

UNIVERSIDADE FEDERAL DE SANTA MARIA  
CENTRO DE TECNOLOGIA  
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

Antônio Mário Kaminski Júnior

**EMPREGO DE REDES NEURAI ARTIFICIAIS PARA PREDIÇÃO DE  
TEMPERATURA DE TOPO DE ÓLEO E PERDA DE VIDA ÚTIL EM  
TRANSFORMADORES DE POTÊNCIA**

Santa Maria, RS  
2020

**Antônio Mário Kaminski Júnior**

**EMPREGO DE REDES NEURAS ARTIFICIAIS PARA PREDIÇÃO DE  
TEMPERATURA DE TOPO DE ÓLEO E PERDA DE VIDA ÚTIL EM  
TRANSFORMADORES DE POTÊNCIA**

Dissertação apresentada ao Curso de Mestrado do Programa de Pós-Graduação em Engenharia Elétrica, área de concentração em Sistemas de Energia da Universidade Federal de Santa Maria (UFSM, RS), como requisito parcial para obtenção do título de **Mestre em Engenharia Elétrica**.

Orientador: Prof. Dr. Tiago Bandeira Marchesan

Santa Maria, RS

2020

Kaminski Júnior, Antônio Mário

Emprego de Redes Neurais Artificiais para Predição de Temperatura de Topo de Óleo e Perda de Vida Útil em Transformadores de Potência / Antônio Mário Kaminski Júnior.- 2020.

103 p.; 30 cm

Orientador: Tiago Bandeira Marchesan

Dissertação (mestrado) - Universidade Federal de Santa Maria, Centro de Tecnologia, Programa de Pós-Graduação em Engenharia Elétrica, RS, 2020

1. Transformadores de Potência 2. Redes Neurais Artificiais 3. Predição de Temperatura I. Bandeira Marchesan, Tiago II. Título.

Sistema de geração automática de ficha catalográfica da UFSM. Dados fornecidos pelo autor(a). Sob supervisão da Direção da Divisão de Processos Técnicos da Biblioteca Central. Bibliotecária responsável Paula Schoenfeldt Patta CRB 10/1728.

---

© 2020

Todos os direitos autorais reservados a Antônio Mário Kaminski Júnior. A reprodução de partes ou do todo deste trabalho só poderá ser feita com autorização por escrito do autor. Endereço: Rua Duque de Caxias, n. 1275, Apartamento 104, Bairro Bonfim, Santa Maria, RS. CEP: 97015-190.


Fone (+55) 55 99643-0146; E-mail: juniorkaminski@hotmail.com.


**Antônio Mário Kaminski Júnior**


**EMPREGO DE REDES NEURAIS ARTIFICIAIS PARA PREDIÇÃO DE  
TEMPERATURA DE TOPO DE ÓLEO E PERDA DE VIDA ÚTIL EM  
TRANSFORMADORES DE POTÊNCIA**

Dissertação apresentada ao Curso de Doutorado do Programa de Pós-Graduação em Engenharia Elétrica, da Universidade Federal de Santa Maria (UFSM, RS), como requisito parcial para obtenção de grau de **Mestre em Engenharia Elétrica**.

**Aprovado em 24 de setembro de 2020:**

  
\_\_\_\_\_  
**Tiago Bandeira Marchesan, Dr. (UFSM) - Videoconferência**  
(Presidente/Orientador)

  
\_\_\_\_\_  
**Vitor Cristiano Bender, Dr. (UFSM) - Videoconferência**

  
\_\_\_\_\_  
**Nelson Knak Neto, Dr. (URI) - Videoconferência**

Santa Maria, RS

2020



*À minha família: Mãe Regina,  
irmã Tassiana, irmão Tiago  
e meu falecido pai Antônio.  
Pela oportunidade da educação que me proporcionaram.*



## AGRADECIMENTOS

Primeiramente agradeço à minha família, mãe Regina, Tassi, Tiago, Taís, Wilfredo e ao pequeno Miguel por todo carinho, apoio e por sempre acreditarem em mim.

À minha namorada, Jacqueline, que esteve ao meu lado durante estes dois anos, por todo incentivo, carinho, paciência e amor. Todos os momentos ao teu lado são especiais, te amo muito.

Aos meus amigos, que estiveram alguns próximos e outros distantes durante a elaboração deste trabalho, muito obrigado por todos momentos que compartilhamos. Vocês são como uma família para mim.

Aos colegas do INRI MT e LabSim, por todo apoio técnico e pela amizade que foi construída ao longo deste percurso. Vocês tiveram grande influência no meu crescimento profissional e sou muito grato por isso.

Aos professores Tiago Marchesan e Vitor Bender pela orientação e contribuições feitas a este trabalho. Agradeço pelo conhecimento compartilhado que enriqueceu este trabalho e contribuiu em minha formação profissional. Obrigado também pela oportunidade de realizar meu mestrado em um âmbito mais próximo do setor privado através de projetos de P&D.

Ao apoio técnico e financeiro da Santo Antônio Energia e Vegoor referente ao projeto de P&D intitulado “Metodologia de predição da capacidade funcional de transformadores elevadores com desbalanço de carga entre enrolamentos considerando dinâmicas térmicas rápidas e vida útil remanescente” nº PD-06683-0318/2018.

Por último, mas não menos importante, à Universidade Federal de Santa Maria, pela educação gratuita e da melhor qualidade possível que me proporcionou durante a graduação e pós-graduação e aos profissionais que tanto trabalham para que isto seja possível.





*“A felicidade não está na estrada que leva a algum lugar.  
A felicidade é a própria estrada.”*

*Bob Dylan – Crônicas Vol. 1*



## RESUMO

### EMPREGO DE REDES NEURAS ARTIFICIAIS PARA PREDIÇÃO DE TEMPERATURA DE TOPO DE ÓLEO E PERDA DE VIDA ÚTIL EM TRANSFORMADORES DE POTÊNCIA

AUTOR: ANTÔNIO MÁRIO KAMINSKI JÚNIOR  
ORIENTADOR: TIAGO BANDEIRA MARCHESAN

A correta predição de temperatura em transformadores de potência possibilita um melhor aproveitamento de sua capacidade nominal, prolongamento da vida útil do equipamento e planejamento estratégico com base nas condições de operação futuras esperadas. A proposição de novos modelos que apresentem uma boa capacidade preditiva é, portanto, de grande interesse aos responsáveis pelos transformadores de potência. O presente trabalho apresenta um método de desenvolvimento de Redes Neurais Artificiais (RNAs), justificando os parâmetros escolhidos com base no comportamento térmico de transformadores, para a predição de temperatura de topo de óleo utilizando o modelo de rede neural NARX, ainda não utilizado para predição de temperatura em transformadores. Todos os conjuntos de dados utilizados para treinamento e testes da capacidade preditiva das RNAs são dados reais de monitoramento, provenientes de cinco transformadores elevadores de uma usina hidrelétrica. Foram realizados testes da capacidade de predição para todos transformadores, combinando redes treinadas de um dos transformadores e aplicadas às entradas de outro, abordando em que situações ocorreram os melhores e piores desempenhos. Após, são apresentados os métodos de cálculo de perda de vida útil de transformadores propostos por normas e realiza-se uma comparação entre o calculado a partir dos dados de monitoramento e a partir dos valores de temperatura fornecidos pela rede neural. Com o intuito de validar a capacidade de predição para cenários futuros esperados, são propostos seis cenários fictícios de longo período de duração e então a vida útil para estes é estimada. Todos os resultados obtidos são satisfatórios, com erros abaixo de 4%, ou 2 °C em valores absolutos, grande parte dos períodos em que foram realizados os testes, capazes de comprovar a capacidade preditiva das RNAs desenvolvidas utilizando o método apresentado não apenas em sua aplicação para monitoramento de temperatura, mas também sob a perspectiva de perda de vida útil.

**Palavras-chave:** Redes Neurais Artificiais, NARX, Predição de Temperatura, Temperatura de Topo de Óleo, Transformadores de Potência, Temperatura de Ponto Mais Quente, Perda de Vida Útil.



## ABSTRACT

### ARTIFICIAL NEURAL NETWORKS APPLICATION FOR TOP OIL TEMPERATURE AND LOSS OF LIFE PREDICTION IN POWER TRANSFORMERS

AUTHOR: ANTÔNIO MÁRIO KAMINSKI JÚNIOR  
ADVISOR: TIAGO BANDEIRA MARCHESAN

The precise temperature prediction in power transformers allows a better use of its nominal capacity, extending the equipment's useful life and strategic planning based on the expected future operating conditions. The proposal of new models that present a good predictive capacity is, therefore, of great interest to those responsible for power transformers. The present work presents a method of developing Artificial Neural Networks (ANNs), justifying the parameters chosen based on the thermal behavior of power transformers, for prediction of top-oil temperature using the NARX neural network model, not yet used for temperature prediction of transformers. All data sets used for training and testing the predictive ability of ANNs are real monitoring data from five elevating transformers in a hydroelectric plant. Tests of prediction ability were performed for all transformers, combining trained networks from one of the transformers and applied to the inputs of others, addressing in which situations the best and worst performances occurred. Afterwards, the methods for calculating the loss of life of transformers proposed by standards are presented and a comparison is made between the one calculated from the monitoring data and from the temperature values provided by the neural network. In order to validate the prediction capacity for expected future scenarios, six fictitious scenarios of long duration are proposed and then their useful life is estimated. All the results obtained are satisfactory, with errors below 4%, or 2 °C on absolute values, most of the periods in which the tests were carried out, capable of proving the predictive capacity of the ANNs developed using the method presented not only in its application for temperature monitoring, but also from the perspective of loss of life.

**Keywords:** Artificial Neural Networks, NARX, Temperature Prediction, Top-Oil Temperature, Power Transformers, Hot-Spot Temperature, Loss of Life.



## LISTA DE ILUSTRAÇÕES

Figura 1 – Estrutura de um neurônio biológico. ....	31
Figura 2 – Estrutura de um neurônio artificial. ....	32
Figura 3 – Função limite rígido. ....	34
Figura 4 – Função linear. ....	34
Figura 5 – Gráfico da função rampa. ....	35
Figura 6– Função sigmoide ....	36
Figura 7 – Rede de alimentação direta de camada única. ....	37
Figura 8 – Rede de alimentação direta com múltiplas camadas. ....	37
Figura 9 – Exemplo de rede recorrente. ....	38
Figura 10 – Esquemático do funcionamento do algoritmo de <i>backpropagation</i> . ....	40
Figura 11– Estrutura de uma rede NARX. ....	44
Figura 12 – Arquiteturas de uma rede NARX. ....	45
Figura 13 – Dados de monitoramento utilizados como parâmetros de entrada das RNAs. .....	50
Figura 14 – Temperatura média (vermelho) utilizada como parâmetro de entrada. ....	51
Figura 15 – Monitoramento da temperatura de topo de óleo TR1. ....	52
Figura 16 – Erro para quantidades de camadas ocultas. ....	53
Figura 17 – Representação simplificada da estrutura das RNAs desenvolvidas. ....	54
Figura 18 – Visão geral da resposta da RNA durante o treinamento. ....	55
Figura 19 – Resposta RNA vs. Dados de monitoramento para o treinamento. ....	56
Figura 20 – Histograma de erros para o treinamento. ....	56
Figura 21 – Correlação das temperaturas RNA e monitoramento para o treinamento. ....	57
Figura 22 – Comparação das temperaturas resposta RNA vs. monitoramento. ....	59
Figura 23 – Comparação das temperaturas resposta RNA vs. monitoramento. ....	59
Figura 24 – Comparação das temperaturas resposta RNA vs. monitoramento. ....	60
Figura 25 – Comparação das temperaturas mais parâmetros de entrada. ....	61
Figura 26 – Comparação das temperaturas mais parâmetros de entrada. ....	61
Figura 27 – Histograma de erros. ....	62
Figura 28 – Correlação temperatura entre resposta RNA e dados de monitoramento. ....	63
Figura 29 – Comparação das temperaturas resposta RNA vs. monitoramento. ....	64
Figura 30 – Comparação das temperaturas resposta RNA vs. monitoramento. ....	64
Figura 31 – Comparação das temperaturas resposta RNA vs. monitoramento. ....	65



Figura 32 – Comparação das temperaturas mais parâmetros de entrada. ....	65
Figura 33 – Comparação das temperaturas mais parâmetros de entrada. ....	66
Figura 34 – Histograma de erros. ....	67
Figura 35 – Correlação temperatura entre resposta RNA e dados de monitoramento. ....	67
Figura 36 – Comparação das temperaturas resposta RNA vs. monitoramento. ....	68
Figura 37 – Comparação das temperaturas resposta RNA vs. monitoramento. ....	69
Figura 38 – Histograma de erros. ....	69
Figura 39 – Correlação temperatura entre resposta RNA e dados de monitoramento. ....	70
Figura 40 – Comparação das temperaturas resposta RNA vs. monitoramento. ....	71
Figura 41 – Comparação das temperaturas resposta RNA vs. monitoramento. ....	71
Figura 42 – Comparação das temperaturas resposta RNA vs. monitoramento. ....	72
Figura 43 – Histograma de erros. ....	73
Figura 44 – Correlação temperatura entre resposta RNA e dados de monitoramento. ....	74
Figura 45 – Comparação das temperaturas resposta RNA vs. monitoramento. ....	75
Figura 46 – Comparação das temperaturas resposta RNA vs. monitoramento. ....	75
Figura 47 – Histograma de erros. ....	76
Figura 48 – Correlação temperatura entre resposta RNA e dados de monitoramento. ....	77
Figura 49 – RNA TR1 aplicada a dados do TR4 mais parâmetros de entrada. ....	79
Figura 50 – RNA TR3 aplicada a dados TR3 mais parâmetros de entrada. ....	80
Figura 51 – Dados de monitoramento TR4. ....	81
Figura 52 – RNA TR1 aplicada em dados do TR5 mais parâmetros de entrada. ....	82
Figura 53 – Temperatura de <i>hot-spot</i> resposta RNA e monitoramento TR1. ....	87
Figura 54 – Comparação perda de vida útil TR1. ....	88
Figura 55 – Comparação perda de vida útil TR2. ....	89
Figura 56 – Comparação perda de vida útil TR3. ....	89
Figura 57 – Comparação perda de vida útil TR4. ....	90
Figura 58 – Comparação perda de vida útil TR5. ....	91
Figura 59 – Parâmetros de entrada da rede neural para o Cenário 5. ....	93
Figura 60 – Parâmetros de entrada da rede neural para o Cenário 4. ....	93
Figura 61 – Temperatura de <i>hot-spot</i> cenários fictícios. ....	94
Figura 62 – Temperatura de <i>hot-spot</i> cenários fictícios. ....	95
Figura 63 – Perda de vida útil cenários fictícios. ....	95
Figura 64 – Temperatura de <i>hot-spot</i> cenários 3 e 6. ....	96

**LISTA DE TABELAS**

Tabela 1 – Resumo dos DMA para todos os testes realizados. ....	77
Tabela 2 – Erro máximo absoluto observado nos testes de capacidade de predição....	78
Tabela 3 – Tempo de vida normal. ....	85
Tabela 4 – Valores utilizados para a obtenção da temperatura de <i>hot-spot</i> .....	86
Tabela 5 – Composição dos conjuntos de monitoramento e resposta RNA.....	86
Tabela 6 – Datas de energização e monitoramento dos transformadores.....	87
Tabela 7 – Resumo dos resultados das comparações. ....	91
Tabela 8 – Condições de operação para os cenários fictícios.....	94
Tabela 9 – Perda de vida útil cenários fictícios. ....	96
Tabela 10 – Taxa de Envelhecimento Relativo pela Temperatura .....	97



**LISTA DE ABREVIATURAS E SIGLAS**

RNA	Redes Neural Artificial
NARX	<i>Non-Linear Autoregressive with Exogenous Inputs Model</i>
OFAF	Transformador com resfriamento óleo forçado e ar forçado
ODAF	Transformador com resfriamento óleo dirigido e ar forçado
DMA	Desvio médio absoluto
EQM	Erro quadrático médio
LM	Levenberg-Marquadt



## LISTA DE SÍMBOLOS

$a$	Saída da rede neural
$a_d$	Resposta desejada pela rede neural
$b_k$	Bias de um neurônio artificial $k$
$d_u$	Número de atrasos considerados na camada de entrada
$d_y$	Número de atrasos considerados na camada de saída
$e$	Erro
$F$	Função de ativação
$k$	Neurônio artificial
$L$	Perda de vida útil
$u$	Vetor de entradas da rede neural
$V$	Taxa de envelhecimento relativo
$v_k$	Saída do somador de um neurônio artificial $k$
$x_m$	Entrada $m$ de um neurônio artificial $k$
$y$	Saída da rede neural
$y_k$	Saída de um neurônio artificial $k$
$w_{km}$	Peso sináptico da entrada $m$ de um neurônio artificial $k$
$\theta_h$	Temperatura de ponto mais quente ( <i>hot-spot</i> )



## SUMÁRIO

<b>AGRADECIMENTOS.....</b>	<b>8</b>
<b>RESUMO .....</b>	<b>12</b>
<b>ABSTRACT.....</b>	<b>14</b>
<b>1. INTRODUÇÃO.....</b>	<b>26</b>
1.1 CONTEXTUALIZAÇÃO.....	26
1.2 MOTIVAÇÃO E JUSTIFICATIVA.....	26
1.3 OBJETIVOS .....	28
1.4 ORGANIZAÇÃO DA DISSERTAÇÃO.....	28
<b>2. REDES NEURAIS ARTIFICIAIS .....</b>	<b>29</b>
2.1 HISTÓRICO .....	29
2.2 O CÉREBRO HUMANO .....	30
2.3 ESTRUTURA DE UMA RNA.....	31
2.4 FUNÇÃO DE ATIVAÇÃO.....	33
2.5 ARQUITETURAS DE REDE .....	36
2.6 APRENDIZADO DE UMA RNA .....	39
2.7 ALGORITMO DE LEVENBERG-MARQUADT .....	41
2.8 AVALIAÇÃO DE DESEMPENHO.....	42
2.9 NARX .....	43
2.10 REDES NEURAIS ARTIFICIAIS NA PREDIÇÃO DE TEMPERATURA EM TRANSFORMADORES .....	45
2.11 CONSIDERAÇÕES DO CAPÍTULO .....	47
<b>3. EMPREGO DE RNAs PARA PREDIÇÃO DE TEMPERATURA EM TRANSFORMADORES DE POTÊNCIA.....</b>	<b>49</b>
3.1 OBJETOS DE ESTUDO .....	49
3.2 DADOS DE MONITORAMENTO.....	50
3.3 DESENVOLVIMENTO DAS RNAs .....	52
3.4 TREINAMENTO.....	54
3.5 TESTES DA CAPACIDADE DE PREDIÇÃO DAS RNAs .....	57
3.6 CONSIDERAÇÕES DO CAPÍTULO .....	82
<b>4. ANÁLISES DE PERDA DE VIDA ÚTIL.....</b>	<b>83</b>
4.1 TAXA DE ENVELHECIMENTO RELATIVO.....	83
4.2 CÁLCULO DA PERDA DE VIDA ÚTIL .....	84
4.3 VIDA DA ISOLAÇÃO.....	84
4.4 DETERMINAÇÃO DA TEMPERATURA DE PONTO MAIS QUENTE.....	85
4.5 COMPARAÇÃO PERDA DE VIDA ÚTIL ENTRE REDE NEURAL E MONITORAMENTO .....	88
4.6 VIDA ÚTIL DE CENÁRIOS FICTÍCIOS .....	92
4.7 CONSIDERAÇÕES DO CAPÍTULO .....	97
<b>5. CONSIDERAÇÕES FINAIS .....</b>	<b>99</b>
5.1 PUBLICAÇÕES DO TRABALHO.....	101
<b>REFERÊNCIAS .....</b>	<b>102</b>





## 1. INTRODUÇÃO

Este capítulo aborda a importância de estudos referentes a predição de temperatura em transformadores de potência e a utilização de Redes Neurais Artificiais como ferramenta para tal. Os objetivos e delimitações do trabalho são apresentados e justificados. Por fim, a organização dos capítulos que compõe este trabalho é apresentada.

### 1.1 CONTEXTUALIZAÇÃO

Transformadores de potência estão presentes em todo o sistema elétrico de potência, da geração à distribuição, integrando todos os seus níveis de tensão. Apresentam a essencial função de assegurar a transmissão de energia elétrica de forma econômica e confiável ao consumidor final. Estes também são os ativos mais caros em uma subestação, portanto sua integridade é importante em termos financeiros, acrescentando-se ao fato de que sua indisponibilidade causa a interrupção do fornecimento de energia e maiores prejuízos à concessionária. Com a finalidade de obter o melhor aproveitamento de sua vida útil e postergar investimentos, os responsáveis por transformadores de potência vêm buscando alternativas para identificar quais equipamentos se aproximam do final de sua vida útil, onde a probabilidade de falha aumenta, e tomar decisões com antecedência (FEIL, 2019).

### 1.2 MOTIVAÇÃO E JUSTIFICATIVA

A vida útil de um transformador é determinada pela sua isolação sólida que, em termos de envelhecimento, é determinada pela temperatura interna de operação do equipamento (VILATHONG et. al., 2007). Os indicadores mais comuns para transformadores imersos em óleo isolante são a temperatura de ponto mais quente e a temperatura de topo de óleo. A temperatura de topo de óleo apresenta maior conveniência em sua medição, enquanto a temperatura de ponto mais quente ocorre no interior dos enrolamentos, e sua medição deve ser feita com a utilização de fibras ópticas, que, além de dispendiosas em termos econômicos, são sensíveis e exigem cuidados especiais durante a construção e manutenção do equipamento.

Portanto, o conhecimento da temperatura de topo de óleo é o indicativo mais difundido para avaliar as condições térmicas de transformadores seja em tempo real ou previstas, permitindo a operação dentro dos limites de temperatura estabelecidos pelas normas ABNT NBR 5356-2 (2007) e IEC 60076-2 (2011) e um maior aproveitamento de sua capacidade nominal. Geralmente, para o cálculo da degradação em transformadores, utiliza-se a

temperatura de topo de óleo de um sistema de monitoramento e calcula-se a temperatura de ponto mais quente pelos modelos propostos por norma.

Diversos modelos para a predição de temperatura em transformadores foram propostos, a exemplo dos modelos propostos pelas normas ABNT NBR 5356-7 (2017) e IEC 60076-7 (2018) amplamente utilizados. Porém, estes modelos apresentam suas limitações, principalmente devido aos efeitos de variação da temperatura ambiente (VILAITHONG et. al., 2007) e pela quantidade necessária de informações.

Redes Neurais Artificiais (RNAs) são uma ferramenta matemática que, de forma simplificada, visa simular o processo de aquisição de conhecimento do cérebro humano e apresentam bons resultados na resolução de problemas que métodos matemáticos mais tradicionais não abrangem. Como tarefas que apresentam um bom desempenho ao serem executadas por RNAs tem-se o estabelecimento de padrões de comportamento entre variáveis de entrada e saída, reconhecimento de padrões, classificação e otimização (PEREIRA et. al., 2018).

As aplicações de RNAs têm encontrado espaço em uma ampla variedade de campos do conhecimento, não apenas na engenharia, matemática e ciências em geral, mas também na literatura, medicina, finanças, administração, entre outros (HAGAN, 2016). Sua aplicação tem se expandido cada vez mais devido a capacidade de resolver problemas complexos, em conjunto com o desenvolvimento de computadores com maior capacidade de processamento e algoritmos mais rápidos (HAGAN, 2016).

No presente trabalho é apresentado o desenvolvimento de uma rede neural para a predição da temperatura de topo de óleo em transformadores de potência considerando como entradas o carregamento do transformador, a temperatura ambiente e o *status* dos trocadores de calor, ou seja, quantos conjuntos de aerotermos (ventiladores e bombas de óleo) estão em operação. Cada uma das entradas apresenta um impacto, de forma distinta, na temperatura de operação do transformador. Para a implementação foi escolhido um modelo de rede neural que represente a dinâmica térmica do transformador, causada tanto pelos seus parâmetros de entrada como pela saída do estado anterior, e seu comportamento não-linear. Este modelo de rede denominado *Non-Linear Autoregressive with Exogenous Inputs Model* (NARX) apresenta bom desempenho em diversas aplicações, em especial predição de variáveis não-lineares (KHALIL, 2012; RUSLAN et. al., 2014; ANDRADE et. al., 2014), incluindo estudos químicos do óleo de transformadores (PEREIRA et. al., 2018). Entretanto, esta arquitetura de rede ainda não foi utilizada para predição de temperatura em transformadores de potência.

### 1.3 OBJETIVOS

O principal objetivo do presente trabalho é o desenvolvimento de RNAs para a predição de temperatura de topo de óleo em um conjunto de transformadores de potência, validando sua capacidade de predição para um longo período em operação. Alguns dos objetivos específicos podem ser considerados como:

- Validar a utilização do modelo de rede neural NARX, que apresenta características interessantes para esta aplicação, para predição de temperatura em transformadores de potência;
- Verificar se as redes neurais desenvolvidas apresentam resultados satisfatórios para um período abrangente em relação aos parâmetros de entrada e saída;
- Analisar a resposta das RNAs desenvolvidas com dados de monitoramento de um transformador apresentam resposta satisfatória quando aplicadas a outros transformadores com modos de resfriamento distintos;
- Realizar uma comparação acerca da perda de vida útil para todos os transformadores sob estudo calculados com dados de temperatura provindos do monitoramento com dados de temperatura obtidos através das RNAs desenvolvidas;
- Validar a capacidade de predição das RNAs sob a perspectiva da perda de vida útil.

### 1.4 ORGANIZAÇÃO DA DISSERTAÇÃO

O Capítulo 2 apresenta uma introdução sobre Redes Neurais Artificiais, seus conceitos, características e habilidades, estrutura de um neurônio artificial, arquiteturas existentes, métodos de treinamento e a arquitetura utilizada neste trabalho. O Capítulo 3 apresenta os transformadores objetos de estudo e os dados de monitoramento utilizados no desenvolvimento deste trabalho, também aborda as características das RNAs desenvolvidas, como ocorreu o processo de treinamento e apresenta os resultados obtidos a partir dos testes realizados para validar sua capacidade de predição.

No Capítulo 4 é apresentada uma análise de vida útil, comparando o que foi calculado utilizando os dados de monitoramento dos transformadores com o calculado através da resposta da rede neural. Também é realizado o cálculo da perda de vida útil para cenários futuros fictícios, validando a capacidade de predição da RNA sob este aspecto. O Capítulo 5 realiza uma síntese dos resultados, abordando as conclusões que podem ser feitas a partir dos resultados obtidos.

## 2. REDES NEURAIS ARTIFICIAIS

Redes Neurais Artificiais têm seu desenvolvimento atribuído às primeiras compreensões da diferença no processamento de informações entre um cérebro humano e um computador digital convencional. O cérebro humano pode ser considerado um computador altamente complexo, não-linear e com processamento paralelo, e possui a capacidade de organizar sua estrutura neuronal para executar determinados tipos de processamentos, despertando o interesse da utilização de um método semelhante para tarefas específicas. A utilização de redes neurais apresenta as seguintes características e habilidades (HAYKIN, 1999; OLIVEIRA, 2005).

- Pode considerar a não-linearidade dos fenômenos físicos;
- Pouco conhecimento estatístico necessário sobre o ambiente no qual a rede está inserida;
- Capacidade de aprender através de exemplos de entrada e saída que representam o ambiente;
- Habilidade de aproximar qualquer relação entre entradas e saídas de natureza contínua;
- Adaptabilidade;
- Generalização
- Tolerância à falhas;
- Informação contextual;
- Armazenamento do conhecimento adquirido através dos pesos sinápticos.

De forma resumida, a rede neural artificial é um método matemático que visa simular o processo de aquisição de conhecimento de um cérebro humano com a intenção de resolver problemas que técnicas mais convencionais não fazem com sucesso. Sua utilização em problemas complexos é fruto de sua capacidade de aprender através de exemplos e generalizar o conhecimento adquirido para novos cenários.

### 2.1 HISTÓRICO

A base do estudo de redes neurais tem início no final do século XIX, onde foram realizados avanços em teorias sobre o processo de aquisição do conhecimento e como é processada a visão humana, porém sem propor nenhum modelo matemático da operação de um neurônio (HAGAN, 2016). O conceito moderno sobre redes neurais, e origem deste campo de pesquisa, surgiu do trabalho entre McCulloch e Pitts (1943), em que demonstraram que uma rede neural composta por neurônios artificiais poderia, a princípio, processar qualquer função lógica ou aritmética (MCCULLOCH, PITTS, 1943; HAGAN, 2016).

A primeira aplicação prática das Redes Neurais Artificiais foi demonstrada na década de 50, gerando uma grande expectativa acerca do potencial desta ferramenta. Entretanto, os modelos de rede neural e algoritmos de treinamento propostos durante as décadas de 50 e 60 apresentavam limitações ao encontrar soluções para apenas algumas classes de problemas e aliada à falta de computadores potentes para o desenvolvimento de algoritmos de treinamento para modelos de redes mais complexos houve uma estagnação da pesquisa neste campo por mais de uma década (HAGAN, 2016).

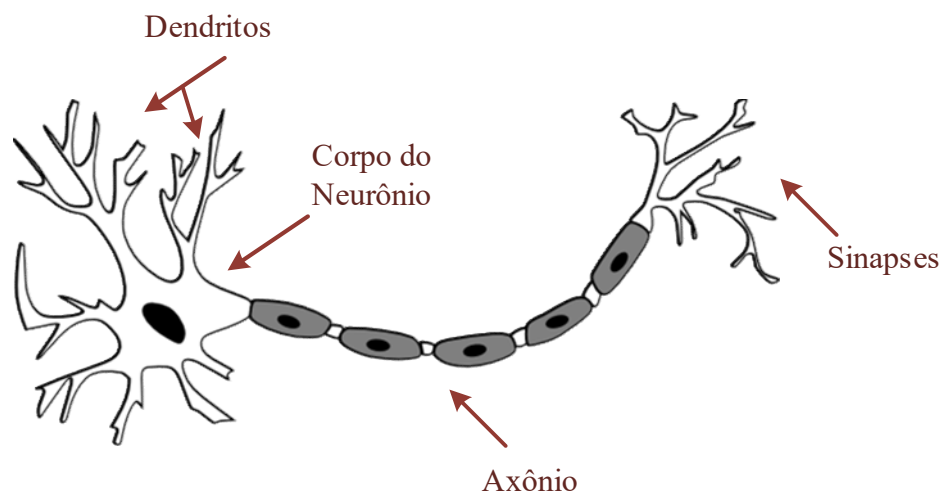
Durante a década de 80, com a popularização dos computadores pessoais e workstations, dois grandes avanços, a utilização de mecânica estatística para explicar o funcionamento de redes recorrentes por Hopfield (1982), e o desenvolvimento do algoritmo de treinamento *backpropagation*. Estas inovações foram realizadas por diversos pesquisadores de forma independente, trazendo solução para os problemas das décadas passadas e revigoraram as pesquisas em RNAs (HOPFIELD, 1982) (HAGAN, 2016).

Desde então novos modelos de redes e algoritmos de treinamento foram propostos. Hoje as RNAs podem ser consideradas ferramentas essenciais em uma grande variedade de campos do conhecimento, e podem oferecer soluções para uma gama de problemas se corretamente utilizadas. Porém, ainda se conhece pouco sobre o funcionamento do cérebro humano, e novas descobertas em conjunto com a constante evolução do poder computacional disponível podem reservar um futuro promissor para este campo do conhecimento.

## 2.2 O CÉREBRO HUMANO

Neurônios biológicos são a célula básica do cérebro humano. Estes que se conectam formando uma rede de neurônios massiva e são responsáveis por orientar todas as atividades do organismo e também por todo o processamento e armazenamento de informações. A estrutura de um neurônio biológico, que poder ser visualizada em Figura 1, é composta pelos dendritos, responsáveis pelo recebimento dos estímulos (estes que podem ser excitação ou inibição e também de diferentes magnitudes), pelo corpo do neurônio, responsável por coletar e combinar as informações recebidas de outros neurônios, e pelos axônios, camada longa e fina responsável por transmitir o impulso até as conexões sinápticas. A sinapse é o terminal de saída do neurônio, onde ele envia o estímulo para outros neurônios através de diferentes concentrações de íons de sódio e potássio. O cérebro humano é formado por bilhões de neurônios que constituem dezenas de trilhões de conexões sinápticas entre si (OLIVEIRA, 2005).

Figura 1 – Estrutura de um neurônio biológico.



Fonte: Autor.

Redes Neurais Artificiais ainda não possuem uma complexidade tão grande como a do cérebro humano, e o funcionamento de um neurônio biológico apresenta uma complexidade muito maior em relação a de um neurônio artificial, entretanto há duas grandes similaridades que valem ser ressaltadas. A primeira é que ambas as redes neurais, biológica e artificial, são compostas por células computacionais básicas que são altamente interconectadas (HAGAN, 2016). A segunda similaridade é que as conexões entre os neurônios definem a função da rede neural (HAGAN, 2016).

Vale ressaltar que a velocidade de transmissão de um neurônio biológico é seis vezes menor se comparada a de um circuito elétrico, todavia o cérebro é capaz de realizar algumas tarefas com maior rapidez em comparação a um computador (HAGAN, 2016). Um dos motivos seria a estrutura massiva de processamento em paralelo de uma rede neural biológica, todos os neurônios operam ao mesmo tempo (HAGAN, 2016). RNAs compartilham desta estrutura, porém sua aplicação é realizada majoritariamente em computadores digitais convencionais, que não são adequados para obter as vantagens desta característica.

### 2.3 ESTRUTURA DE UMA RNA

Um neurônio artificial é uma representação matemática da estrutura e funcionamento de um neurônio biológico, mesmo que de forma primitiva. São elementos de processamento divididos em diversas camadas interconectadas, que recebem um conjunto de informações de

entrada e o propagam até a camada de saída onde se obtém um novo estímulo. A estrutura de um neurônio artificial pode ser dividida em três elementos básicos (HAYKIN, 1999).

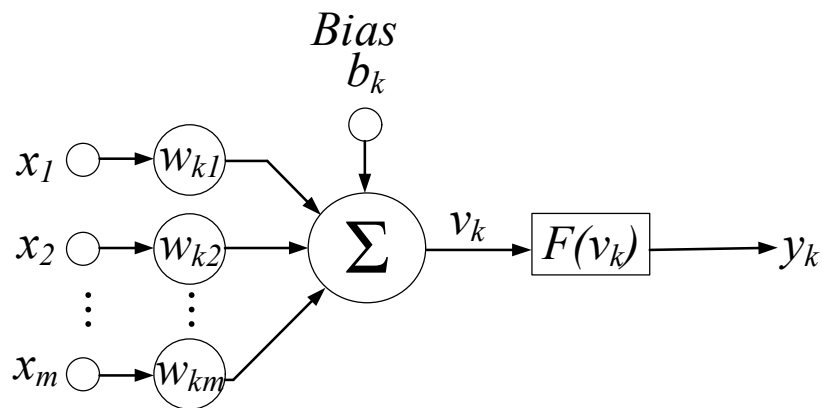
I. Um conjunto de conexões sinápticas, compostas pelos nós de entrada que são multiplicados pelos seus respectivos pesos sinápticos. Esta operação constitui um combinador linear;

II. Um somador, responsável por somar as entradas já ponderadas pelos pesos sinápticos;

III. Uma função de ativação, responsável por limitar, ou condicionar, a amplitude do sinal de saída do neurônio.

Conforme Figura 2, um neurônio denominado  $k$  recebe um sinal de entrada denominado  $x_{km}$  e o atribui a um peso sináptico  $w_{km}$  (OLIVEIRA, 2005). As entradas já ponderadas pelos pesos sinápticos são então combinadas pela ação do somador, resultando em  $v_k$ . Enquanto a função de ativação  $F$  condiciona o sinal de saída. O elemento  $b_k$ , bias, é uma constante que tem como objetivo somar ou subtrair a entrada da função de ativação. A função de ativação é escolhida para satisfazer alguma especificação do problema que os neurônios estão tentando resolver (HAGAN, 2016).

Figura 2 – Estrutura de um neurônio artificial.



Fonte: (OLIVEIRA, 2005).

Matematicamente um neurônio pode ser representado pela equação (1) (HAYKIN, 1999).

$$y_k = F(v_k + b_k) \quad (1)$$

Onde:



$$v_k = \sum_{j=1}^m w_{kj}x_j \quad (2)$$

Em que:

$x_1, x_2, x_2, \dots, x_m \rightarrow$  Vetor de entradas;

$w_{k1}, w_{k2}, w_{k2}, \dots, w_{km} \rightarrow$  Pesos sinápticos;

$v_k \rightarrow$  Saída do somador;

$b_k \rightarrow$  Bias;

$F(\ ) \rightarrow$  Função de ativação;

$y_k \rightarrow$  Saída do neurônio  $k$ .

Um neurônio artificial por si só apresenta muitas limitações. Entretanto, um conjunto interconectado destes elementos apresenta a capacidade de resolver problemas complexos (MORIGGI, 2018). A este conjunto de neurônios interconectados em forma de rede dá-se o nome de Rede Neural Artificial, que pode ser diferenciada sob três aspectos: Sua função de ativação, sua arquitetura e seu algoritmo de treinamento (MORIGGI, 2018).

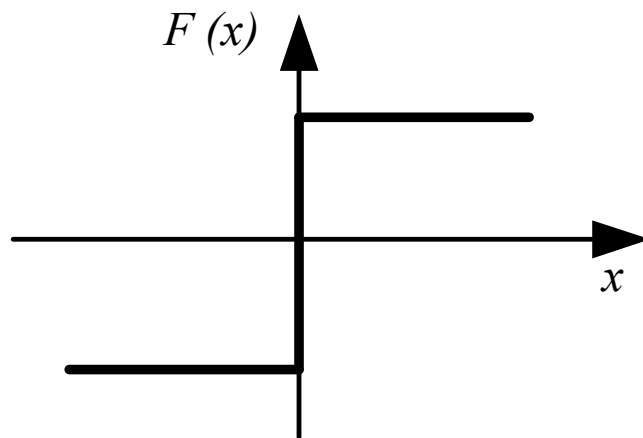
## 2.4 FUNÇÃO DE ATIVAÇÃO

A função de ativação é responsável pelo condicionamento do sinal de saída de um neurônio a partir do sinal proveniente da ação do combinador linear. Quanto aos tipos de função de ativação existe uma certa variedade, as três mais comumente utilizadas são a função limite rígido, função linear e função sigmoide (HAGAN, 2016; HAYKIN, 1999).

A função limite rígido, ou *threshold*, é utilizada em neurônios criados para classificar as entradas em duas categorias distintas (HAGAN, 2016). Esta função assume um valor  $y$  em sua saída quando seu argumento é igual ou maior do que zero e um valor  $-y$  quando o argumento é menor do que zero, conforme equação (3) e que pode ser visualizado em Figura 3.

$$f(x) = \begin{cases} -y, & \text{se } x < 0; \\ +y, & \text{se } x \geq 0; \end{cases} \quad (3)$$

Figura 3 – Função limite rígido.

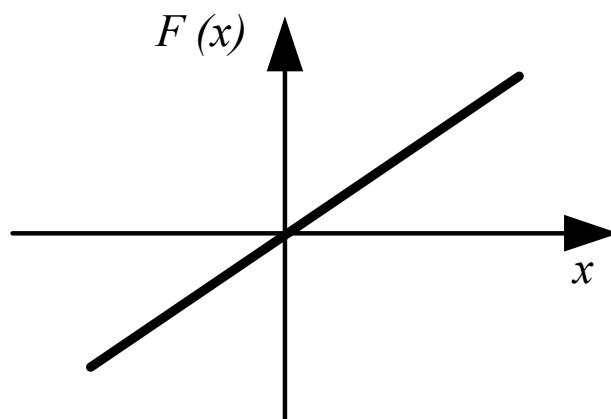


Fonte: Autor.

A função linear apresenta a característica de que sua saída é linearmente correspondente ao valor de entrada conforme equação (4) e Figura 4.

$$f(x) = ax \quad (4)$$

Figura 4 – Função linear.

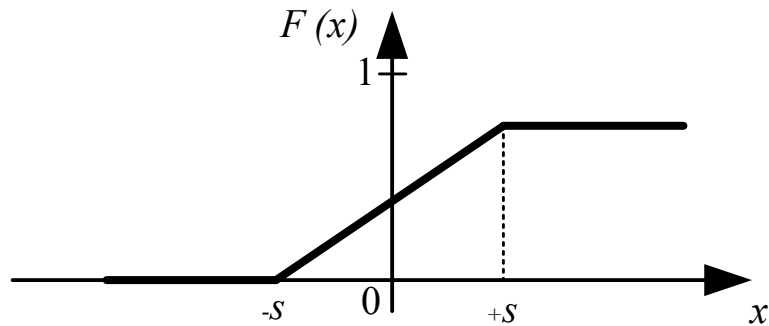


Fonte: Autor.

Uma combinação entre as duas funções apresentadas pode ser realizada para criar a função rampa. Esta função apresenta um comportamento linear em uma faixa entre  $s$  e  $-s$  e apresenta e saturada fora dela conforme equação (5) e pode ser visualizado na Figura 5.

$$f(x) = \begin{cases} 0, & \text{se } x \leq -s \\ ax, & \text{se } -s < x < +s \\ 1, & \text{se } x \geq +s \end{cases} \quad (5)$$

Figura 5 – Gráfico da função rampa.



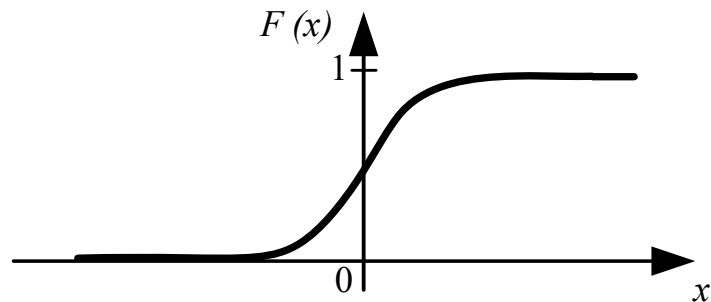
Fonte: Autor.

A última das principais funções de ativação é a função sigmoide. Esta apresenta um gráfico em forma de “S”, conforme Figura 6, e pode ser considerada uma função estritamente crescente com um balanceamento adequado entre um comportamento linear e não-linear (HAYKIN, 1999) e sua utilização pode, virtualmente, aproximar qualquer função desejada (HAGAN, 2016).

$$f(x) = \frac{1}{1 + e^{(-x/T)}} \quad (6)$$

O  $T$  visto na equação (6) corresponde a inclinação da função sigmoide, quando  $T$  tende a zero a função assume uma característica de função limite rígido (OLIVEIRA, 2005).

Figura 6– Função sigmoide



Fonte: Autor.

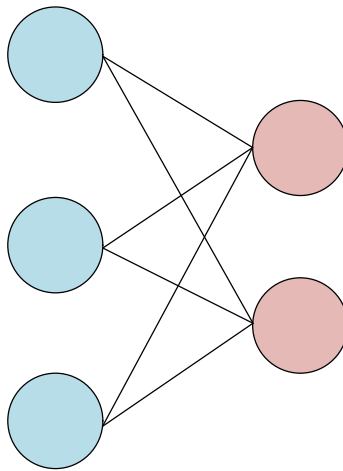
A função sigmoide é, com vantagem, a mais utilizada na construção de RNAs (HAYKIN, 1999), principalmente em redes de múltiplas camadas treinadas utilizando o algoritmo de *backpropagation* (HAGAN, 2016).

## 2.5 ARQUITETURAS DE REDE

A utilização de apenas um neurônio na resolução de problemas pode não apresentar a resposta desejada, portanto geralmente utilizam-se vários em paralelo, em uma estrutura denominada camada (HAGAN, 2016). A forma como as camadas de uma RNA estão organizadas denomina-se arquitetura de rede, ou topologia de rede (HAYKIN, 1999). Em relação ao padrão de conexões e a propagação dos dados até a camada de saída podem ser identificados três arquiteturas básicas de redes: alimentação direta de uma única camada, alimentação direta com múltiplas camadas e redes recorrentes (HAYKIN, 1999).

Em uma rede de alimentação direta de camada única tem-se um conjunto de nós de entrada que se projeta diretamente para a camada de saída, conforme Figura 7. O fluxo de informações apresenta apenas uma direção, sem realimentação. É denominada camada única por conta da camada de saída, a camada de entrada não é contabilizada pois nela não ocorre nenhuma computação (HAYKIN, 1999).

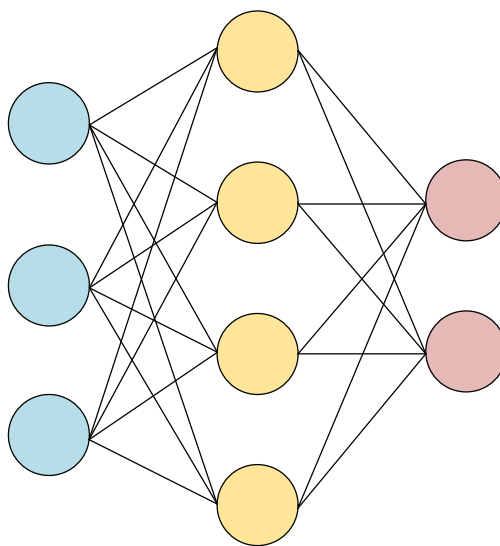
Figura 7 – Rede de alimentação direta de camada única.



Fonte: Adaptado de (HAYKIN, 1999).

Redes de alimentação direta com múltiplas camadas caracterizam-se pela presença de neurônios ocultos. Estes neurônios estão presentes em camadas localizadas entre a camada de entrada e de saída, de forma útil, possibilitando maior capacidade computacional e melhor desempenho na aproximação de funções complexas devido ao maior número de conexões sinápticas (HAYKIN, 1999). O termo oculto provém do fato de estas camadas não serem acessíveis diretamente a partir da entrada ou saída da rede.

Figura 8 – Rede de alimentação direta com múltiplas camadas.

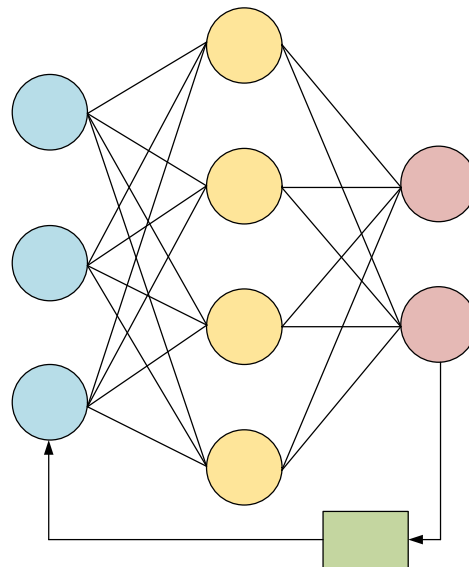


Fonte: Adaptado de (OLIVEIRA, 2005).

Nesta configuração, a saída de cada camada é a entrada da camada adjacente. Pode-se visualizar na Figura 8 que os neurônios de uma camada estão todos interconectados aos neurônios das outras camadas. Esta arquitetura de rede é mais poderosa do que a de apenas uma camada, podendo ser treinada, utilizando uma função de ativação sigmoide na primeira camada e uma função de ativação linear na segunda, para aproximar a maioria das funções consideravelmente bem (HAGAN, 2016).

As redes recorrentes distinguem-se por apresentarem um laço de realimentação conectando as camadas de saída com a de entrada, envolvendo todas as outras camadas. Esta propriedade acrescenta um comportamento dinâmico à RNA, pois a saída não depende mais apenas dos dados de entrada, mas também de seu valor atual.

Figura 9 – Exemplo de rede recorrente.



Fonte: Autor.

O exemplo visualizado na Figura 9 é uma das possibilidades para esta arquitetura de rede, pode-se também aplicar uma estrutura sem camadas ocultas, ou laços de realimentação provenientes de uma das camadas ocultas no lugar das saídas, ou não apresentar entradas externas utilizando apenas a realimentação. A inserção de laços de realimentação impacta profundamente o aprendizado e o desempenho de uma rede neural (HAYKIN, 1999). Entre as vantagens encontram-se a possibilidade de representar comportamentos temporais (HAGAN, 2016) e a capacidade de assumir um comportamento dinâmico não-linear (HAYKIN, 1999).

A escolha da arquitetura a ser utilizada depende de diversos fatores (MORIGGI, 2018).

- Grau de complexidade do problema;
- Dimensão do espaço de entrada;
- Característica dinâmica ou estática do problema;
- Conhecimento que se tem sobre o problema;
- Representatividade dos dados.

## 2.6 APRENDIZADO DE UMA RNA

Uma característica importante das RNAs é sua capacidade de aprender sobre o ambiente no qual ela será inserida e melhorar seu desempenho através do aprendizado (HAYKIN, 1999). Também chamado de processo de treinamento, é o período no qual a rede busca estabelecer a intensidade das conexões sinápticas entre os neurônios (ajuste dos pesos sinápticos) para obter a resposta que melhor represente o ambiente desejado.

O processo de treinamento constitui, de forma generalizada, a seguinte sequência de eventos (HAYKIN, 1999):

- I. A rede neural é estimulada pelo ambiente;
- II. A rede muda seus parâmetros de acordo com os estímulos;
- III. A rede neural responde de uma maneira diferente aos estímulos por conta das mudanças em sua estrutura.

O conjunto de procedimentos, bem definidos, que constituem a adaptação dos parâmetros de uma RNA para que ela desempenhe satisfatoriamente uma determinada função é denominado algoritmo de aprendizagem ou algoritmo de treinamento (OLIVEIRA, 2005).

Existe uma grande variedade de algoritmos de treinamento e cada um apresenta suas vantagens e desvantagens que também dependem da aplicação. Basicamente, os algoritmos de treinamento diferem-se pela maneira que buscam ajustar os pesos sinápticos (HAYKIN, 1999).

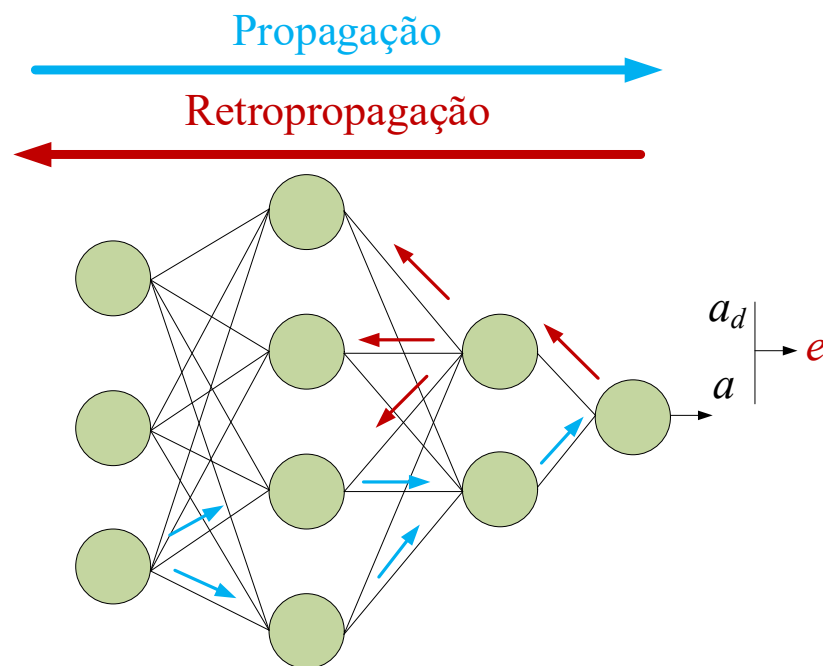
O processo de treinamento pode ser realizado de duas formas: com ou sem professor. O treinamento com professor, ou treinamento supervisionado, constitui um problema de otimização no ajuste dos pesos sinápticos e requer a apresentação do conjunto de dados diversas vezes à rede, até que o erro apurado seja satisfatório (MORIGGI, 2018). Neste tipo de treinamento são fornecidas para a rede um conjunto de informações sobre o ambiente e também a resposta desejada, a diferença entre a resposta desejada e obtida (erro) é a orientação para o ajuste dos pesos sinápticos de forma a minimizá-la.

O algoritmo de *backpropagation* e suas variações são o método de treinamento supervisionado mais utilizado. O processo de treinamento de uma RNA por este algoritmo apresenta três etapas (MORIGGI, 2018):

- Propagação dos parâmetros de entrada;
- Retropropagação do erro associado;
- Ajuste dos pesos sinápticos.

Conforme Figura 10, na fase de propagação um conjunto de informações de entrada é aplicado a rede e percorre pelas camadas, sendo ponderado pelos pesos sinápticos atuais. Ao chegar na camada de saída obtém-se a resposta  $a$  da rede para esta configuração de pesos sinápticos. Ao compará-la com a saída desejada  $a_d$  tem-se um erro  $e$  para cada elemento de saída.

Figura 10 – Esquemático do funcionamento do algoritmo de *backpropagation*.



Fonte: Adaptado de (MORIGGI, 2018).

A segunda fase é a de retropropagação, onde o erro  $e$  percorre da camada de saída para a camada adjacente no sentido contrário ao da fase anterior. Os pesos das conexões sinápticas são multiplicados pelos erros correspondentes para encontrar um valor de erro estimado para cada um dos neurônios da camada, e os pesos são ajustados pelo gradiente decrescente (MORIGGI, 2018). Isto é repetido até que todos os pesos sinápticos da rede tenham sido



ajustados. Geralmente é utilizado como critério de parada um número máximo de iterações, mas algumas variações deste algoritmo apresentam critérios de parada específicos.

Alguns problemas podem surgir durante o treinamento, como a existência de mínimos locais, tempo de treinamento indeterminado ou *overfitting* (MORIGGI, 2018). Por *overfitting* entende-se quando após um ciclo a rede passar a piorar seu desempenho durante o treinamento, pois a generalização do ambiente já foi concluída e, ao continuar o treinamento, a rede passa a sobrepor os pesos sinápticos adequados com informações sobre ruídos e peculiaridades do conjunto de dados.

Os processos de treinamento sem professor não apresentam uma comparação entre a saída obtida e desejada. Estes podem ser divididos em aprendizado por reforço e aprendizagem não-supervisionada.

A aprendizagem por reforço é geralmente utilizada em ambientes desconhecidos, fornecendo base para que haja uma interação da rede com o ambiente e que desenvolva a habilidade para a realização de tarefas a partir de sua própria experiência (OLIVEIRA, 2005). Dessa forma, ao realizar uma ação é realizada uma observação crítica do desempenho da rede, se esta ação acarreta em degradação do desempenho ela é desencorajada, enquanto ações avaliadas como positivas são reforçadas. Este paradigma de aprendizado é apropriado para situações dinâmicas em que não há como fornecer um conjunto de pares entrada-saída para o treinamento (OLIVEIRA, 2005).

Por fim, a aprendizagem não-supervisionada é utilizada quando se busca obter padrões significativos ou características nos dados de entrada sem supervisão (OLIVEIRA, 2005). Isso implica que a rede deve achar atributos estatísticos relevantes, desenvolvendo uma representação própria dos estímulos fornecidos a sua entrada, portanto deve haver redundância nos dados de entrada para que seja possível identificar padrões (OLIVEIRA, 2005).

## 2.7 ALGORITMO DE LEVENBERG-MARQUADT

O algoritmo de Levenberg-Marquadt (LM) é uma variação do algoritmo de *backpropagation*, que na verdade se trata de uma modificação do método de Newton, designado a minimizar a soma dos quadrados de outras funções não-lineares, aplicada ao treinamento de RNAs (HAGAN, 2016). Este algoritmo é adequado para o treinamento de redes neurais em que a avaliação de desempenho durante o treinamento é realizada pelo erro quadrático médio (HAGAN, 2016).

Uma particularidade do treinamento com o algoritmo de Levenberg-Marquadt é que o conjunto de dados é dividido em três. O primeiro constitui a parte que será utilizada para o

treinamento da rede de fato, realizando o ajuste dos pesos sinápticos. O segundo é o conjunto de validação, este é responsável pela parada antecipada do treinamento melhorando a eficiência e o tempo de treinamento da rede. Ou seja, quando o erro passa a aumentar o treinamento é interrompido, minimizando o problema de *overfitting*. O terceiro e último conjunto, denominado de teste, fornece uma estimativa de precisão da rede neural treinada ao aplicar um conjunto de dados que não participaram do processo de treinamento.

## 2.8 AVALIAÇÃO DE DESEMPENHO

A avaliação de desempenho das redes é realizada de duas formas: durante o treinamento e durante os testes realizados após o treinamento. Durante o treinamento é utilizado o erro quadrático médio (EQM), enquanto para os testes a serem realizados após o treinamento, aplicando conjuntos de dados desconhecidos à rede até então com intuito de validar seu desempenho é utilizado o desvio médio absoluto (DMA) e o coeficiente de determinação ( $R^2$ ).

O erro  $e_i$  de uma rede neural para um ciclo ou iteração  $n$  é dado pela equação (7). Onde  $d_i$  é a saída desejada e  $y_j$  é a resposta da rede neural.

$$e_j(n) = d_j(n) - y_j(n) \quad (7)$$

O EQM é uma grandeza utilizada durante o treinamento com o algoritmo de *backpropagation* para avaliar o desempenho da rede neural por conta da sua característica de avaliar todo o universo de amostras em uma única grandeza. Este parâmetro, conforme equação (8), é representado pelo somatório dos quadrados dos erros de todo o conjunto de amostras ( $N$ ) dividido pelo número de amostras. Quanto mais próximo de zero for seu valor, melhor o desempenho calculado.

$$EQM = \frac{1}{N} \sum_{j=1}^N (e_j)^2 \quad (8)$$

O DMA, dado pela equação (9), é uma medida de erro muito comum para avaliar erros de previsão. Ele fornece o somatório do módulo dos erros de todas as amostras de um conjunto, dividido pelo número total de amostras. O valor fornecido é sempre positivo, e quanto menor este valor melhor o desempenho da rede neural.

$$DMA = \frac{1}{N} \sum_{j=1}^N |e_j| \quad (9)$$

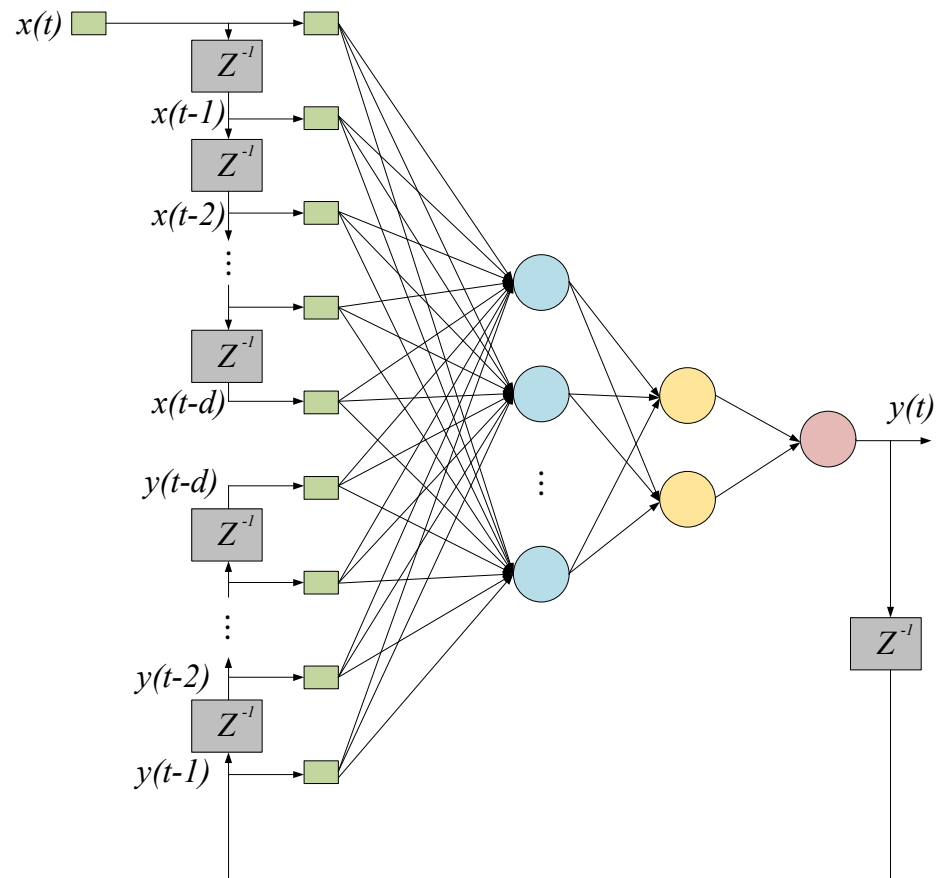
## 2.9 NARX

A previsão de eventos requer a utilização de um modelo de rede neural dinâmica. Devido a não-linearidade do comportamento térmico, e a influência da temperatura anterior nas trocas de calor que determinarão o próximo estado, além da influência de grandezas que variam com o tempo na temperatura de um transformador, o modelo de rede NARX apresenta-se como a alternativa mais viável. Este modelo apresenta um laço de realimentação conectando suas diversas camadas. Uma vantagem da utilização deste modelo em relação a outros modelos dinâmicos de redes neurais é que o modelo NARX apresenta dinâmica tanto na camada de saída como na camada de entrada, enquanto outros modelos apresentam dinâmica apenas em uma das camadas. Em resumo, as entradas e as saídas do estado atual têm influência direta no valor de saída do próximo estado, representando o que ocorre em um transformador real devido a variações da temperatura ambiente ou de seu carregamento. A equação que descreve o comportamento deste modelo de rede é descrita em equação (10).

$$y(n) = f(y(n-1), y(n-2), \dots, y(n-d_y), u(n-1), u(n-2), \dots, u(n-d_u)) \quad (10)$$

Onde  $u$  é o vetor de entradas e  $y$  o vetor de saídas. A arquitetura desta rede pode ser visualizada na Figura 11.

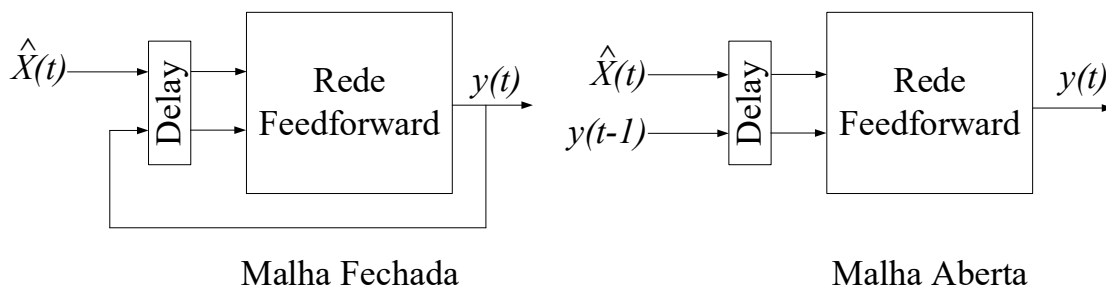
Figura 11– Estrutura de uma rede NARX.



Fonte: (ANDRADE et. al., 2014).

Apesar de se tratar de uma rede com feedback, ou malha fechada, este modelo de rede pode ser representado em malha aberta se considerado o valor desejado de saída do estado anterior  $y(n-1)$ , este que se encontra disponível durante o treinamento, como uma nova entrada no lugar do laço de realimentação, conforme Figura 12. Dessa forma, pode-se considerar a rede como de alimentação direta e utilizar um algoritmo de treinamento *backpropagation* de implementação mais simples, ao invés de *backpropagation* dinâmico, melhorando a eficiência e precisão do treinamento, pois a rede baseia-se em um estado de saída anterior  $y(n-1)$  sem erro.

Figura 12 – Arquiteturas de uma rede NARX.



Fonte: (HAGAN, 2016).

Após o treinamento a rede é utilizada em malha fechada para a realização dos testes de sua capacidade de predição. Ou seja, o valor de saída fornecido pela resposta da rede é quem terá influência na determinação do próximo valor de temperatura e não mais o valor de saída desejado.

## 2.10 REDES NEURAIS ARTIFICIAIS NA PREDIÇÃO DE TEMPERATURA EM TRANSFORMADORES

Uma das aplicações de RNAs que apresenta bons resultados é a predição de comportamentos não-lineares, incluindo predição de temperatura em transformadores (SU et. al., 2018) (WEI et. al., 2017) (HUANG et. al., 2016) (VILAITHONG et. al., 2007) (HE et. al., 2000) (PRADHAN, RAMU, 2004). Uma breve apresentação sobre o que foi desenvolvido por estes autores é descrita nos próximos parágrafos.

He et. al. (2000) apresentam comparações entre quatro diferentes arquiteturas de redes neurais com diferentes configurações para a predição de temperatura de topo de óleo em quatro transformadores de 22 MVA. As arquiteturas utilizadas foram uma rede de alimentação direta (FFNN), uma rede com processamento temporal (TPNN), uma rede com realimentação (RNN). Estas três apresentam propriedades não-lineares, e o modelo *Linear AutoRegressive with Exogenous Inputs* (ARX). Dentre as quatro arquiteturas apenas a rede com realimentação apresenta dinâmica tanto na camada de entrada quanto na camada de saída, enquanto ARX e TPNN apresentam dinâmica apenas na camada de entrada e FFNN não apresenta dinâmica. São consideradas como entradas das redes neurais apenas o carregamento dos transformadores e a temperatura ambiente, onde foram utilizados para o treinamento dados de cinco dias de operação e outros cinco dias de monitoramento para realização dos testes de capacidade de

predição das redes neurais desenvolvidas. O modelo RNN apresentou os menores erros, tanto em média para o conjunto de amostras como o menor erro máximo observado.

Pradhan e Ramu (2004) apresentam um método de combinação linear ótima para oito RNAs de arquiteturas e configurações diferentes para a predição da temperatura de topo de óleo e temperatura de topo mais quente. Basicamente, o melhor resultado, de acordo com o treinamento, para determinadas informações de entrada é selecionado e a resposta dos restantes descartada.

É proposto por Vilaithong et. al. (2007) uma rede de alimentação direta com múltiplas camadas para a predição da temperatura de topo de óleo. A rede é treinada para três algoritmos de treinamento *backpropagation* e configurações diferentes. Uma grande quantidade de dados é utilizada durante o treinamento, e os objetos de estudo são três transformadores de elevada potência (40, 150 e 850 MVA). Por utilizar um modelo de rede neural estático, é proposto utilizar como entrada, juntamente do carregamento, temperatura ambiente e *status* do sistema de refrigeração a ar, o último estado da temperatura de topo de óleo, simulando uma dinâmica na camada de saída.

Em Huang et. al. (2016) é proposto um modelo de predição da temperatura de topo de óleo baseado no modelo *Kernel-based extreme learning machine* (KELM), este que seria um tipo de rede de alimentação direta com bom desempenho em regressão não-linear e classificação. Entretanto, apresenta dinâmica apenas na camada de entrada. O objeto de estudo é um transformador de 50 MVA, onde foram utilizados para treinamento um conjunto de dados de um mês de monitoramento e um conjunto de cinco meses para a realização dos testes de validação da capacidade de predição da rede neural desenvolvida. São consideradas apenas duas variáveis de entrada: a corrente de carga e temperatura ambiente.

Wei et. al. (2017) propõe um modelo baseado em redes neurais para determinação da temperatura de ponto mais quente para um transformador de 250 MVA, onde foram utilizadas apenas 80 amostras de dados de monitoramento para treinamento e 20 para os testes de validação da capacidade preditiva. A principal proposta é uma otimização do algoritmo de treinamento Levenberg-Marquadt para uma rede de alimentação direta com múltiplas camadas. As entradas utilizadas foram temperatura ambiente, carregamento, temperatura de topo de óleo e estado anterior da elevação de temperatura de topo de óleo, enquanto as saídas são compostas pela temperatura de ponto mais quente, constante do óleo e expoente do óleo. A adição do estado anterior da elevação de temperatura de topo de óleo como uma das entradas fornece propriedades dinâmicas à rede neural em sua camada de entrada, apesar de o modelo utilizado ser estático.

A predição de temperatura de topo de óleo utilizando uma rede de Elman é proposta por Su et. al. (2018). Este modelo seria uma modificação de uma rede de alimentação direta com múltiplas camadas com a adição de uma “camada de recebimento”, adicionando dinâmica à camada de entrada. São utilizados como parâmetros de entrada a temperatura da base de óleo, temperatura ambiente, corrente de carga e potência ativa. Foram utilizadas 300 amostras de um período de 24 horas, onde 270 constituíram o conjunto de treinamento e 30 para os testes da capacidade de predição.

Dentre os trabalhos apresentados, apenas Vilaithong et. al. (2007) faz utilização de um grande conjunto de dados. Isto é um fator importante para o treinamento e para a validação da capacidade preditiva da rede neural, pois busca-se a maior abrangência das condições de operação possíveis para o transformador. Outros dois pontos que fazem exceção apenas ao trabalho de Vilaithong et. al. (2007) são:

- A utilização de poucas entradas, não levando em consideração a influência dos trocadores de calor, ou de muitas entradas que acabam por ser redundantes;
- Arquiteturas de RNAs com dinâmica apenas na camada de entrada, não levando em consideração a influência do estado atual da temperatura para a determinação do seguinte.

Portanto, pode-se notar que Vilaithong et. al. (2007) apresenta aspectos semelhantes aos abordados neste trabalho. Entretanto, podem ser identificados dois problemas: (1) Ao inserir como um dos parâmetros de entrada a saída do estado anterior, cria-se uma maior complexidade em questões de estruturar a RNA para predição de temperatura em longos períodos de tempo, além de impossibilitar a consideração da influência de estados anteriores (como por exemplo, para uma amostragem de 15 minutos é recomendável levar em consideração dois estados anteriores para uma melhor representação da dinâmica do transformador), sendo que existem modelos de RNAs que possibilitam fazer isso de forma mais simples. (2) Não é considerada a dinâmica na camada de entrada, desta forma a rede pode não reconhecer/adquirir certos tipos de comportamento que acontecem por influência de variações em algum dos parâmetros de entrada.

## 2.11 CONSIDERAÇÕES DO CAPÍTULO

O presente capítulo apresentou uma breve introdução ao histórico de redes neurais artificiais, suas características, vantagens e aplicações. Foram apresentados também os

conceitos que envolvem esta ferramenta, o modelo de rede utilizado no presente trabalho e as características provenientes de sua utilização.

Discorreu-se sobre trabalhos encontrados na literatura que fizeram uso de RNAs com um propósito semelhante, justificando a escolha pelo modelo de rede NARX em contrapartida às limitações dos modelos utilizados nos trabalhos citados, e identificados os parâmetros estruturais utilizados no desenvolvimento das RNAs para predição de temperatura em transformadores. Dessa forma foram apresentados os conceitos necessários para a compreensão e justificou-se os critérios utilizados para o desenvolvimento das RNAs no Capítulo 3.



### **3. EMPREGO DE RNAS PARA PREDIÇÃO DE TEMPERATURA EM TRANSFORMADORES DE POTÊNCIA**

#### **3.1 OBJETOS DE ESTUDO**

Os transformadores cujos dados de monitoramento foram utilizados para o desenvolvimento deste trabalho têm suas características apresentadas abaixo. Estes cinco equipamentos constituem o conjunto de treze transformadores em operação na Usina Hidrelétrica Santo Antônio, situada no Rio Madeira em Porto Velho – RO, Brasil.

- TR1  
Tensão: 13,8/525 kV  
Potência Nominal: 330 MVA  
Modo de Resfriamento: ODAF
- TR2  
Tensão: 13,8/525 kV  
Potência Nominal: 330 MVA  
Modo de Resfriamento: ODAF
- TR3  
Tensão: 13,8/525 kV  
Potência Nominal: 330 MVA  
Modo de Resfriamento: OFAF
- TR4  
Tensão: 13,8/525 kV  
Potência Nominal: 330 MVA  
Modo de Resfriamento: OFAF
- TR5  
Tensão: 13,8/525 kV  
Potência Nominal: 330 MVA  
Modo de Resfriamento: OFAF

Todos os transformadores citados estão ligados à quatro unidades geradoras de 70 MW. Estas que são de cinco pás para o TR1 e TR2 e de quatro pás para os transformadores TR3, TR4 e TR5.

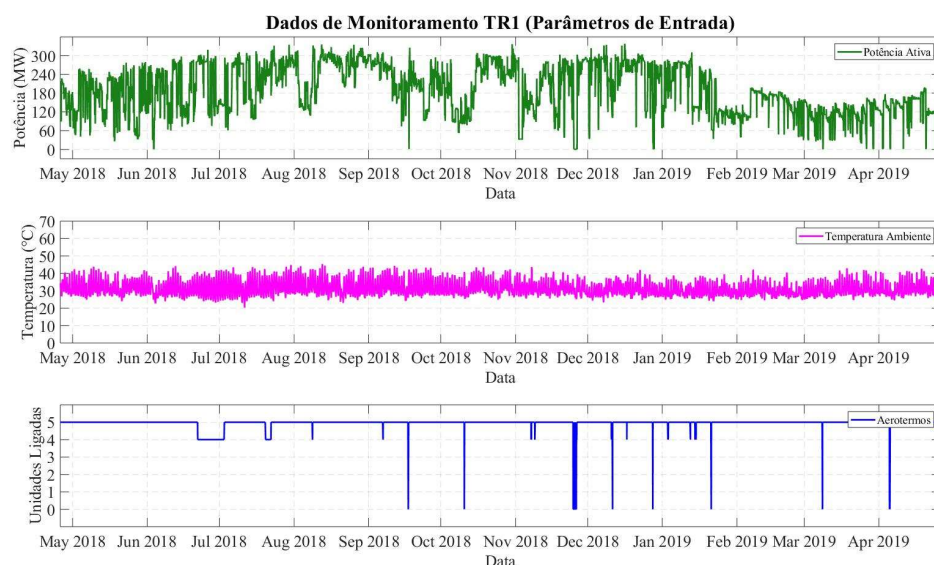
### 3.2 DADOS DE MONITORAMENTO

Os dados de monitoramento constituem dois períodos diferentes de um ano para cada um dos transformadores. O primeiro sendo de 06 de Novembro de 2016 a 06 de Novembro de 2017, e o segundo de 26 de Abril de 2018 a 26 de Abril de 2019.

As grandezas monitoradas são a potência ativa do transformador, a temperatura ambiente, o número de aerotermos em operação (conjunto ventiladores e bombas de óleo) e a temperatura de topo de óleo. A amostragem é de 10 minutos, resultando, para um período de um ano, em cerca de 52 mil pontos para cada uma das grandezas.

A Figura 13 ilustra a exemplo do TR1 os dados de monitoramento utilizados como parâmetros de entrada para o desenvolvimento das redes neurais para um dos períodos de um ano. O número de conjunto de aerotermos ligados está representado em azul, e pode variar de 0 a 5 unidades ligadas. Seu impacto na temperatura de operação do transformador está relacionado à velocidade de escoamento do óleo e à eficiência das trocas térmicas com o ambiente através da circulação do óleo pelos radiadores com o auxílio dos ventiladores. Portanto, é esperado que quanto maior o número de unidades ligadas menor seja a temperatura de operação do transformador.

Figura 13 – Dados de monitoramento utilizados como parâmetros de entrada das RNAs.



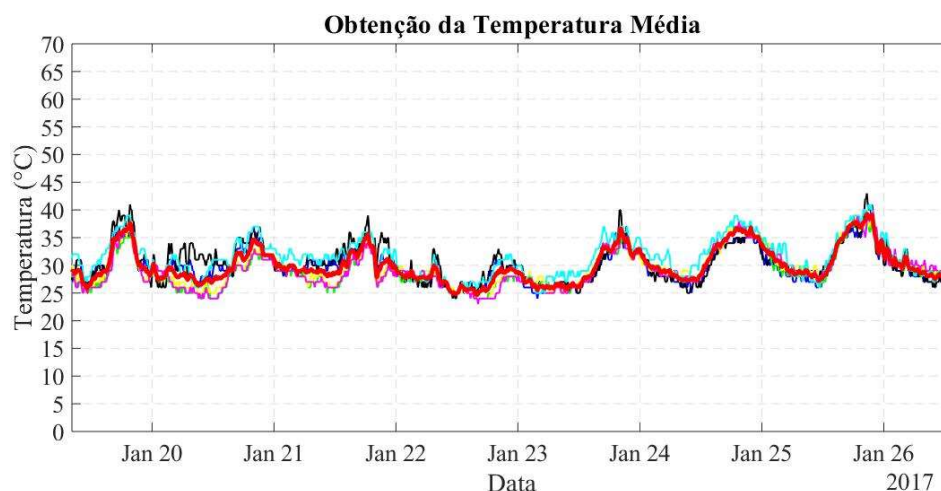
Fonte: Autor.

A potência ativa do transformador, ilustrada em verde na Figura 13, é a grandeza utilizada para expressar a quantidade de corrente que circula pelos enrolamentos. Como

transformadores são energizados a um nível de tensão praticamente constante, um aumento na potência resulta em maiores correntes e, conseqüentemente, em maiores perdas ôhmicas nos enrolamentos, resultando em elevação da temperatura caso a transferência de calor ao ambiente externo não aumente nas mesmas proporções. O transformador TR1 opera com um nível de potência considerável o ano inteiro, o mesmo não ocorre para todos os transformadores. Por conta da sazonalidade do fluxo do Rio Madeira e por se tratar de uma usina fio d'água, alguns transformadores operam com carga reduzida, a exemplo do TR2 e TR5, ou encontram-se desligados, como o TR3 e TR4, durante o período de julho a dezembro. Alguns transformadores também apresentam maiores variações nos níveis de potência observados devido a paradas para manutenção ou limpeza por conta dos desafios operacionais enfrentados na usina.

A temperatura ambiente, representada em rosa na Figura 13, tem grande influência nas trocas térmicas que ocorrem entre o transformador e o ambiente durante a dissipação do calor gerado pelas perdas. Por se tratar de uma localização de clima quente, a maior parte do ano a temperatura se mantém entre 30 e 40 °C, entretanto é observada a ocorrência de variações de mais de 10 °C durante um mesmo dia. Foram observados dados de monitoramento da temperatura ambiente para doze transformadores, cada um apresenta um sensor junto de sua localização, e constatou-se diferenças nas medições para mesmos instantes de tempo. Por conta disso foram excluídos os casos extremos e uma média foi obtida a partir das seis medições remanescentes para os períodos, conforme Figura 14, esta que será utilizada para o desenvolvimento do estudo para todos os cinco transformadores.

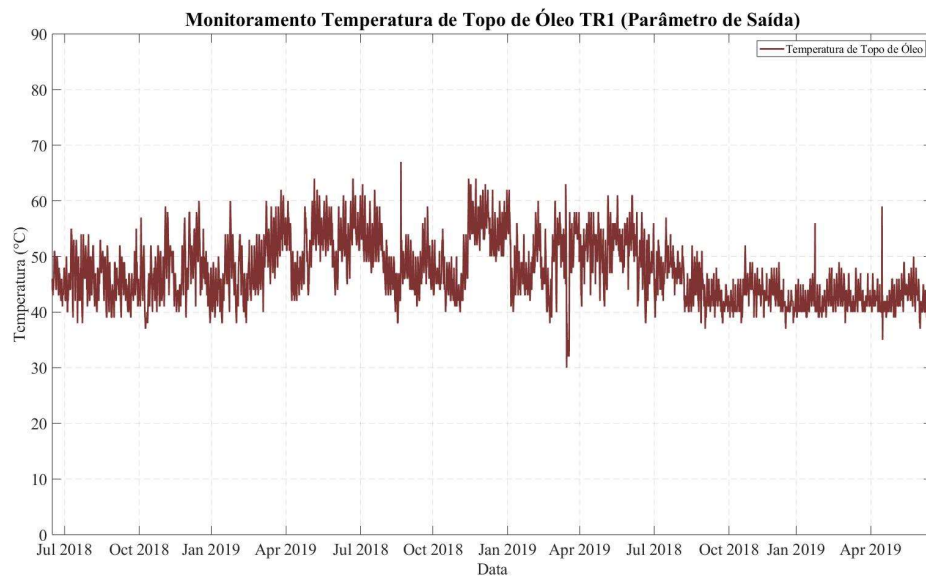
Figura 14 – Temperatura média (vermelho) utilizada como parâmetro de entrada.



Fonte: Autor.

A temperatura de topo de óleo, cujos valores medidos podem ser visualizados na Figura 15 para um dos períodos de um ano, é o parâmetro de saída das redes neurais a serem desenvolvidas. Esta grandeza será utilizada como critério de comparação para ajuste dos pesos sinápticos durante o processo de treinamento e também como base para comparação à resposta da rede durante os testes de capacidade de predição.

Figura 15 – Monitoramento da temperatura de topo de óleo TR1.



Fonte: Autor.

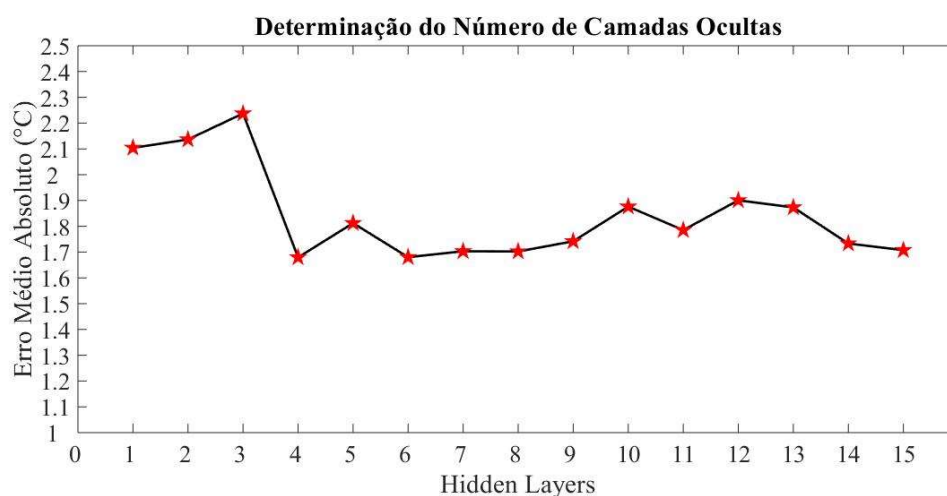
### 3.3 DESENVOLVIMENTO DAS RNAS

O desenvolvimento das RNAs foi realizado através da *toolbox* de redes neurais do *software* MATLAB. O modelo NARX foi escolhido como arquitetura de rede, como justificado anteriormente. O número de estados anteriores que interferem na obtenção da temperatura de topo de óleo pela rede utilizado é de três, ou seja, 30 minutos. Isto é realizado através da observação da influência das grandezas na dinâmica da temperatura do transformador a partir dos dados de monitoramento e este intervalo se encontra de acordo com outros estudos (HUANG et. al., 2016).

Para a determinação do número de camadas ocultas não existe um método específico, e geralmente é levado em consideração o conhecimento que se tem sobre a complexidade do problema e a realização de testes preliminares. Portanto, para uma rede NARX, utilizando o algoritmo de treinamento Levenberg-Marquadt, o número de atrasos nas camadas de entrada e

saída igual a três, treinou-se uma RNA para diversas quantidades de camadas ocultas e avaliou-se o erro obtido. Os dados utilizados para o treinamento foram referentes ao período 1 (Novembro de 2016 a Novembro de 2017) do transformador TR1 e, após o treinamento, aplicou-se às entradas da rede os dados de monitoramento referentes ao período 2 (Abril de 2018 a Abril de 2019) do transformador TR5. A resposta da rede neural foi então comparada com os dados de monitoramento de temperatura de topo de óleo referente ao mesmo transformador e mesmo período das entradas aplicadas para se obter o erro médio absoluto.

Figura 16 – Erro para quantidades de camadas ocultas.

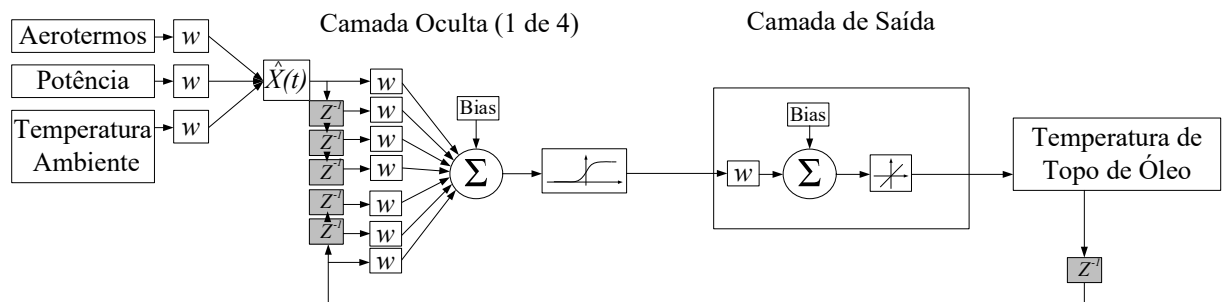


Fonte: Autor.

O número de camadas ocultas que apresentou o menor erro foi quatro, conforme Figura 16, com cerca de 1,67 °C. Hagan (2016) afirma que duas ou três camadas ocultas são suficientes para aproximar qualquer função desejada, desde que os outros parâmetros da RNA desenvolvida estejam bem estruturados. Foi observado durante a realização deste teste que, apesar de apresentarem um bom desempenho em relação ao desvio médio do valor desejado, o aumento do número de camadas ocultas acarreta no aumento do erro máximo, em módulo, observado e no tempo de treinamento da rede.

Uma representação simplificada, não levando em consideração as várias conexões entre os neurônios das diversas camadas e sim o fluxo de informações, da estrutura da rede neural desenvolvida pode ser visualizado na Figura 17.

Figura 17 – Representação simplificada da estrutura das RNAs desenvolvidas.



Fonte: Autor.

### 3.4 TREINAMENTO

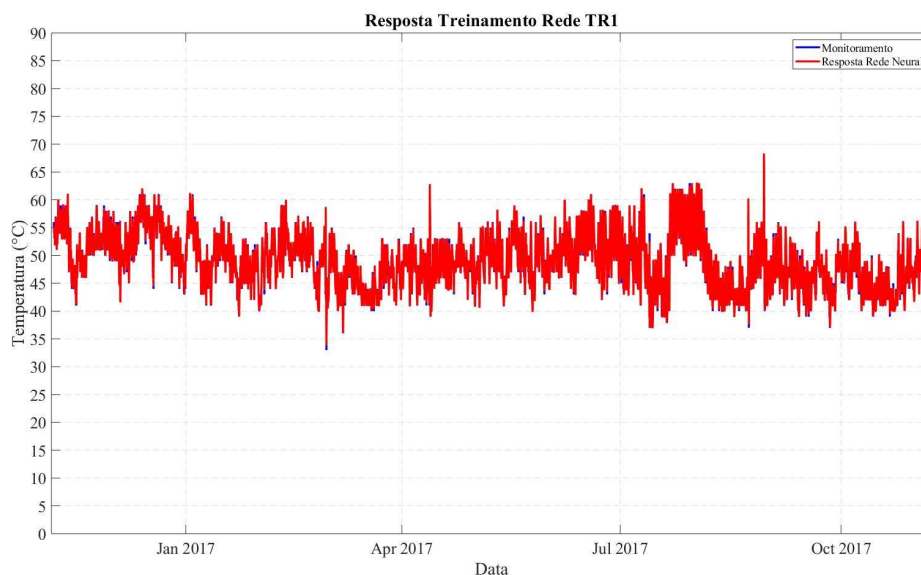
Como a arquitetura de RNA utilizada torna possível a realização do treinamento supervisionado em malha aberta. Este é realizado pelo algoritmo de Levenberg-Marquadt, também utilizado em He et. al. (2000), Huang et. al. (2016) e Vilaithong et. al. (2007). Todos os treinamentos realizados utilizam um dos períodos de um ano de monitoramento de um transformador. Dentro disso o conjunto de dados é dividido em três, são os conjuntos de treinamento, teste e validação. O conjunto de treinamento, responsável pelo ajuste dos pesos sinápticos, recebeu 90% dos dados que compõe o período de um ano. O conjunto de validação é responsável pela parada antecipada do treinamento com o intuito de evitar o *overfitting* e recebeu a parcela de 5%. O conjunto de teste é utilizado para uma avaliação prévia, indicando se o treinamento obteve sucesso ao representar o conjunto de dados de treinamento e recebeu os 5% restantes. As porcentagens foram escolhidas de forma a dedicar uma amostra abrangente ao treinamento, enquanto um período significativo de aproximadamente três semanas é fornecido ao conjunto de validação e também ao conjunto de teste.

A parada do treinamento ocorre, para todos os casos presentes neste trabalho, por conta do conjunto de validação. Após algumas iterações do treinamento é realizado um teste com o conjunto de validação, se o desempenho da rede melhora em relação ao treinamento este prossegue. Se em seis tentativas seguidas de testes com o conjunto de validação o desempenho da rede permanecer estagnado o treinamento é interrompido e a rede assume os pesos sinápticos estabelecidos para o melhor desempenho.

A Figura 18 apresenta a resposta da rede, em comparação aos dados de monitoramento de temperatura de topo de óleo, obtida durante o treinamento. Como o treinamento baseia-se em um estado anterior de temperatura de topo de óleo sem erro, os desempenhos de todos os treinamentos realizados neste trabalho são semelhantes. Também são semelhantes outros

aspectos do treinamento, como o tempo, que fica entre 5 a 10 segundos, com cerca de 40 iterações.

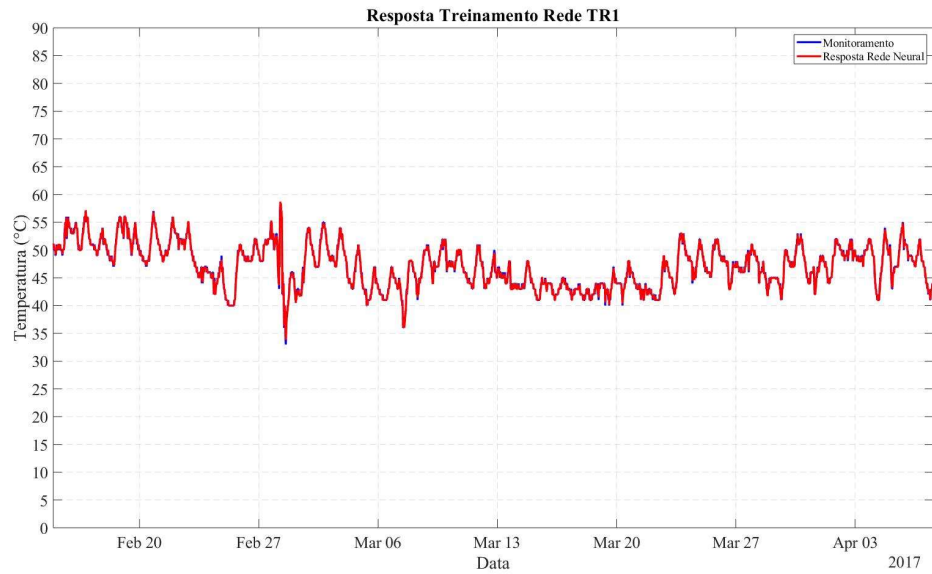
Figura 18 – Visão geral da resposta da RNA durante o treinamento.



Fonte: Autor.

A Figura 19 apresenta com mais detalhes o comportamento da resposta da rede durante o treinamento. A avaliação de desempenho durante o treinamento é realizada pelo Erro Quadrático Médio, muito utilizado durante o treinamento de redes neurais principalmente pela sua característica em transformar erros de um período em apenas uma grandeza positiva, e foi observado um valor de 0,19 °C.

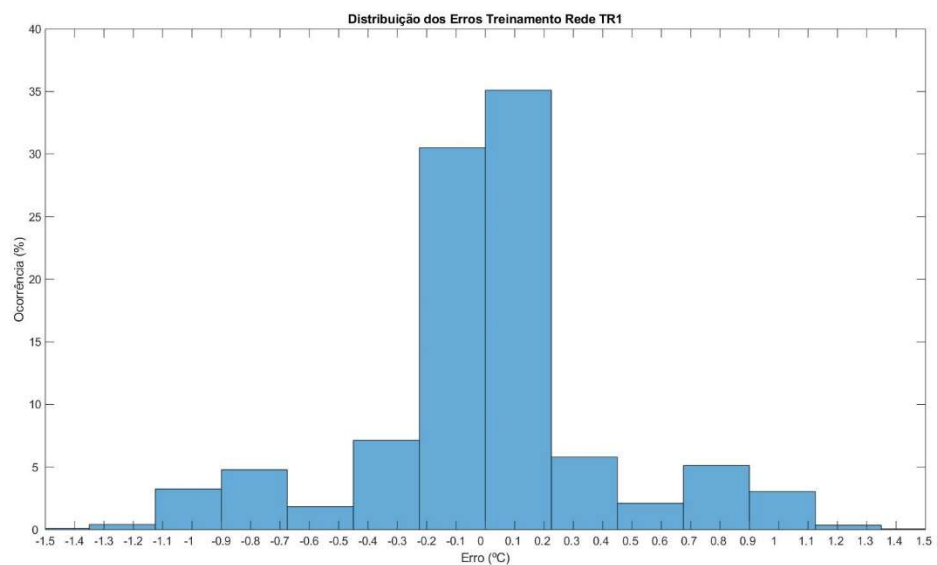
Figura 19 – Resposta RNA vs. Dados de monitoramento para o treinamento.



Fonte: Autor.

O Desvio Médio Absoluto observado durante o treinamento foi de 0,283 °C, e sua distribuição em valores percentuais pode ser visualizada no histograma da Figura 20. Pode-se observar que 65% dos erros encontram-se abaixo de 0,2 °C, enquanto quase a totalidade dos erros observados encontra-se abaixo de 1 °C.

Figura 20 – Histograma de erros para o treinamento.

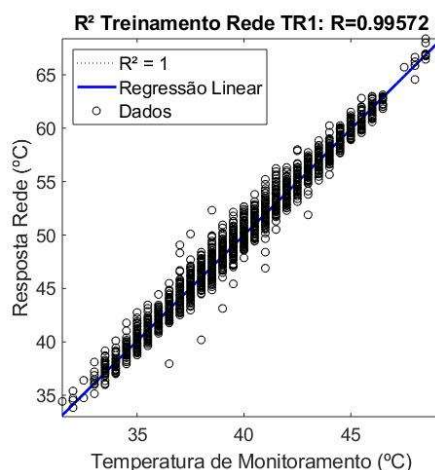


Fonte: Autor.



A Figura 21 representa a correlação entre a resposta da rede durante o treinamento e os dados de monitoramento com um coeficiente de determinação de mais de 0,99. Pode-se observar que em praticamente todas as situações a resposta encontra-se próxima do valor desejado, confirmando o bom desempenho apresentado.

Figura 21 – Correlação das temperaturas RNA e monitoramento para o treinamento.



Fonte: Autor.

O bom desempenho durante o processo de treinamento indica que os pesos sinápticos obtidos durante esta etapa representam bem o ambiente que se está buscando representar, neste caso o comportamento da temperatura em transformadores de potência. Entretanto, não serve como indicativo da capacidade preditiva da rede neural desenvolvida, pois podem ser encontradas novas situações de operação que não foram abordadas no conjunto de dados de treinamento. Portanto, é necessária a realização de testes nas mais diversas condições para avaliação do desempenho desta metodologia.

### 3.5 TESTES DA CAPACIDADE DE PREDIÇÃO DAS RNAS

Com o intuito de avaliar a validade do método utilizado para diversos cenários, foi treinada uma rede neural com dados de monitoramento de cada um dos transformadores e aplicados a cada uma delas os dados de entrada de cada um dos cinco transformadores. Desta forma pode-se observar em que circunstâncias se obtém os melhores e piores desempenhos de acordo com as particularidades que o conjunto de dados de monitoramento de cada um dos transformadores apresenta.

Todos os testes apresentados nesta seção são realizados com um conjunto de dados que não participou da etapa de treinamento, ou seja, é desconhecido à rede. Nos casos em que há comparação entre rede treinada e aplicada a um mesmo transformador cada uma das etapas corresponde a um dos períodos de monitoramento, por exemplo no caso do TR1 a rede é treinada com dados do período 1 e, após, são aplicados os dados de entrada do período 2. Então, a resposta da rede neural é comparada com os dados de monitoramento de temperatura de topo de óleo para avaliação de desempenho.

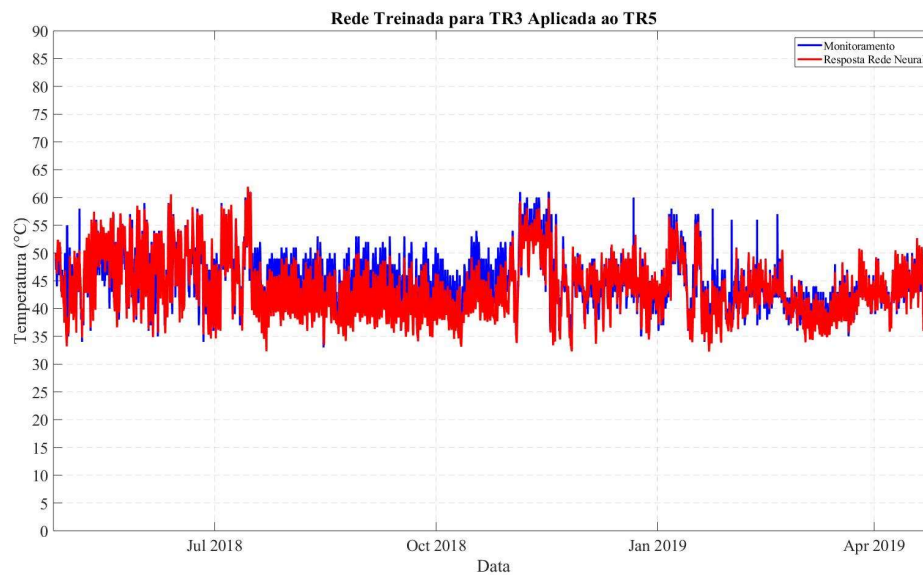
Os casos demonstrados foram escolhidos com o intuito de apresentar as mais diversas situações observadas e também abordar os cinco transformadores sob estudo. Portanto, como primeiro caso tem-se uma RNA treinada com dados do TR3 e aplicados os dados de entrada do TR5. Na sequência é apresentada uma RNA cujo treinamento foi realizado com dados do TR1 e aplicados os dados de entrada do TR5. O terceiro caso apresenta uma RNA treinada com dados do TR2 e aplicados dados de entrada do TR1, enquanto o quarto caso apresenta a situação contrária à vista no primeiro: uma rede treinada com dados do TR5 e aplicadas as entradas do TR3. Como quinto e último caso apresentado tem-se uma RNA treinada com dados de monitoramento do TR1 e aplicados os dados de entrada do TR4.

Após os testes de capacidade de predição é apresentada uma visão geral dos resultados para todas comparações realizadas, através da Tabela 1 e Tabela 2. Ao final do Capítulo 3 são discutidas as situações que ocasionam os erros máximos observados e uma elevação do erro médio.

Vale ressaltar novamente que, apesar de o treinamento ser realizado em malha aberta, utilizando o conjunto de dados de monitoramento de temperatura de topo de óleo referentes ao período como uma das entradas para a diminuição do erro e ajuste mais preciso dos pesos sinápticos, após esta etapa utiliza-se a rede em malha fechada. Com isso, são fornecidos apenas os parâmetros de entrada e os valores anteriores de temperatura de topo de óleo são fornecidos pelo laço de realimentação.

A Figura 22 apresenta uma visualização geral da aplicação dos dados de monitoramento do transformador TR5 em uma rede neural treinada com os dados do transformador TR3. Em azul tem-se os valores medidos pelo monitoramento do TR5 e em vermelho a resposta obtida pela rede neural desenvolvida.

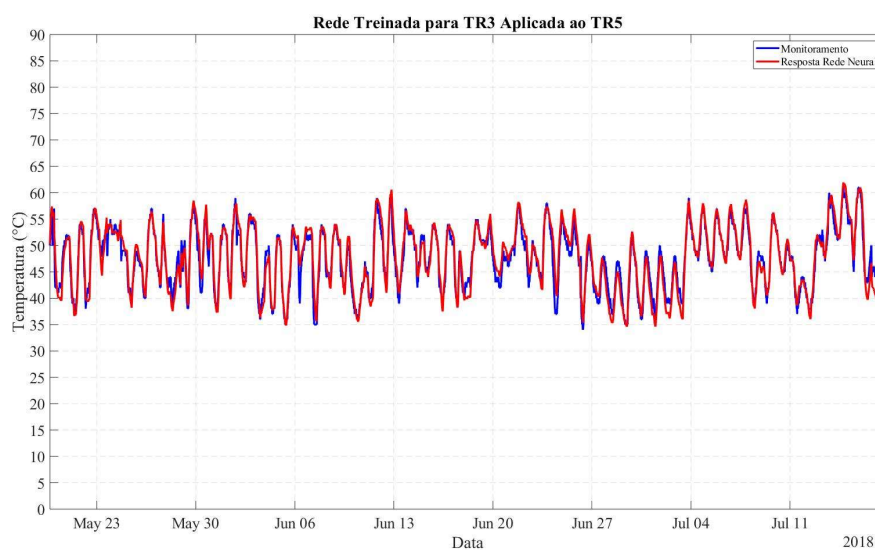
Figura 22 – Comparação das temperaturas resposta RNA vs. monitoramento.



Fonte: Autor.

O erro médio absoluto observado para o período de um ano é de 1,71 °C e o erro máximo absoluto foi de 20,6 °C. Pode-se visualizar pela Figura 23 que a resposta da rede segue o mesmo perfil da temperatura medida pelo monitoramento durante todo o período. A Figura 23 representa um período de aproximadamente 2 meses onde o erro se mantém em valores ínfimos, abaixo de 1 °C em praticamente todos os instantes e nunca maior que 5 °C, demonstrando um excelente desempenho para esta aplicação.

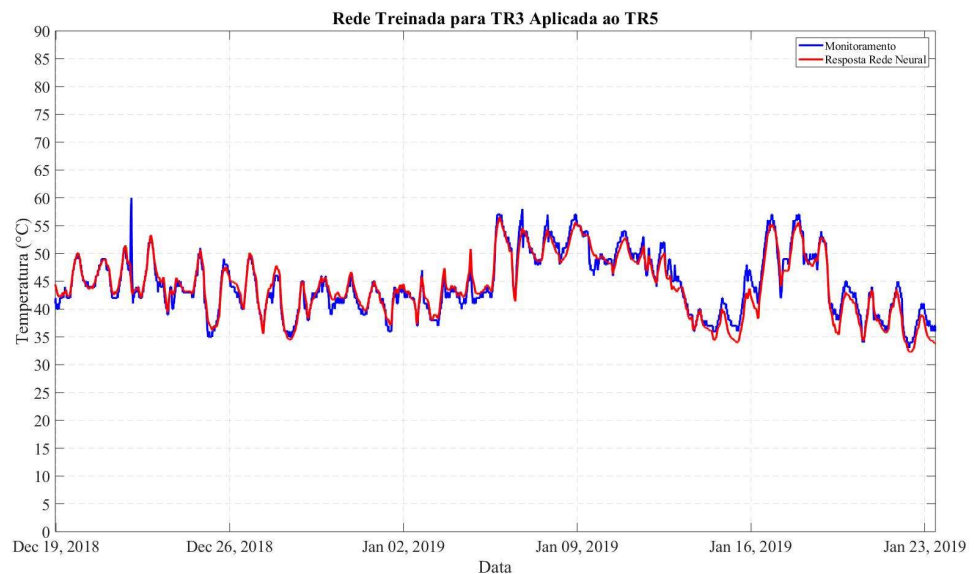
Figura 23 – Comparação das temperaturas resposta RNA vs. monitoramento.



Fonte: Autor.

A Figura 24 mostra um bom comportamento para grandes variações de temperatura para um período pouco maior que um mês. Entretanto, este caso apresenta um pico de temperatura nos dados de monitoramento, comportamento este que a rede neural não obteve sucesso ao replicar. Isto ocorre pelas poucas ocorrências deste tipo de comportamento em comparação ao período total utilizado para treinamento, e serão discutidos ao final deste capítulo.

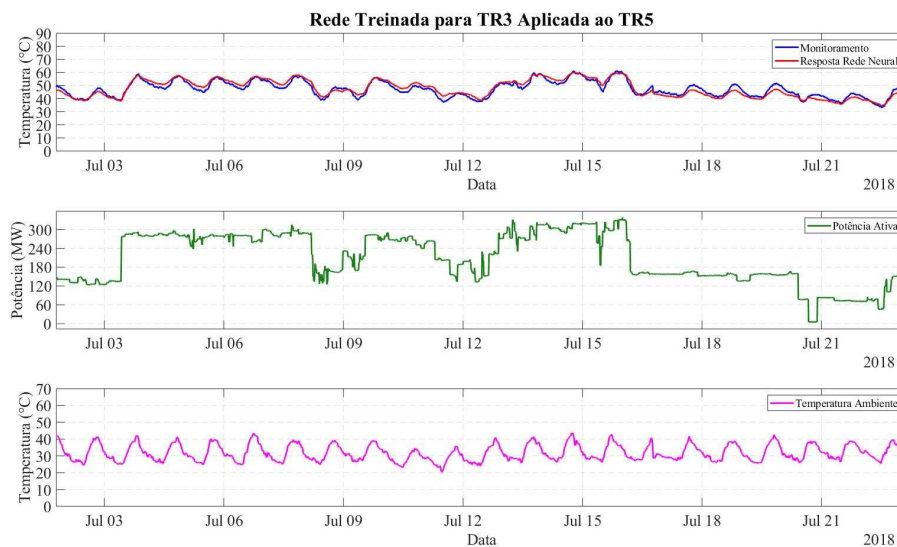
Figura 24 – Comparação das temperaturas resposta RNA vs. monitoramento.



Fonte: Autor.

Com o intuito de avaliar o comportamento da rede em comparação aos dados de monitoramento, a Figura 25 acrescenta a visualização da potência e a temperatura ambiente. Pode-se perceber que durante o período a temperatura prevista apresenta valores muito próximos dos dados de monitoramento para grandes variações da potência de operação do transformador. O período visualizado também apresenta variações de temperatura ambiente que podem ultrapassar 10 °C durante o período de um dia, e o desempenho da rede é satisfatório ao replicar a influência destas variações na temperatura de topo de óleo.

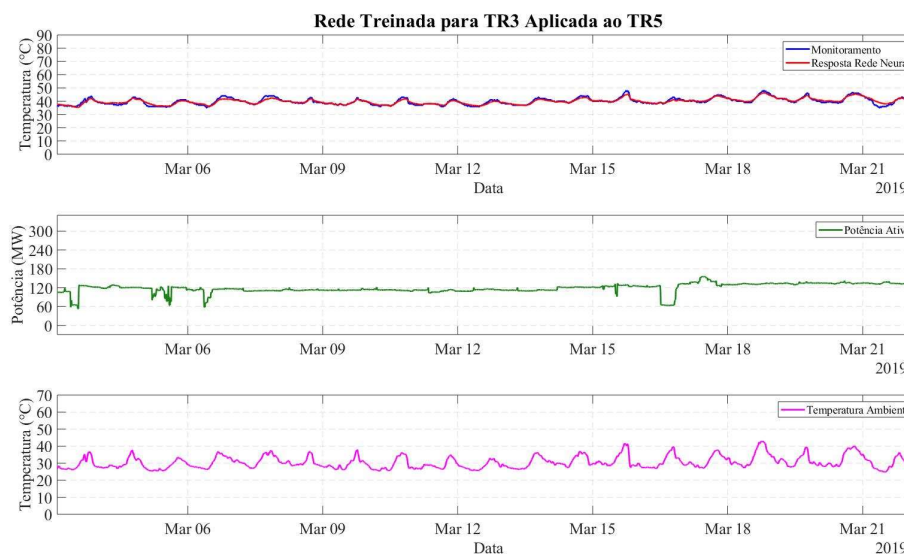
Figura 25 – Comparação das temperaturas mais parâmetros de entrada.



Fonte: Autor.

O desempenho em relação a variações na temperatura ambiente pode ser visualizado com maior clareza na Figura 26, que ilustra um período sem grandes variações na potência do transformador. Pode ser visto que neste caso a rede neural apresenta uma resposta ainda mais próxima dos dados de monitoramento do que visto na Figura 25, comprovando a sua capacidade de representar a influência da temperatura ambiente na temperatura de topo de óleo do transformador.

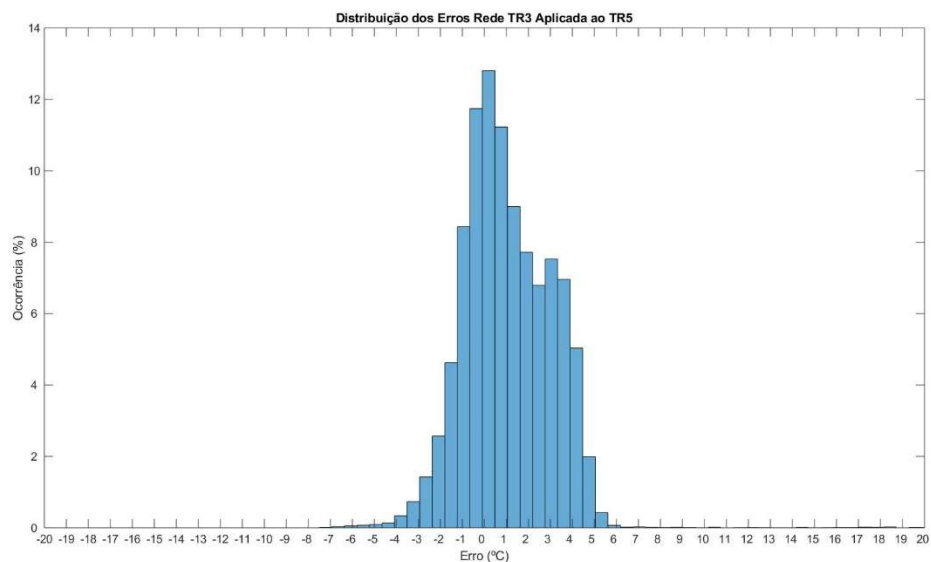
Figura 26 – Comparação das temperaturas mais parâmetros de entrada.



Fonte: Autor.

O histograma que ilustra a distribuição dos erros para o período de um ano do caso avaliado pode ser visto na Figura 27. Aproximadamente 65 % das ocorrências de erro situam-se, em valores absolutos, abaixo dos 2 °C. Pode-se notar uma concentração de cerca de 25% dos erros em torno de 3 e 4 °C, estas ocorrências representam o período de julho a dezembro em que o TR5 opera com potência reduzida, enquanto o TR3 não apresenta este comportamento por longos períodos no conjunto de dados que constituíram a etapa de treinamento, diminuindo a precisão ao determinar a temperatura. Estas situações serão comentadas com maiores detalhes ao final do presente capítulo.

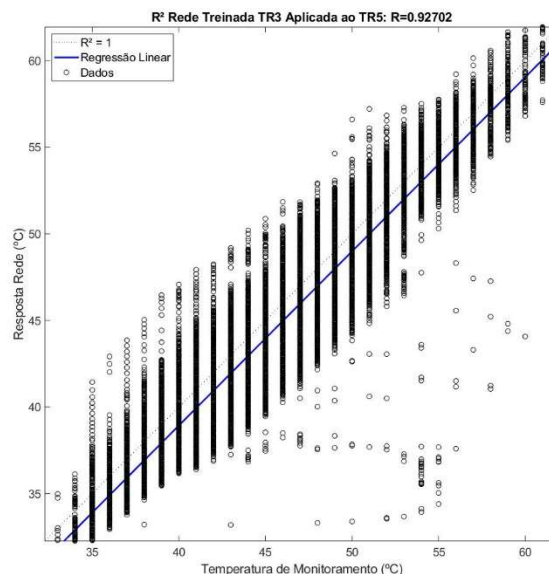
Figura 27 – Histograma de erros.



Fonte: Autor.

A Figura 28 apresenta a correlação entre os dados de monitoramento e a resposta da rede neural. Pode-se visualizar que apenas uma pequena parcela dos dados encontra-se distantes dos valores pretendidos, resultando em um coeficiente de determinação de 0,92 e comprovando a capacidade preditiva da rede neural desenvolvida para esta aplicação.

Figura 28 – Correlação temperatura entre resposta RNA e dados de monitoramento.

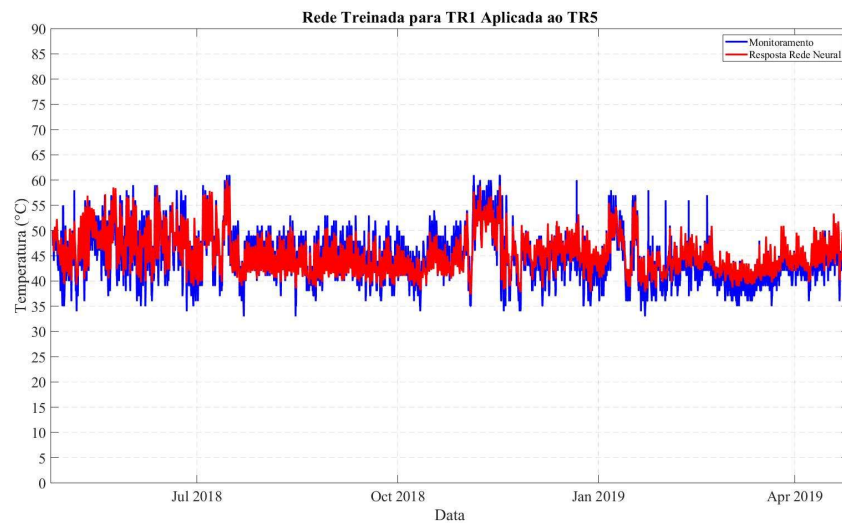


Fonte: Autor.

Apesar de o período com carga reduzida e os picos de temperatura apresentados nos dados de monitoramento elevarem o DMA deste caso, o desempenho avaliado foi muito satisfatório para a grande maioria das condições de operação. Os dois equipamentos comparados apresentam as mesmas características construtivas e tipo de refrigeração, portanto é concebível que a resposta da rede neural para condições que foram abordadas no conjunto de dados de treinamento apresente erros de ordem muito pequena. Ressalta-se a importância de, para obter a melhor capacidade preditiva possível, escolher um conjunto de dados de treinamento abrangente e que represente bem o ambiente que se deseja aplicar.

O próximo caso a ser avaliado é para a rede neural treinada com dados de monitoramento do transformador TR1 aplicada ao TR5. A visão geral da comparação entre a resposta da rede neural e a temperatura de topo de óleo provinda do sistema de monitoramento pode ser vista na Figura 29. O DMA apresentado foi de 1,67 °C, enquanto o erro máximo absoluto foi de 15,13 °C, sendo considerado em termos de desempenho geral o melhor caso nestes dois aspectos.

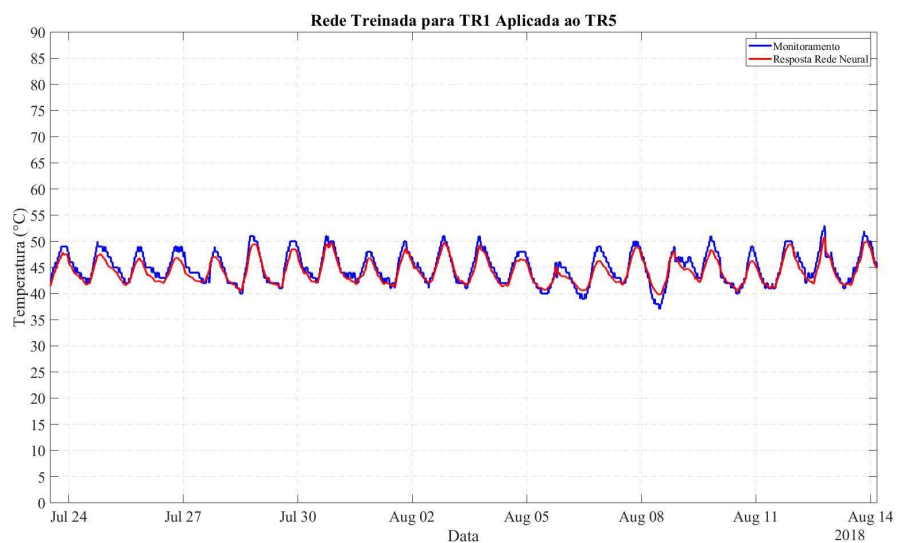
Figura 29 – Comparação das temperaturas resposta RNA vs. monitoramento.



Fonte: Autor.

A comparação entre a rede e o monitoramento para um período de aproximadamente três semanas pode ser visualizada na Figura 30. Este pode também ser considerado muito satisfatório, com erros muito pequenos durante o período e representando bem as variações de temperatura vistas no monitoramento do transformador, entretanto, pode-se notar uma leve diferença presente em grande parte dos pontos não observada para o caso anterior. Esta diferença ocorre por conta da diferença no tipo de resfriamento entre os transformadores para o qual a rede foi treinada e realizado o teste da capacidade preditiva.

Figura 30 – Comparação das temperaturas resposta RNA vs. monitoramento.

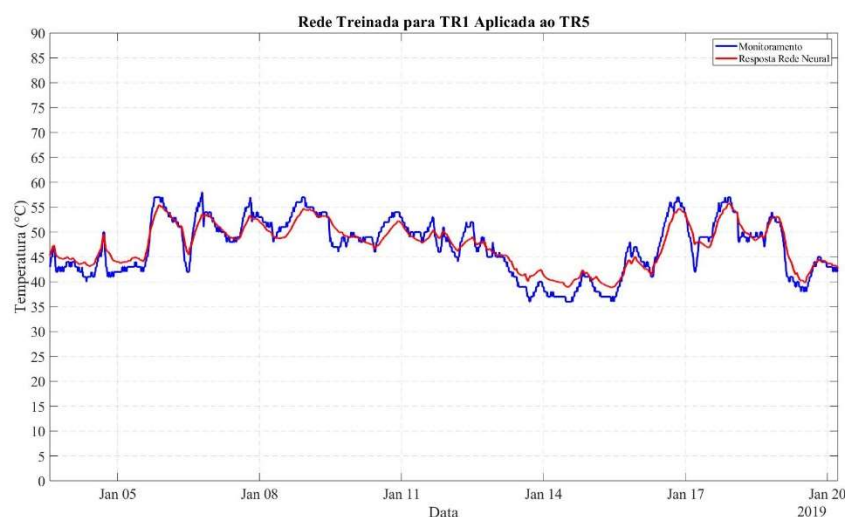


Fonte: Autor.



Para um período com maiores variações da temperatura de topo de óleo a referida diferença é mais visível, conforme Figura 31. Todavia, o desempenho permanece dentro de uma margem de erro satisfatória e a rede representa bem o comportamento do transformador.

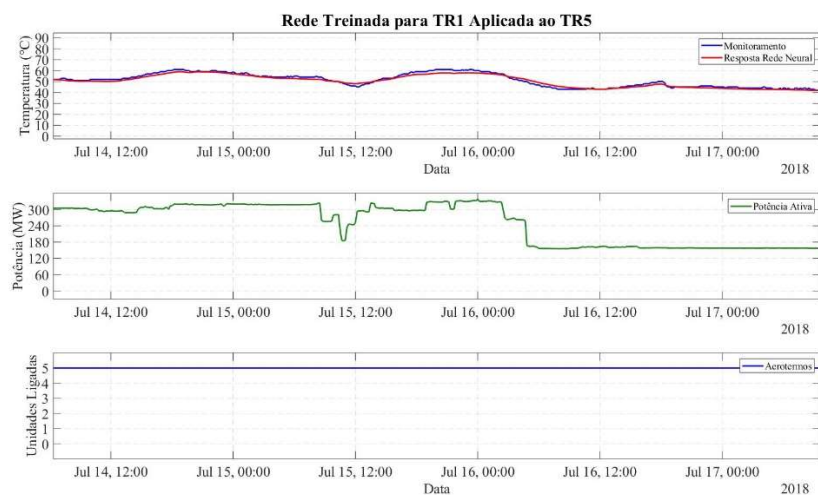
Figura 31 – Comparação das temperaturas resposta RNA vs. monitoramento.



Fonte: Autor.

A Figura 32 ilustra, sob a perspectiva das variações de potência, o comportamento da rede neural em comparação à temperatura de monitoramento. Pode-se observar que o carregamento do transformador vai de seu valor nominal para metade, enquanto a rede neural apresenta uma resposta adequada à ocorrência.

Figura 32 – Comparação das temperaturas mais parâmetros de entrada.

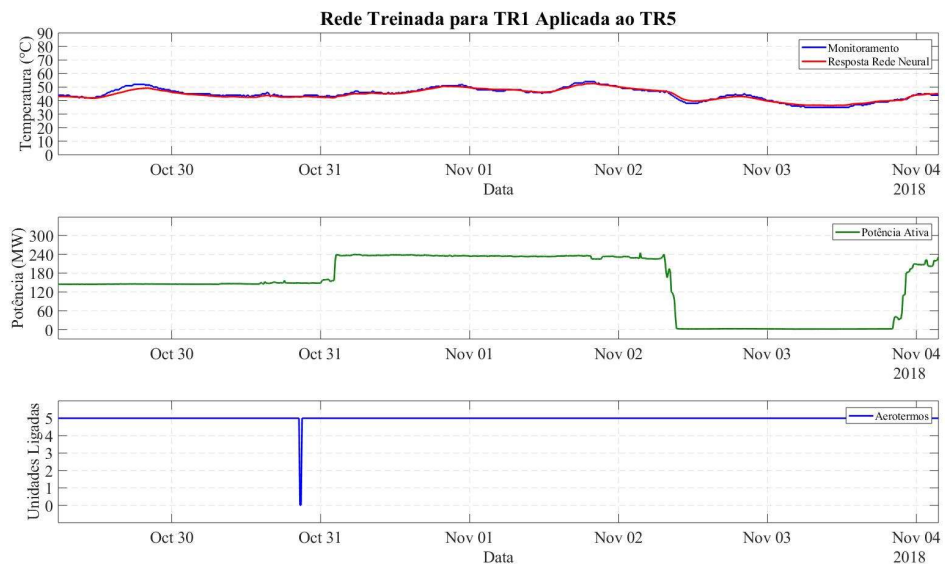


Fonte: Autor.

Na Figura 33 pode-se observar que o carregamento vai de 50% de seu valor nominal para 75% e, após, para zero. Novamente, o desempenho da rede durante as variações representa bem o comportamento medido no transformador.

Neste caso pode-se observar que o número de aerotermos em operação vai de todas unidades ligadas para nenhuma durante um ponto de amostragem sem efeitos na temperatura de topo de óleo. Isso pode ser causado por erro de medição nos *status* dos aerotermos ou um desligamento muito rápido, de apenas alguns poucos minutos e que foi registrado. Essas ocorrências influenciam de forma que durante o treinamento a rede não adquira a característica dos transitórios que causam as elevações de temperatura registradas no monitoramento e, conseqüentemente, representam os erros máximos observados durante os testes.

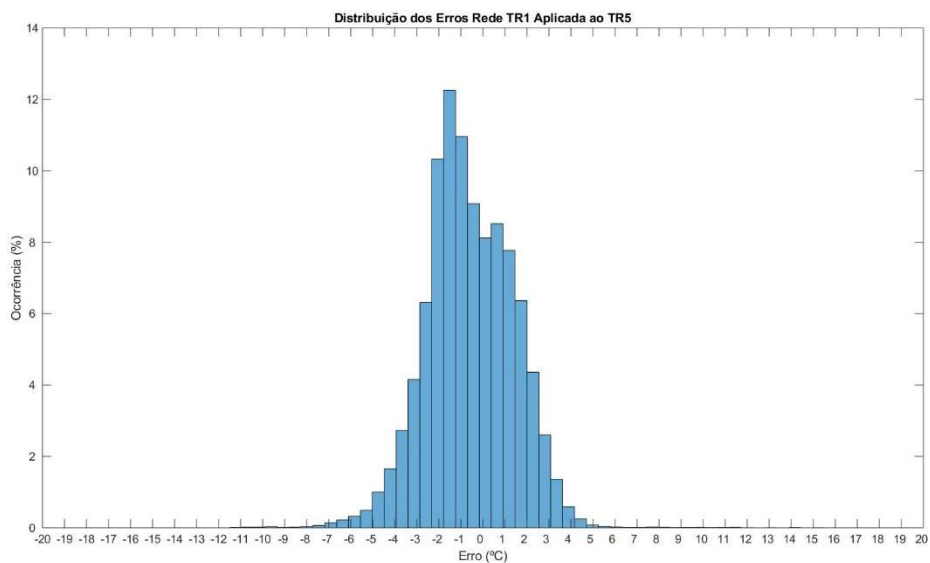
Figura 33 – Comparação das temperaturas mais parâmetros de entrada.



Fonte: Autor.

Em comparação ao caso visto na Figura 23, este apresentou um melhor desempenho geral apesar das pequenas diferenças de temperatura causadas pelo método de resfriamento empregado nos transformadores TR1 e TR5. Isto se deve pela rede apresentar uma resposta melhor ao período com carga reduzida que ocorre na operação do TR5, por conta da abrangência das condições de operação no conjunto de dados do TR1 que constituíram a etapa de treinamento. O histograma da Figura 34 ilustra a distribuição dos erros para o teste realizado para o período de um ano, onde pode-se visualizar que cerca de 72% dos erros observados estão, em módulo, abaixo de 2 °C e cerca de 22% encontram-se entre 2 e 4 °C.

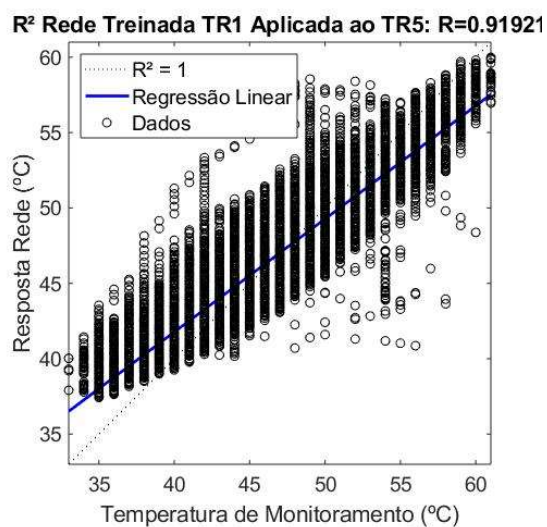
Figura 34 – Histograma de erros.



Fonte: Autor.

A boa correspondência entre a resposta da rede neural e os dados de monitoramento pode ser visualizada na Figura 35. O coeficiente de determinação calculado foi de 0,919, e corrobora com o desempenho satisfatório desta aplicação.

Figura 35 – Correlação temperatura entre resposta RNA e dados de monitoramento.

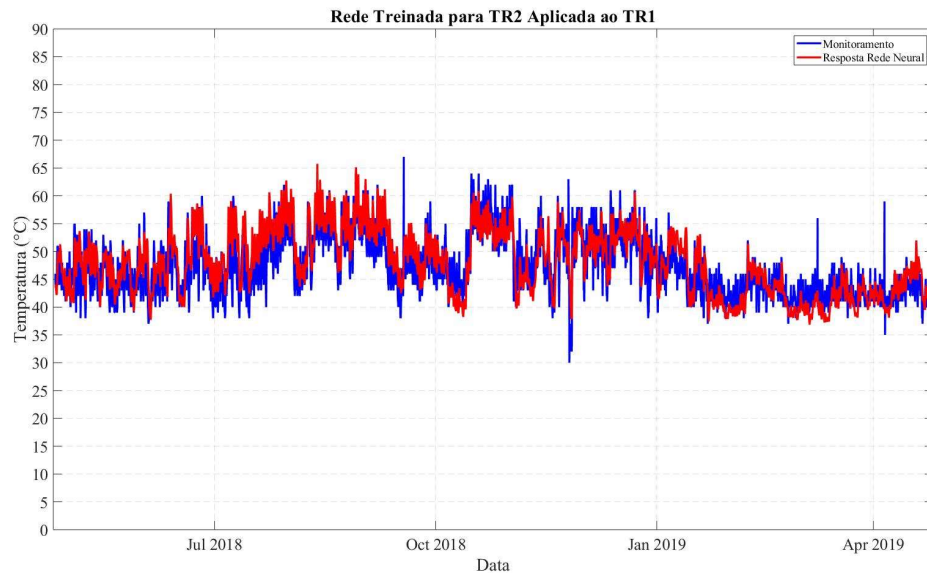


Fonte: Autor.

O próximo caso refere-se à aplicação das entradas de monitoramento de um período de um ano do transformador TR1 em uma rede treinada com o conjunto de dados de

monitoramento de um ano do transformador TR2. O DMA para o período foi de 2,38 °C e o erro máximo observado foi de 18,51 °C.

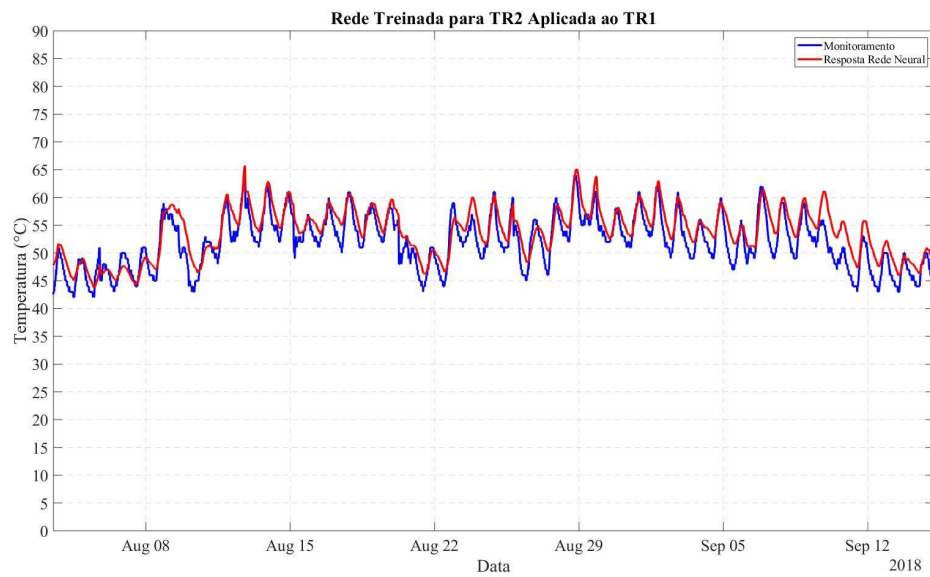
Figura 36 – Comparação das temperaturas resposta RNA vs. monitoramento.



Fonte: Autor.

O desempenho mantém-se em valores satisfatórios, conforme a Figura 37 que representa um período de aproximadamente seis semanas, a maior parcela do período total. Entretanto, apesar de os dois transformadores apresentarem o mesmo tipo de resfriamento, existe uma pequena diferença nos valores de temperatura da rede e monitoramento em quase todos instantes. Isso se deve ao fato de que o conjunto de dados do TR2 utilizados para treinamento não abrange a mesma variedade de condições de operação que os dados do TR1, causando inclusive, em alguns pequenos períodos, um deslocamento semelhante ao que foi discutido quando se tem a operação com carga reduzida e, conseqüentemente, elevando o DMA.

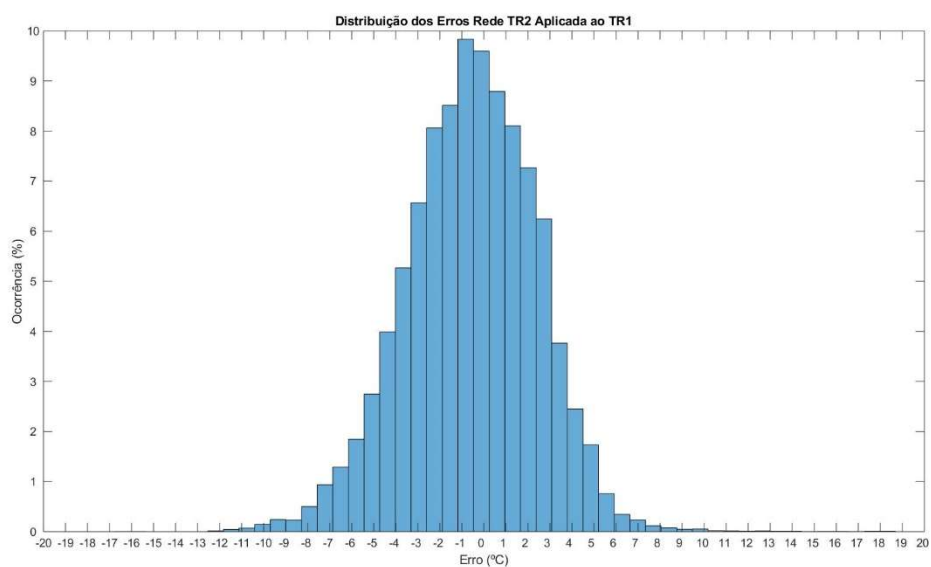
Figura 37 – Comparação das temperaturas resposta RNA vs. monitoramento.



Fonte: Autor.

A distribuição dos erros pode ser observada no histograma da Figura 38, onde 60% das ocorrências de erro encontram-se abaixo de 2 °C, ou seja, abaixo do DMA. Aproximadamente 29.5 % encontram-se entre 2 e 4 °C, parcela um pouco maior do que a observada nos casos anteriores.

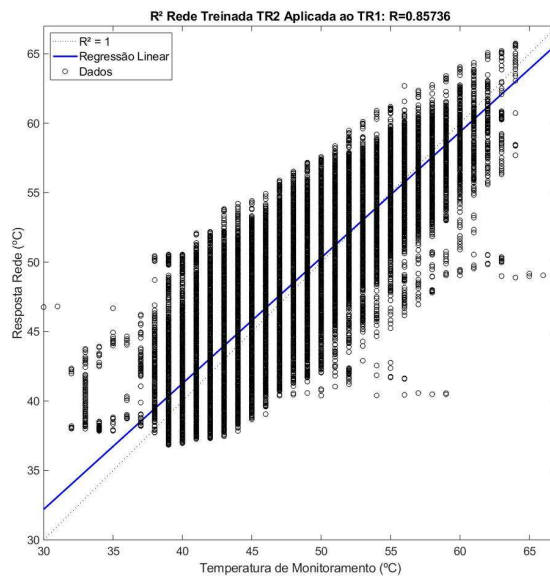
Figura 38 – Histograma de erros.



Fonte: Autor.

A correlação entre a resposta da rede e a temperatura medida no transformador pode ser vista na Figura 39. Pode-se visualizar que em poucas ocasiões a resposta encontra-se muito distante do valor pretendido, resultando em um coeficiente de determinação de 0,85, corroborando com o fato de que a rede neural representa bem o comportamento térmico do transformador para um mesmo conjunto de entradas.

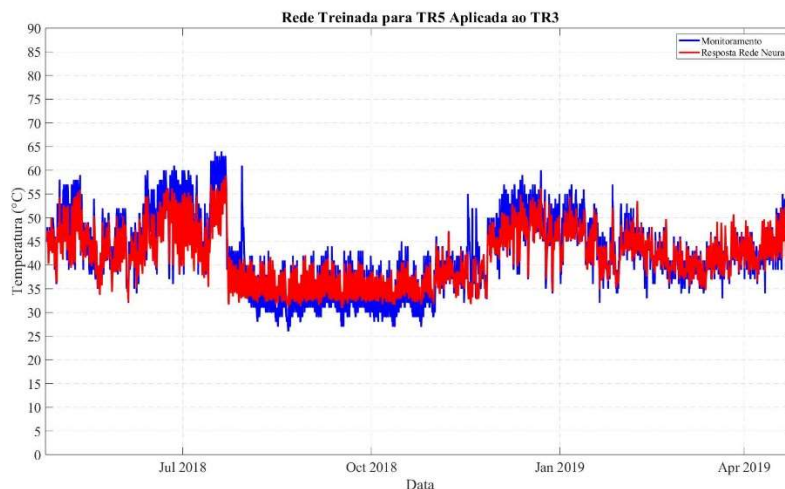
Figura 39 – Correlação temperatura entre resposta RNA e dados de monitoramento.



Fonte: Autor.

Como próximo caso temos a situação inversa a apresentada no primeiro. A rede neural teve seu treinamento realizado com dados de monitoramento do transformador TR5 e aplicada para prever a temperatura de topo de óleo do transformador TR3. A Figura 40 representa o período total de um ano no qual foi realizado este teste. O DMA para o período foi de 2,25 °C e o erro máximo absoluto de 21,45 °C.

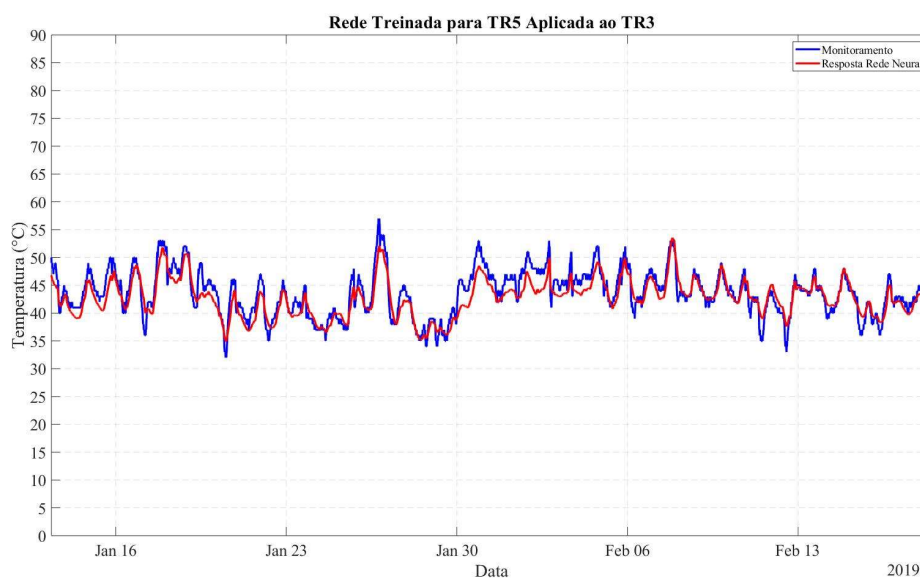
Figura 40 – Comparação das temperaturas resposta RNA vs. monitoramento.



Fonte: Autor.

A Figura 41 representa a comparação entre a resposta da rede e a temperatura de topo de óleo medida no transformador para um período pouco maior que um mês. Pode-se notar que mesmo para grandes variações de temperatura a resposta é muito satisfatória e representa bem o comportamento do transformador real. O DMA é cerca de 0,55 °C maior do que a aplicação inversa, novamente, por conta da maior abrangência das condições de operação presentes nos dados de monitoramento do transformador TR3.

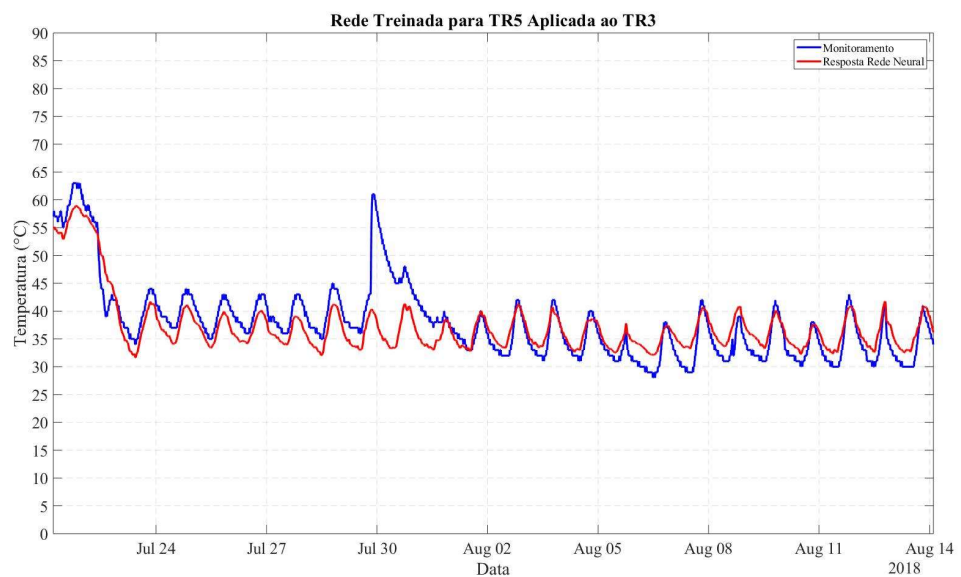
Figura 41 – Comparação das temperaturas resposta RNA vs. monitoramento.



Fonte: Autor.

Pode-se visualizar na Figura 42 o comportamento da rede durante o desligamento do TR3, ressaltando que o transformador TR5 apresenta apenas a condição de operação com carga reduzida em seus dados de monitoramento durante o período de sazonalidade visto na usina. A pequena diferença de temperatura vista ao longo deste período, somada ao pico de temperatura que não tem seu comportamento reproduzido pela rede e causa o erro máximo para o caso, colaboram para a elevação do DMA.

Figura 42 – Comparação das temperaturas resposta RNA vs. monitoramento.

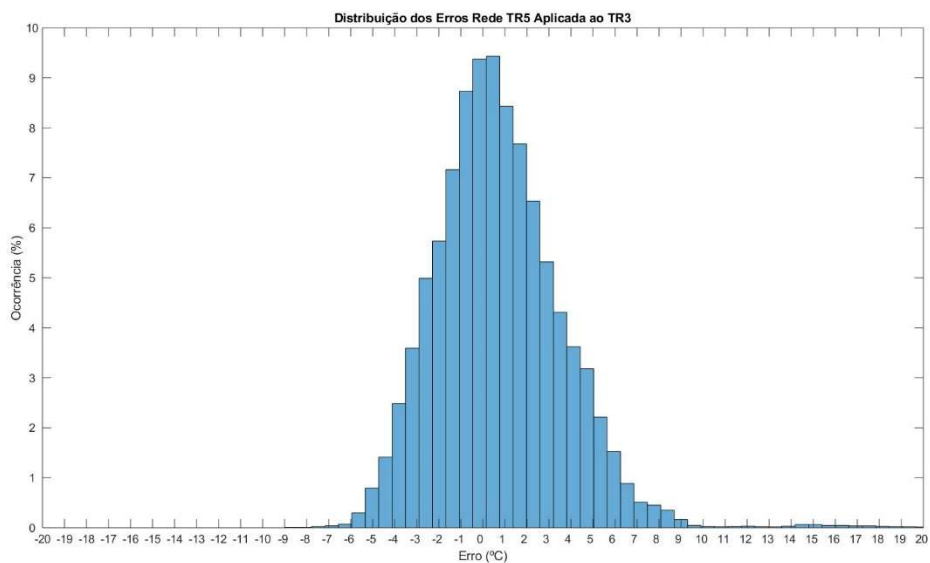


Fonte: Autor.

O histograma de erros da Figura 43 representa a distribuição dos erros em porcentagem para o período de um ano. Cerca de 55% dos erros observados encontram-se abaixo de 2 °C e 26,5 % entre 2 e 4 °C.



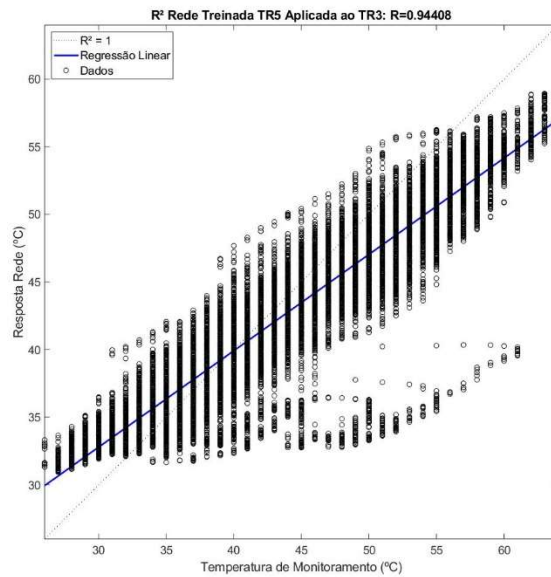
Figura 43 – Histograma de erros.



Fonte: Autor

A correlação entre a resposta da rede e a temperatura de monitoramento está exposta na Figura 44. O coeficiente de determinação calculado é de 0,94, poucos pontos podem ser ditos como distantes do valor pretendido, demonstrando um bom desempenho. Pode-se observar que para os menores valores da temperatura de monitoramento a rede apresenta uma resposta um pouco maior, enquanto para os maiores valores de temperatura de monitoramento a resposta da rede é menor que o valor pretendido, isto ocorre pela já discutida abrangência dos dados de treinamento em relação ao ambiente em que se vai aplicar. Ou seja, durante a etapa de treinamento estas condições de operação não foram vistas, portanto, a rede não está condicionada a representá-las com precisão.

Figura 44 – Correlação temperatura entre resposta RNA e dados de monitoramento.

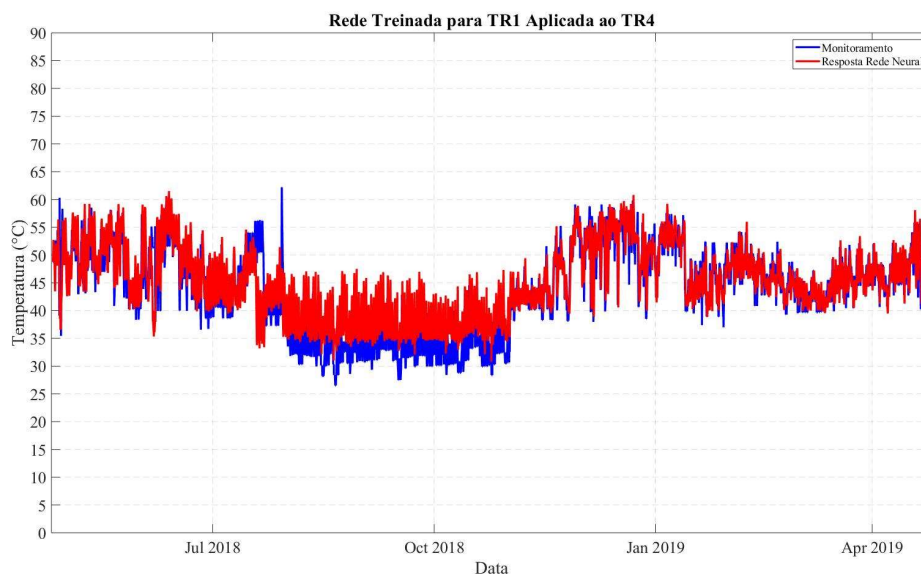


Fonte: Autor.

O próximo caso, visto em seu período total de um ano na Figura 45, é a aplicação dos dados de monitoramento do transformador TR4 em uma rede neural treinada com dados do TR1. Os casos que incluem o TR4 são os mais delicados, pois seu monitoramento de temperatura de topo de óleo é proveniente de um sistema de medição distinto dos outros, este que registra apenas variações de 2 °C. Portanto, isto ocasiona algumas diferenças no comportamento dinâmico.

O DMA calculado para o período foi de 3,2 °C e o erro máximo absoluto de 20 °C. Pode-se visualizar que durante o período em que o TR4 permanece desligado a rede apresenta uma diferença maior de temperatura, causado por este comportamento não ter sido apresentado a rede durante a etapa de treinamento, ocasionando uma elevação do DMA.

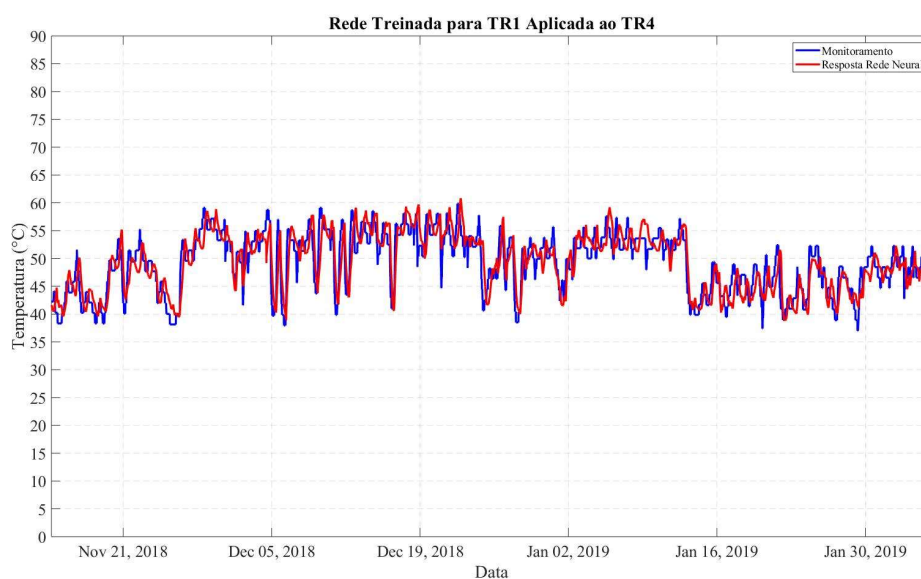
Figura 45 – Comparação das temperaturas resposta RNA vs. monitoramento.



Fonte: Autor.

Pode-se visualizar na Figura 46 que a rede representa bem as variações de temperatura, com algumas diferenças na dinâmica conforme já explicado. Em conjunto com o período em que o transformador permanece sem carga e os picos de temperatura, esta diferença ocasiona um DMA maior do que os casos anteriores.

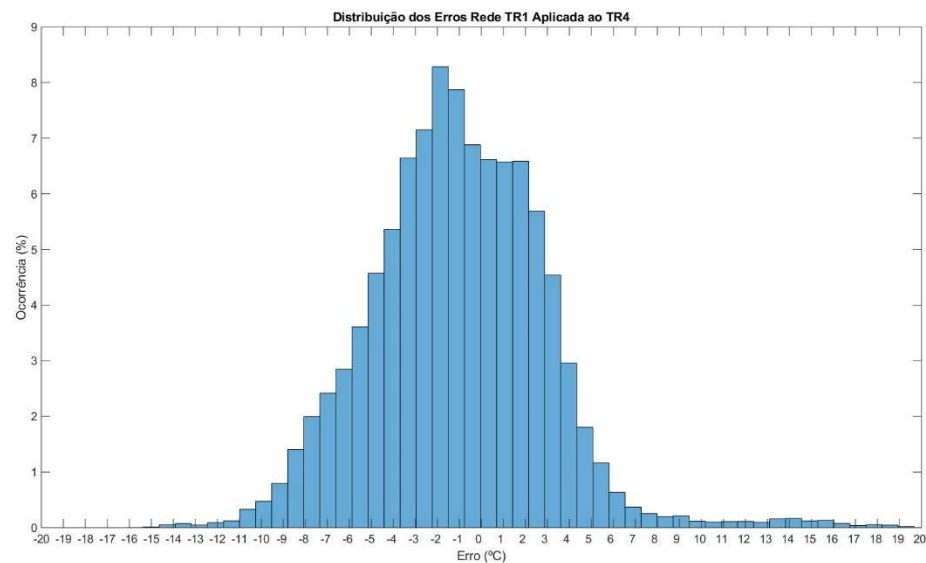
Figura 46 – Comparação das temperaturas resposta RNA vs. monitoramento.



Fonte: Autor.

A distribuição dos erros pode ser vista no histograma da Figura 47. Este caso apresenta uma concentração maior de erros acima dos 4 °C pelos motivos previamente demonstrados. Cerca de 42,5% dos erros encontram-se abaixo de, em módulo, 2 °C, enquanto aproximadamente 31,5 % encontram-se entre 2 e 4 °C, concentração maior nesta faixa do que para os casos anteriores.

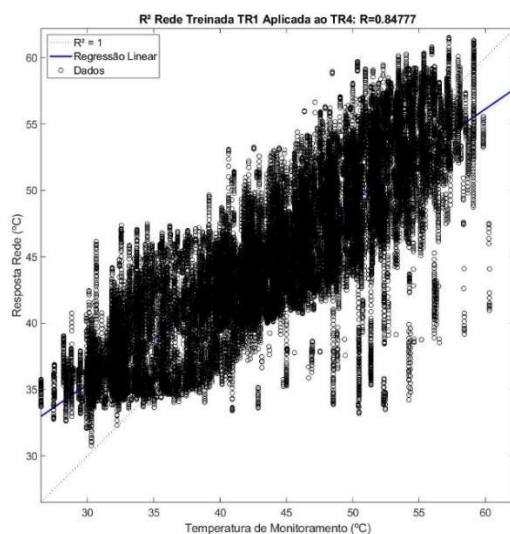
Figura 47 – Histograma de erros.



Fonte: Autor.

A correlação entre a temperatura de topo de óleo medida no transformador e a resposta da rede pode ser visualizada na Figura 48, resultando em um coeficiente de determinação de 0,84. Ela possui um aspecto diferente por conta do monitoramento de temperatura do TR4 considerar duas casas decimais. Apesar do desempenho estar abaixo dos outros casos, a rede neural ainda representou bem o comportamento térmico do transformador real, com poucos pontos distantes do pretendido.

Figura 48 – Correlação temperatura entre resposta RNA e dados de monitoramento.



Fonte: Autor.

A Tabela 1 apresenta um resumo do DMA observado para todas as combinações entre os transformadores sob estudo. Nos casos que envolvem o mesmo transformador o treinamento foi realizado com os dados do período de um ano e o teste foi realizado aplicando as entradas de monitoramento do outro período. Pode ser observado que as redes treinadas com dados dos transformadores TR1 e TR3 apresentam um melhor desempenho geral, devido a maior abrangência das condições de operação em relação aos outros, apesar de em alguns casos, conforme já discutido, o conjunto de entradas correspondente ao transformador para o qual está sendo realizado o teste da capacidade preditiva apresentar operação com potência reduzida ou sem carga durante longos períodos, não ser representado com a precisão do restante do conjunto e elevarem o DMA. Os piores casos avaliados tem um DMA mais elevado também por conta desta falta de abrangência no conjunto de dados de treinamento.

Tabela 1 – Resumo dos DMA para todos os testes realizados.

DMA		Entradas Aplicadas				
		TR1	TR2	TR3	TR4	TR5
Rede Treinada	TR1	1,948	2,335	3,183	3,199	1,672
	TR2	2,379	1,736	3,835	3,538	2,568
	TR3	2,874	2,438	2,005	3,244	1,717
	TR4	2,858	4,020	4,942	2,993	3,961
	TR5	3,310	2,708	2,252	3,803	1,992

Fonte: Autor.

Os erros máximos absolutos observados para todos os testes realizados podem ser vistos na Tabela 2. Erros desta ordem, como pôde ser visto nos histogramas apresentados, ocorrem com uma frequência muito baixa dentro do período de um ano dos testes e todos representam elevações súbitas da temperatura do monitoramento por períodos muito breves, na maioria das vezes acompanhados pelo desligamento de aerotermos. O caso mais crítico é da aplicação dos dados de um período do TR4 em uma rede treinada com dados do outro período do mesmo transformador, pois foi o único teste em que se utilizou o período referente a Novembro de 2016 até Novembro de 2017 do TR4. Este período apresenta uma única ocorrência da elevação de temperatura de topo de óleo para além dos 100 °C que será apresentada mais adiante.

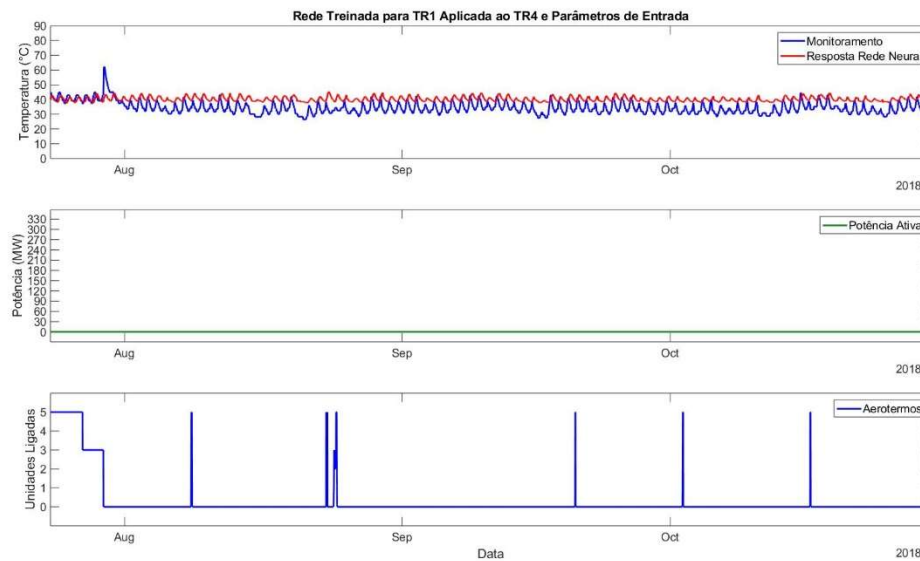
Tabela 2 – Erro máximo absoluto observado nos testes de capacidade de predição.

Erro Máximo		Entradas Aplicadas				
		TR1	TR2	TR3	TR4	TR5
Rede Treinada	TR1	19,06	16,60	19,80	20,00	15,13
	TR2	18,51	19,76	18,43	21,07	15,52
	TR3	23,24	21,30	22,00	25,75	20,60
	TR4	20,25	19,34	19,95	52,69	17,16
	TR5	21,21	22,42	21,45	23,89	19,51

Fonte: Autor.

Por mais que a rede reconheça a operação na potência reduzida ou desligada, durante os longos períodos em que permanece nestas condições há variações no status dos aerotermos, estas variações podem não compor situações vistas no conjunto utilizado para treinamento. A Figura 49 apresenta o caso da rede neural treinada com dados do transformador TR1 aplicada ao transformador TR4. Quando há 5 aerotermos a rede apresenta boa resposta, pois os exemplos que o TR1 tem de potência zerada são períodos curtos em que não há desligamento dos ATs, porém quando todos são desligados ocorre uma diferença de temperatura entre a rede e monitoramento que ocasiona uma elevação do DMA para o período total. Isso ocorre também em outros casos, como por exemplo para a rede treinada com dados do TR3 aplicada ao TR5.

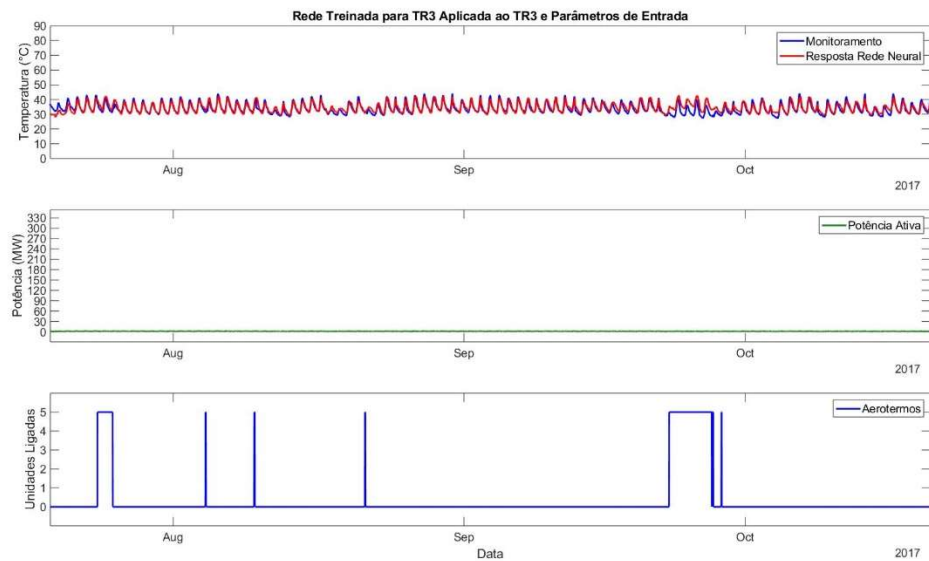
Figura 49 – RNA TR1 aplicada a dados do TR4 mais parâmetros de entrada.



Fonte: Autor.

A Figura 50 apresenta uma rede treinada para um período do TR3 e aplicada ao outro que não participou da etapa de treinamento. Aqui pode-se ver que a resposta da RNA é muito satisfatória durante o período. Entretanto, ao ocorrer variações no monitoramento dos aerotermos a RNA responde de maneira diferente aos dados de monitoramento de temperatura. Isto ocorre por conta de o período em que o transformador permanece desligado utilizado para treinamento é em sua grande maioria sem operação dos aerotermos, sendo esta a característica que a RNA adquire para determinar a temperatura de topo de óleo. Portanto, a rede apenas responde com a precisão desejada para esta situação.

Figura 50 – RNA TR3 aplicada a dados TR3 mais parâmetros de entrada.



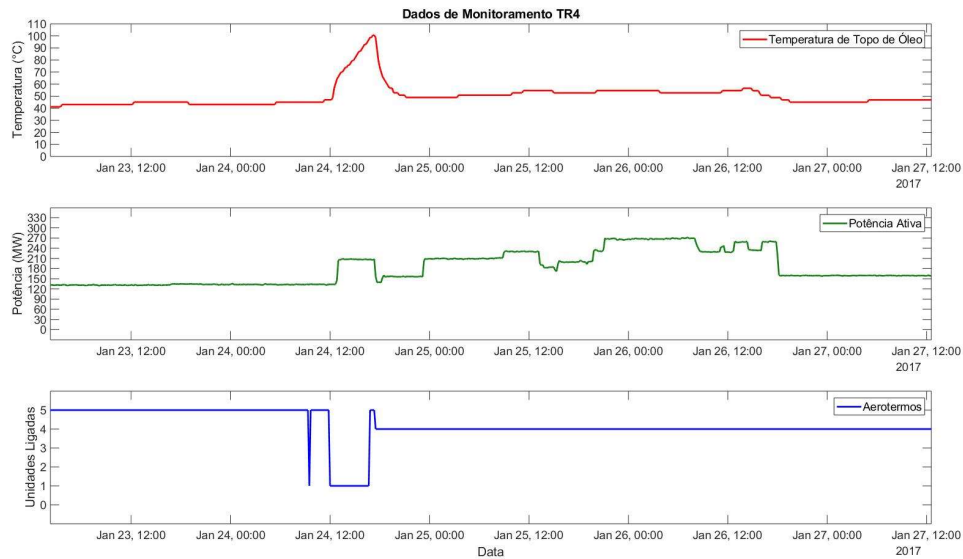
Fonte: Autor.

Em Wei et. al. (2017) e Vilaithong et. al. (2007) as redes neurais desenvolvidas para predição de temperatura em transformadores também apresentam seus pontos críticos quando a potência vai a zero. Portanto, vale ressaltar novamente que a abrangência do conjunto de dados utilizados para o treinamento deve representar o máximo de condições de operações possíveis que o ambiente que se quer predizer a temperatura pode apresentar.

A Figura 51 apresenta a situação em que ocorreu o maior dos erros máximos apresentados na Tabela 2. Por algumas horas apenas um conjunto de aeroterms permaneceu em operação com o transformador sob carga, resultando em uma elevação da temperatura a níveis muito altos. Ocorrências como essa não são muito frequentes nos conjuntos de dados abordados neste trabalho, e nestas proporções, elevando a tais níveis de temperatura, é um acontecimento isolado e por isso não foi replicado pela rede.



Figura 51 – Dados de monitoramento TR4.

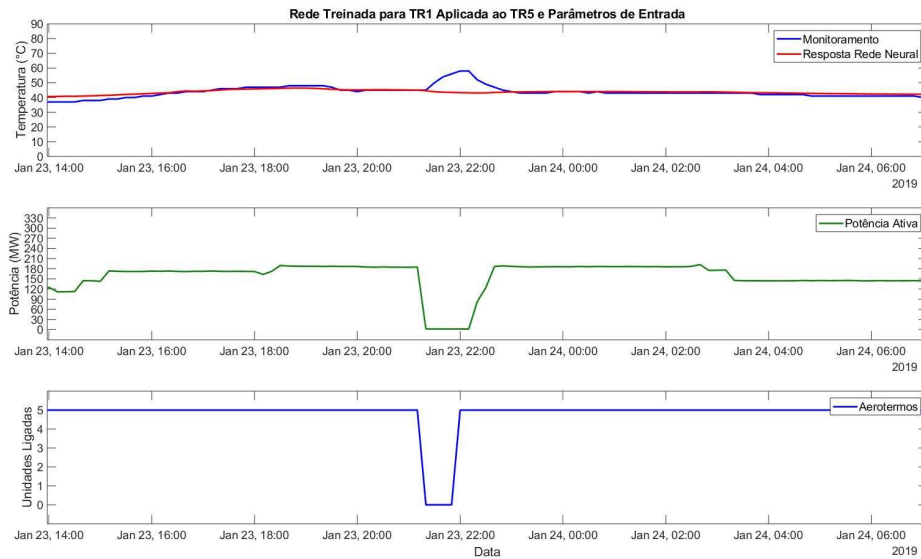


Fonte: Autor.

Os casos que constituem todos os outros erros máximos observados são causados por ocorrências como as da Figura 52. Estes eventos são causados quando ocorre o desligamento do transformador e de todos conjuntos de aeroterms por um curto período de tempo e a temperatura de topo de óleo sofre uma elevação. Este comportamento pode indicar que ao desligar os conjuntos de aeroterms a circulação do óleo por convecção natural causou um acúmulo de óleo de temperaturas mais elevadas na parte superior do tanque, refletindo na medição.

Porém, esta característica não é reproduzida pelas redes neurais por constituir uma parcela muito pequena do conjunto de dados utilizados durante os treinamentos. Além de que existem algumas ocorrências de desligamentos de todos conjuntos de aeroterms sem impactos visíveis na temperatura, colaborando com a não aquisição desta característica.

Figura 52 – RNA TR1 aplicada em dados do TR5 mais parâmetros de entrada.



Fonte: Autor.

### 3.6 CONSIDERAÇÕES DO CAPÍTULO

O presente capítulo apresentou como foi realizado o desenvolvimento das RNAs utilizadas para predição de temperatura e uma série de testes da capacidade de predição. Os resultados dos testes foram apresentados, abordando todas as variedades de comportamentos observados e os melhores e piores desempenhos tiveram suas causas discorridas.

Desta forma, conforme os resultados apresentados, pode-se ressaltar o excelente desempenho das RNAs desenvolvidas para predição de temperatura em transformadores de potência. Todo o estudo foi desenvolvido com dados reais de monitoramento, e por conta das dificuldades de operação e manutenção que existem em uma usina hidrelétrica, é compreensível a presença de algumas anomalias que causam alguns períodos com uma queda de desempenho na precisão da resposta da rede, como no caso do *status* dos conjuntos de aerotermos. Os testes foram realizados com um conjunto de dados desconhecido à rede, que não participou da etapa de treinamento, com o intuito de avaliar a capacidade preditiva das RNAs desenvolvidas. Os resultados demonstraram a capacidade das redes neurais desenvolvidas através desta metodologia de representar as mais variadas condições e operação dos transformadores desde que inclusas no conjunto de dados utilizados na etapa de treinamento.

#### 4. ANÁLISES DE PERDA DE VIDA ÚTIL

A predição de temperatura em transformadores de potência não tem sua importância atribuída apenas para determinar se o transformador irá operar dentro dos limites de temperatura para as condições de operação futuras, mas também para observar a degradação do papel isolante causada pela temperatura e sua respectiva perda de vida útil.

Este capítulo apresenta uma análise da capacidade preditiva das RNAs desenvolvidas sob a perspectiva da perda de vida útil, realizando uma comparação entre o obtido a partir dos dados de monitoramento de temperatura e da resposta da rede neural. Após, é realizado o cálculo da perda de vida útil a partir da resposta da RNA para seis cenários fictícios com condições de operação distintas que levam em consideração a sazonalidade vista na usina e se prolongam até uma projeção ao ano 2050.

Não existe um único método para determinação da vida útil remanescente em transformadores de potência, e os critérios referentes ao fim da vida útil também podem ser distintos. Porém, por ser um critério útil aos responsáveis pelos transformadores, é apropriado focar no envelhecimento e condições da isolação sólida (ABNT NBR 5356-7, 2017).

##### 4.1 TAXA DE ENVELHECIMENTO RELATIVO

Embora o envelhecimento seja influenciado pela temperatura, umidade, conteúdo de oxigênios e ácidos, o cálculo proposto pelas normas ABNT NBR 5356-7 e IEC 60076-7 leva em consideração apenas a temperatura ao determinar a degradação do papel isolante (ABNT, 2017; IEC, 2018). Contudo, a distribuição de temperaturas em um transformador não é uniforme, então utiliza-se o ponto onde ocorre a maior degradação da isolação, a temperatura de ponto mais quente, ou temperatura de *hot-spot* (ABNT, 2017; IEC, 2018).

Para papel termoestabilizado, definido como papel à base de celulose modificado quimicamente para reduzir a taxa na qual se decompõe, o cálculo da taxa de envelhecimento é realizado conforme equação (10) (ABNT NBR 5356-7, 2017). Em que  $\theta_h$  é a temperatura de ponto mais quente expressa em °C. Ou seja, o envelhecimento relativo é igual a 1 quando a temperatura de ponto mais quente é de 110 °C absoluto, variando de forma exponencial e crescente.

$$V = e^{\left(\frac{15000}{110 + 273} - \frac{15000}{\theta_h + 273}\right)} \quad (10)$$

## 4.2 CÁLCULO DA PERDA DE VIDA ÚTIL

A perda de vida  $L$  para um determinado período de tempo pode ser calculada pelas equações (11) ou (12) (ABNT NBR 5356-7, 2017).

$$L = \int_{t_1}^{t_2} V \quad (11)$$

$$L = \sum_{n=1}^N V_n t_n \quad (12)$$

Em que:

$V_n$  → Taxa de envelhecimento relativo durante o  $n$ ésimo intervalo;

$t_n$  →  $n$ ésimo intervalo de tempo;

$n$  → Número de cada intervalo;

$N$  → Número total de intervalos durante o período considerado.

## 4.3 VIDA DA ISOLAÇÃO

Para cenários com baixas concentrações de oxigênio e umidade a norma propõe os limites de vida útil presentes na Tabela 3. Este tempo de vida normal ocorre para uma temperatura de referência de 110 °C. Ou seja, para o utilizado neste trabalho que é o de 150 000 horas, se o transformador hipoteticamente operar durante todo este tempo à temperatura de referência desde o início de sua vida útil, o papel isolante deve atingir um grau de polimerização de 200 após 150 000 horas ou 17,12 anos. Por grau de polimerização entende-se como a estimativa do comprimento médio das moléculas de celulose. Sua diminuição faz com que o papel perca suas propriedades mecânicas e se torne suscetível à ruptura quando submetido a esforços eletromecânicos causados, por exemplo, por curto circuito.

Tabela 3 – Tempo de vida normal.

Base	Vida Normal da	
	Isolação	
	Horas	Anos
50% da resistência à tração inicial	60000	7,42
25% da resistência à tração inicial	135000	15,41
Grau de polimerização remanescente de 200	150000	17,12
Interpretação dos dados de ensaio relativos à vida operacional do transformador de distribuição	180000	20,55

Fonte: (ABNT NBR 5356-7, 2017).

#### 4.4 DETERMINAÇÃO DA TEMPERATURA DE PONTO MAIS QUENTE

Existem diversos modelos para a determinação da temperatura de ponto mais quente em transformadores de potência, entretanto todos apresentam suas limitações. As próprias normas ABNT NBR 5356-7 (2017) e IEC 60076-7 (2018) apresentam mais de uma opção de cálculo para o ponto mais quente.

No presente trabalho, os dados de monitoramento de temperatura resultam da medição direta da temperatura de topo de óleo do tanque. Para uma melhor precisão a temperatura é referida ao óleo adjacente ao enrolamento que apresenta a temperatura de ponto mais quente, o acréscimo da temperatura do óleo neste ponto pode ser até 15 °C mais elevada em relação ao topo de óleo do tanque (ABNT NBR 5356-7, 2017).

Contudo, a temperatura de ponto mais quente, e que deteriora o papel isolante, ocorre nos enrolamentos, portanto é necessário saber o gradiente de temperatura entre óleo e enrolamentos. Este gradiente é obtido através dos ensaios de elevação de temperatura, e é determinado pela diferença da temperatura média dos enrolamentos, calculada através da medição das resistências à quente, e temperatura média do óleo, que é a média das temperaturas de base e topo de óleo.

Os valores do gradiente enrolamento-óleo são fornecidos pelos relatórios de ensaios nos manuais dos equipamentos, e são distintos de acordo com o tipo de resfriamento empregado no transformador e também para os diferentes enrolamentos deste. Para os transformadores com resfriamento ODAF, TR1 e TR2, foi utilizado o gradiente enrolamento-óleo para o enrolamento de alta tensão, enquanto para os outros transformadores, de resfriamento OFAF, foi utilizado o gradiente do enrolamento de baixa tensão que apresentou o maior valor. Os valores utilizados podem ser visualizados na Tabela 4.

Tabela 4 – Valores utilizados para a obtenção da temperatura de *hot-spot*.

<b>Transformador</b>	<b>Gradiente Enrolamento-Óleo</b>	<b>Topo de óleo do tanque para óleo adjacente</b>	<b>Elevação <i>hot-spot</i> para topo de óleo do tanque</b>
TR1	19,7 °C	+ 15 °C	34,7 °C
TR2	19,7 °C	+ 15 °C	34,7 °C
TR3	22,72 °C	+ 15 °C	37,72 °C
TR4	22,72 °C	+ 15 °C	37,72 °C
TR5	22,72 °C	+ 15 °C	37,72 °C

Fonte: Autor.

Os dados de temperatura de topo de óleo utilizados para a obtenção da temperatura de ponto mais quente e posterior cálculo da vida útil são provenientes de outro sistema de monitoramento que possui um período maior do que o utilizado no Capítulo 3 para os treinamentos e testes da capacidade de predição. Ao utilizar um período de monitoramento maior tem-se uma estimativa mais precisa da degradação do transformador. Porém, este não fornece todas as entradas necessárias para obter a resposta da RNA durante todo o período que é de aproximadamente quatro anos, variando para cada transformador.

Ambos os sistemas apresentam a temperatura de topo de óleo com medição direta, apenas com amostragens diferentes. Como se pôde obter a resposta da rede neural apenas para os dois períodos de um ano cada utilizados no Capítulo 3, por conta da disponibilidade dos parâmetros utilizados como entrada da RNA, os dados de temperatura utilizados para a comparação da perda de vida útil foram divididos em cinco períodos vistos na Tabela 5.

Tabela 5 – Composição dos conjuntos de monitoramento e resposta RNA.

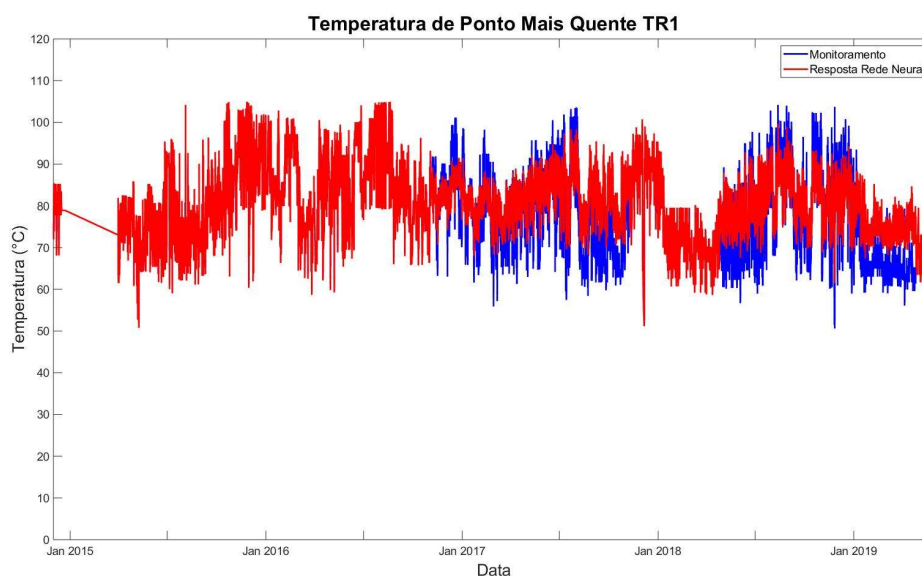
<b>Período</b>	<b>Conjunto Monitoramento</b>	<b>Conjunto Resposta RNA</b>
Início Monitoramento até 06/11/2016	Dados de Monitoramento	Dados de Monitoramento
06/11/2016 até 06/11/2017	Dados de Monitoramento	Resposta Rede Neural
06/11/2017 até 26/04/2018	Dados de Monitoramento	Dados de Monitoramento
26/04/2018 até 26/04/2019	Dados de Monitoramento	Resposta Rede Neural
26/04/2019 até o final do monitoramento	Dados de Monitoramento	Dados de Monitoramento

Fonte: Autor.

Portanto, conforme pode também ser visualizado na Figura 53, que ilustra a temperatura de ponto mais quente em termos absolutos, nas comparações da perda de vida útil para o período sob análise, a perda de vida útil referente à resposta da rede tem apenas dois períodos de um

ano cada em que se difere da temperatura provinda do monitoramento. Por apresentar o melhor desempenho nos testes de capacidade de predição, foi utilizada uma RNA treinada com dados do transformador TR3 para todos os transformadores, exceto ao calcular a perda de vida útil do TR3, que então utilizou-se uma rede treinada com dados do TR5.

Figura 53 – Temperatura de *hot-spot* resposta RNA e monitoramento TR1.



Fonte: Autor.

Como alguns dos transformadores entraram em operação antes do início do período de monitoramento, foi considerada uma perda de vida útil inicial. Esta perda de vida útil foi calculada estimando que durante o tempo sem monitoramento o transformador apresentou um perfil de temperatura de ponto mais quente semelhante a um período de igual tamanho dentro dos dados de monitoramento.

Tabela 6 – Datas de energização e monitoramento dos transformadores

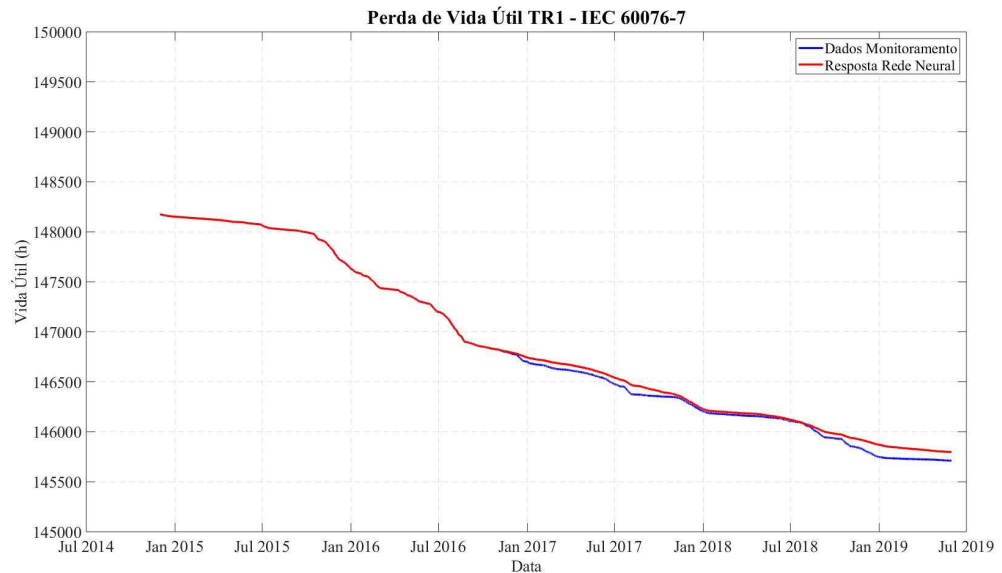
Transformador	Data Energização	Data Início Monitoramento	Tempo sem monitoramento	Perda de Vida Útil Inicial (horas)
TR1	16/12/2011	30/11/2014	2 anos e 11 meses	1826
TR2	28/10/2015	28/10/2015	-	0
TR3	04/04/2013	04/05/2015	2 anos e 1 mês	406
TR4	23/05/2014	24/09/2015	1 ano e 4 meses	569
TR5	24/09/2015	24/09/2015	-	0

Fonte: Autor.

#### 4.5 COMPARAÇÃO PERDA DE VIDA ÚTIL ENTRE REDE NEURAL E MONITORAMENTO

Todos os gráficos que apresentam a perda de vida útil para o período estão apresentados nos limites de 145 000 a 150 000 horas para uma melhor visualização. A Figura 54 apresenta a evolução da perda de vida útil do transformador TR1 para um período de quatro anos e meio, já levando em consideração um valor inicial que representa o período desde a energização do transformador até o início do monitoramento.

Figura 54 – Comparação perda de vida útil TR1.



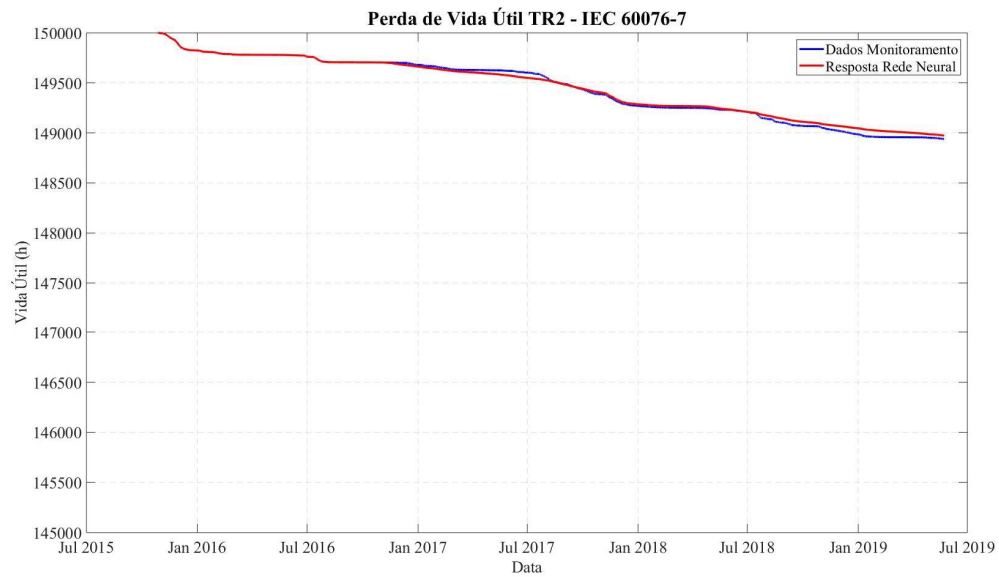
Fonte: Autor.

Pode-se perceber que a perda de vida útil calculada a partir da resposta da rede neural empregada apresenta valores muito próximos ao calculado a partir dos dados de monitoramento. A perda de vida útil para a rede durante o período foi de 2377 horas, enquanto para os dados de monitoramento foi de 2464 horas, resultando em uma diferença de apenas 87 horas ao final do período para a comparação realizada.

Para a comparação realizada ao transformador TR2 a perda de vida útil obtida através da resposta da rede neural foi de 1028 horas para o período, enquanto para os dados de monitoramento foi de 1063 horas, resultando em uma diferença de 35 horas. Conforme Figura 55, pode-se visualizar que ambos os casos apresentam uma evolução muito semelhante.



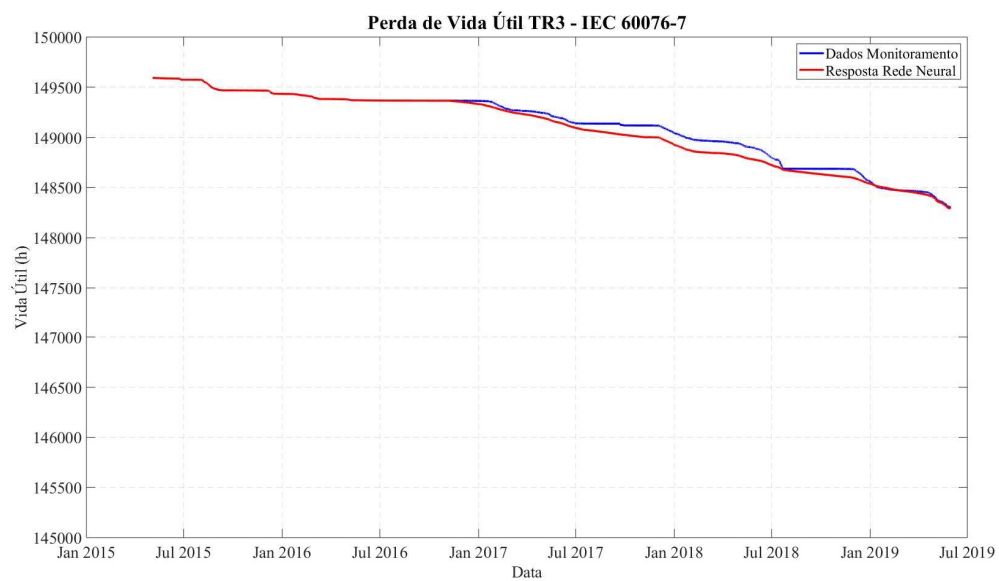
Figura 55 – Comparação perda de vida útil TR2.



Fonte: Autor.

Em relação ao transformador TR3, visto na Figura 56, obteve-se uma perda de vida útil para o período de 1306 horas para a rede e de 1292 horas para os dados de monitoramento. Resultando em uma diferença de apenas 14 horas na vida útil final.

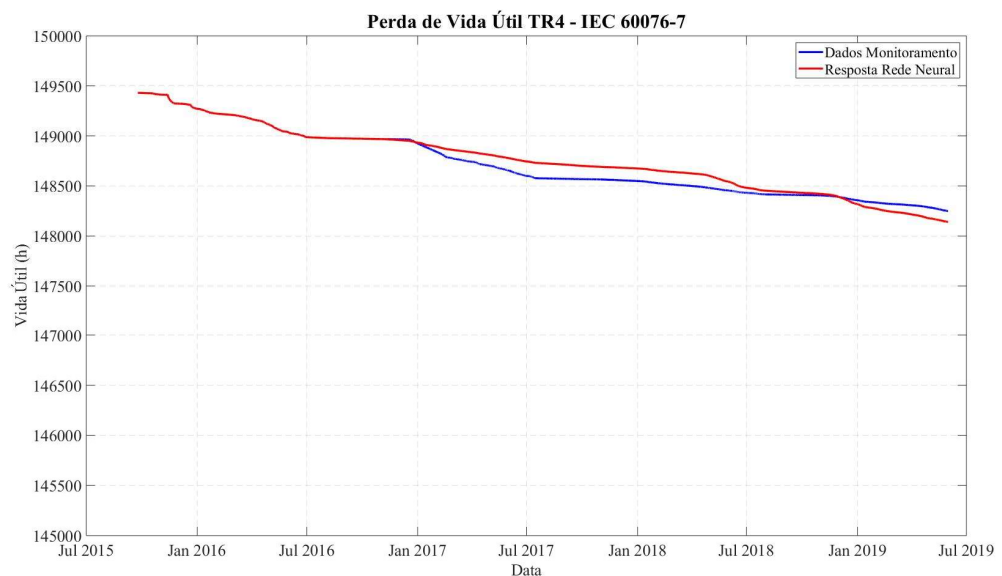
Figura 56 – Comparação perda de vida útil TR3.



Fonte: Autor.

O caso do TR4 apresentou a maior diferença entre a comparação realizada, de 109 horas. Onde a perda de vida útil para o período calculada a partir da resposta da rede neural foi de 1291 horas e para os dados de monitoramento de 1182 horas, conforme Figura 57.

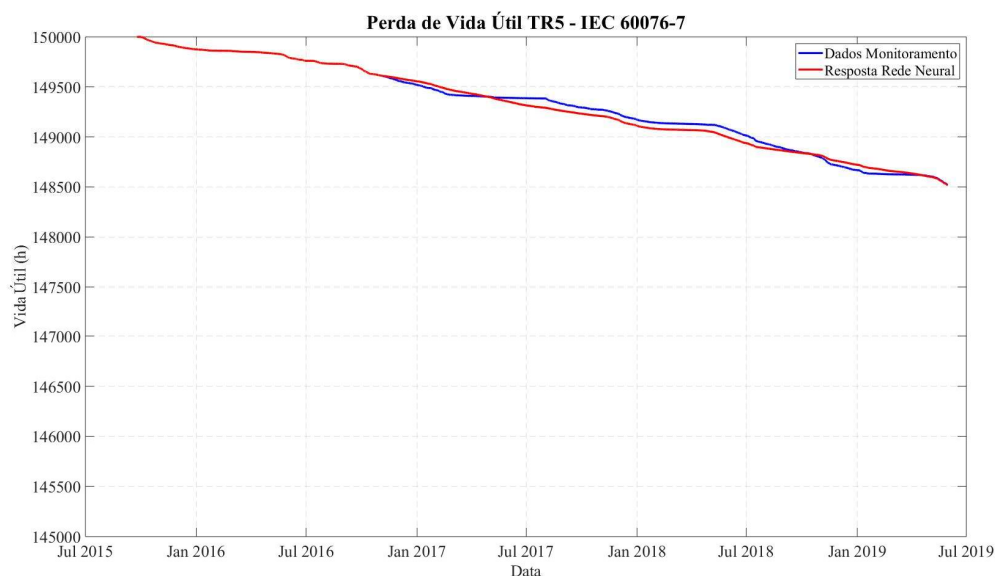
Figura 57 – Comparação perda de vida útil TR4.



Fonte: Autor.

O último dos transformadores é o TR5, conforme Figura 58. A perda de vida útil para o período calculada para a rede neural foi de 1484 horas, enquanto para os dados de monitoramento foi de 1480 horas. A diferença de apenas 4 horas na vida útil final comprova a boa capacidade das RNAs desenvolvidas de acordo com a metodologia do presente trabalho em prever a temperatura de topo de óleo em transformadores e a possibilidade de sua aplicação para o cálculo da perda de vida útil.

Figura 58 – Comparação perda de vida útil TR5.



Fonte: Autor.

A Tabela 7 apresenta uma síntese dos resultados observados para todos os cinco transformadores. O pior caso observado, para o transformador TR4, e o melhor caso, para o TR5, ocorrem de acordo com o desempenho da capacidade de predição da temperatura das RNAs desenvolvidas no capítulo anterior.

Tabela 7 – Resumo dos resultados das comparações.

Transformador	Período de Monitoramento (h)	Período Resposta Rede (h)	Diferença (h)	Final Monitoramento (h)	Final Resposta Rede (h)
TR1	2464	2377	87	4290	4203
TR2	1063	1028	35	1063	1028
TR3	1292	1306	14	1698	1712
TR4	1182	1291	109	1751	1860
TR5	1480	1484	4	1480	1484

Fonte: Autor.

Vale ressaltar que o transformador TR1 é o único que não apresenta um período de operação com potência reduzida ou sem carga e também é o equipamento a mais tempo em operação. Por conta disso apresenta uma maior deterioração em relação aos outros.

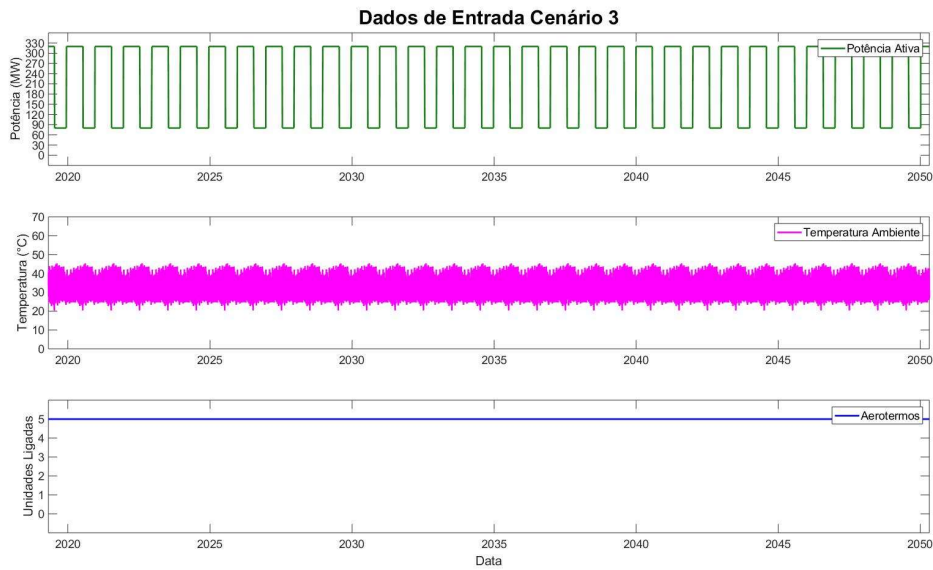
#### 4.6 VIDA ÚTIL DE CENÁRIOS FICTÍCIOS

Com o intuito de validar a capacidade preditiva das RNAs sob a perspectiva de vida útil, foram considerados seis cenários fictícios a serem aplicados ao transformador TR1 do instante em que terminam os dados de monitoramento até 2050, resultando em um período de 31 anos exatos. A escolha do TR1 foi realizada com base na sua maior degradação vista nas comparações, portanto os dados de entrada dos cenários foram aplicados a uma rede neural treinada com dados do TR1 para uma melhor representação.

A temperatura ambiente utilizada para a construção dos parâmetros de entrada dos cenários fictícios é a temperatura média referente ao período de um ano dos conjuntos de dados utilizados para treinamento e testes da capacidade de predição das RNAs vistos anteriormente. O período de um ano de temperatura ambiente teve então seu perfil replicado aos 31 anos totais do período. Este mesmo conjunto será utilizado nos seis cenários.

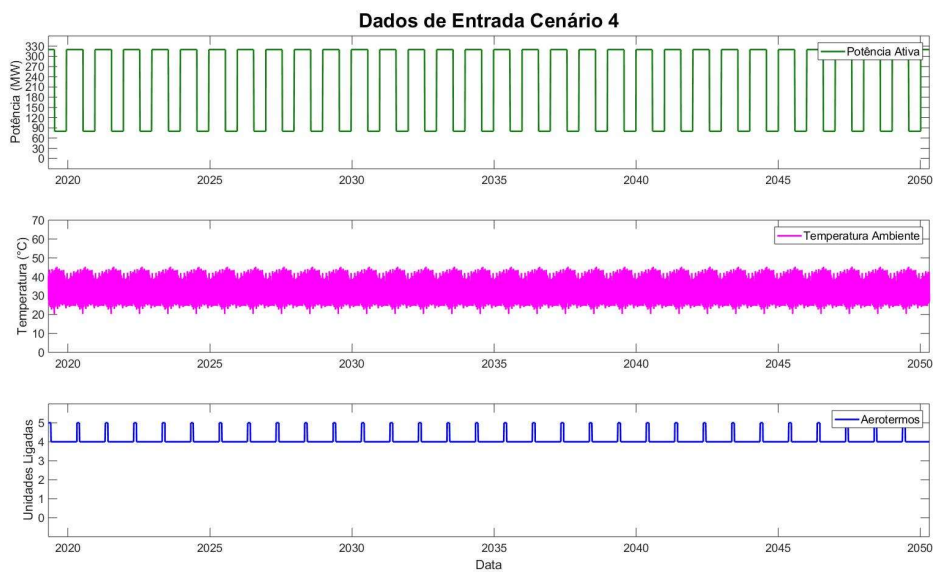
Os cenários 1 e 2 apresentam uma potência constante durante todo seu período de 320 MW, enquanto o Cenário 1 opera somente com 5 conjuntos de aerotermos ligados o Cenário 2 opera com 5 conjuntos durante 10% do tempo e 4 conjuntos ligados durante o restante do período. Os parâmetros de entrada para os cenários 3 e 4 estão ilustrados nas Figura 59 e Figura 60 respectivamente, onde leva-se em consideração a sazonalidade vista na usina para os dados de monitoramento durante o período de julho a dezembro. Portanto, durante este período nestes cenários o transformador opera com potência de 80 MW, e no restante do ano a 320 MW.

Figura 59 – Parâmetros de entrada da rede neural para o Cenário 5.



Fonte: Autor.

Figura 60 – Parâmetros de entrada da rede neural para o Cenário 4.



Fonte: Autor.

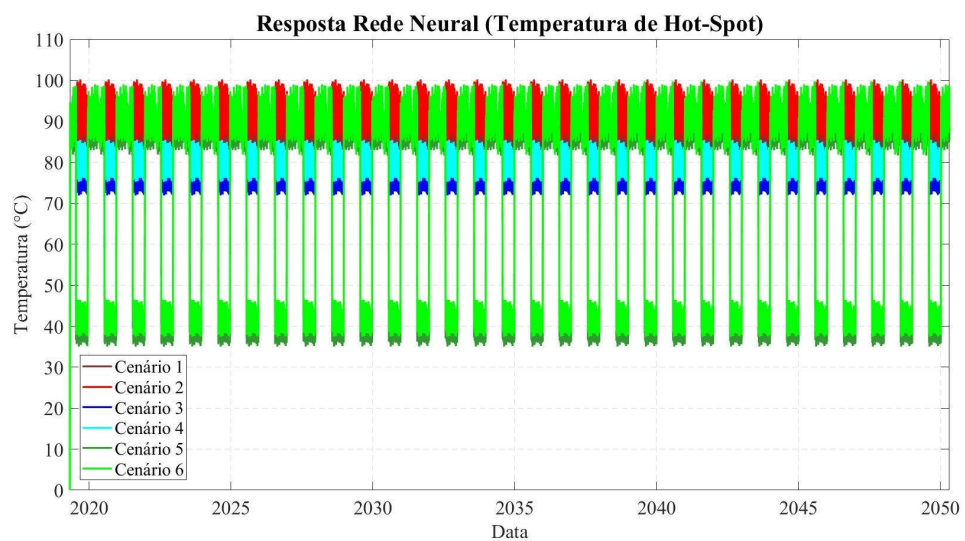
O mesmo padrão de comportamento dos conjuntos de aeroterms para o Cenário 1 é utilizado aos cenários 3 e 5, enquanto o mesmo padrão do Cenário 2 é aplicado aos cenários 4 e 6. Em relação a potência dos cenários 5 e 6 leva-se em consideração a sazonalidade sem carga. A Tabela 8 resume os parâmetros de entrada para os seis cenários.

Tabela 8 – Condições de operação para os cenários fictícios.

Cenário	Potência (Jul – Dez)	Potência (Dez – Jul)	Temperatura Ambiente	Aeroterms
1	320 MW	320 MW	Média	5 (100%)
2	320 MW	320 MW	Média	5 (10%) / 4 (90%)
3	80 MW	320 MW	Média	5 (100%)
4	80 MW	320 MW	Média	5 (10%) / 4 (90%)
5	0 MW	320 MW	Média	5 (100%)
6	0 MW	320 MW	Média	5 (10%) / 4 (90%)

Fonte: Autor.

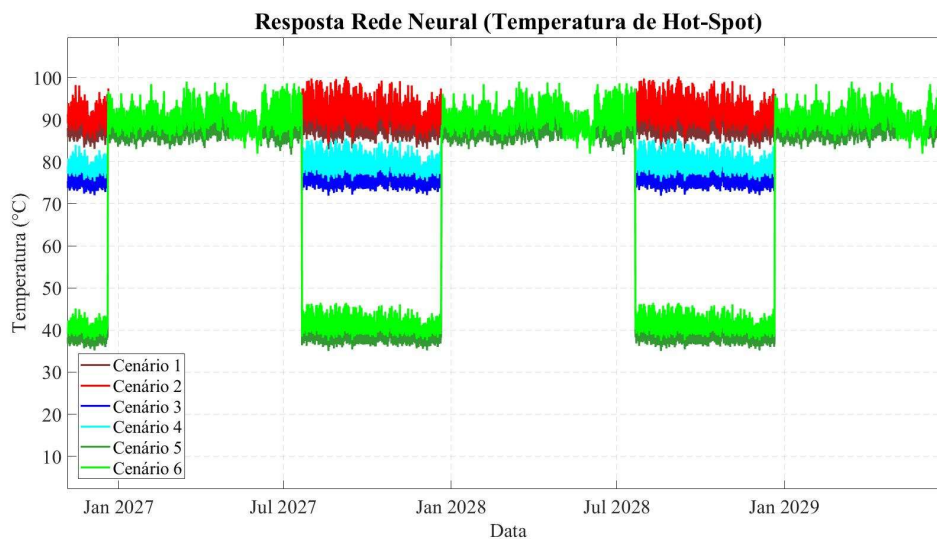
A resposta obtida através da aplicação das entradas dos cenários fictícios na rede neural, já acrescida da elevação de temperatura de ponto mais quente em relação ao topo de óleo, para os instantes em que há carregamento, para o período total pode ser vista na Figura 61. Foi utilizado o mesmo valor para o cálculo da perda de vida útil no transformador TR1 de 34,7 °C.

Figura 61 – Temperatura de *hot-spot* cenários fictícios.

Fonte: Autor.

Um período de pouco mais de dois anos pode ser visualizado na Figura 62, ilustrando a temperatura de *hot-spot* obtida para todos os cenários. Pode-se observar que durante os instantes que todos transformadores operam com 320 MW de potência e 5 conjuntos de aeroterms ligados suas temperaturas são as mesmas, quando há diferença apenas no número de aeroterms em operação entre os cenários de mesma potência há sempre uma pequena elevação de temperatura para o cenário de 4 aeroterms. Em nenhum instante a temperatura de ponto mais quente para os cenários ultrapassa o valor de 100 °C.

Figura 62 – Temperatura de *hot-spot* cenários fictícios.

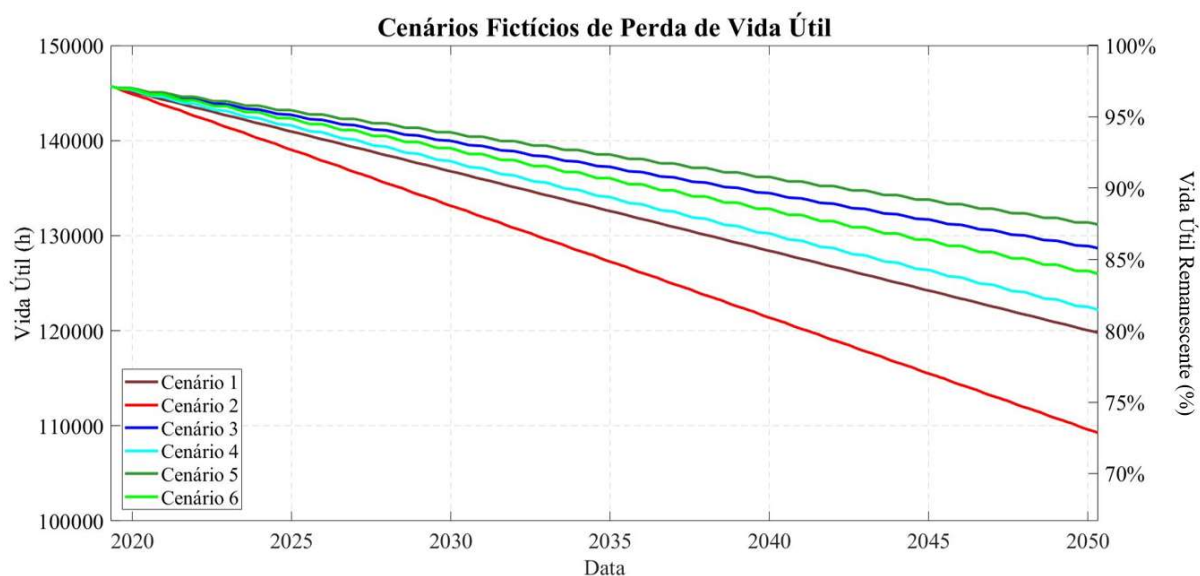


Fonte: Autor.

A Figura 63 apresenta a perda de vida útil calculada através da equação (12) para os seis cenários durante o período. O valor inicial de vida útil é atribuído ao valor de 4290 horas calculados para os dados de monitoramento do TR1, ou seja, 145710 horas, cerca de 97% do valor base inicial.

O cenário com maior degradação é o cenário 2 que tem perda de 36435 horas durante o período e chega a vida útil final de 109275 horas, cerca de 73% de vida útil remanescente segundo o limite estabelecido. A Tabela 9 resume os resultados apresentados.

Figura 63 – Perda de vida útil cenários fictícios.



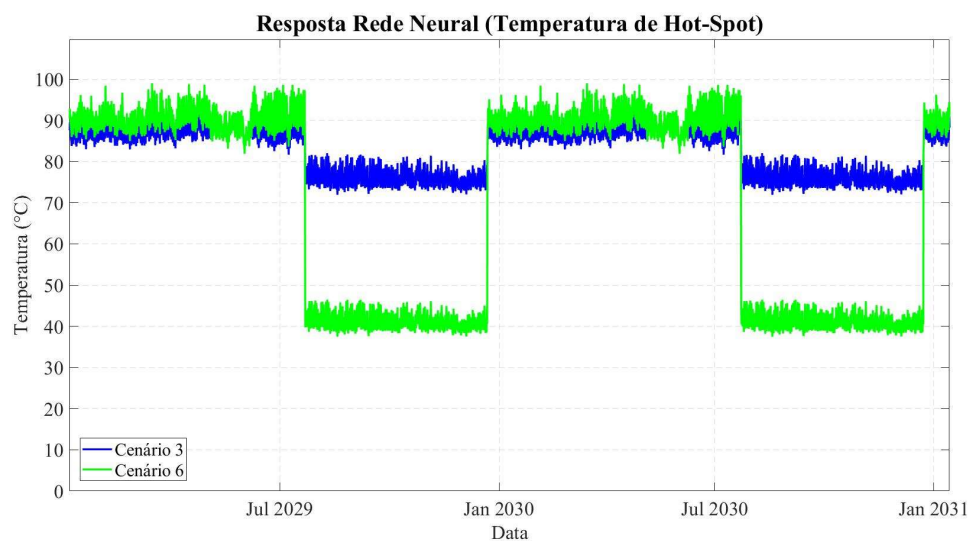
Fonte: Autor.

Tabela 9 – Perda de vida útil cenários fictícios.

Cenário	Perda de Vida Útil no Período (horas)	Vida Útil Remanescente (horas)	Vida Útil Remanescente (%)
1	25907	119803	79,87
2	36435	109275	72,85
3	17038	128672	85,78
4	23527	122183	81,45
5	14548	131162	87,44
6	19742	125968	83,98

Fonte: Autor.

Pode ser observado que o Cenário 6 apresenta uma maior perda de vida útil em comparação ao Cenário 3, sendo que durante um período a potência do Cenário 3 é maior. Entretanto, durante o período referente em que ambos operam com potência de 320 MW o Cenário 6 possui 4 conjuntos de aerotermos em operação, enquanto o Cenário 3 apresenta os 5 ligados durante todo o período considerado. A diferença de temperatura de *hot-spot* causada pela operação com 4 aerotermos ao invés de 5, conforme pode ser visualizado na Figura 64, por menor que seja se comparada a diferença que estes apresentam durante o período de julho a dezembro está mais próxima da temperatura de referência utilizada para o cálculo da taxa de envelhecimento e portanto resulta em uma taxa de envelhecimento maior, conforme Tabela 10 e equação (10), e consequente perda de vida útil.

Figura 64 – Temperatura de *hot-spot* cenários 3 e 6.

Fonte: Autor.



Tabela 10 – Taxa de Envelhecimento Relativo pela Temperatura

$\theta_h$ (°C)	Taxa de Envelhecimento
80	0,036
86	0,073
92	0,145
98	0,282
104	0,536
110	1,0
116	1,83
122	3,29

Fonte: (ABNT NBR 5356-7, 2017).

#### 4.7 CONSIDERAÇÕES DO CAPÍTULO

O presente capítulo apresentou uma análise da perda de vida útil para os dados de monitoramento transformadores sob estudo e realizou comparações com a temperatura prevista pelas RNAs desenvolvidas. Os resultados apresentados mostraram que a perda de vida útil estimada a partir da temperatura da rede neural assume valores muito próximos do calculado a partir dos dados de monitoramento e comprova a capacidade desta metodologia de ser utilizada para este propósito.

Com o intuito de simular a capacidade de utilizar a rede neural para avaliar cenários futuros foram propostos seis cenários fictícios, que vão até 2050. Os resultados apresentados são compatíveis com a perda de vida útil observada para os dados de monitoramento dos transformadores, de acordo com as condições de operação de cada cenário.

Entretanto, a perda vida útil calculada para os transformadores é muito menor do que seu tempo em operação. Isso se deve pelo fato de a estimativa utilizada considerar apenas a temperatura de ponto mais quente como fator deteriorante da isolação sólida, e como os transformadores sob estudo são conectados a unidades geradoras e seu projeto é conservador, não operam em sobrecarga e as temperaturas de operação são normalmente mais baixas.

Contudo, como o objetivo destas análises é apenas comprovar a capacidade de utilizar a temperatura prevista pelas redes neurais desenvolvidas para estimar a vida útil dos transformadores, mesmo que não represente a vida útil real, o difundido cálculo proposto pelas normativas cumpre seu papel. Caso deseja-se determinar a temperatura para cenários em que se tem previsões para concentração de umidade e oxigênio e tenha-se em mãos um modelo de

perda de vida útil que leve em consideração estes fatores, o método apresentado neste trabalho para o desenvolvimento das RNAs pode ser utilizado.

## 5. CONSIDERAÇÕES FINAIS

Os transformadores de potência são equipamentos essenciais ao assegurar o funcionamento adequado do sistema elétrico de potência, de forma confiável, segura e econômica. Sendo os ativos mais caros de uma subestação, é de grande interesse o conhecimento das condições atuais e futuras destes equipamentos em termo de degradação, evitando falhas ou sinistros, que acarretam em grandes prejuízos técnicos e financeiros. Nesta linha, o presente trabalho apresentou um método para a predição de temperatura de topo de óleo em transformadores de potência empregando Redes Neurais Artificiais.

Inicialmente, após a contextualização e motivação que justificam este trabalho, foi realizada uma introdução às características e capacidades das RNAs como ferramenta e uma revisão dos aspectos mais importantes que compõe sua estrutura. Após, alguns comentários sobre outros trabalhos que empregam RNAs para o mesmo propósito são realizados, ressaltando as características que foram ou deixaram de ser levadas em consideração, e quais mais se assemelham aos objetivos aqui apresentados.

No Capítulo 3 foram apresentados os conjuntos de dados de monitoramento referentes aos objetos de estudo, explicados os parâmetros utilizados para a arquitetura de RNA desenvolvida neste trabalho, comentários a respeito do treinamento e realizados testes de capacidade de predição para uma série de RNAs com conjuntos de dados que não participaram da etapa de treinamento. Os resultados apresentados para todos os casos apresentam um desempenho muito satisfatório das redes neurais desenvolvidas, onde na maior parte do período utilizado para o teste o erro entre a resposta da rede e o monitoramento é ínfimo. Os instantes em que ocorrem os maiores erros, responsáveis pelo aumento do erro médio absoluto geral, foram apresentados e as suas causas discorridas. Constatou-se que pelas diferentes condições de operação impostas aos transformadores o conjunto de dados utilizados para o treinamento da rede neural nem sempre apresenta a abrangência necessária para replicar de forma precisa os comportamentos atribuídos a outro transformador. Entretanto, mesmo para transformadores com diferentes tipos de resfriamento, em situações que foram abordadas pelo conjunto de dados que participaram dos dados de treinamento o desempenho das RNAs consolida esta ferramenta para a utilização em predição de temperatura em transformadores.

O Capítulo 4 apresentou a metodologia para estimar a vida útil de transformadores proposta por norma. Foram então realizadas comparações entre a perda de vida útil calculada para os dados de monitoramento e a temperatura resposta para a rede neural utilizada. Os resultados novamente foram muito satisfatórios, podendo ser visualizada nos gráficos da

evolução da vida útil remanescente dos transformadores que se mantém em valores próximos durante o período considerado. Como a capacidade de predição da temperatura já havia sido comprovada, era esperado que ao realizar o cálculo da vida útil considerando apenas da temperatura de operação os resultados fossem também satisfatórios.

Porém, como foi visto durante a perda de vida útil calculada para os cenários fictícios, quando próximos da temperatura de referência, poucos graus podem determinar uma grande diferença na vida útil final para longos períodos de tempo. Levando então em consideração que os pontos críticos observados para as RNAs desenvolvidas são períodos de desligamento ou operação com potência reduzida, são condições de operação com temperaturas baixas, não apresentando um grande impacto no cálculo da perda de vida útil.

Em relação aos cenários fictícios, foram considerados seis cenários com condições de operação inspiradas no que pôde ser visto nos dados de monitoramento. Utilizando-se então o transformador TR1 como base e utilizando uma rede neural treinada com os dados dele calculou-se a degradação referente às condições de operação até o ano de 2050, onde foi discutido sobre os motivos de os transformadores não chegarem ao final de sua vida útil neste período, sendo que o estimado por norma é de 35 anos de operação. A degradação observada para os seis cenários esteve de acordo com o esperado para os parâmetros de entrada utilizados, comprovando-se mais uma vez a capacidade do método apresentado neste trabalho de ser aplicado não apenas na predição da temperatura de operação dos transformadores como para prever a consequente degradação e possibilitando um melhor planejamento pelos responsáveis.

Portanto, os resultados validam a utilização de RNAs desenvolvidas sob o método apresentado neste trabalho para predição de temperatura de topo de óleo em transformadores de potência e previsão da perda de vida útil para cenários futuros. Ressalta-se a importância da escolha de um conjunto de dados para treinamento cuja abrangência contenha a maior variedade possível de comportamentos que os transformadores sob estudo possam apresentar, a partir disso pode-se obter respostas mais precisas.

O método apresentado resulta em diversos benefícios no âmbito de operação e manutenção em subestações de energia e usinas geradoras. Ao ser empregado para a previsão de temperatura e perda de vida útil dá-se aos operadores uma estimativa de quais transformadores podem processar maior parcela da demanda prevista, facilitando a manutenção preventiva e a possibilidade de traçar cenários de degradação aos transformadores também em curto prazo. Em conjunto com análises do óleo isolante pode-se calcular o Health-Index e compor uma ferramenta importante na gestão de ativos e planejamento estratégico a longo prazo.

Novos avanços podem ser realizados nesta linha. A metodologia apresentada é válida para a predição de temperatura de ponto mais quente caso se possuam dados de monitoramento provindos de medição direta, e não calculada como foi realizado nas análises de perda de vida útil. Também podem ser realizados avanços na etapa de treinamento, utilizando outros métodos que apresentem um bom desempenho, como o algoritmo de LM utilizado, mas que se possa reforçar o treinamento com outros conjuntos de dados para as condições de operação em que houve os piores desempenhos.

### 5.1 PUBLICAÇÕES DO TRABALHO

KAMINSKI, A. M.; MARCHESAN, T. B.; BENDER, V. C.; MEDEIROS, L. H.; OLIVEIRA, M. M.; BUENO, D. M.; NETO, J. B. F.; WILHELM, H. M. Emprego de Rede Neural Artificial para Predição de Temperatura de Topo de Óleo em Transformador de Potência. SBSE 2020, Santo André, Brasil, 2020.

## REFERÊNCIAS

- ABNT. **NBR 5356-2: Aquecimento**. Associação Brasileira de Normas Técnicas, 2007.
- ABNT. **NBR 5356-7: Guia de carregamento para transformadores imersos em líquido isolante**. Associação Brasileira de Normas Técnicas, 2017.
- ANDRADE, L. C. M. de et. al. **Very short-term load forecasting based on NARX recurrent neural networks**. Power & Energy Society General Meeting. Washington DC, EUA, 2014.
- FEIL, D. L. P. **Substituição de Transformadores de Potência em Subestações de Energia: Uma Estratégia Global**. 2019. Tese (Doutorado em Engenharia Elétrica). Universidade Federal de Santa Maria (UFSM), Santa Maria, 2019.
- HAGAN, M. T.; DEMUTH, H. B. **Neural network design**. Segunda Edição, Editora Martin Hagan. Stillwater, Oklahoma, EUA, 2016.
- HAYKIN, S. **Neural networks and learning machines**. Prentice Hall, Terceira Edição. Upper Saddle River, Nova Jersey, EUA, 2009.
- HAYKIN, S. **Neural Networks: A Comprehensive Foundation**. Prentice Hall. Segunda Edição. Upper Saddle River, Nova Jersey, EUA, 1999.
- HE, Q.; SI, J.; TYLAVSKY, D. J. **Prediction of top-oil temperature for transformers using neural networks**. IEEE Transactions on Power Delivery, Vol. 15, 2000.
- HOPFIELD, J. J. **Neural Networks and Physical Systems with Emergent Collective Computational Abilities**. Proceedings of the National Academy of Sciences. Vol. 79, p. 2554 – 2558. 1982.
- HUANG, H. et. al. **Transformer top-oil modeling based on Kernel-based extreme learning machine**. ICEEA 2016. Kuala Lumpur, Malásia, 2016.
- IEC. **Std 60076-2: Temperature rise for liquid-immersed transformers**. International Electrotechnical Commission, 2011.
- IEC. **Std 60076-7: Loading guide for mineral-oil-immersed power transformers**. International Electrotechnical Commission, 2018.
- KHALIL, R. A. **Comparison of four neural network learning methods based on genetic algorithm for non-linear dynamic systems identification**. Al-Rafidain Engineering Vol. 20, 2012.
- MCCULLOCH, W.; PITTS, W. **A Logical Calculus of the Ideas Immanent in Nervous Activity**. Bulletin of Mathematical Biophysics. Vol. 5, p. 115 – 133. 1943.

MORIGGI, T. **Predição de Performance de Cimentos Compostos por Meio da Aplicação de Redes Neurais Artificiais Visando a Garantia da Qualidade.** 2018. Dissertação (Mestrado em Engenharia de Produção, Universidade Federal do Paraná, Curitiba, 2018).

OLIVEIRA, A. R. de. **Redes neurais artificiais aplicadas na detecção, classificação e localização de defeitos em linhas de transmissão.** Dissertação de Mestrado, Departamento de Engenharia Elétrica, Universidade Federal de Juiz de Fora. Juiz de Fora, Minas Gerais, Brasil, 2005.

PEREIRA, F. H. **Nonlinear autoregressive neural network models for prediction of transformer oil-dissolved gas concentrations.** *Energies* Vol. 11, 2018.

PRADHAN, M. K.; RAMU, T. S. **On-line monitoring of temperature in power transformers using optimal linear combination of ANNs.** IEEE International Symposium in Electrical Insulation. Indianapolis, Indiana, EUA, 2004.

RUSLAN, F. A. et. al. **Flood water level modeling and prediction using NARX neural network: Case study at Kelang River.** CSPA2014, Kuala Lumpur, Malásia, 2014.

SU, X. et. al. **Application of Elman neural network in top-oil temperature prediction of transformer.** ICHVE 2018. Atenas, Grécia, 2018.

VILAITHONG, R.; TENBOHLEN, S. E STIRL, T., **Neural network for transformer top-oil temperature prediction.** Symposium on High Voltage Engineering. Ljubljana, Eslovênia, 2007.

WEI, B. G. et. al. **A method of optimized neural network by LM algorithm to transformer winding hot spot temperature forecasting.** NEFES 2017. Kunming, China, 2017.