelas 9 e 10 no que se

refere a comparação entre as previsões de direção realizadas pelos modelos e o modelo da persistência. Neste aspecto destacaram-se o modelo W7 no parque 01 e também o modelo W7 no parque 02 apesar de neste parque os modelos não se comportarem bem no quesito previsão de direções.

A Figura 109 de forma mais clara, destaca a comparação realizada pela Figura 107, retirando do escopo gráfico o modelo (W8), possibilitando uma melhor percepção de disputa de desempenho dos modelos ao longo do horizonte de previsão de direções.

De forma semelhante a Figura 110 propõe o mesmo artifício da Figura 108 só que evidenciando as comparações para as previsões de direções de vento.

A Tabela 11 mostra um resumo médio dos erros de previsão cometidos pelos modelos previsores de dados de vento propostos para o parque 01 e apresentando de forma sucinta a média de desempenho para as perspectivas inicial (12h iniciais) e perspectiva final, todo o horizonte de previsão (24h).

Tabela 11 – Tabela resumo das médias de MAE para 12h e 24h referentes a análise do Conjunto de Simulação para os modelos previsores de dados de vento no Parque 01.

Média de Desempenho dos Modelos previsores de dados de vento para o Parque 01								
Cód.	(W1)	(W2)	(W3)	(W4)	(W5)	(W6)	(W7)	(W8)
MAE Médio 12h[m/s]/[°]	1,2572	1,2371 / 13,8222	1,2616	1,2798	1,3627 / 13,6426	1,3106	1,2674 / 11,8427	1,9437 / 27,4029
MAE Médio 24h[m/s]/[°]	1,3318	1,3191 / 13,8465	1,3691	1,3568	1,4143 / 14,5083	1,3513	1,2733 / 13,1565	1,9988 / 26,8774

Fonte: O Autor (2021).

A Tabela 12 apresenta um resumo médio dos erros de previsão cometidos pelos modelos previsores de dados de vento propostos para o parque 02 e apresentando de forma sucinta a média de desempenho para as perspectivas inicial (12h iniciais) e perspectiva final, todo o horizonte de previsão (24h).

Tabela 12 – Tabela resumo das médias de MAE para 12h e 24h referentes a análise do Conjunto de Simulação para os modelos previsores de dados de vento no Parque 02.

Média de Desempenho dos Modelos previsores de dados de vento para o Parque 02								
Cód.	(W1)	(W2)	(W3)	(W4)	(W5)	(W6)	(W7)	(W8)
MAE Médio 12h[m/s]/[°]	1,1756	1,1776 / 3,9424	1,1476	1,1605	1,1909 / 4,5568	1,2350	1,2617 / 3,9356	3,0363 / 2,9410
MAE Médio 24h[m/s]/[°]	1,2033	1,2180 / 4,0455	1,1894	1,2036	1,2354 / 4,5880	1,2616	1,2705 / 3,9847	2,9891 / 3,3166

Fonte: O Autor (2021).

As Tabelas 13 e 14 mostram que no Parque 01 os maiores ganhos médios horários das previsões de dados de vento em relação ao modelo da persistência nas 12h iniciais do

horizonte são obtidos pelo modelo W2 com **(36,3533%)** na previsão de velocidades e pelo modelo W7 com **(56,7830%)** na previsão de direções. Para a análise de todo o horizonte de previsão o destaque é para o modelo W7 com **(36,2968%)** para a previsão de velocidades e com **(51,0500%)** para a previsão de direções.

Tabela 13 – Aferição dos Ganhos percentuais relativos aos modelos previsores de dados de vento atuantes no Parque 01 em relação ao modelo da persistência.

Ganhos (%) dos Modelos previsores de dados de vento para o Parque 01							
Cód.	(W1)	(W2)	(W3)	(W4)	(W5)	(W6)	(W7)
MAE Médio 12h[m/s] / [°]	35,3192	36,3533 / 49,5594	35,0929	34,1565	29,8914 / 50,2148	32,5719	34,7945 / 56,7830
MAE Médio 24h[m/s] / [°]	33,3700	34,0054 / 48,4827	31,5039	32,1193	29,2425 / 46,0204	32,3944	36,2968 / 51,0500

Fonte: O Autor (2021).

Tabela 14 – Aferição dos Ganhos percentuais relativos aos modelos previsores de dados de vento atuantes no Parque 02 em relação ao modelo da persistência.

Ganhos (%) dos Modelos previsores de dados de vento para o Parque 02							
Cód.	(W1)	(W2)	(W3)	(W4)	(W5)	(W6)	(W7)
MAE Médio 12h[m/s] / [°]	61,2818	61,2160 / -34,0496	62,2040	61,7791	60,7779 / -54,9405	59,3255	58,4461 / -33,8184
MAE Médio 24h[m/s] / [°]	59,7437	59,2519 / -21,9773	60,2088	59,7337	58,6698 / -38,3344	57,7933	57,4956 / -20,1441

Fonte: O Autor (2021).

Já para o parque 02 os maiores ganhos médios horários das previsões de dados de vento em relação ao modelo da persistência nas 12h iniciais do horizonte são obtidos pelo modelo W3 com **(62,2040%)** na previsão de velocidades e pelo modelo W7 **(-33,8184%)** na previsão de direções. Para a análise de todo o horizonte de previsão o destaque é para o modelo W3 com **(60,2088%)** para a previsão de velocidades e para o modelo W7 com **(-20,1441%)** para a previsão de direções.

Desta forma nota-se que uma competição benéfica mostrando o bom desempenho das ferramentas de IA utilizadas e desenvolvidas nesta tese.

7.1.2 DISCUSSÃO DOS RESULTADOS RELATIVOS AOS MODELOS APROXIMADORES DE CURVAS DE POTÊNCIA

Conforme arguido anteriormente nesta tese, foram desenvolvidos doze modelos de curvas de potência, a Tabela 15 procura resumir estes modelos indicando uma codificação para os mesmos podendo ser: RNA, FIS (Mamdani), FIS (Sugeno), RNA adequada por autoencoder, FIS adequada pelo ANFIS e curva polinomial.

Tabela 15 – Codificação dos Modelos de curvas de potência desenvolvidos.

Modelo	Cod.	Modelo	Cod.	Modelo	Cod.
CP1RNA	(A)	CP2RNA AET2	(E)	CP2Fuzzy (Sugeno)	(I)
CP2RNA	(B)	CP1Fuzzy (Mamdani)	(F)	CP1Fuzzy (Sugeno-ANFIS)	(J)
CP1RNA AET1	(C)	CP2Fuzzy (Mamdani)	(G)	CP2Fuzzy (Sugeno-ANFIS)	(K)
CP1RNA AET2	(D)	CP1Fuzzy (Sugeno)	(H)	CP Média Polinomial	(L)

Fonte: O Autor (2021).

A seguir serão mostradas as respectivas Tabelas 16 e 17 com os índices de desempenho (MAE, NMAE e RMSE), relativos a estas aproximações, para ambos os parques.

Tabela 16 – Aferição dos Desempenhos relativos às curvas de potência atuantes no parque 01.

Desempenho das Curvas de potência para o Parque 01												
Cód.	(A)	(B)	(C)	(D)	(E)	(F)	(G)	(H)	(I)	(J)	(K)	(L)
MAE [MW]	1,8439	1,5920	3,5523	1,6840	1,5402	1,7275	1,8706	2,0195	1,8208	1,6708	1,5696	3,1808
NMAE [%]	2,6132	2,2563	5,0345	2,3866	2,1828	2,4483	2,6511	2,8621	2,5805	2,3679	2,2245	4,5079
RMSE [MW]	3,5800	2,0938	4,2712	2,1216	2,0174	2,2138	2,3680	2,4690	2,3126	2,1139	2,0032	3,7568

Fonte: O Autor (2021).

Tabela 17 – Aferição dos Desempenhos relativos às curvas de potência atuantes no parque 02.

Desempenho das Curvas de potência para o Parque 02												
Cód.	(A)	(B)	(C)	(D)	(E)	(F)	(G)	(H)	(I)	(J)	(K)	(L)
MAE [MW]	3,2755	3,3360	3,2660	3,2109	3,5506	3,7879	3,5854	3,5675	3,5368	3,2692	3,1884	3,3653
NMAE [%]	2,5996	2,6476	2,5921	2,5483	2,8179	3,0063	2,8455	2,8314	2,8070	2,5946	2,4899	2,6709
RMSE [MW]	4,2666	4,3714	4,2436	4,2006	4,7607	5,1762	4,6717	4,7555	4,6631	4,2602	4,1319	4,3355

Fonte: O Autor (2021).

As Tabelas 16 e 17 mostram o desempenho com relação ao erro de aproximação da curva de potência relativa aos modelos de CP desenvolvidos. Na Tabela 16 fica evidente no parque 01 que o melhor modelo aproximado de CP considerando o MAE e o NMAE é o modelo (E) com (1,5402MW) e com (2,1828%) para o NMAE. Já considerando o RMSE é o modelo (K) com um RMSE de (2,0032MW).

Analisando a Tabela 17 referenciando o parque 02, o melhor modelo aproximado de CP considerando o MAE, o NMAE e o RMSE é o modelo (K), com respectivamente (3,1884MW), (2,4899%) e (4,1319MW)

Ainda sobre Tabela 16 percebe-se uma considerável vantagem do modelo CP2 RNA AET2 no que se refere ao MAE e NMAE e uma competição com o modelo CP2Fuzzy (Sugeno ANFIS) quando se trata o RMSE. Outro ponto interessante é que os modelos que mais se aproximam desse desempenho são os modelos a RNA adequadas por AEs. Ainda nesse aspecto nos modelos a RNAs foram consideradas duas camadas ocultas, o que de certa forma agrega uma melhor performance as mesmas.

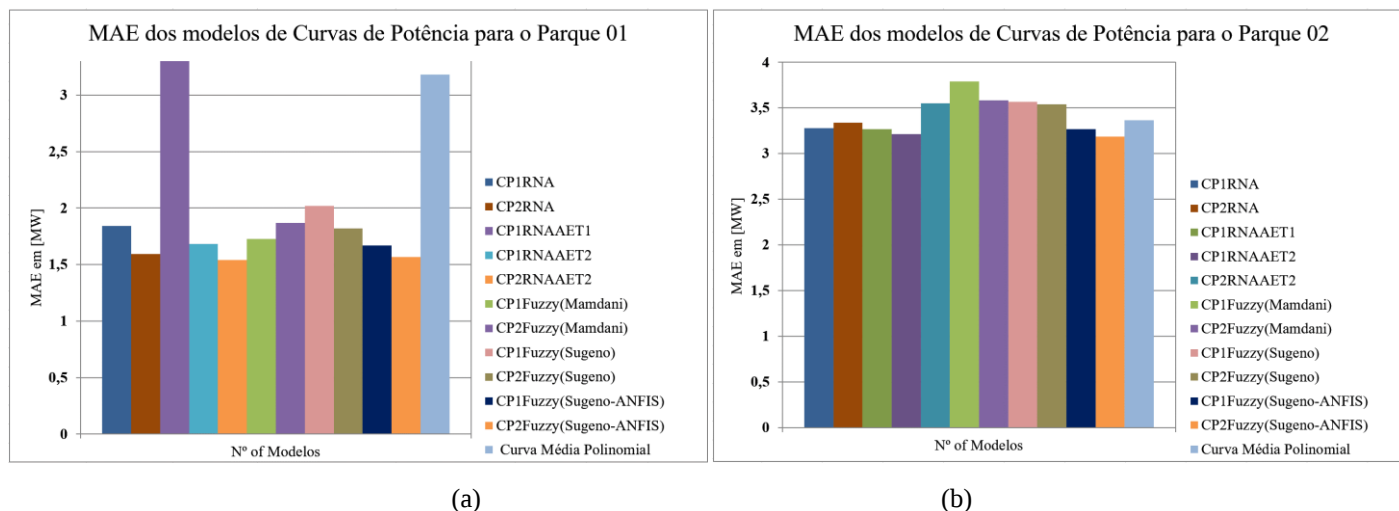
Devido à característica neural do ANFIS há uma melhora nas inferências (Fuzzy – Neurais) fator perceptível no desempenho destas CPs.

Resumindo a Tabela 17 percebe-se novamente uma considerável vantagem do modelo CP2Fuzzy (Sugeno ANFIS), em relação aos outros modelos de CPs abordados. Novamente o ANFIS melhora o processo da aprendizagem, o que aumenta o desempenho destas CPs. Outro fator é o número de entradas, os resultados mostram vantagens dos modelos com duas entradas em relação aos modelos com uma entrada.

Além disso existe uma boa correlação da segunda entrada (direção média) com a primeira entrada (velocidade média), e esta é uma diferença de abordagem deste trabalho em relação a vários outros em que não haja o uso de outra variável que não a velocidade como entrada no modelo da curva de potência.

A Figura 111 mostra um diagrama de barras que expressa de forma bastante notória o desempenho do MAE para modelos aproximadores de curvas de potência inferentes nos parques em estudo.

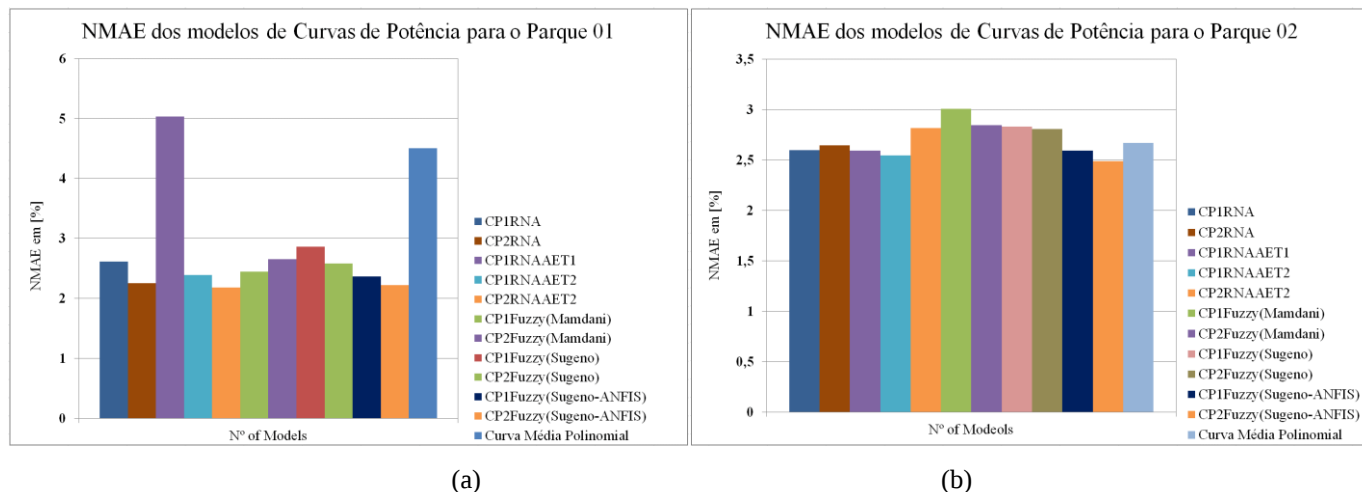
Figura 111 – MAE do Conjunto de Simulação para os modelos de curvas de Potência atuantes (a) no Parque 01 e (b) no Parque 02.



Fonte: O Autor (2021).

A Figura 112 mostra um diagrama de barras que expressa de forma bastante notória o desempenho do NMAE dos modelos aproximadores de curvas de potência inferentes nos parques em estudo.

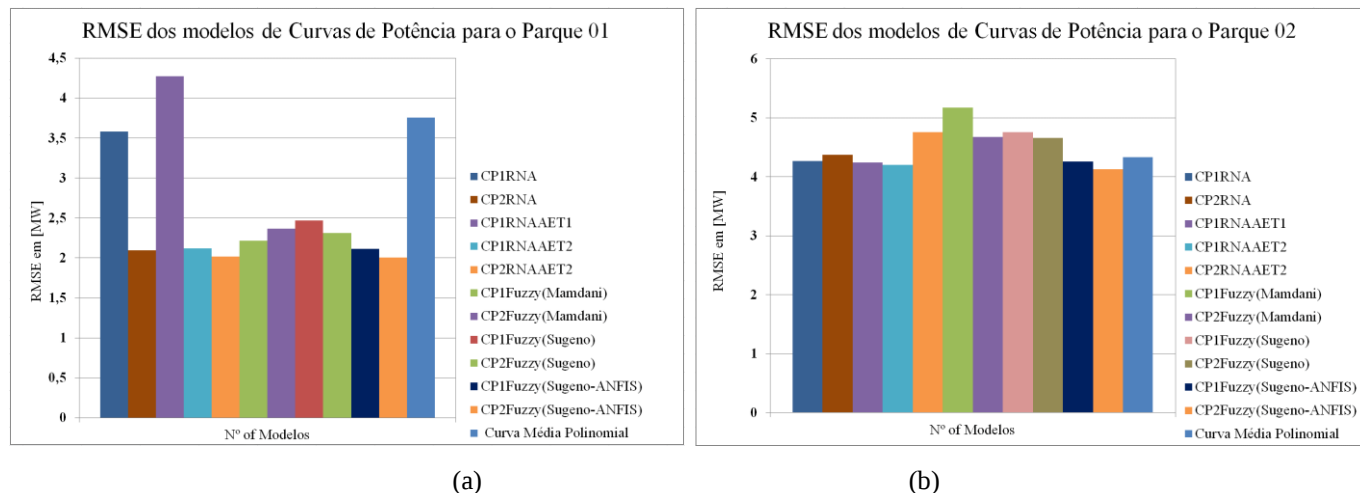
Figura 112 – NMAE do Conjunto de Simulação para os modelos de curvas de Potência atuantes (a) no Parque 01 e (b) no Parque 02.



Fonte: O Autor (2021).

A Figura 113 mostra um diagrama de barras que expressa de forma bastante notória o desempenho do RMSE dos modelos aproximadores de curvas de potência inferentes nos parques em estudo.

Figura 113 – RMSE do Conjunto de Simulação para os modelos de curvas de Potência atuantes (a) no Parque 01 e (b) no Parque 02.



Fonte: O Autor (2021)

As Tabelas 18 e 19 mostram os ganhos dos modelos de CPs propostos em relação a o modelo da curva média:

Tabela 18 – Aferição dos Ganhos percentuais relativos aos modelos de curvas de potência atuantes no parque 01 em relação ao modelo de referência.

Ganhos (%) dos modelos das Curvas de potência para o Parque 01 em relação ao modelo da CP média											
Cód.	(A)	(B)	(C)	(D)	(E)	(F)	(G)	(H)	(I)	(J)	(K)
MAE [MW]	42,0303	49,9497	-11,6795	47,0573	51,5782	45,6898	41,1909	36,5097	42,7565	47,4723	50,6539
NMAE [%]	42,0307	49,9479	-11,6817	47,0574	51,5783	45,6887	41,1899	36,5092	42,7561	47,4722	50,6533
RMSE [MW]	4,7061	44,2664	-13,6925	43,5264	46,3000	41,0722	36,9676	34,2792	38,4423	43,7314	46,6780

Fonte: O Autor (2021)

Tabela 19 – Aferição dos Ganhos percentuais relativos aos modelos de curvas de potência atuantes no parque 02 em relação ao modelo de referência.

Ganhos (%) dos modelos das Curvas de potência para o Parque 02 em relação ao modelo da CP média											
Cód.	(A)	(B)	(C)	(D)	(E)	(F)	(G)	(H)	(I)	(J)	(K)
MAE [MW]	2,6684	0,8707	2,9507	4,5880	-5,5062	-12,5576	-6,5403	-6,0084	-5,0961	2,8556	5,2566
NMAE [%]	2,6695	0,8724	2,9503	4,5902	-5,5038	-12,5576	-6,5371	-6,0092	-5,0957	2,8567	6,7767
RMSE [MW]	1,5892	-0,8280	2,1197	3,1115	-9,8074	-19,3911	-7,7546	-9,6875	-7,5562	1,7368	4,6961

Fonte: O Autor (2021)

Nas Tabelas 18 e 19 os ganhos relativos dos modelos de curvas de potências em relação ao modelo de referência da curva média ficam evidenciados, os resultados expressam uma vantagem notória dos modelos da CPRNA AET# e dos modelos de inferência pré-estabelecida respectivamente os modelos **CPFuzzy 1 (Sugeno-ANFIS)** e **CP2Fuzzy (Sugeno ANFIS)**. Assim com relação ao parque 01 o Modelo (E) **CP2RNA AET2** foi o que obteve

melhores ganhos. para os índices **MAE**, **NMAE**, respectivamente **51,5782%** e **51,5783%**, já em relação ao **RMSE** o melhor Modelo foi o **(K) CP2Fuzzy (Sugeno-ANFIS) 46,6780%**.

Em relação ao parque 02 o Modelo **(K): CP2Fuzzy (Sugeno-ANFIS)** também obteve os melhores resultados sendo os ganhos descritos respectivamente para os índices **MAE**, **NMAE** e **RMSE**, dados por **5,2566%**; **6,7767%** e **4,6961%**.

Também é possível notar a presença de ganhos negativos que expressam um desempenho baixo dos modelos em relação ao modelo de referência (Curva média polinomial). Assim pode-se tomar a pré-inferência dos modelos Fuzzy e adequada pelo AE as redes neurais como fatores cruciais na melhora do desempenho dos modelos desenvolvidos.

Objetivando mostrar o desempenho dos modelos Previsores de Potência as Tabelas 7.17 a 7.24 mostram a tendência média horária dos índices de MAE e NMAE para as previsões de Potência realizadas pelos modelos previsores de potência desta Tese. Tais modelos compostos segundo o que indica a Tabela 6.4.

Devido ao número bastante elevado de modelos compostos as Tabelas foram divididas em duas partes em que a primeira ilustra as 12h iniciais do horizonte e a segunda ilustra as 12h finais deste horizonte. Este fato foi resultado do alto grau de experimentação que procurou averiguar o casamento entre as várias heurísticas propostas neste trabalho. Nesta Tese foram propostos 32 modelos previsores de geração novos compostos por modelos que foram descritos previamente ao longo deste trabalho.

Além disso na primeira parte além da perspectiva inicial do horizonte também é fornecida a média das 12h iniciais e na segunda parte além da perspectiva final também é fornecida a média dos índices de MAE e NMAE relativo as 24h do horizonte.

Tabela 20 – Desempenho em MW dos Modelos previsores nas primeiras 12h do período de Simulação para o Parque 01

MAE em [MW] dos modelos previsores de potência desenvolvidos para as primeiras 12h do horizonte

Horizonte	1h	2h	3h	4h	5h	6h	7h	8h	9h	10h	11h	12h	Média 12h
M22	6,5043	7,6273	8,8238	9,7948	10,4683	10,4166	11,2478	10,6256	11,3844	12,4607	11,1138	10,6187	10,0905
M23	6,4490	7,5613	8,7463	9,7026	10,4294	10,3516	11,1789	10,6232	11,3691	12,4066	11,0793	10,6369	10,0445
M24	6,3346	7,4450	8,6305	9,4642	10,2013	10,1143	10,7589	10,3899	11,0090	12,0154	10,8383	10,5766	9,8148
M25	6,5236	7,5968	8,8032	9,7570	10,5187	10,4431	11,2231	10,6748	11,4435	12,5032	11,1127	10,6260	10,1021
M2A	5,8998	7,7929	8,7854	9,3179	9,2709	9,6941	10,0239	9,8416	9,7320	9,9805	10,2379	10,2072	9,2320
M21	5,8588	7,7722	8,7104	9,2458	9,2294	9,6522	9,9603	9,8019	9,6925	9,9525	10,1700	10,2085	9,1879
M2B	5,9964	7,7258	8,6085	9,2407	9,2029	9,6204	9,9123	9,7209	9,6286	9,8773	10,0077	10,1580	9,1416
M2C	5,8478	7,8026	8,7610	9,2398	9,2170	9,6903	9,9875	9,8529	9,7182	9,9621	10,1924	10,2429	9,2095
M32	5,7693	7,5183	9,0558	9,2236	9,6292	9,6787	10,7104	10,4162	10,3543	10,9924	10,5929	10,3846	9,5271
M33	6,2172	7,4501	8,6400	8,9015	9,3900	9,5508	10,0326	10,0376	9,9032	10,2516	9,9576	9,8343	9,1805
M34	5,7560	7,5227	9,0668	9,2390	9,6399	9,6893	10,7367	10,4277	10,3728	11,0162	10,6199	10,4020	9,5407
M35	5,9028	7,7442	9,3176	9,5285	9,8974	9,9129	11,0694	10,7081	10,6890	11,4056	10,9794	10,7058	9,8217
M36	5,7732	7,5265	9,0657	9,2336	9,6349	9,6827	10,7166	10,4238	10,3642	11,0070	10,6031	10,3946	9,5355
M1A	5,4983	7,4311	8,7862	8,9065	9,4945	9,6892	10,2733	10,4579	10,8289	10,5869	10,3588	10,3675	9,3899
M11	5,6445	7,4160	8,4725	9,0997	9,3069	9,6817	9,9762	10,1543	10,5411	10,3602	10,3718	10,3538	9,2816
M20	5,5026	7,4453	8,8018	8,9222	9,5143	9,7008	10,2885	10,4741	10,8467	10,6048	10,3676	10,3753	9,4037
M1B	7,2230	8,9997	9,9104	10,3682	10,5708	10,6101	10,8074	11,0655	11,2746	11,3077	11,4589	11,3888	10,4154
M1C	7,0889	9,1156	9,9973	10,3647	10,4686	10,5574	10,7367	11,0620	11,2857	11,1966	11,2644	11,3425	10,3734
M42	5,6589	7,3422	8,4734	9,2037	9,1397	9,6670	10,3529	10,1416	10,3875	10,2614	10,5191	10,5541	9,3085
M43	5,5618	7,1059	8,1509	8,8480	8,9041	9,1654	9,7037	9,5010	9,6366	9,5308	9,6989	9,7818	8,7991
M44	6,7684	8,2371	9,1795	9,8289	9,7855	10,2396	10,8216	10,7510	10,9060	10,8609	11,0551	11,0765	9,9592
M45	5,9810	7,6363	8,7548	9,4929	9,4481	10,0048	10,7354	10,5456	10,7609	10,6482	10,9038	10,9576	9,6558
M46	5,6673	7,3481	8,4839	9,2114	9,1413	9,6736	10,3678	10,1536	10,4032	10,2773	10,5390	10,5699	9,3197
M52	5,6788	7,4284	8,3737	8,6375	9,5467	10,0244	10,9143	11,6378	11,3027	11,7936	11,6094	10,9803	9,8273
M53	6,1560	7,4163	8,3784	9,0470	8,9223	9,3010	9,9867	10,4956	10,2698	10,5070	10,3987	10,0383	9,2431
M54	5,5944	7,3795	8,3527	8,6256	9,5648	10,0523	10,9421	11,6804	11,3382	11,8319	11,6473	11,0093	9,8349
M55	5,6601	7,5202	8,5307	8,7277	9,8529	10,4268	11,3669	12,1535	11,7910	12,3248	12,1054	11,3841	10,1537
M56	5,6020	7,3954	8,3496	8,6289	9,5573	10,0332	10,9211	11,6518	11,3124	11,8081	11,6247	10,9909	9,8229
M62	6,5454	7,6477	8,4122	8,9564	9,3797	9,7381	10,1221	10,3763	10,5211	10,6232	10,8276	10,5740	9,4770
M64	6,5225	7,6242	8,3802	8,9341	9,3394	9,7122	10,0847	10,3361	10,4820	10,5646	10,7672	10,5367	9,4403
M65	6,5814	7,6802	8,4479	9,0163	9,3819	9,7070	10,0277	10,2281	10,3285	10,4156	10,6795	10,4898	9,4153
M66	6,5442	7,6523	8,4081	8,9662	9,3805	9,7575	10,1416	10,3859	10,5225	10,6133	10,8271	10,5897	9,4824
Mod.Per	5,0207	8,0623	10,3003	12,0580	13,6638	15,0691	16,2990	17,2662	17,9586	18,4599	18,7995	18,9322	14,3241

Fonte: O Autor (2021)

Tabela 21 – Desempenho em MW dos Modelos previsores nas primeiras 12h do período de Simulação para o Parque 02

MAE em [MW] dos modelos previsores de potência desenvolvidos para as primeiras 12h do horizonte

Horizonte	1h	2h	3h	4h	5h	6h	7h	8h	9h	10h	11h	12h	Média 12h
M22	7,4592	9,1055	9,7760	10,5090	10,8250	11,1463	11,0447	11,1739	11,5161	11,4015	11,1301	11,1584	10,5205
M23	7,1820	9,0144	9,7153	10,6071	11,0668	11,4922	11,3265	11,5704	11,9386	11,5949	11,3010	11,4579	10,6889
M24	7,3351	8,8714	9,6152	10,4222	10,7423	11,1750	11,0389	11,2170	11,6638	11,5382	11,2595	11,4099	10,5240
M25	7,4217	9,0337	9,7008	10,4468	10,7854	11,1858	11,0623	11,1234	11,5075	11,3982	11,1617	11,1930	10,5017
M2A	7,0420	9,2892	9,9456	10,3706	10,4264	10,9891	11,0763	11,3881	10,9757	11,3602	11,6890	11,7395	10,5243
M21	7,0584	9,1480	9,8625	10,1728	10,4345	10,8423	11,0523	11,2862	11,1725	11,4254	11,9046	12,0866	10,5372
M2B	7,1328	9,1240	9,8684	10,1969	10,3347	10,8668	10,9444	11,2562	11,0120	11,3531	11,7757	11,8660	10,4776
M2C	7,0105	9,2001	9,9039	10,2842	10,3761	10,9029	11,0408	11,3057	10,9815	11,3338	11,7058	11,7193	10,4804
M32	6,9391	8,7735	9,4768	10,1177	10,3068	10,6254	10,7814	11,0365	11,2817	11,0811	11,1253	11,4247	10,2475
M33	6,8876	8,6621	9,3861	10,0118	10,2279	10,5969	10,6998	10,9787	11,2464	11,0648	11,1320	11,4480	10,1952
M34	6,9315	8,5858	9,3078	9,9451	10,2318	10,6774	10,6810	11,0097	11,3037	11,1567	11,2047	11,5943	10,2191
M35	7,2524	8,6618	9,3603	9,9733	10,3410	10,6958	10,8432	11,0997	11,4058	11,2735	11,2790	11,6260	10,3177
M36	6,9435	8,7678	9,4696	10,1161	10,3053	10,6251	10,7812	11,0328	11,2813	11,0855	11,1248	11,4311	10,2470
M1A	7,2312	9,2054	10,1577	10,3894	10,5700	10,6059	10,8088	11,0415	11,1539	11,1849	11,1922	11,2453	10,3989
M11	7,1339	9,1473	10,0732	10,3414	10,4852	10,5567	10,7467	11,0074	11,1688	11,1490	11,1782	11,2481	10,3530
M20	7,0889	9,1156	9,9973	10,3647	10,4686	10,5574	10,7367	11,0620	11,2857	11,1966	11,2644	11,3425	10,3734
M1B	7,2230	8,9997	9,9104	10,3682	10,5708	10,6101	10,8074	11,0655	11,2746	11,3077	11,4589	11,3888	10,4154
M1C	7,2251	9,1954	10,1490	10,3941	10,5742	10,5991	10,8058	11,0453	11,1583	11,1906	11,1979	11,2521	10,3989
M42	6,9669	8,7086	9,7192	10,0911	10,4427	10,5202	10,7140	10,8774	10,7815	11,0518	10,7382	10,7062	10,1098
M43	6,9246	8,6492	9,7163	10,0608	10,4023	10,5016	10,6733	10,8639	10,7776	11,0776	10,7514	10,7451	10,0953
M44	6,9985	8,6599	9,7914	10,1368	10,4203	10,5555	10,7100	10,9210	10,8603	11,1913	10,9008	10,9167	10,1719
M45	7,1483	8,6308	9,7688	10,1066	10,3913	10,6220	10,7614	10,8784	10,8053	11,1529	10,9846	11,0030	10,1878
M46	6,9641	8,7018	9,7173	10,0923	10,4355	10,5179	10,7128	10,8785	10,7877	11,0515	10,7390	10,7070	10,1088
M52	7,4968	9,6917	10,9433	11,3536	12,0013	12,1019	12,2537	12,1538	11,9365	12,3403	12,5201	12,2391	11,4193
M53	7,5229	9,6149	10,9715	11,3786	12,0507	12,1124	12,2550	12,0589	11,8585	12,2965	12,5420	12,1934	11,4046
M54	7,6652	9,5310	11,0647	11,4670	12,1814	12,1759	12,3089	12,0011	11,7910	12,3073	12,6197	12,1788	11,4410
M55	7,6950	9,4002	10,6641	11,2045	11,9113	11,9977	12,1857	12,0403	11,7933	12,0898	12,1768	12,1151	11,2728
M56	7,4985	9,6942	10,9422	11,3517	12,0065	12,0951	12,2518	12,1501	11,9352	12,3422	12,5203	12,2392	11,4189
M62	8,0476	10,3670	11,1159	11,5811	11,8791	12,1653	12,3640	12,3382	12,4328	12,5589	12,4657	12,1448	11,6217
M64	8,0856	10,1474	11,0460	11,5432	11,9241	12,1677	12,5097	12,4924	12,6849	12,7334	12,7576	12,3358	11,7023
M65	8,0526	10,1298	10,9099	11,3932	11,7356	12,0154	12,3437	12,2932	12,4620	12,5260	12,4863	12,1405	11,5407
M66	7,9957	10,2431	11,0605	11,5328	11,8308	12,0736	12,3434	12,3135	12,4314	12,4931	12,4823	12,1133	11,5761
Mod.Per	8,9653	15,5186	20,5351	25,1778	29,5092	33,7270	37,9104	41,6583	44,7131	46,4698	46,8647	47,2012	33,1875

Fonte: O Autor (2021)

Tabela 22 – Desempenho em MW dos Modelos previsores nas 12h finais do período de Simulação para o Parque 01

MAE em [MW] dos modelos previsores de potência desenvolvidos para 12h finais do horizonte

Horizonte	13h	14h	15h	16h	17h	18h	19h	20h	21h	22h	23h	24h	Média 24h
M22	10,6910	11,0905	11,0567	10,3705	10,9645	10,6293	11,8980	10,9794	11,0662	11,0809	10,8571	10,0699	10,4933
M23	10,6701	11,0869	10,9956	10,2829	10,9157	10,5586	11,8987	10,9301	11,0418	11,0464	10,7995	10,0131	10,4489
M24	10,5510	10,9147	10,9060	10,1250	10,6848	10,3534	11,6337	10,7204	10,8727	10,8338	10,6782	9,9217	10,2489
M25	10,7120	11,1328	11,0316	10,3498	10,9703	10,6294	11,9657	10,9671	11,0803	11,1107	10,8494	10,0412	10,5027
M2A	10,1643	10,2823	10,4256	10,4186	10,8822	10,8785	10,1389	10,3241	10,2593	10,6159	11,0612	10,7384	9,8739
M21	10,1326	10,2363	10,3864	10,3183	10,7835	10,8049	10,0898	10,2843	10,2509	10,5955	11,0420	10,7058	9,8285
M2B	10,0821	10,1617	10,3764	10,1822	10,5874	10,6672	10,0657	10,1983	10,3138	10,5549	10,9344	10,6655	9,7704
M2C	10,1729	10,2709	10,4254	10,3459	10,8077	10,8336	10,1194	10,3116	10,2450	10,6047	11,0372	10,7240	9,8505
M32	10,8527	10,5212	11,0026	10,5931	10,8202	11,2860	10,6657	10,6691	10,4738	11,1284	10,7053	10,4964	10,1475
M33	10,1508	10,0093	10,1808	9,9548	10,0710	10,3787	10,1007	9,9359	9,9170	10,3882	10,1467	10,0067	9,6420
M34	10,8856	10,5435	11,0306	10,6152	10,8453	11,3175	10,6900	10,6958	10,4888	11,1526	10,7199	10,5098	10,1660
M35	11,2880	10,8927	11,4251	10,9418	11,2148	11,7667	10,9767	11,0669	10,7745	11,5450	11,0379	10,8234	10,4839
M36	10,8632	10,5318	11,0140	10,6041	10,8288	11,2979	10,6752	10,6805	10,4836	11,1419	10,7166	10,5065	10,1571
M1A	10,5202	10,6733	10,6318	10,7354	10,5066	10,5629	10,2751	10,3588	10,6016	10,4003	10,5525	10,6258	9,9635
M11	10,0831	10,5691	10,3087	10,4167	10,2206	10,1324	9,9402	9,8147	10,0907	9,9850	10,3837	10,3733	9,7374
M20	10,5345	10,6838	10,6486	10,7509	10,5183	10,5763	10,2864	10,3810	10,6254	10,4160	10,5635	10,6472	9,9782
M1B	11,2465	11,2702	11,5149	11,2473	11,0415	11,0633	10,9593	11,1259	11,2932	11,3149	11,3052	11,6638	10,8346
M1C	11,2250	11,2258	11,4806	11,1285	10,8674	10,9950	10,9630	11,2289	11,4009	11,3153	11,3978	11,6945	10,8085
M42	10,4046	10,3936	10,4961	10,7086	10,8248	10,9688	10,8650	11,0739	11,1246	11,3439	11,3900	11,7525	10,1270
M43	9,4995	9,6029	9,5376	9,9160	9,9578	9,9830	10,2199	10,2653	10,2849	10,2298	10,3388	10,6169	9,4184
M44	10,9841	10,9644	11,0947	11,1979	11,3135	11,4674	11,3854	11,6457	11,7069	11,8550	11,9570	12,2409	10,7218
M45	10,8325	10,7781	10,9310	11,1019	11,2486	11,3840	11,1963	11,4539	11,4924	11,7551	11,8259	12,1791	10,5020
M46	10,4203	10,4061	10,5092	10,7231	10,8330	10,9787	10,8770	11,0861	11,1407	11,3594	11,4082	11,7787	10,1399
M52	10,8653	10,3127	10,3968	10,4551	10,1933	10,1069	10,0721	10,3857	10,2602	10,8546	10,8231	10,6767	10,1388
M53	10,0603	9,9084	9,9869	10,0550	10,2194	10,0488	9,8484	9,9273	9,7802	10,1499	10,1884	10,2055	9,6373
M54	10,8919	10,3255	10,4144	10,4744	10,2039	10,1148	10,0855	10,4059	10,2757	10,8748	10,8415	10,6859	10,1505
M55	11,2639	10,5727	10,6799	10,7419	10,4285	10,3184	10,2996	10,6861	10,5452	11,1869	11,1640	10,9546	10,4452
M56	10,8763	10,3164	10,4019	10,4638	10,2016	10,1133	10,0803	10,3951	10,2672	10,8661	10,8358	10,6735	10,1403
M62	10,4776	10,4290	10,5390	10,4328	10,3047	10,4168	10,4529	10,4148	10,3388	10,1260	10,1073	10,1100	9,9114
M64	10,4506	10,4102	10,5350	10,4369	10,2933	10,3825	10,4191	10,3783	10,2951	10,0826	10,0748	10,0827	9,8802
M65	10,4368	10,4601	10,6066	10,5405	10,4036	10,4675	10,4932	10,3931	10,3036	10,0597	10,0553	10,0722	9,8865
M66	10,4988	10,4536	10,5792	10,4731	10,3231	10,4013	10,4213	10,3751	10,2883	10,0598	10,0544	10,0783	9,9081
Mod.Per	18,7812	18,4273	17,9706	17,3486	16,5973	15,7027	14,8337	13,9723	13,0561	12,2111	11,6222	11,4221	15,1621

Fonte: O Autor (2021)

Tabela 23 – Desempenho em MW dos Modelos previsores nas 12h finais do período de Simulação para o Parque 02

MAE em [MW] dos modelos previsores de potência desenvolvidos para 12h finais do horizonte													
Horizonte	13h	14h	15h	16h	17h	18h	19h	20h	21h	22h	23h	24h	Média 24h
M22	11,0113	11,1647	11,4933	11,4230	11,4739	11,4566	11,3471	11,5252	11,6150	11,9122	11,4722	11,2093	10,9729
M23	11,2292	11,2683	11,3941	11,1938	11,3305	11,2257	11,2371	11,3575	11,5265	11,8996	11,3820	11,3631	11,0281
M24	11,2236	11,2144	11,4370	11,2691	11,4351	11,3304	11,2758	11,3867	11,5285	11,8109	11,3867	11,2647	10,9521
M25	11,0180	11,1372	11,4287	11,2837	11,3667	11,3672	11,2752	11,4426	11,5555	11,8412	11,3940	11,1840	10,9298
M2A	11,4716	11,4534	11,4926	10,9560	10,9559	10,8460	10,8820	11,2067	11,4331	11,4192	11,5206	11,4860	10,8923
M21	11,7415	11,6605	11,7211	11,1072	11,1844	11,0307	11,2556	11,3476	11,8245	11,7431	11,8371	11,8127	11,0297
M2B	11,5291	11,5198	11,5411	11,1300	11,0906	10,9519	11,0415	11,2084	11,5175	11,4921	11,6547	11,5727	10,9159
M2C	11,4610	11,4373	11,4946	10,9722	10,9566	10,8581	10,9372	11,1968	11,4781	11,4250	11,5404	11,4723	10,8748
M32	11,2036	11,4140	11,2832	10,8909	10,5601	10,6919	10,6347	10,9602	11,2955	11,5408	11,3934	11,3719	10,6754
M33	11,2268	11,4630	11,3206	10,9119	10,5696	10,7190	10,6520	10,9860	11,3360	11,5874	11,4116	11,3609	10,6620
M34	11,3505	11,6236	11,4887	11,0354	10,6455	10,8532	10,7706	11,1213	11,4942	11,7032	11,5057	11,4058	10,7345
M35	11,4172	11,6796	11,5082	11,0678	10,7733	10,9823	10,7645	11,0814	11,3891	11,4915	11,4752	11,3676	10,7837
M36	11,2092	11,4177	11,2822	10,8939	10,5563	10,6986	10,6339	10,9598	11,2953	11,5357	11,3869	11,3633	10,6749
M1A	11,0849	11,0652	11,2897	10,9164	10,6493	10,6897	10,7172	11,0200	11,0231	10,9984	11,1781	11,4294	10,7020
M11	11,1055	11,0902	11,3258	10,9651	10,6987	10,7750	10,7851	11,0483	11,1276	11,0945	11,2388	11,5103	10,7084
M20	11,2250	11,2258	11,4806	11,1285	10,8674	10,9950	10,9630	11,2289	11,4009	11,3153	11,3978	11,6945	10,8085
M1B	11,2465	11,2702	11,5149	11,2473	11,0415	11,0633	10,9593	11,1259	11,2932	11,3149	11,3052	11,6638	10,8346
M1C	11,0826	11,0614	11,2931	10,9160	10,6546	10,6935	10,7183	11,0223	11,0255	10,9974	11,1749	11,4209	10,7020
M42	10,6330	10,8507	11,0866	10,9100	10,7608	10,5915	10,6571	11,0563	11,1783	11,0608	11,0649	10,9429	10,5046
M43	10,6528	10,8970	11,1452	10,9672	10,8242	10,6568	10,7440	11,1332	11,2753	11,1572	11,1443	11,0228	10,5318
M44	10,8078	11,0526	11,3242	11,1552	11,0019	10,8675	10,9657	11,3729	11,4790	11,4102	11,3650	11,2546	10,6716
M45	10,8755	11,1490	11,3004	11,1428	10,9575	10,7429	10,8443	11,2074	11,3423	11,3905	11,3315	11,1931	10,6554
M46	10,6298	10,8484	11,0847	10,9102	10,7555	10,5921	10,6569	11,0537	11,1753	11,0579	11,0614	10,9457	10,5032
M52	12,4670	12,3794	12,2878	11,8078	11,6684	11,3615	11,6676	11,6761	11,5277	11,4079	11,5783	11,5901	11,6022
M53	12,3138	12,3364	12,2521	11,7993	11,5787	11,2914	11,5377	11,5784	11,6163	11,4627	11,6425	11,6236	11,5787
M54	12,0979	12,3111	12,2605	11,8525	11,5001	11,2500	11,4102	11,4672	11,8257	11,6192	11,8238	11,7430	11,6022
M55	12,2276	12,4934	12,2177	11,8399	11,4627	11,3418	11,3381	11,4086	11,5858	11,7027	11,7633	11,6544	11,5129
M56	12,4610	12,3750	12,2822	11,8021	11,6616	11,3650	11,6608	11,6670	11,5288	11,4110	11,5827	11,5910	11,6006
M62	11,9277	11,8260	11,8175	11,7465	11,4424	11,4558	11,5291	11,1156	11,2353	11,4198	11,5199	11,5461	11,5851
M64	12,2885	12,1363	12,1707	12,0971	11,8222	11,9229	12,0769	11,2281	11,3296	11,5394	11,7502	11,7817	11,7738
M65	11,9911	11,8662	11,8480	11,7879	11,4750	11,5573	11,6432	11,0142	11,1338	11,4646	11,5511	11,4984	11,5550
M66	11,9611	11,8038	11,8242	11,7446	11,4133	11,4868	11,5533	11,1006	11,1757	11,3643	11,4877	11,5149	11,5560
Mod.Per	47,6021	47,2250	45,8759	43,3612	39,9521	35,8091	31,4874	27,4332	23,3468	19,1516	15,3361	12,9712	32,8251

Fonte: O Autor (2021)

Tabela 24 – Desempenho percentual dos Modelos previsores nas primeiras 12h do período de Simulação para o Parque 01

NMAE em [%] dos modelos previsores de potência desenvolvidos para as primeiras 12h do horizonte													
Horizonte	1h	2h	3h	4h	5h	6h	7h	8h	9h	10h	11h	12h	Média 12h
M22	9,2181	10,8096	12,5053	13,8815	14,8361	14,7628	15,9408	15,0590	16,1343	17,6597	15,7509	15,0492	14,3006
M23	9,1397	10,7162	12,3955	13,7509	14,7809	14,6707	15,8431	15,0556	16,1127	17,5830	15,7019	15,0750	14,2354
M24	8,9776	10,5513	12,2314	13,4130	14,4576	14,3343	15,2479	14,7250	15,6023	17,0287	15,3604	14,9896	13,9099
M25	9,2455	10,7664	12,4762	13,8279	14,9074	14,8002	15,9058	15,1286	16,2181	17,7199	15,7493	15,0595	14,3171
M2A	7,7494	11,0443	12,4510	13,2057	13,1390	13,7387	14,2062	13,9479	13,7926	14,1447	14,5095	14,4660	13,0329
M21	8,3033	11,0150	12,3447	13,1034	13,0803	13,6794	14,1160	13,8916	13,7365	14,1050	14,4133	14,4679	13,0214
M2B	8,4984	10,9492	12,2003	13,0963	13,0426	13,6343	14,048	13,7767	13,646	13,9984	14,1832	14,3962	12,9558
M2C	8,2877	11,0582	12,4164	13,0950	13,0626	13,7334	14,1546	13,9639	13,7730	14,1186	14,4450	14,5166	13,0521
M32	8,1764	10,6552	12,8342	13,0720	13,6468	13,7170	15,1791	14,7622	14,6745	15,5788	15,0127	14,7173	13,5022
M33	8,8112	10,5585	12,2448	12,6155	13,3078	13,5358	14,2185	14,2256	14,0352	14,5289	14,1123	13,9375	13,0110
M34	8,1575	10,6614	12,8498	13,0938	13,6620	13,7320	15,2164	14,7784	14,7007	15,6126	15,0508	14,7420	13,5215
M35	8,3656	10,9753	13,2052	13,5042	14,0269	14,0489	15,6879	15,1759	15,1488	16,1643	15,5604	15,1727	13,9197
M36	8,1820	10,6669	12,8483	13,0862	13,6549	13,7227	15,1880	14,7730	14,6885	15,5994	15,0270	14,7316	13,5140
M1A	7,7923	10,5316	12,4521	12,6225	13,4560	13,7319	14,5596	14,8212	15,3471	15,0041	14,6808	14,6931	13,3077
M11	7,9996	10,5103	12,0075	12,8964	13,1900	13,7212	14,1386	14,3911	14,9392	14,6829	14,6993	14,6738	13,1542
M20	7,7985	10,5517	12,4742	12,6448	13,4840	13,7483	14,5812	14,8442	15,3723	15,0294	14,6933	14,7042	13,3272
M1B	8,3614	10,5374	12,4221	12,6589	13,4633	13,7652	14,5832	14,8070	15,3093	14,9579	14,6663	14,6750	13,3506
M1C	7,8014	10,5388	12,4639	12,6265	13,4623	13,7352	14,5660	14,8300	15,3606	15,0161	14,6910	14,7033	13,3163
M42	8,0200	10,4056	12,0088	13,0438	12,9531	13,7004	14,6725	14,3730	14,7215	14,5427	14,9081	14,9576	13,1923
M43	7,8823	10,0707	11,5518	12,5396	12,6192	12,9895	13,7524	13,4651	13,6573	13,5073	13,7456	13,8631	12,4703
M44	9,5924	11,6738	13,0095	13,9299	13,8683	14,5118	15,3368	15,2367	15,4564	15,3924	15,6676	15,6980	14,1145
M45	8,4764	10,8225	12,4076	13,4537	13,3902	14,1792	15,2146	14,9456	15,2507	15,0909	15,4532	15,5295	13,6845
M46	8,0318	10,4140	12,0236	13,0547	12,9553	13,7098	14,6936	14,3901	14,7437	14,5653	14,9363	14,9800	13,2082
M52	8,0481	10,5278	11,8675	12,2413	13,5299	14,2069	15,4681	16,4935	16,0185	16,7142	16,4532	15,5617	13,9276
M53	8,7245	10,5106	11,8742	12,8217	12,6449	13,1818	14,1534	14,8747	14,5547	14,8909	14,7374	14,2266	13,0996
M54	7,9286	10,4584	11,8377	12,2244	13,5555	14,2464	15,5075	16,5539	16,0688	16,7686	16,5070	15,6027	13,9383
M55	8,0216	10,6579	12,0900	12,3691	13,9638	14,7772	16,1096	17,2244	16,7106	17,4671	17,1562	16,1339	14,3901
M56	7,9393	10,4810	11,8333	12,2292	13,5449	14,2194	15,4777	16,5133	16,0322	16,7348	16,4749	15,5766	13,9214
M62	9,2763	10,8385	11,9220	12,6934	13,2932	13,8012	14,3454	14,7056	14,9108	15,0555	15,3453	14,9858	13,4311
M64	9,2439	10,8053	11,8767	12,6618	13,2360	13,7645	14,2924	14,6486	14,8554	14,9725	15,2597	14,9330	13,3791
M65	9,3273	10,8847	11,9726	12,7782	13,2963	13,7571	14,2117	14,4956	14,6379	14,7614	15,1354	14,8666	13,3437
M66	9,2747	10,8451	11,9162	12,7072	13,2944	13,8287	14,3730	14,7193	14,9128	15,0415	15,3446	15,0080	13,4388
Mod.Per	7,1155	11,4262	14,598	17,0890	19,3648	21,3564	23,0995	24,4703	25,4515	26,1619	26,6433	26,8313	20,3006

Fonte: O Autor (2021)

Tabela 25 – Desempenho percentual dos Modelos previsores nas primeiras 12h do período de Simulação para o Parque 02

NMAE em [%] dos modelos previsores de potência desenvolvidos para as primeiras 12h do horizonte													
Horizonte	1h	2h	3h	4h	5h	6h	7h	8h	9h	10h	11h	12h	Média 12h
M22	5,9200	7,2266	7,7587	8,3405	8,5913	8,8463	8,7657	8,8681	9,1398	9,0488	8,8334	8,8559	8,3496
M23	5,7000	7,1543	7,7106	8,4183	8,7832	9,1208	8,9893	9,1829	9,4751	9,2023	8,9690	9,0936	8,4833
M24	5,8215	7,0408	7,6311	8,2716	8,5257	8,8691	8,7610	8,9024	9,2570	9,1573	8,9361	9,0554	8,3524
M25	5,8902	7,1696	7,6990	8,2911	8,5598	8,8776	8,7796	8,8281	9,1329	9,0462	8,8585	8,8833	8,3347
M2A	5,5889	7,3724	7,8934	8,2307	8,2749	8,7215	8,7907	9,0381	8,7108	9,0160	9,2770	9,3171	8,3526
M21	5,6019	7,2603	7,8274	8,0737	8,2814	8,6050	8,7717	8,9573	8,8671	9,0678	9,4481	9,5926	8,3628
M2B	5,6610	7,2413	7,8320	8,0928	8,2022	8,6244	8,6860	8,9335	8,7397	9,0104	9,3458	9,4174	8,3155
M2C	5,5639	7,3016	7,8602	8,1620	8,2350	8,6531	8,7626	8,9728	8,7155	8,9951	9,2903	9,3010	8,3178
M32	5,5072	6,9631	7,5212	8,0299	8,1800	8,4328	8,5567	8,7591	8,9537	8,7946	8,8296	9,0672	8,1329
M33	5,4663	6,8747	7,4492	7,9459	8,1174	8,4102	8,4919	8,7133	8,9257	8,7816	8,8350	9,0857	8,0914
M34	5,5012	6,8141	7,3872	7,8929	8,1205	8,4741	8,4770	8,7379	8,9712	8,8545	8,8926	9,2018	8,1104
M35	7,0734	8,4481	9,1293	9,7272	10,0859	10,4319	10,5756	10,8258	11,1244	10,9953	11,0007	11,3392	10,0631
M36	5,5107	6,9586	7,5156	8,0287	8,1788	8,4326	8,5565	8,7562	8,9534	8,7980	8,8292	9,0723	8,1325
M1A	5,7391	7,3059	8,0617	8,2456	8,3889	8,4174	8,5784	8,7631	8,8523	8,8769	8,8827	8,9249	8,2531
M11	5,6618	7,2598	7,9946	8,2075	8,3216	8,3784	8,5291	8,7360	8,8641	8,8484	8,8716	8,9271	8,2167
M20	5,6261	7,2346	7,9344	8,2259	8,3084	8,3789	8,5212	8,7794	8,9569	8,8862	8,9400	9,0020	8,2328
M1B	5,7325	7,1426	7,8654	8,2287	8,3896	8,4207	8,5773	8,7821	8,9481	8,9743	9,0944	9,0387	8,2662
M1C	5,7342	7,2979	8,0547	8,2493	8,3922	8,4120	8,5760	8,7661	8,8558	8,8814	8,8872	8,9303	8,2531
M42	5,5293	6,9116	7,7137	8,0088	8,2879	8,3494	8,5032	8,6329	8,5568	8,7713	8,5224	8,4970	8,0237
M43	5,4957	6,8644	7,7114	7,9848	8,2558	8,3346	8,4709	8,6221	8,5536	8,7918	8,5328	8,5279	8,0121
M44	5,5543	6,8730	7,7709	8,0451	8,2701	8,3774	8,5000	8,6674	8,6193	8,8820	8,6514	8,6641	8,0729
M45	6,9719	8,4179	9,5278	9,8572	10,1349	10,3599	10,4959	10,6100	10,5387	10,8777	10,7135	10,7315	9,9364
M46	5,5271	6,9062	7,7121	8,0098	8,2822	8,3476	8,5022	8,6338	8,5616	8,7711	8,5230	8,4977	8,0229
M52	5,9498	7,6918	8,6851	9,0108	9,5248	9,6047	9,7251	9,6459	9,4734	9,7939	9,9366	9,7135	9,0630
M53	5,9705	7,6309	8,7075	9,0307	9,5641	9,6130	9,7262	9,5706	9,4115	9,7591	9,9540	9,6773	9,0513
M54	6,0835	7,5643	8,7815	9,1008	9,6678	9,6634	9,7689	9,5247	9,3579	9,7677	10,0156	9,6657	9,0802
M55	7,5051	9,1683	10,4009	10,9280	11,6174	11,7016	11,8850	11,7432	11,5023	11,7915	11,8763	11,8162	10,9947
M56	5,9512	7,6938	8,6843	9,0093	9,5290	9,5993	9,7237	9,6429	9,4724	9,7954	9,9368	9,7136	9,0626
M62	6,3870	8,2278	8,8222	9,1914	9,4279	9,6550	9,8127	9,7922	9,8673	9,9674	9,8934	9,6387	9,2236
M64	6,4172	8,0535	8,7666	9,1613	9,4635	9,6569	9,9283	9,9146	10,0674	10,1059	10,1251	9,7903	9,2876
M65	6,3909	8,0395	8,6587	9,0422	9,3139	9,5360	9,7966	9,7565	9,8905	9,9413	9,9098	9,6353	9,1593
M66	6,3458	8,1294	8,7782	9,1530	9,3895	9,5822	9,7964	9,7726	9,8662	9,9152	9,9066	9,6137	9,1874
Mod.Per	7,1153	12,3163	16,2977	19,9824	23,4200	26,7675	30,0876	33,0621	35,4866	36,8808	37,1942	37,4613	26,3393

Fonte: O Autor (2021)

Tabela 26 – Desempenho percentual dos Modelos previsores nas 12h finais do período de Simulação para o Parque 01

NMAE em [%] dos modelos previsores de potência desenvolvidos para 12h finais do horizonte													
Horizonte	13h	14h	15h	16h	17h	18h	19h	20h	21h	22h	23h	24h	Média 24h
M22	15,1516	15,7179	15,6700	14,6974	15,5392	15,0641	16,8622	15,5604	15,6833	15,7042	15,3870	14,2713	14,8715
M23	15,1220	15,7127	15,5834	14,5733	15,4701	14,9640	16,8632	15,4905	15,6488	15,6554	15,3054	14,1910	14,8085
M24	14,9532	15,4687	15,4564	14,3496	15,1429	14,6732	16,4877	15,1933	15,4092	15,3540	15,1334	14,0613	14,5251
M25	15,1814	15,7778	15,6344	14,6680	15,5474	15,0644	16,9581	15,5429	15,7034	15,7465	15,3762	14,2307	14,8848
M2A	14,4052	14,5724	14,7755	14,7655	15,4226	15,4174	14,3691	14,6317	14,5398	15,0453	15,6762	15,2188	13,9681
M21	14,3603	14,5073	14,7200	14,6235	15,2828	15,3130	14,2996	14,5753	14,5279	15,0164	15,6491	15,1726	13,9293
M2B	14,2887	14,4015	14,7058	14,4306	15,0048	15,1179	14,2655	14,4533	14,6171	14,9587	15,4967	15,1155	14,7380
M2C	14,4174	14,5563	14,7753	14,6626	15,3170	15,3538	14,3416	14,6139	14,5196	15,0293	15,6423	15,1984	13,9605
M32	15,3808	14,9110	15,5933	15,0129	15,3347	15,9949	15,1158	15,1205	14,8439	15,7716	15,1719	14,8759	10,1475
M33	14,3860	14,1855	14,4285	14,1082	14,2730	14,7090	14,3151	14,0815	14,0547	14,7226	14,3803	14,1818	9,6420
M34	15,4274	14,9426	15,6329	15,0442	15,3704	16,0395	15,1502	15,1585	14,8651	15,8058	15,1926	14,8949	14,4076
M35	15,9977	15,4376	16,1920	15,5070	15,8940	16,6762	15,5566	15,6844	15,2700	16,3619	15,6433	15,3393	14,8582
M36	15,3956	14,9260	15,6094	15,0285	15,3469	16,0118	15,1293	15,1367	14,8577	15,7907	15,1879	14,8902	14,3950
M1A	14,9095	15,1266	15,0678	15,2146	14,8904	14,9701	14,5622	14,6809	15,0249	14,7396	14,9554	15,0592	14,1206
M11	14,2901	14,9789	14,6099	14,7629	14,4850	14,3600	14,0876	13,9097	14,3008	14,1511	14,7161	14,7014	13,8001
M20	14,9299	15,1415	15,0915	15,2366	14,9069	14,9891	14,5782	14,7123	15,0586	14,7618	14,9710	15,0896	14,1414
M1B	14,8302	15,1043	15,0445	15,1742	14,8547	14,9295	14,5325	14,6303	15,0091	14,7149	14,9616	15,0464	14,1266
M1C	14,9232	15,1341	15,0790	15,2280	14,9055	14,9840	14,5725	14,6903	15,0351	14,7514	14,9599	15,0707	14,1304
M42	14,7458	14,7301	14,8754	15,1766	15,3413	15,5453	15,3983	15,6943	15,7661	16,0769	16,1424	16,6560	14,3523
M43	13,4631	13,6095	13,5170	14,0533	14,1125	14,1483	14,4839	14,5483	14,5761	14,4981	14,6525	15,0467	13,3481
M44	15,5671	15,5391	15,7238	15,8700	16,0339	16,2519	16,1358	16,5047	16,5914	16,8014	16,9459	17,3483	15,1953
M45	15,3522	15,2751	15,4918	15,7340	15,9418	16,1338	15,8678	16,2328	16,2873	16,6598	16,7601	17,2606	14,8838
M46	14,7680	14,7478	14,8941	15,1972	15,3530	15,5593	15,4152	15,7116	15,7890	16,0989	16,1681	16,6931	14,3706
M52	15,3987	14,6155	14,7346	14,8173	14,4463	14,3238	14,2746	14,7189	14,5411	15,3836	15,3389	15,1314	14,3690
M53	14,2578	14,0426	14,1537	14,2502	14,4833	14,2415	13,9575	14,0694	13,8608	14,3847	14,4394	14,4636	13,6583
M54	15,4364	14,6336	14,7597	14,8446	14,4613	14,3350	14,2935	14,7475	14,5631	15,4121	15,3650	15,1443	14,3857
M55	15,9636	14,9841	15,1359	15,2238	14,7797	14,6236	14,5970	15,1447	14,9451	15,8545	15,8220	15,5252	14,8034
M56	15,4143	14,6207	14,7419	14,8297	14,4580	14,3328	14,2861	14,7323	14,5511	15,3998	15,3569	15,1268	14,3711
M62	14,8492	14,7804	14,9362	14,7857	14,6042	14,7630	14,8142	14,7602	14,6525	14,3509	14,3244	14,3283	14,0468
M64	14,8109	14,7537	14,9305	14,7915	14,5881	14,7145	14,7664	14,7085	14,5906	14,2893	14,2783	14,2895	14,0026
M65	14,7913	14,8243	15,0321	14,9383	14,7443	14,8349	14,8714	14,7294	14,6027	14,2570	14,2507	14,2747	14,0115
M66	14,8793	14,8152	14,9932	14,8428	14,6302	14,7410	14,7695	14,7039	14,5809	14,2571	14,2494	14,2834	14,0421
Mod.Per	26,6174	26,1157	25,4685	24,5870	23,5223	22,2544	21,0228	19,8020	18,5036	17,306	16,4714	16,1877	20,8984

Fonte: O Autor (2021)

Tabela 27 – Desempenho percentual dos Modelos previsores nas 12h finais do período de Simulação para o Parque 02

NMAE em [%] dos modelos previsores de potência desenvolvidos para 12h finais do horizonte

Horizonte	13h	14h	15h	16h	17h	18h	19h	20h	21h	22h	23h	24h	Média 24h
M22	8,7392	8,8609	9,1217	9,0659	9,1062	9,0926	9,0057	9,1470	9,2183	9,4542	9,1049	8,8963	8,7087
M23	8,9120	8,9431	9,0430	8,8840	8,9925	8,9093	8,9183	9,0139	9,1481	9,4441	9,0333	9,0183	8,7525
M24	8,9076	8,9003	9,0770	8,9438	9,0755	8,9924	8,9490	9,0371	9,1496	9,3737	9,0371	8,9402	8,6922
M25	8,7445	8,8390	9,0704	8,9553	9,0212	9,0216	8,9486	9,0814	9,1710	9,3978	9,0428	8,8762	8,6744
M2A	9,1045	9,0900	9,1211	8,6952	8,6952	8,6079	8,6365	8,8942	9,0739	9,0629	9,1433	9,1159	8,6447
M21	9,3187	9,2544	9,3025	8,8152	8,8765	8,7545	8,9330	9,0061	9,3845	9,3199	9,3945	9,3751	8,7537
M2B	9,1501	9,1427	9,1596	8,8334	8,8020	8,6919	8,7631	8,8956	9,1408	9,1207	9,2497	9,1847	8,6634
M2C	9,0960	9,0772	9,1227	8,7081	8,6957	8,6176	8,6803	8,8863	9,1096	9,0674	9,1590	9,1050	8,6308
M32	8,8918	9,0587	8,9549	8,6436	8,3810	8,4857	8,4402	8,6986	8,9647	9,1593	9,0424	9,0253	8,4726
M33	8,9101	9,0976	8,9846	8,6603	8,3886	8,5071	8,4539	8,7191	8,9968	9,1964	9,0568	9,0166	8,4619
M34	9,0084	9,2251	9,1180	8,7582	8,4488	8,6136	8,5481	8,8265	9,1224	9,2883	9,1315	9,0522	8,5194
M35	11,1354	11,3914	11,2243	10,7947	10,5074	10,7113	10,4989	10,8080	11,1081	11,2079	11,1920	11,0871	10,5176
M36	8,8962	9,0617	8,9541	8,6459	8,3780	8,4910	8,4396	8,6983	8,9645	9,1553	9,0372	9,0185	8,4721
M1A	8,7975	8,7819	8,9601	8,6638	8,4518	8,4839	8,5057	8,7460	8,7485	8,7289	8,8715	9,0709	8,4936
M11	8,8139	8,8018	8,9887	8,7024	8,4911	8,5516	8,5596	8,7685	8,8314	8,8052	8,9197	9,1352	8,4987
M20	8,9087	8,9094	9,1116	8,8321	8,6249	8,7262	8,7008	8,9118	9,0483	8,9804	9,0459	9,2814	8,5781
M1B	5,7325	7,1426	7,8654	8,2287	8,3896	8,4207	8,5773	8,7821	8,9481	8,9743	9,0944	9,0387	8,2662
M1C	8,7957	8,7789	8,9628	8,6635	8,4560	8,4869	8,5066	8,7479	8,7504	8,7281	8,8689	9,0642	8,4936
M42	8,4389	8,6116	8,7989	8,6587	8,5403	8,4060	8,4580	8,7748	8,8717	8,7784	8,7817	8,6848	8,3370
M43	8,4546	8,6484	8,8454	8,7041	8,5907	8,4578	8,5270	8,8359	8,9487	8,8549	8,8447	8,7483	8,3586
M44	8,5776	8,7719	8,9874	8,8533	8,7317	8,6250	8,7029	9,0261	9,1103	9,0557	9,0198	8,9322	8,4695
M45	10,6072	10,8739	11,0216	10,8678	10,6871	10,4779	10,5767	10,9308	11,0624	11,1094	11,0518	10,9169	10,3925
M46	8,4364	8,6099	8,7974	8,6589	8,5361	8,4065	8,4578	8,7728	8,8693	8,7761	8,7789	8,6871	8,3359
M52	9,8945	9,8249	9,7523	9,3713	9,2607	9,0170	9,2600	9,2667	9,1490	9,0539	9,1891	9,1985	9,2081
M53	9,7729	9,7908	9,7239	9,3645	9,1894	8,9614	9,1569	9,1892	9,2193	9,0974	9,2401	9,2251	9,1894
M54	9,6015	9,7707	9,7306	9,4067	9,1271	8,9286	9,0557	9,1010	9,3855	9,2216	9,3839	9,3198	9,2081
M55	11,9259	12,1851	11,9162	11,5478	11,1799	11,0619	11,0584	11,1271	11,3000	11,4140	11,4731	11,3668	11,2288
M56	9,8897	9,8214	9,7478	9,3668	9,2552	9,0198	9,2546	9,2595	9,1498	9,0564	9,1927	9,1992	9,2069
M62	9,4664	9,3857	9,3790	9,3226	9,0813	9,0919	9,1501	8,8219	8,9169	9,0633	9,1428	9,1636	9,1945
M64	9,7527	9,6320	9,6593	9,6008	9,3827	9,4626	9,5848	8,9112	8,9917	9,1583	9,3256	9,3506	9,3443
M65	9,5167	9,4176	9,4032	9,3555	9,1072	9,1725	9,2406	8,7414	8,8364	9,0989	9,1676	9,1257	9,1706
M66	9,4930	9,3681	9,3843	9,3211	9,0582	9,1165	9,1693	8,8100	8,8696	9,0193	9,1172	9,1388	9,1714
Mod.Per	37,7794	37,4802	36,4094	34,4137	31,7080	28,4199	24,9900	21,7724	18,5292	15,1997	12,1715	10,2946	26,0517

Fonte: O Autor (2021)

Analisando os dados de MAE presentes nas Tabelas 20 e 22, inerentes ao parque 01, podem-se notar em duas perspectivas médias, uma inicial e outra final, o comportamento relativo ao erro de previsão de dados de potência. Os dados mostram que para as 12h iniciais do horizonte, 4 modelos apresentaram resultados relevantes, sendo os mesmos os modelos M21, M2B, M2C e M33.

Respectivamente ao citado no parágrafo anterior os resultados com relação ao índice de MAE foram: 9,1879MW; 9,1416MW; 9,2095MW e **9,1805MW**, desta forma o melhor modelo nas primeiras 12h é o modelo M33. Com relação a previsão da primeira hora quatro modelos também se destacaram sendo os mesmos, os modelos M1A, M20, M43 e Mper. Respectivamente ao citado no período anterior os resultados com relação ao índice de MAE da primeira hora foram: 5,4983MW; 5,5026MW; 5,5618MW e **5,0207MW**, desta forma o melhor modelo na hora 1 é o modelo MPer.

Esta característica foi explicada anteriormente quando se fez a análise das previsões de dados de vento, como o modelo é composto pelos modelos previsores de dados de vento, percebe-se então que o problema persiste, sendo esta característica da homogeneidade da base de dados do parque 01, conforme fora explicado anteriormente, porém este erro se modifica drasticamente nas horas posteriores, tendo o modelo MPer um péssimo desempenho médio ao longo do horizonte. Assim conforme fora feito anteriormente, analisando a segunda hora da previsão o modelo que se destaca é o M43 com MAE de **7,1043MW**.

Para a perspectiva completa do horizonte 24h, quatro modelos também se destacam performando com muita proximidade, sendo os mesmos os modelos M11, M53, M43 e M33 os quais apresentaram índices de MAE respectivos de 9,7374MW; 9,6373MW; **9,4184MW**, e 9,6420MW, desta forma a melhor previsão no horizonte é realizada pelo modelo M43.

Portanto após as análises relativas aos desempenhos apresentados nas Tabelas 20 e 22 fica comprovado que para o Parque 01 as novas heurísticas propostas também apresentaram um desempenho satisfatório em relação a previsão de dados de Potência. Especificamente as composições influenciadas pelas heurísticas AET1 e AET2, caso dos modelos M43, M45, e as composições com a heurística LSTM2, caso do modelo M53.

Analisando os dados de MAE presentes nas Tabelas 21 e 23, inerentes ao parque 02, podem-se notar em duas perspectivas médias, uma inicial e outra final, o comportamento relativo ao erro de previsão de dados de potência. Os dados mostram que para as 12h iniciais do horizonte, quatro modelos apresentaram resultados relevantes, sendo os mesmos os modelos M46, M43, M42 e M33.

Respectivamente ao citado no parágrafo anterior os resultados com relação ao índice de MAE foram: 10,1088MW; **10,0953MW**; 10,1098MW e 10,1952MW, desta forma o melhor modelo nas primeiras 12h é o modelo **M43**. Com relação a previsão da primeira hora quatro modelos também se destacaram sendo os mesmos, os modelos M32, M33, M34 e M43. Respectivamente ao citado no período anterior os resultados com relação ao índice de MAE da primeira hora foram: 6,9391MW; **6,8876MW**; 6,9315MW e 6,9246MW, desta forma o melhor modelo na hora 1 é o modelo **M33**.

Para a perspectiva completa do horizonte 24h, quatro modelos também se destacam performando com muita proximidade, sendo os mesmos os modelos M42, M43, M45 e **M46** os quais apresentaram índices de MAE respectivos de 10,5046MW; 10,5318MW; 10,6564MW, e **10,5032MW**, desta forma o melhor do horizonte é o modelo M43.

Portanto após as análises relativas aos desempenhos apresentados nas Tabelas 21 e 23 fica comprovado que para o Parque 02 as novas heurísticas propostas também apresentaram um desempenho satisfatório em relação a previsão de dados de Potência. Especificamente as composições influenciadas pelas heurísticas AET1 e AET2, caso do modelo M43, M46, M45, as composições com modelos Fuzzy Sugeno (ANFIS), caso do Modelo M46.

Analisando os dados de NMAE presentes nas Tabelas 24 e 26, inerentes ao parque 01, podem-se notar em duas perspectivas médias, uma inicial e outra final, o comportamento relativo ao erro de previsão de dados de potência. Os dados mostram que para as 12h iniciais do horizonte, quatro modelos apresentaram resultados relevantes, os modelos M2B, M21, M33 e M43.

Respectivamente ao citado no parágrafo anterior os resultados com relação ao índice de NMAE foram: 12,9558%; 13,0214%; 13,0110% e **12,4703%**, desta forma o melhor modelo nas primeiras 12h é o modelo M43. Com relação a previsão da primeira hora quatro modelos também se destacaram sendo os mesmos, os modelos M2A, M1A, M20 e Mper. Respectivamente ao citado no período anterior os resultados com relação ao índice de NMAE da primeira hora foram: 7,7494%; 7,7923%; 7,7985% e **7,1155%**, desta forma o melhor modelo na hora 1 é o modelo **MPer**.

Esta característica foi explicada anteriormente quando se fez a análise das previsões de dados de vento, como o modelo é composto pelos modelos previsores de dados de vento, percebe-se então que o problema persiste, sendo esta característica da homogeneidade da base de dados do parque 01, conforme fora explicado anteriormente, porém este erro se modifica drasticamente nas horas posteriores, tendo o modelo MPer um péssimo desempenho médio ao

longo do horizonte. Assim conforme fora feito anteriormente, analisando a segunda hora da previsão o modelo que se destaca é o **M43** com NMAE de **10,0707%**.

Para a perspectiva completa do horizonte 24h, quatro modelos também se destacam performando com muita proximidade, sendo os mesmos os modelos M32, **M33**, M43 e M53 os quais apresentaram índices de NMAE respectivos de 10,1475%; **9,6420%**; 13,3481%, e 13,6583%, desta forma a melhor previsão no horizonte é realizada pelo modelo M43.

Assim após as análises relativas aos desempenhos do NMAE apresentados nas Tabelas 24 e 26 fica comprovado que para o Parque 01 as novas heurísticas propostas também apresentaram um desempenho satisfatório em relação a previsão de dados de Potência. Especificamente as composições influenciadas pelas heurísticas AET1 e AET2, caso dos modelos M21, M33 e M43, e as composições com a heurística LSTM1, caso do modelo M53.

Verificando os dados de NMAE presentes nas Tabelas 25 e 27, referentes ao parque 02, podem-se notar em duas perspectivas médias, uma inicial e outra final, o comportamento relativo ao erro de previsão de dados de potência. Os dados mostram que para as 12h iniciais do horizonte, quatro modelos apresentaram resultados relevantes, sendo os mesmos os modelos M42, M43, M44 e M46.

Respectivamente ao citado no parágrafo anterior os resultados com relação ao índice de NMAE foram: 8,0237%; **8,0121%**; 8,0729% e 8,0229%, desta forma o melhor modelo nas primeiras 12h é o modelo **M43**. Com relação a previsão da primeira hora quatro modelos também se destacaram sendo os mesmos, os modelos M33, M34, M36 e M43. Respectivamente ao citado no período anterior os resultados com relação ao índice de NMAE da primeira hora foram: **5,4663%**; 5,5012%; 5,5107% e 5,4957%, desta forma o melhor modelo na hora 1 é o modelo **M33**.

Para a perspectiva completa do horizonte 24h, quatro modelos também se destacam performando também muito próximos, sendo os mesmos os modelos M1B, M42, M44 e M46 os quais apresentaram índices de NMAE respectivos de **8,2662%**; 8,3370%; 8,3586%, e 8,3359%, desta forma o melhor do horizonte é o modelo **M1B**.

Portanto após as análises relativas aos desempenhos do NMAE apresentados nas Tabelas 25 e 27 fica comprovado que para o Parque 02 as novas heurísticas propostas também apresentaram um desempenho satisfatório em relação a previsão de dados de Potência. Especificamente as composições influenciadas pelas heurísticas AET1 e AET2, caso dos modelos M21, M33 e M43, e as composições com a heurística LSTM1, caso do modelo M53.

As Tabelas 28 e 29 mostradas a seguir, exibem os ganhos relativos ao desempenho de cada modelo predictor de potência desenvolvido em comparação ao modelo da persistência, calculados pela Equação 7.5.

Tabela 28 – Aferição dos Ganhos percentuais relativos aos índices de MAE dos modelos previsores de potência atuantes nos parques 01 e 02 em relação ao modelo da persistência.

G(%)MAE	Parque 01		Parque 02	
Modelo	Ganho(%) 12h	Ganho(%) 24h	Ganho(%) 12h	Ganho(%) 24h
M22	29,5559	28,7450	68,2998	66,5864
M23	29,8771	29,0465	67,7924	66,4183
M24	31,4807	30,4046	68,2893	66,6497
M25	29,4750	28,6811	68,3565	66,7176
M2A	35,5493	32,9510	68,2884	66,8318
M21	35,8572	33,2593	68,2495	66,4134
M2B	36,1804	33,6538	68,4291	66,7599
M2C	35,7064	33,1099	68,4206	66,8851
M32	33,4892	31,0931	69,1224	67,4923
M33	35,9089	34,5257	69,2800	67,5331
M34	33,3942	30,9675	69,2080	67,3123
M35	31,4325	28,8088	68,9109	67,1625
M36	33,4305	31,0279	69,1239	67,4938
M1A	34,4470	32,3426	68,6662	67,4113
M11	35,2031	33,8779	68,8045	67,3918
M20	34,3507	32,2428	68,7431	67,0870
M1B	27,2877	26,4274	68,6165	67,0075
M1C	27,5810	26,6046	68,6662	67,4113
M42	35,0153	31,2323	69,5373	68,0124
M43	38,5715	36,0441	69,5810	67,9296
M44	30,4726	27,1933	69,3502	67,5039
M45	32,5907	28,6859	69,3023	67,5532
M46	34,9371	31,1447	69,5403	68,0167
M52	31,3934	31,1522	65,5916	64,6701
M53	35,4718	34,5577	65,6359	64,7417
M54	31,3403	31,0728	65,5262	64,6701
M55	29,1147	29,0716	66,0330	64,9420
M56	31,4241	31,1420	65,5928	64,6750
M62	33,8389	32,6964	64,9817	64,7222
M64	34,0951	32,9082	64,7388	64,1476
M65	34,2697	32,8655	65,2258	64,8138
M66	33,8012	32,7188	65,1191	64,8108

Fonte: O Autor (2021)

Como pode-se perceber na Tabela 28 o melhor modelo predictor de potência sobre análise das perspectivas do horizonte, inicial (12h) e total (24h) é respectivamente o Modelo **M43** em 12h com um ganho MAE percentual de **38,5715%** e **36,0441%**, para o parque 01. Já

considerando o Parque 02 o melhor modelo predictor de potência sobre análise das perspectivas do horizonte, inicial (12h) e total (24h) é respectivamente o Modelo **M43** em 12h com um ganho MAE percentual de **69,5810%** e o Modelo **M46** com **36,0441%**, ambos os ganhos calculados em relação ao modelo da persistência.

Tabela 29 – Aferição dos Ganhos percentuais relativos aos índices de NMAE dos modelos previsoires de potência atuantes nos parques 01 e 02 em relação ao modelo da persistência.

<i>G(%)NMAE</i>	Parque 01		Parque 02	
Modelo	Ganho(%) 12h	Ganho(%) 24h	Ganho(%) 12h	Ganho(%) 24h
M22	29,5559	28,8390	68,2999	66,5715
M23	29,8771	29,1405	67,7923	66,4033
M24	31,4805	30,4966	68,2892	66,6348
M25	29,4746	28,7754	68,3564	66,7031
M2A	35,8006	33,1619	68,2885	66,8171
M21	35,8572	33,3475	68,2497	66,3987
M2B	27,4013	29,4779	68,4293	66,7454
M2C	35,7060	33,1982	68,4206	66,8705
M32	33,4888	51,4437	69,1226	67,4777
M33	35,9084	53,8625	69,2801	67,5188
M34	33,3937	31,0588	69,2080	67,2981
M35	31,4322	28,9027	61,7944	59,6280
M36	33,4307	31,1191	69,1241	67,4797
M1A	34,4469	32,4321	68,6662	67,3971
M11	35,2030	33,9658	68,8044	67,3776
M20	34,3508	32,3326	68,7433	67,0728
M1B	34,2356	32,4034	68,6165	68,2700
M1C	34,4045	32,3853	68,6662	67,3971
M42	35,0154	31,3235	69,5372	67,9982
M43	38,5719	36,1286	69,5812	67,9153
M44	30,4726	27,2896	69,3504	67,4896
M45	32,5908	28,7802	62,2754	60,1082
M46	34,9370	31,2359	69,5402	68,0025
M52	31,3933	31,2435	65,5914	64,6545
M53	35,4720	34,6443	65,6358	64,7263
M54	31,3406	31,1636	65,5261	64,6545
M55	29,1150	29,1649	58,2575	56,8980
M56	31,4238	31,2335	65,5929	64,6591
M62	33,8390	32,7853	64,9816	64,7067
M64	34,0952	32,9968	64,7386	64,1317
M65	34,2696	32,9542	65,2257	64,7985
M66	33,8011	32,8078	65,1191	64,7954

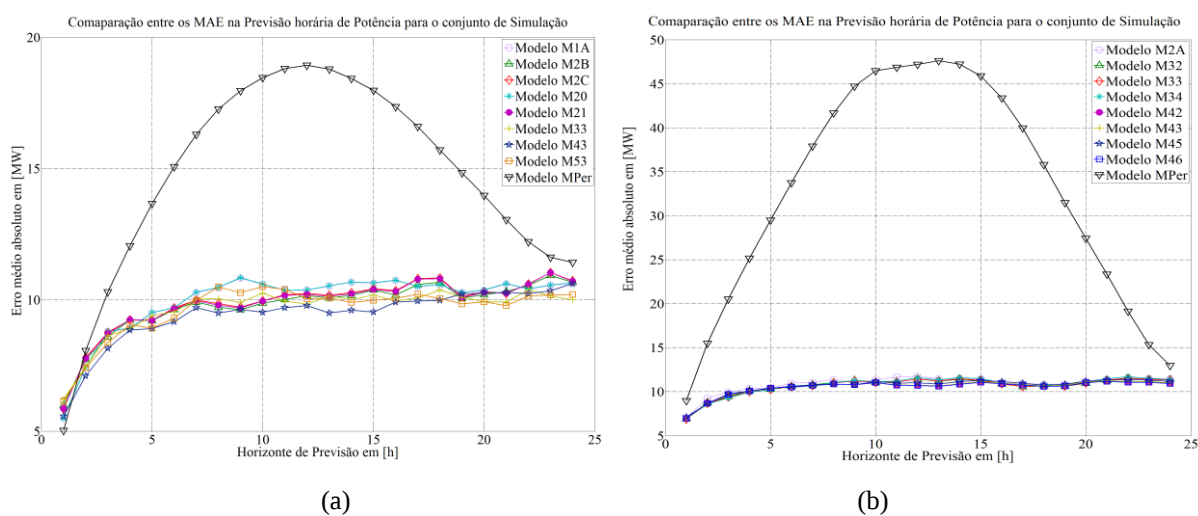
Fonte: O Autor (2021)

De forma semelhante, pode-se perceber na Tabela 29 o melhor modelo predictor de potência sobre análise das perspectivas do horizonte, inicial (12h) e total (24h) são respectivamente os Modelos **M43** em 12h com um ganho NMAE percentual de **38,5719%** e o Modelo **M33** com **53,8625%**, para o parque 01. Já considerando o Parque 02 os melhores modelos predictors de potência sobre análise das perspectivas do horizonte, inicial (12h) e total (24h) são respectivamente os Modelos **M43** em 12h com um ganho NMAE percentual de **69,5812%** e o Modelo **M1B** com **68,2700%**, ambos os ganhos NMAE calculados em relação ao modelo da persistência.

Com o intuito de mostrar ao longo do horizonte o que fora analisado pelas Tabelas 20 a 27 de acordo com os índices de MAE e NMAE atestados, escolheram-se os melhores modelos em termos de resultados para a 1h de previsão, a perspectiva inicial da previsão (12h) e para todo o horizonte previsão (24h). Pelas análises realizadas anteriormente no Parque 01 destacaram-se os Modelos: M2B, M2C, M20, M21, M33, M43, M53 e o modelo da Persistência. Já no Parque 02 o destaque foi para os Modelos M2A, M32, M33, M34, M42, M43, M45, M46 e o modelo da Persistência.

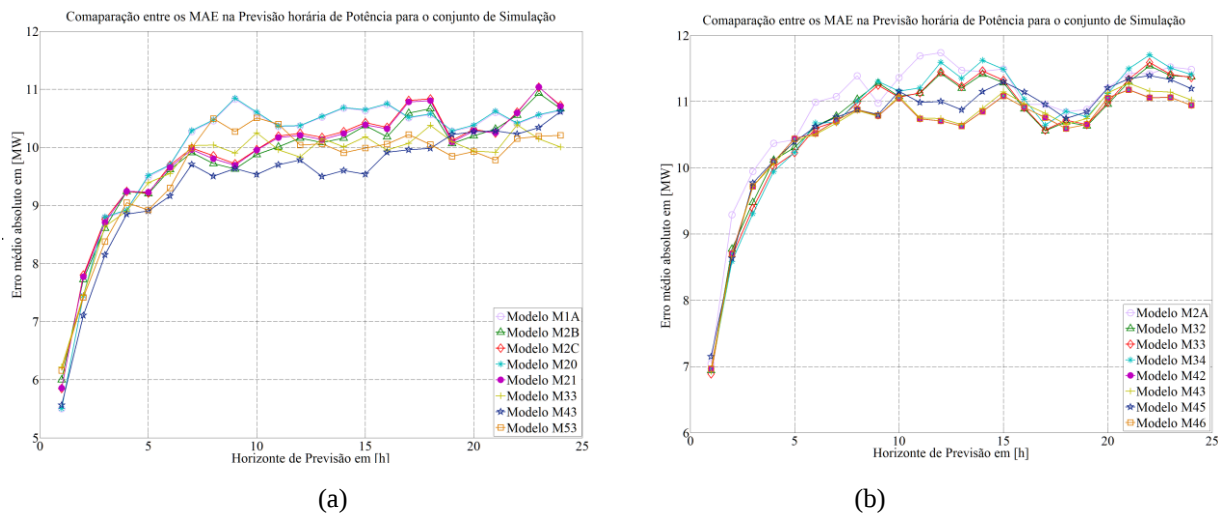
As Figuras 114 a 117 mostram as distribuições do erro médio absoluto MAE e do erro médio absoluto normalizado NMAE ao longo do horizonte de previsão utilizando em parte da comparação o modelo da persistência.

Figura 114 – Comparação do MAE de Potência em (MW) do Conjunto de Simulação para os modelos predictors de dados de vento com o modelo da Persistência, atuantes (a) no Parque 01 e (b) no Parque 02.



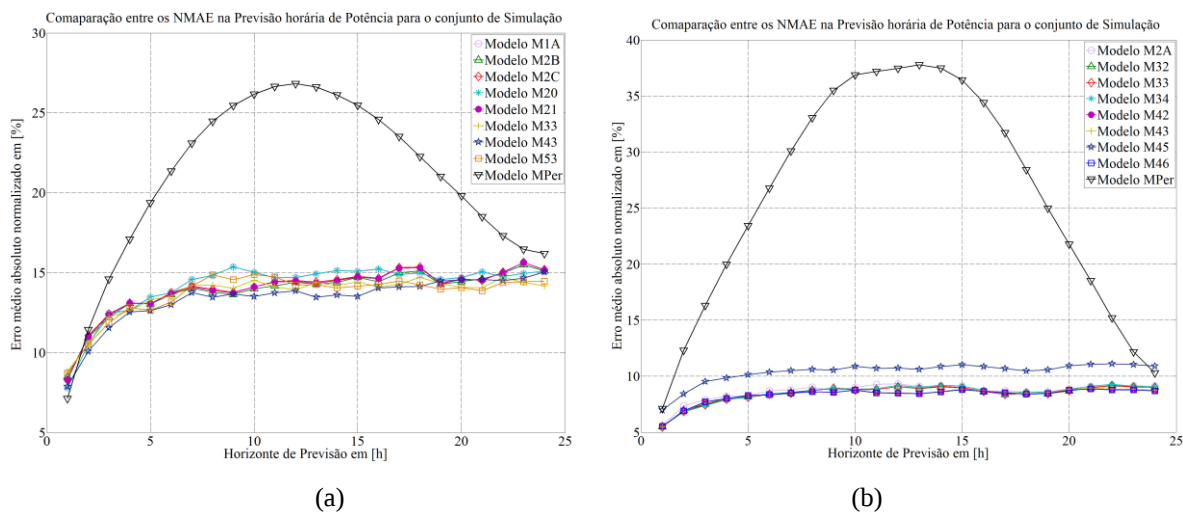
Fonte: O Autor (2021).

Figura 115 – Comparação do MAE de Potência em (MW) do Conjunto de Simulação para os modelos previsores de dados de vento sem o modelo da Persistência, atuantes (a) no Parque 01 e (b) no Parque 02.



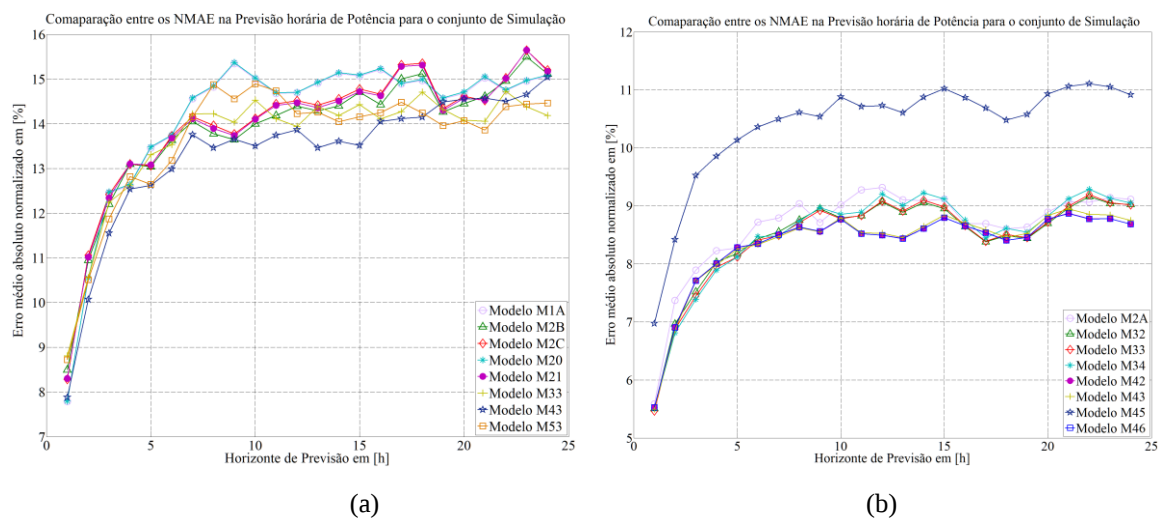
Fonte: O Autor (2021).

Figura 116 – Comparação do NMAE(%) de Potência do Conjunto de Simulação para os modelos previsores de dados de vento com o modelo da Persistência atuantes (a) no Parque 01 e (b) no Parque 02.



Fonte: O Autor (2021).

Figura 117 – Comparação do NMAE(%) de Potência do Conjunto de Simulação para os modelos previsores de dados de vento sem o modelo da Persistência atuantes (a) no Parque 01 e (b) no Parque 02.



Fonte: O Autor (2021).

8 CONCLUSÕES E TRABALHOS FUTUROS

Esta seção apresenta conclusões e comentários em resumo a respeito dos modelos e heurísticas tratados nesta tese, além de sugestões de implementações e análise para trabalhos futuros que tomem como partida este trabalho.

8.1 Conclusões

Modelos previsores de geração eólica de curto prazo (horizonte de 24h) foram desenvolvidos com base em estudos relativos à previsão de dados de ventos e de geração eólica. Para isso, foram necessários estudos e aplicações que incidiram no entendimento e no desenvolvimento de heurísticas que utilizaram ferramentas de IA. O objetivo destas ferramentas é melhorar o desempenho dos modelos de IA desenvolvidos anteriormente em ALBUQUERQUE, (2015).

Os estudos das novas heurísticas contribuíram de forma significativa no desenvolvimento das novas técnicas de aprendizagem profunda propostas, fato comprovado pela melhora eficaz no desempenho dos modelos previamente desenvolvidos. Tais estudos permitem ao trabalho contribuir no estabelecimento de procedimentos para escolha de modelos previsores de dados de vento e geração, de acordo com as características da previsão, base de dados, horizonte sobre estudo, e ferramenta de IA a ser aplicada.

No que se refere as técnicas de *deep learning* apresentadas e discutidas neste trabalho, as mesmas se mostraram eficazes na melhora da exatidão seja de forma parcial ou total no que se refere ao tipo de modelagem sobre estudo e também enfatizando às bases de dados referentes aos parques eólicos sobre estudo.

No capítulo referente a análise e discussão de resultados ficou claramente demonstrada a eficiência em acurácia das técnicas e heurísticas propostas, não havendo, porém, melhor heurística, mas sim uma melhor composição das heurísticas através da combinação entre os modelos desenvolvidos.

Inicialmente, na proposição da previsão de dados relativos ao vento, os novos modelos propostos se destacaram dependendo, no entanto, da faixa considerada ao horizonte de previsão. Desta forma foi demonstrado que, em perspectiva inicial (12 primeiras horas), o destaque foi para os modelos previamente desenvolvidos em ALBUQUERQUE (2015) com destaque para o Modelo W2(RNA2) seguido de perto pelo novo modelo W7(LSTM2-AET2) composto com as novas heurísticas AE e LSTM, relativas ao parque 01.

No caso do parque 02, nesta mesma perspectiva inicial, os modelos da nova heurística AE já se diferenciam, destacando-se em relação aos modelos prévios, sendo o melhor modelo o W3(RNA1-AET1), seguido de perto pelo modelo W4(RNA1-AET2).

Analisando o horizonte completo (24h) a vantagem é toda dos novos modelos propostos, superando os modelos prévios, sendo o melhor modelo para o parque 01 o W7 e o modelo W3 para o parque 02. Outro ponto bastante interessante é com relação a previsão de dados de direção do vento, tendo o erro uma melhora significativa, porém os valores para um dos parques (parque 02), não superam o modelo da persistência, mas são os que mais se aproximam. Este destaque da persistência é devido a uma característica da base de dados explicada no capítulo anterior. Tendo o modelo W7 os melhores desempenhos quando comparado com o modelo da persistência W8 para ambos os parques.

Em ambos os parques foi atestado que as novas heurísticas propostas AE e LSTM, se destacaram com relação a resiliência de desempenho ao longo do horizonte sendo assim uma boa solução quanto a precisão da previsão quando se avalia o médio e longo prazo do horizonte. Caso exemplificado pelo modelo W7 tanto em questão de velocidades quanto em questão de direções em ambos os parques.

Os modelos de curva de potência desenvolvidos, também utilizaram duas das novas heurísticas propostas: a AET#, e a SUGENO-ANFIS, esta última, propicia um pré-treinamento (Inferência Pré estabelecida), uma das principais contribuições desta tese, definida da conversão de um sistema MAMDANI sintonizado em um sistema SUGENO inicial, através da transformada *mam2sug*, que é posteriormente sintonizado pelo ANFIS. Esta contribuição gerou nos modelos de CP FIS desenvolvidos uma melhor eficácia e pode ser considerada como a característica inicial de aprendizagem profunda, pois o FIS final é proposto segundo uma disposição melhorada a partir de uma perspectiva inicial menos desordenada do que a perspectiva inicial do histórico real.

Os resultados relativos aos modelos de CP desenvolvidos demonstraram que os melhores modelos de CP foram os modelos os ligados as novas heurísticas destacando-se respectivamente por índice de aferição. Sendo os destaques para o parque 01 em MAE e NMAE o modelo (E), CP2RNA AET2, e em RMSE o modelo (K), CP2 SUGENO-ANFIS. Já para o parque 02 de forma semelhante os destaques dos mesmos índices de desempenho foram todos para o modelo (CP2 SUGENO-ANFIS).

Os previsores de geração compostos pelos modelos desenvolvidos de previsão de dados de vento e pelos modelos de curvas de potência estabeleceram um levantamento de competitividade entre as heurísticas propostas ao longo do horizonte. Isto pôde ser notado

pelo número de modelos relativo ao número de combinações que foram testadas nesta tese. Além disso, devido ao grande número de modelos previsores de geração propostos a pesquisa é evidenciada sobre o comportamento dos modelos ao longo do horizonte atestando a indicação em termos de desempenho para cada modelo desenvolvido.

No caso do **parque 01** o destaque foi para o **Modelo 43**, com um ganho MAE (%) percentual de **36,0441%** e para o **parque 02** o destaque foi para o **Modelo 46** com um ganho MAE (%) de **68,0167%**, ganhos estes relativos ao MAE do modelo da Persistência.

Desta forma verifica-se novamente o bom desempenho das novas heurísticas propostas AET1 e SUGENO ANFIS, que combinadas adequam bem os valores de dados de vento previstos a o modelo de curva de potência aproximada, respondendo assim com um valor previsto de potência aproximado pelo modelo de CP, com uma precisão considerável quando comparado com os modelos propostos previamente em ALBUQUERQUE (2015).

Assim nesta tese, as novas propostas foram efetivas em seu papel, otimizando o desempenho da previsão de geração adequada por uma previsão de dados eólicos e impactando positivamente em todas as modelagens propostas. Tendo essas novas técnicas, técnicas compostas pelos modelos de IA desenvolvidos e as novas heurísticas propostas, papel fundamental na pesquisa da melhora no desempenho da previsão de geração eólica fundamentada por aplicações de ferramentas de inteligência artificial e de Deep Learning.

8.2 Propostas para trabalhos futuros

Como próximos Passos de investigação das questões discutidas nesta tese (Trabalhos Futuros), tem-se:

- Investigar o desempenho dos modelos *deep learning* LSTM com outros tipos de Heurísticas LSTM, enfatizando as abordadas nesta tese como o modelo LSTM GRU, LSTM com acoplamento de entradas e LSTM com recorrência independente;
- Investigar o comportamento de um modelo de LSTM implementado na modelagem da curva de potência, criando uma arquitetura composta por redes de Aprendizagem Profunda distintas.
- Desenvolver alterações nos algoritmos que melhorem a percepção de aprendizagem das redes por variável de entrada, modificando os algoritmos

ADAM, SGDM, RMSprop.

- Aproveitar o caráter composto entre os modelos previsores de dados de vento e geração ilustrados nesta tese e desenvolver um modelo previsor auto seletivo que adeque o melhor modelo de acordo com a melhor perspectiva do horizonte. Assim, dependendo da hora sobre avaliação de previsão, o modelo principal faria a escolha de qual modelo adotar.

REFERÊNCIAS

- ACKERMANN, T.; SDER, L.; “Wind energy technology and current status: A review.” *Renew. Sustain. Energy Rev.* vol. 4. no. 4. pp. 315–374. 2000.
- ACKLEY, D. H.; HINTON, G. E.; SEJNOWSKI, T. J. “A Learning Algorithm for Boltzmann Machines”. In *Cognitive Science* 9, 147-169 (1985).
- AFONSO MARQUES, A. C.;. *Comparação Entre Técnicas de Inteligência Artificial e Funções de Pedotransferência para Estimar a Condutividade Hidráulica Saturada de uma Área do Agreste Pernambuco*. 2014. Tese (Doutorado em Engenharia Elétrica) –. Centro de Tecnologia e Geociências, Universidade Federal de Pernambuco, Recife.
- AGÊNCIA NACIONAL DE ENERGIA ELÉTRICA - ANEEL, Banco de Informações de Geração-Capacidade de Geração do Brasil.
Disponível em: <http://www.aneel.gov.br/siga>.
Acesso em: 26 fev. 2021.
- AKDAĞ, S. A.; GÜLER, Ö.; “Comparison of Wind Turbine Power Curve Models”. In *International Renewable Energy Congress November 5-7, 2010*, pages 215-219. Sousse, Tunisia, ID161/ ©IREC2010.
- ALBADI, M. H.; EL-SAADANY, E. F.; “Wind Turbines Capacity Factor Modeling—A Novel Approach”. In *IEEE Transactions on Power Systems*, Vol. 24, N3. August 2009.
- ANSOFF, H. Igor. *Estratégia empresarial*. São Paulo: McGraw-Hill, 1977.
- ASSOCIAÇÃO BRASILEIRA DE ENERGIA EÓLICA – ABEEOLICA (2018).
Disponível em: <http://abeeolica.org.br/noticias/brasil-e-o-oitavo-pais-do-mundo-em-producao-de-energia-eolica-epoca-negocios/>
Acesso em: 02 jul. 2018.
- BAPTISTA, D.; DIAS, F. M. “Energy production forecasting for a wind farm composed of turbines with different features” in *2017 International Conference in Energy and Sustainability in Small Developing Economies (ES2DE)*, 10-12 July 2017, Funchal, Portugal.
- BARRY, D., October 25, 2006. The Irish experience. In: *Talk on the Workshop for Best Practice in the Use of Short-term Forecasting of Wind Power*, Delft (NL).
- BENGIO, Y.; SIMARD, P. and FRASCONI, P.; Learning long-term dependencies with gradient descent is difficult. *IEEE Transactions on Neural Networks*, 5(2):157–166, 1994.
- BENGIO, Y.; LAMBLIN, P.; POPOVICI, D.; LAROCHELLE, H.; "Greedy layer-wise training of deep networks." *Advances in neural information processing systems* 19 (2007): 153.
- BENGIO, Y.; “Deep learning of representations for unsupervised and transfer learning” *Unsupervised and Transfer Learning Challenges in Machine Learning*, Volume 7 (2012): 19.

BESSA, R.J., MATOS, M.A., COSTA, I.C., BREMERMAN, L., FRANCHIN, I.G., PESTANA, R., MACHADO, N., WALDL, H-P.; WICHMANN, C., October 2012. Reserve setting and steady-state security assessment using wind power uncertainty forecast: a case study. *IEEE Transactions on Sustainable Energy* 3 (4), 827 e 837.

BEYER, H.G., HEINEMANN, D., MELLINGHOFF, H., MÖNNICH, K., WALDL, H.-P., March 1e5, 1999. Forecast of regional power output of wind turbines. In: *Proceedings of the European Wind Energy Conference, Nice, France, ISBN 1 902916 00 X*, pp. 1070 e 1073.

BITAR, E. Y.; RAJAGOPAL, R.; KHARGONEKAR, P. P.; “Bringing Wind Energy to Market”. In *IEEE TRANSACTIONS ON POWER SYSTEMS*, VOL. 27, NO. 3, August 2012.

BOSSANYI, E.A., 1985. Short-term wind prediction using Kalman filters. *Wind Engineering* 9 (1), 1 e 8.

BOULARD, H.; KAMP, Y., "Auto-association by multilayer perceptrons and singular value decomposition. “*Biological cybernetics* 59.4-5 (1988): 291-294.

CABEZON, D., MARTI, I., SAN ISIDRO, M.J., PEREZ, I., 2004. Comparison of methods for power curve modeling. In: *Proceedings to Global Windpower 2004, Chicago (USA)*.

CARRILLO, C.; MONTAÑO, A. F. O.; CIDRÁS, J.; DÍAZ-DORADO, E.; “Review of power curve modelling for wind turbines”. *Renewable and Sustainable Energy Reviews* 21 (2013) 572–581.

CHO, K.; MERRIENBOER, V. B.; BAHDANAU, D. and BENGIO, Y. “On the properties of neural machine translation: Encoder-decoder approaches.” *arXiv preprint arXiv:1409.1259*, 2014.

CIRESAN, D. C. et. al., "Deep, big, simple neural nets for handwritten digit recognition." *Neural computation* 22.12 (2010): 3207-3220.

COLAK, I.; SAGIROGLU, S.; YESILBUDAK, M.; “Data mining and wind power prediction: A literature review”. *Renewable Energy* 46 (2012) 241–247.

COLLINS, J., PARKES, J., TINDAL, A., 2009. Short term forecasting for utility-scale wind farms – the power model challenge. *Wind Engineering* 33 (3), 247 e 258.

COLLOBERT, R. et. al.; "Natural language processing (almost) from scratch." *The Journal of Machine Learning Research* 12 (2011): 2493-2537.

COSTA, A., CRESPO, A., NAVARRO, J., LIZCANO, G., MADSEN, H., FEITOSA, E., 2007. A review on the young history of the wind power short-term prediction. *Renewable and Sustainable Energy Reviews* 12 (6), 1725e1744.

DAYAN, P. et. al., "The helmholtz machine." *Neural computation* 7.5 (1995): 889-904.

DE ALBUQUERQUE, J. C.; *Avaliação de curvas de potência em modelos de previsão de geração eólica em curto prazo*. 2015. Dissertação (Mestrado em Engenharia Elétrica) – Centro de Tecnologia e Geociências, Universidade Federal de Pernambuco, Recife.

DE AQUINO, R. R. B., DE ALBUQUERQUE, J. C., NETO, O. N., FERREIRA, A. A., NETO, A. C., LYRA, M.M., CARVALHO JR., M. A. "Assessment of power curves in models of wind power forecasting." in Neural Networks (IJCNN). The 2016 International Joint Conference on. vol. no. pp.3915-3922. 24-29 Jul. 2016.

DE AQUINO, R. R. B.; DE ALBUQUERQUE, J. C.; NETO, O. N.; LIRA, M. M. S.; FERREIRA, A. A.; CODECEIRA N. A. Influence of Deep Learning on Precision Improvement in Predictive Models of Wind Power Generation. In: International Conference on Computational Science and Computational Intelligence, 2017, Las Vegas, USA.

DUTTON, A.G., KARINIOTAKIS, G., HALLIDAY, J.A., NOGARET, E., 1999. Load and wind power forecasting methods for the optimal management of isolated power systems with high Wind penetration. *Wind Engineering* 23 (2), 69 e 87.

ERNST, B., OAKLEAF, B., AHLSTROM, M.L., LANGE, M., MEØHRLÉN, C., LANGE, B., FOCKEN, U., ROHRIG, K., 2007. Predicting the wind. *IEEE Power & Energy Magazine* 5 (6), 78 e 89.

FELLOWS, A., HILL, D., September 10 e 14, 1990. Wind and load forecasting for integration of Wind power into a meso-scale electrical grid. In: *Proceedings of the European Community Wind Energy Conference*, pp. 636 e 640. Madrid, Spain.

FILDES, R. et. al., "Forecasting and operational research: a review." *Journal of the Operational Research Society* 59.9 (2008): 1150-1172.

FOCKEN, U., LANGE, M., WALDL, H.-P., June 2e6, 2001a. Previento e a wind power prediction system with an innovative upscaling algorithm. In: *Proceedings of the European Wind Energy Conference*, Copenhagen, Denmark, ISBN 3-936338-09-4, pp. 826 e 829.

FOCKEN, U., JAHN, J., SCHALLER, M., October 14 e 15, 2009. Transformer congestion forecast based on highly localized wind power. In: *Talk on the 8th International Workshop on Large-Scale Integration of Wind Power into Power Systems as Well as on Transmission Networks for Offshore Wind Farms*, Bremen (DE).

GERS, F., A.; SCHMIDHUBER, J. and CUMMINS, F.; Learning to forget: Continual prediction with LSTM. *Neural Computation*,12(10):2451–2471, 2000.

GERS, F., A.; SCHRAUDOLPH, N., N.; SCHMIDHUBER, J.; "Learning Precise Timing with LSTM Recurrent Networks." *On Journal of Machine Learning Research* 3 (2002) 115-143.

GIEBEL, G., 2001. On the Benefits of Distributed Generation of Wind Energy in Europe. PhD thesis from the Carl von Ossietzky Universitat Oldenburg. *Fortschr. -Ber. VDI Reihe 6* Nr. 444. Düsseldorf, VDI Verlag, ISBN 3-18-344406-2.

GIEBEL, G., KARINIOTAKIS, G. N., BROWNSWORD, R. The State of the art in Short-Term Prediction of Wind Power. *Literature Overview*. 36 p. 2003.

Disponível:

<http://orbit.dtu.dk/files/128933990/GGiebelEtAl_StateOfTheArtInShortTermPrediction_ANEMOSplus_2011.pdf>. Acesso em: 02 jul. 2018.

GIEBEL, G., CLINE, J.; FRANK, H.P., SHAW, W., PINSON, P., HODGE, B.-M., KARINIOTAKIS, G., MADSEN, J., MÖHRLEN, C., 2016. Wind power forecasting: IEA Wind Task 36 & future research issues. *Journal of Physics: Conference Series* 753 (3), 032042 (Online).

GILL, S.; STEPHEN, B.; GALLOWAY, S. J.; “Wind Turbine Condition Assessment Through Power Curve Copula Modeling.” in *IEEE Transactions on Sustainable Energy*. Vol. 3. N 1. January 2014.

GLOBAL WIND STATISTIC – GLOBAL WIND ENERGY CONCIL, (2017). Arquivo disponível em: <http://gwec.net/global-figures/graphs/> Acesso em: 02 jul. 2018.

GOTTSCHALL, J.; PEINKE, J. “How to improve the estimation of power curves for wind turbines”. In *IOP Publishing Environmental Research Letters* 3 (2008) 015005 (7pp).

HAYKIN, S. *Redes neurais: princípios e prática*. 2ª ed. Porto Alegre: Bookman, 2001.

HINTON, G. E. and SEJNOWSKI, T. J. (1986) “Learning and relearning in Boltzmann machines”. In Rumelhart, D. E. and McClelland, J. L., editors, *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*. Volume 1: Foundations, MIT Press, Cambridge, MA.

HINTON, G. "Training products of experts by minimizing contrastive divergence." *Neural computation* 14.8 (2002): 1771-1800.

HINTON, G.; OSINDERO, S. and TEH, Y-W. "A fast-learning algorithm for deep belief nets". *Neural computation* 18.7 (2006): 1527-1554.

HOCHREITER, S.; *Untersuchungen zu dynamischen neuronalen Netzen*. Diploma thesis, Institut für Informatik, Lehrstuhl Prof. Brauer, Technische Universität München, 1991. URL <http://www7.informatik.tu-muenchen.de/~Ehochreit>.

HOCHREITER, S. and SCHMIDHUBER, J.; “Long short-term memory. *Neural Computation*”, 9(8):1735–1780, 1997.

HRASKO, R. & PACHECO, A. & KROHLING, R. (2015). “Information Technology and Quantitative Management (ITQM 2015) Time Series Prediction using Restricted Boltzmann Machines and Backpropagation”. *Procedia Computer Science*. 55. 10.1016/j.procs.2015.07.104.

IEC61400-12-1. Wind turbines - Part 12-1: power performance measurements of electricity producing wind turbines – 2005.

JAMES WILCZAK, et al., October 2015. The wind forecast improvement project (WFIP): a public-private partnership addressing wind energy forecast needs. AMS, Bulletin of the

American Meteorology Society 96 (10). <http://dx.doi.org/10.1175/BAMS-D-14-00107.1>.

JANG, J.S. ANFIS: Adaptive network based fuzzy inference systems. *IEEE Transactions on Systems, Man and Cybernetics*, v.23, n.3:665-685, 1993.

JANG, JYH-SHING R. & CHUEN-TSAI SUN (1995), *Neuro-Fuzzy Modeling and Control*, *Proceedings of the IEEE*, 83(3), 378–406.

JANG, SUN, MIZUTANI (1997) – *Neuro-Fuzzy and Soft Computing* – Prentice Hall, pp 335–368, ISBN 0-13-261066-3.

JANGAMSHETTI, S. H.; RAU, V. G.; “Normalized Power Curves as a Tool for Identification of Optimum Wind Turbine Generator Parameters.” In *IEEE Transactions on Energy Conversion*, Vol. 16, N3. September 2001.

JIN, T.; TIAN, Z.; “Uncertainty Analysis for Wind Energy Production with Dynamic Power Curves”. in *Probabilistic Methods Applied to Power Systems (PMAPS)*. 2010 IEEE 11th International Conference on. vol. no. pp.745-750. 14-17 Jun. 2010.

KARINIOTAKIS, G.N., December 1996. *Contribution to the Development of a Control System for Autonomous Wind-Diesel Power Systems (Ph.D. thesis in Energetics)*. Ecole Des Moines de Paris, France.

KARINIOTAKIS, G.N., STAVRAKAKIS, G.S., NOGARET, E.F., December 1996. Wind power forecasting using advanced neural network models. *IEEE Transactions on Energy Conversion* 11 (4), 762 e 767.

KARINIOTAKIS, G.N., NOGARET, E., DUTTON, A.G., HALLIDAY, J.A., ANDROUTSOS, A., March 1 e 5, 1999. Evaluation of advanced wind power and load forecasting Meghods for the optimal management of isolated power systems. In: *Proceedings of the European Wind Energy Conference*, Nice, France, ISBN 1 902916 00 X, pp. 1082 e 1085.

KARINIOTAKIS, *Renewable Energy Forecast*, Woodhead Publishing Series in Energy, Copyright © 2017 Elsevier Ltd. All rights reserved. ISBN: 978-0-08-100505-7

KHALFALLAHA, M. G.; KOLIUBB, A. M.(2007a); “Wind turbines power curve variability” in *The Ninth Arab International Conference on Solar Energy (AICSE-9)*, Kingdom of Bahrain. *Desalination* 209 (2007) 230–237.

KHALFALLAHA, M. G.; KOLIUBB, A. M.(2007b); “Suggestions for improving wind turbines power curves” in *The Ninth Arab International Conference on Solar Energy (AICSE-9)*, Kingdom of Bahrain. *Desalination* 209 (2007) 221–229.

KHODAYAR, M.; TESHNEHLAB, M.; “Robust Deep Neural Network for Wind Speed Prediction”. In *2015 4th Iranian Joint Congress on Fuzzy and Intelligent Systems (CFIS)*, 9-11 Sept. 2015, Zahedan, Iran.

KINGMA, D., P.; BA, J., L.; “Adam: A Method for Stochastic Optimization”. In 3rd 2015 International Conference on Learning Representations (ICLR), 7-9 May 2015, San Diego, CA, United States.

KOSKO, B. Fuzzy engineering New Jersey: Prentice Hall, 1997. 549 p. ISBN 0-13-124991-6.

KOUTNÍK, J.; GREFF, K.; GOMEZ, F.; SCHMIDHUBER, J.; “On Neural and Evolutionary Computing (cs.NE); Machine Learning (cs.LG): A Clockwork RNN.” arXiv:1402.3511 [cs.NE], 2014.

KUSIAK, A., ZHENG, H.; SONG, Z.; “On-line monitoring of power curves” Renewable Energy 34 (2009) 1487–1493.

KUSIAK, A., ZHENG, H., SONG, Z., 2009a. Wind farm power prediction: a data-mining approach. Wind Energy 12 (3), 275 e 293.

KUSIAK, A., ZHENG, H., SONG, Z., March 2009b. Short-term prediction of wind farm power: a data mining approach. IEEE Transactions on Energy Conversion 24 (1), 125 e 136.

KUSIAK, A.; VERMA, A.; “Monitoring Wind Farms with Performance Curves” in IEEE Transactions on Sustainable Energy. Vol. 4, N 1, July 2013.

LANDBERG, L., 1994. Short-term Prediction of Local Wind Conditions (Ph.D. thesis). Risø-R- 702(EN), Risø National Laboratory, Roskilde, Denmark, ISBN 87-550-1916-1.

LANGE, M., FOCKEN, U., December 15 e 16, 2005a. State-of-the-Art in wind power prediction in Germany and international developments. In: Prediction of Wind Power and Reducing the Uncertainty for Grid Operators, Second Workshop of International Feed-in Cooperation. Berlin (DE).

LANGE, M., FOCKEN, U., 2005b. Physical Approach to Short-term Wind Power Prediction. Springer-Verlag, Berlin.

LIMA, F. G., PERERA, L. C. J., KIMURA, H., SILVA FILHO, A. C. Aplicação de redes neurais na análise e na concessão de crédito ao consumidor. RAUSP, v.44, n.1, p. 34-45, mar. 2009.

LOUKA, P., GALANIS, G., SIEBERT, N., KARINIOTAKIS, G., KATSAFADOS, P., PYTHAROULIS, I., KALLOS, G., 2008. Improvements in wind speed forecasts for wind power prediction purposes using Kalman filtering. Journal of Wind Engineering and Industrial Aerodynamics 96, 2348 e 2362.

LUDERMIR, T. B., BRAGA, A. P., CARVALHO, A. C. P. L. F. Redes Neurais Artificiais: Teorias e Aplicações. 1ª ed. Rio de Janeiro: LTC, 2000.

LYDIA, M.; SELVAKUMAR, A. I.; KUMAR, S. S.; KUM, G. E. P.; “Advanced Algorithms for Wind Turbine Power Curve Modeling.” in IEEE Transactions on Sustainable Energy. Vol. 4, N 3, July 2013.

LYDIA, M.; KUMAR, S. S.; SELVAKUMAR, A. I.; KUMAR, G. E. P.; “A comprehensive review on wind turbine power curve modeling techniques.” *Renew. Sustain. Energy Rev.* vol. 30. pp. 452–460. 2014.

MADSEN, H. (Ed.), 1995. *Wind Power Prediction Tool in Control Dispatch Centres*. ELSAM, Skaerbaek, Denmark, ISBN 87-87090-25-2.

MARVUGLIA, A.; MESSÍNEO, A. “Monitoring of wind farms’ power curves using machine learning techniques”. *Applied Energy* 98 (2012) 574–583.

MCCARTHY, E., March 23, 1998. Wind speed forecasting in the central California wind resource area. In: Paper Presented in the EPRI-DOE-NREL Wind Energy Forecasting Meeting. Burlingame, CA.

MCCULLOCH, W.S. & PITTS, W. *Bulletin of Mathematical Biophysics* (1943) 5: 115. <https://doi.org/10.1007/BF02478259>.

MÖNNICH, K., 2000. *Vorhersage der Leistungsabgabe netzeinspeisender Windkraftanlagen zur Unterstuetzung der Kraftwerkseinsatzplanung* (Ph.D. thesis). Carl von Ossietzky Universitat Oldenburg.

MONTEIRO, C., BESSA, R., MIRANDA, V., BOTTERUD, A., WANG, J., CONZELMANN, G., November 2009. *Wind Power Forecasting: State-of-the-Art 2009*. Argonne National Laboratory ANL/DIS 10-1. See also: <http://www.dis.anl.gov/projects/windpowerforecasting.html>.

MORENO, P., BENITO, L., BORÉN, R., CABRÉ, M., June 16e20, 2003. Short-term wind forecast. Results of first year planning maintenance at a wind farm. In: Poster Presented on the European Wind Energy Conference and Exhibition, Madrid (ES).

NIELSEN, T.S., MADSEN, H., 1996. *Using Meteorological Forecasts in On-line Predictions of Wind Power*. ELSAM, Skaerbaek, Denmark.

NOGARET, E., STAVRAKAKIS, G., BONIN, J.C., KARINIOTAKIS, G., PAPADIAS, B., CONTAXIS, G., PAPADOPOULOS, M., HATZIARGYRIOU, N., PAPATHANASSIOU, S., GAROPOULOS, J., KARAGOUNIS, E., HALLIDAY, J., DUTTON, G., PEDAS-LOPES, J., ANDROUTSOS, A., PLIGOROPOULOS, P., October 10 e 14, 1994. Development and implementation of an advanced control system for médium size wind-diesel systems. In: *Proceedings of the EWEC '94 in Thessaloniki*, pp. 599 e 604.

NORGAARD, P., HOLTINEN, H. “A Multi-Turbine Power Curve Approach”. In *Proceedings of Nordic Wind power. Conference NWPC'04, 1.-2.3.2004*. Chalmers University of Technology, Gothenburg, Sweden. 5p. Submitted to *wind Energy* 2005.

OLAOFE, Z. O.; FOLLY, K. A.; “Wind energy analysis based on turbine and developed site power curves: A case-study of Darling City.” *Renew. Sustain. Energy Rev.* vol. 53. pp. 306–318. 2013.

PIELKE, R.A., KENNEDY, E., January 1980. *Mesoscale Terrain Features*. report Uva Env. Sci meso 1980-1 University of Virginia Dept of Environmental Science.

PINSON, P., MADSEN, H., GIEBEL, G., KARINIOTAKIS, G., VON BREMEN, L., MARTÍ, I., May 7 e 10, 2007a. Forecasting of wind generation e recent advances and future challenges. In: Proceedings of the European Wind Energy Conference and Exhibition, Milan (IT).

PINSON, P., CHRISTENSEN, L.E.A., MADSEN, H., SØRENSEN, P.E., DONOVAN, M.H., JENSEN, L.E., 2008c. Regime-switching modelling of the fluctuations of offshore wind generation. *Journal of Wind Engineering and Industrial Aerodynamics* 96 (12), 2327 e 2347.

RANZATO, M.A.; POULTNEY, C.; CHOPRA, S.; LECUN, Y.; “Efficient learning of sparse representations with an energy-based model”. In: *Advances in neural information processing systems*. [S.l.: s.n.], 2006. p.1137–1144.

RENANI, E. T.; ELIAS, M. F. M.; RAHIM, N. A. “Wind power prediction Using enhanced parametric Wind power curve modeling” in 4th IET Clean Energy and Technology Conference (CEAT 2016), 14-15 Nov. 2016, Kuala Lumpur, Malaysia.

REZENDE, S. O. *Sistemas inteligentes: fundamentos e aplicações*. Barueri: Manole, 2005. 525 p. 2. ed. ISBN 85-204-1683-7.

SÀNCHEZ, I., USAOLA, J., RAVELO, O., VELASCO, C., DOMÍNGUEZ, J., LOBO, M.G., GONZÁLEZ, G., SOTO, F., June 2002. SIPREÓLICO wind power prediction system based on flexible combination of dynamic models. Application to the Spanish power system. In: *Proceedings of the World Wind Energy Conference in Berlin, Germany*.

SARIKAYA, R., HINTON, G., DEORAS, A. (2014). Application of Deep Belief Networks for Natural Language Understanding. *Audio, Speech, and Language Processing, IEEE/ACM Transactions on*. 22. 778-784. 10.1109/TASLP.2014.2303296.

SEIDE, F.; GANG, L.; YU, D.; "Conversational Speech Transcription Using Context Dependent Deep Neural Networks." *Interspeech*. 2011.

SCHLECHTINGEN, M.; SANTOS, I. F.; ACHICHE, S.; “Using Data-Mining Approaches for Wind Turbine Power Curve Monitoring: A Comparative Study.” in *IEEE Transactions on Sustainable Energy*. Vol. 4. N 3. January 2013.

ŠIMIĆ, Z.; MIKULIČIĆ, V.; E “Small Wind Off-Grid System Optimization Regarding Wind Turbine Power Curve”. *IEEE Africon 2007* pages 1-6. November 2007.

SKAMAROCK, W.C., December 2004. Evaluating mesoscale NWP models using kinetic energy spectra. *Monthly Weather Review* 132, 3019 e 3032.

STIEBLER, M. *Wind Energy Systems for Electric Power Generation Green Energy and Technology*. 1ª ed. Berlin Heidelberg: © Springer-Verlag, 2008.

STEPHEN, B.; GALLOWAY, S. J.; MCMILLAN, D.; HILL, D. C.; Infield, D. G. “A Copula Model of Wind Turbine Performance”. In *IEEE Transactions on Power Systems*, Vol. 26, N2. May 2011.

STILL, D., GRAINGER, B., June 2 e 6, 2001. Demanding seas e the UK's first offshore wind farm. In: Proceedings of the European Wind Energy Conference, Copenhagen, Denmark, ISBN 3 936338-09-4, pp. 169e171. Note: This statement was done during the talk.

SUHUA, L.; ZHIHENG, L.; YAOWU, W. "Clustering Analysis of the Wind Power Output Based on Similarity Theory". DRPT2008 6-9 April 2008 Nanjing China.

THAPAR, V.; AGNIHOTRI, G.; SETHI, V. K.; "Critical analysis of methods for mathematical modelling of wind turbines". Renewable Energy 36 (2011) 3166-3177.

TORRES, J.L., GARCIA, A., De BLAS, M., De Francisco, A., 2005. Forecast of hourly average Wind speed with ARMA models in Navarre (Spain). Solar Energy 79 (1), 65 e 77.
TRIVELLATO, F.; BATTISTI, L.; MIORI, G.; "The ideal power curve of small wind turbines from field data." J. Wind Eng. Ind. Aerodyn. 107-108 (2012) 611-620.

TROEN, I., LANDBERG, L., September 10 e 14, 1990. Short-term prediction of local wind conditions. In: Proceedings of the European Community Wind Energy Conference, Madrid (ES), ISBN 0-9510271-8-2, pp. 76 e 78.

TROEN, I., PETERSEN, E.L., 1998. European Wind Atlas. Published for the EU Commission DGXII by Risø National Laboratory, Denmark, ISBN 87-550-1482-8.

ULUYOL, O.; PARTHASARATHY, G.; FOSLIEN, W.; KIM, K.; "Power Curve Analytic for Wind Turbine Performance Monitoring and Prognostics." Annual Conference of the Prognostics and Health Management Society, 2011.

ÜSTÜNTAS, T., SAHIN, A.D.; Wind turbine power curve estimation based on cluster center Fuzzy logic modeling. J. Wind Eng. Ind. Aerodyn. 96 (2008) 611-620.

VINCENT, P.; LAROCHELLE, H., LAJOIE, I.; BENGIO, Y. MANZAGOL, P. A.; "Stacked Denoising Autoencoders: Learning Useful Representations in a Deep Network with a Local Denoising Criterion". In Journal of Machine Learning Research 11 (2010) 3371-3408.

VINCENT, C., GIEBEL, G., PINSON, P., MADSEN, H., February 2010b. Resolving non-stationary spectral information in wind speed time series using the Hilbert-Huang transform. Journal of Applied Meteorology and Climatology 49 (2), 253 e 267.

WANG, D.; KALATHIL, D.; POOLLA, K.; GUAN, X., "Coordination of Wind Power and Flexible Load through Demand". In 2015 IEEE 54th Annual Conference on Decision and Control (CDC), December 15-18, 2015. Osaka, Japan.

YAN, J.; ZHANG, H.; LIU, Y.; HAN, S.; LI, L.; LU, Z.; "Forecasting the High Penetration of Wind Power on Multiple Scales Using Multi-to-Multi Mapping". Published in IEEE Transactions on Power Systems (2017). Sponsored by IEEE Power & Energy Society, Volume: 33, Issue: 3, May 2018.

YAO, K.; COHN, T.; VYLOMOVA, K.; DUH, K.; DYER C.; "On Neural and Evolutionary Computing: Depth-Gated LSTM." arXiv preprint, arXiv:1508.03790, 2015.

ZAMANI, M.H.; RIAHY, G.H.; ARDAKANI, A. J. "Modifying power curve of Variable Speed Wind Turbines by Performance Evaluation of Pitch-Angle and Rotor Speed Controllers". 2007 IEEE Canada Electrical Power Conference pages 347-352.