

UNIVERSIDADE FEDERAL FLUMINENSE
ESCOLA DE ENGENHARIA
MESTRADO EM ENGENHARIA ELÉTRICA E DE TELECOMUNICAÇÕES

SISTEMA INTELIGENTE PARA ESTIMAÇÃO DE POTENCIAL DE PROTEÇÃO
CATÓDICA BASEADO EM REDES NEURAS ARTIFICIAIS

Rodrigo Machado da Silva Soares

Niterói
Outubro de 2019

UFF - PROGRAMA DE PÓS-GRADUAÇÃO DE ENGENHARIA ELÉTRICA E DE
TELECOMUNICAÇÕES

SISTEMA INTELIGENTE PARA ESTIMAÇÃO DE POTENCIAL DE PROTEÇÃO
CATÓDICA BASEADO EM REDES NEURAIS ARTIFICIAIS

Rodrigo Machado da Silva Soares

Dissertação apresentada ao Programa de Pós-Graduação
em Engenharia Elétrica e de Telecomunicações da Universidade
Federal Fluminense, como requisito parcial para obtenção do
título de Mestre em Engenharia Elétrica e de Telecomunicações.

Orientador: Prof. Márcio Zamboti Fortes, Dr.
Coorientador: Prof. Vitor Hugo Ferreira, DSc..

Niterói

Outubro de 2019

S586s Silva soares, Rodrigo Machado da
Sistema Inteligente para Estimação de Potencial de
Proteção Catódica Baseado em Redes Neurais Artificiais /
Rodrigo Machado da Silva soares ; Marcio Zamboti Fortes,
orientador ; Vitor Hugo Ferreira, coorientador. Niterói, 2019.
128 p. : il.

Dissertação (mestrado)-Universidade Federal Fluminense,
Niterói, 2019.

DOI: <http://dx.doi.org/10.22409/PPGEET.2019.m.12131872739>

1. Proteção catódica. 2. Aprendizado de Máquina. 3.
Produção intelectual. I. Fortes, Marcio Zamboti, orientador.
II. Ferreira, Vitor Hugo, coorientador. III. Universidade
Federal Fluminense. Escola de Engenharia. IV. Título.

CDD -

RODRIGO MACHADO DA SILVA SOARES

**SISTEMA INTELIGENTE PARA ESTIMAÇÃO DE POTENCIAL DE PROTEÇÃO
CATÓDICA BASEADO EM REDES NEURAS ARTIFICIAIS**

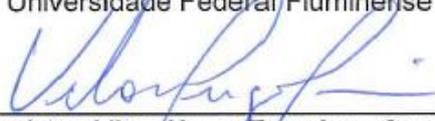
Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Engenharia Elétrica e de Telecomunicações da Universidade Federal Fluminense, como requisito parcial para a Obtenção do Grau de Mestre em Engenharia Elétrica e de Telecomunicações.

Área de concentração: Sistemas de Energia Elétrica

BANCA EXAMINADORA



Prof. Marcio Zamboti Fortes - Orientador
Universidade Federal Fluminense - UFF



Prof. Dr. Vitor Hugo Ferreira - Coorientador
Universidade Federal Fluminense - UFF



Prof. Dr. André Abel Augusto
Universidade Federal Fluminense - UFF



Prof. Dr. Bruno Henrique Dias
Universidade Federal de Juiz de Fora – UFJF

Niterói

Outubro de 2019

Dedicatória

Dedico este trabalho à Deus, pois sem Ele nada sou e a minha esposa por todo apoio, paciência, compreensão e companheirismo.

Agradecimentos

A Deus, pelas oportunidades que surgiram em minha vida. Hoje eu posso dar passos outrora inimagináveis.

A minha esposa Karolina, por todo apoio dado nesses anos de estudos, incentivo para que ingressasse neste desafio, compreensão com as ausências e amor incondicional dedicado em todos os momentos e principalmente nestes últimos meses.

A minha família de sangue e de coração, sinto pela ausência e agradeço a compreensão ao longo desses anos.

A todos os professores com os quais pude interagir e aprender um pouco mais. Agradeço, principalmente, ao professor Márcio Zamboti e ao professor Vitor Hugo, exemplos de dedicação, didática e paciência, muito obrigado por todo o ensinamento ao longo desses anos.

Aos meus amigos de mestrado da UFF, que foram fundamentais nessa caminhada. Sem vocês esse desafio seria muito mais árduo e faria menos sentido.

Agradeço a todos os gestores de minha empresa, pelo incentivo e apoio nesses últimos dois anos.

A todos que participaram e torceram por mim de alguma forma para a realização deste trabalho. Muito obrigado!

Resumo

A proteção catódica é uma técnica utilizada para preservação de estruturas metálicas imersas em um eletrólito, cuja função é minimizar a corrosão a valores desprezíveis. Para proteção de grandes estruturas, é utilizada uma corrente contínua externa conectando o anodo (estrutura a ser sacrificada) ao catodo (estrutura a ser protegida). O controle para verificação do bom funcionamento deste sistema se dá através da inspeção em campo ou por meio de monitoração via *datalogger* remoto. A periodicidade da inspeção de um elemento em campo se dará de acordo com a importância deste para um determinado sistema.

As estruturas metálicas protegidas catódicamente estão normalmente dispostas nos subsolos de grandes cidades, sendo assim, estão sujeitas a fatores não controlados, como o crescimento desordenado, ocupação irregular de espaços públicos, obras não informadas por outras concessionárias de serviços públicos, falta de segurança pública, dentre diversos outros fatores que possivelmente irão dificultar a inspeção dos elementos de proteção catódica.

Para minimizar os custos com inspeções em campo, os elementos de proteção catódica são categorizados com o objetivo de definir maiores e menores periodicidades de inspeção. Os elementos com maior periodicidade são os mais críticos em relação aos fatores não controlados citados no parágrafo anterior.

Tendo uma boa base de dados de inspeções anteriores, alinhada com uma metodologia adequada para estimação de um dado faltante em um determinado instante, levará ao processo de manutenção do sistema de proteção catódica uma maior confiabilidade na análise completa da inspeção, melhorando os índices da manutenção preventiva, segurança operacional e respeito ao meio ambiente.

O objetivo principal desta dissertação é, a partir das bases de dados de dois sistemas de proteção catódica de gasodutos de distribuição de gás natural distintos, criar uma rede neural que, através de duas metodologias baseadas em Redes Neurais Artificiais (RNA), possa estimar o dado faltante de um dos elementos do sistema em um determinado instante. Após a análise dos resultados obtidos, uma das metodologias será a escolhida a partir da comparação do erro associado. Parte das bases de dados não foram utilizadas no treinamento com o propósito de testar os resultados e embasar a escolha da metodologia mais adequada aos objetivos deste trabalho.

Palavras-chave: Aprendizado de máquina, redes neurais, regressão, rede Bayesiana, proteção catódica.

Abstract

Cathodic protection is a technique used for the preservation of metallic structures immersed in an electrolyte, whose function is to minimize corrosion to negligible values. For protection of large structures, an external direct current is used connecting the anode (structure to be sacrificed) to the cathode (structure to be protected). The control to verify the good functioning of this system is given through the field inspection or through remote datalogger monitoring. The periodicity of the inspection of an element in the field will be according to the importance of this element for a given system.

The cathodically protected metal structures are normally arranged in the subsoil of large cities and are therefore subject to uncontrolled factors, such as disordered growth, irregular occupation of public spaces, works not reported by other utilities, lack of public safety, among many other factors that possibly will make it difficult to inspect the cathodic protection elements.

To minimize costs with field inspections, cathodic protection elements are categorized with the purpose of defining higher and lower inspections periodicities. The elements with the greatest frequency are the most critical in relation to the uncontrolled factors mentioned in the previous paragraph.

The main objective of this dissertation is, from the databases of two systems of cathodic protection of distinct natural gas distribution pipelines, to create a neural network that, through two methodologies based on Artificial Neural Networks (RNA or ANN), can estimate the missing data from one of the system elements at a given time. After analyzing the results obtained, one of the methodologies will be chosen from the comparison of the associated error. Part of the databases were not used in the training with the purpose of testing the results and to base the choice of the most adequate methodology to the objectives of this work.

Keywords: Machine learning, neural networks, regression, Bayesian network, cathodic protection.

Sumário

Lista de Figuras	xii
Lista de Tabelas	xvi
Lista de Abreviaturas.....	xvii
Capítulo 1 - Introdução.....	19
1.1 Motivação	22
Objetivos.....	23
1.2.....	23
1.3 Objetivos Específicos do Trabalho	23
1.4 Publicação decorrente da pesquisa.....	24
1.5 Estrutura do Documento	24
Capítulo 2 - Corrosão e Métodos de Mitigação	25
2.1 Corrosão.....	25
2.1.1 Polarização	28
2.1.2 Potencial do eletrodo.....	29
2.1.3 Passividade	30
2.1.4 Diagrama de Pourbaix.....	31
2.2 Métodos para inibição da corrosão	33
2.2.1 Critérios para combate e controle da corrosão	33
2.2.2 Revestimento de dutos enterrados.....	35
2.2.3 Proteção Catódica.....	36
2.2.3.1 Proteção catódica galvânica	37
2.2.3.2 Proteção catódica por corrente impressa	38
2.2.4 Pontos de medição de potencial	39
2.2.5 Interferência elétrica.....	40
2.2.6 Inspeção do sistema de proteção catódica.....	41

Capítulo 3 - Redes Neurais.....	44
3.1 Redes <i>perceptrons</i> de Múltiplas Camadas (MLP)	45
3.1.1 Normalização dos dados de entrada	50
3.2 Rede <i>Bayesiana</i>	52
3.2.1 Regressão	53
3.2.2 Método Automático de Treinamento de MLPs utilizando rede <i>Bayesiana</i>	58
Capítulo 4 - Aplicação de RNAs na estimação de dados faltantes de proteção catódica.....	61
4.1 Estrutura dos sistemas de proteção catódica	62
4.1.1 Estrutura caso 01	63
4.1.2 Estrutura caso 02	64
4.2 Desenvolvimento da RNA	66
4.2.1 Metodologia 01	66
4.2.2 Metodologia 02	68
4.3 Metodologia para comparação entre as saídas e erro associado.	69
Capítulo 5 - Resultados	71
5.1 Resultados Caso 01	71
5.1.1 Caso1 – Resultado PT01	72
5.1.2 Caso1 – Resultado PT02	75
5.1.3 Caso1 – Resultado PT03	77
5.1.4 Caso1 – Resultado PT04	79
5.1.5 Caso1 – Resultado PT05	81
5.1.6 Caso1 – Resultado PT06	83
5.1.7 Caso1 – Resultado PT07	85
5.1.8 Caso1 – Resultado PT08	87
5.1.9 Caso1 – Resultado PT09	89
5.1.10 Resultado Geral – Caso 01	91
5.2 Resultados Caso 02	93
5.2.1 Caso2 – Resultado PT01	94

5.2.2	Caso2 – Resultado PT02	96
5.2.3	Caso2 – Resultado PT03	98
5.2.4	Caso2 – Resultado PT04	100
5.2.5	Caso2 – Resultado PT05	102
5.2.6	Resultado Geral – Caso 02	104
5.3	Considerações gerais.....	106
5.3.1	PT01 – Caso 02	107
5.3.2	PT02 – Caso 02	108
5.3.3	PT03 – Caso 02	110
5.3.4	PT04 – Caso 02	111
5.3.5	PT05 – Caso 02	113
Capítulo 6 - Conclusões		116
Trabalhos Futuros		117
6.1		117
Capítulo 7 - Bibliografia.....		119

Lista de Figuras

FIGURA 1– CICLO DA CORROSÃO [10]	25
FIGURA 2– CÉLULA E CORROSÃO SIMPLES [10]	27
FIGURA 3 – POTENCIAL DO ELETRODO	29
FIGURA 4 – CURVA DE POLARIZAÇÃO ANÓDICA [15]	31
FIGURA 5 – DIAGRAMA DE POURBAIX [17]	33
FIGURA 6 – SISTEMA DE PROTEÇÃO CATÓDICA POR ANODO DE SACRIFÍCIO.....	38
FIGURA 7 – SISTEMA DE PROTEÇÃO CATÓDICA POR ANODO DE SACRIFÍCIO.....	39
FIGURA 8 – SISTEMA DE PROTEÇÃO CATÓDICA POR CORRENTE IMPRESSA	40
FIGURA 9 – DIAGRAMA ESQUEMÁTICO DE UM NEURÓNIO - ADAPTADO DE [27]	45
FIGURA 10 – GRAFO ARQUITETURAL DE UM PERCEPTRON DE MÚLTIPLAS CAMADAS COM DUAS CAMADAS OCULTAS - ADAPTADO DE [26]	47
FIGURA 11 – DIREÇÕES DOS FLUXOS DE SINAL BÁSICO DE UM PERCEPTRON DE MÚLTIPL CAMADAS - ADAPTADO DE [26].....	
FIGURA 12 – FUNÇÃO TANGENTE HIPERBÓLICA [33]	51
FIGURA 13 – FUNÇÃO LOGÍSTICA [33].....	52
FIGURA 14 – ESTRUTURA PARA TREINAMENTO	62
FIGURA 15– IMAGEM DOS ELEMENTOS DE PROTEÇÃO CATÓDICA DO CASO 1 DISPOSTOS NO MAPA	63
FIGURA 16 – GRÁFICO COM OS DADOS DO CASO 01.	64
FIGURA 17 – IMAGEM DOS ELEMENTOS DE PROTEÇÃO CATÓDICA DO CASO 2 DISPOSTOS NO MAPA	65
FIGURA 18 – GRÁFICO COM OS DADOS DO CASO 02.	65
FIGURA 19 – ESTRUTURA DA REDE NEURAL ARTIFICIAL.	67
FIGURA 20 – GRÁFICO TUBO SOLO X TEMPO: PT01, CASO 01	73
FIGURA 21 – GRÁFICO DISPERSÃO: PT01, CASO 01	73
FIGURA 22 – GRÁFICO HISTOGRAMA: PT01, CASO 01, METODOLOGIA 01.....	74
FIGURA 23 – GRÁFICO HISTOGRAMA: PT01, CASO 01, METODOLOGIA 02.....	74
FIGURA 24 – GRÁFICO TUBO SOLO X TEMPO: PT02, CASO 01	75
FIGURA 25 – GRÁFICO DISPERSÃO: PT02, CASO 01	75
FIGURA 26 – GRÁFICO HISTOGRAMA: PT02, CASO 01, METODOLOGIA 01.....	76
FIGURA 27 – GRÁFICO HISTOGRAMA: PT02, CASO 01, METODOLOGIA 02.....	76

FIGURA 28 – GRÁFICO TUBO SOLO X TEMPO: PT03, CASO 01	77
FIGURA 29 – GRÁFICO DISPERSÃO: PT03, CASO 01	77
FIGURA 30 – GRÁFICO HISTOGRAMA: PT03, CASO 01, METODOLOGIA 01.....	78
FIGURA 31 – GRÁFICO HISTOGRAMA: PT03, CASO 01, METODOLOGIA 02.....	78
FIGURA 32 – GRÁFICO TUBO SOLO X TEMPO: PT04, CASO 01	79
FIGURA 33 – GRÁFICO DISPERSÃO: PT04, CASO 01	79
FIGURA 34 – GRÁFICO HISTOGRAMA: PT04, CASO 01, METODOLOGIA 01.....	80
FIGURA 35 – GRÁFICO HISTOGRAMA: PT04, CASO 01, METODOLOGIA 02.....	80
FIGURA 36 – GRÁFICO TUBO SOLO X TEMPO: PT05, CASO 01	81
FIGURA 37 – GRÁFICO DISPERSÃO: PT05, CASO 01	81
FIGURA 38 – GRÁFICO HISTOGRAMA: PT05, CASO 01, METODOLOGIA 01.....	82
FIGURA 39 – GRÁFICO HISTOGRAMA: PT05, CASO 01, METODOLOGIA 02.....	82
FIGURA 40 – GRÁFICO TUBO SOLO X TEMPO: PT06, CASO 01	83
FIGURA 41 – GRÁFICO DISPERSÃO: PT06, CASO 01	83
FIGURA 42 – GRÁFICO HISTOGRAMA: PT06, CASO 01, METODOLOGIA 01.....	84
FIGURA 43 – GRÁFICO HISTOGRAMA: PT06, CASO 01, METODOLOGIA 02.....	84
FIGURA 44 – GRÁFICO TUBO SOLO X TEMPO: PT07, CASO 01	85
FIGURA 45 – GRÁFICO DISPERSÃO: PT07, CASO 01	85
FIGURA 46 – GRÁFICO HISTOGRAMA: PT07, CASO 01, METODOLOGIA 01.....	86
FIGURA 47 – GRÁFICO HISTOGRAMA: PT07, CASO 01, METODOLOGIA 02.....	86
FIGURA 48 – GRÁFICO TUBO SOLO X TEMPO: PT08, CASO 01	87
FIGURA 49 – GRÁFICO DISPERSÃO: PT08, CASO 01	87
FIGURA 50 – GRÁFICO HISTOGRAMA: PT08, CASO 01, METODOLOGIA 01.....	88
FIGURA 51 – GRÁFICO HISTOGRAMA: PT08, CASO 01, METODOLOGIA 02.....	88
FIGURA 52 – GRÁFICO TUBO SOLO X TEMPO: PT09, CASO 01	89
FIGURA 53 – GRÁFICO DISPERSÃO: PT09, CASO 01	89
FIGURA 54 – GRÁFICO HISTOGRAMA: PT09, CASO 01, METODOLOGIA 01.....	90
FIGURA 55 – GRÁFICO HISTOGRAMA: PT09, CASO 01, METODOLOGIA 02.....	90
FIGURA 56 – ERRO ABSOLUTO MÉDIO PARA OS RESULTADOS DO CASO 1	91
FIGURA 57 – ERRO PERCENTUAL MÉDIO PARA OS RESULTADOS DO CASO 1	91
FIGURA 58 – COMPARAÇÃO DA FUNÇÃO R ² PARA OS RESULTADOS DO CASO 1	92
FIGURA 59 – COMPARAÇÃO DO DESVIO PADRÃO PARA OS RESULTADOS DO CASO 1.....	92
FIGURA 60 – GRÁFICO TUBO SOLO X TEMPO: PT01, CASO 02	94

FIGURA 61 – GRÁFICO DISPERSÃO: PT01, CASO 02	94
FIGURA 62 – GRÁFICO HISTOGRAMA: PT01, CASO 02, METODOLOGIA 01	95
FIGURA 63 – GRÁFICO HISTOGRAMA: PT01, CASO 02, METODOLOGIA 02	95
FIGURA 64 – GRÁFICO TUBO SOLO X TEMPO: PT02, CASO 02	96
FIGURA 65 – GRÁFICO DISPERSÃO: PT02, CASO 02	96
FIGURA 66 – GRÁFICO HISTOGRAMA: PT02, CASO 02, METODOLOGIA 01	97
FIGURA 67 – GRÁFICO HISTOGRAMA: PT02, CASO 02, METODOLOGIA 02	97
FIGURA 68 – GRÁFICO TUBO SOLO X TEMPO: PT03, CASO 02	98
FIGURA 69 – GRÁFICO DISPERSÃO: PT03, CASO 02	98
FIGURA 70 – GRÁFICO HISTOGRAMA: PT03, CASO 02, METODOLOGIA 01	99
FIGURA 71 – GRÁFICO HISTOGRAMA: PT03, CASO 02, METODOLOGIA 02	99
FIGURA 72 – GRÁFICO TUBO SOLO X TEMPO: PT04, CASO 02	100
FIGURA 73 – GRÁFICO DISPERSÃO: PT04, CASO 02	100
FIGURA 74 – GRÁFICO HISTOGRAMA: PT04, CASO 02, METODOLOGIA 01	101
FIGURA 75 – GRÁFICO HISTOGRAMA: PT04, CASO 02, METODOLOGIA 02	101
FIGURA 76 – GRÁFICO TUBO SOLO X TEMPO: PT05, CASO 02	102
FIGURA 77 – GRÁFICO DISPERSÃO: PT05, CASO 02	102
FIGURA 78 – GRÁFICO HISTOGRAMA: PT05, CASO 02, METODOLOGIA 01	103
FIGURA 79 – GRÁFICO HISTOGRAMA: PT05, CASO 02, METODOLOGIA 02	103
FIGURA 80 – ERRO ABSOLUTO MÉDIO PARA OS RESULTADOS DO CASO 2	104
FIGURA 81 – ERRO PERCENTUAL MÉDIO PARA OS RESULTADOS DO CASO 2	104
FIGURA 82 – COMPARAÇÃO DA FUNÇÃO R^2 PARA OS RESULTADOS DO CASO 2	105
FIGURA 83 – COMPARAÇÃO DO DESVIO PADRÃO PARA OS RESULTADOS DO CASO 2	105
FIGURA 84 – HISTOGRAMA DO ERRO EM CONJUNTO COM A CURVA DA DISTRIBUIÇÃO NORMAL – PT01	107
FIGURA 85 – VARIÂNCIA ESTIMADA PARA O PT01, CASO 02	108
FIGURA 86 – HISTOGRAMA DO ERRO EM CONJUNTO COM A CURVA DA DISTRIBUIÇÃO NORMAL – PT02	109
FIGURA 87 – VARIÂNCIA ESTIMADA PARA O PT02, CASO 02	109
FIGURA 88 – HISTOGRAMA DO ERRO EM CONJUNTO COM A CURVA DA DISTRIBUIÇÃO NORMAL – PT03	110
FIGURA 89 – VARIÂNCIA ESTIMADA PARA O PT02, CASO 02	111

FIGURA 90 – HISTOGRAMA DO ERRO EM CONJUNTO COM A CURVA DA DISTRIBUIÇÃO NORMAL – PT04	112
FIGURA 91 – VARIÂNCIA ESTIMADA PARA O PT02, CASO 02	113
FIGURA 92 – HISTOGRAMA DO ERRO EM CONJUNTO COM A CURVA DA DISTRIBUIÇÃO NORMAL – PT05	114
FIGURA 93 – VARIÂNCIA ESTIMADA PARA O PT02, CASO 02	114

Lista de Tabelas

TABELA 1 - GRAU DE CORROSÃO DO SOLO EM FUNÇÃO DA RESISTIVIDADE ELÉTRICA.	28
TABELA 2 - POTENCIAIS PADRÃO DE REDUÇÃO [16].	29
TABELA 3 - PERIODICIDADE MÍNIMA DE MEDIÇÕES E VERIFICAÇÕES DE ROTINA [18].	42
TABELA 4 – NÍVEIS DE CONFIANÇA PARA VALORES DE A	58
TABELA 5 – RESULTADOS E VARIAÇÃO DO NÚMERO DE NEURÓNIOS APLICADOS NA METODOLOGIA 01	67
TABELA 6 – RESULTADOS E VARIAÇÃO DO NÚMERO DE NEURÓNIOS APLICADOS NA METODOLOGIA 02	68
TABELA 7 – RESULTADOS GERAIS OBTIDOS NAS METODOLOGIAS 01 E 02 PARA O CASO 01	71
TABELA 8 – RESULTADOS CONSIDERANDO APENAS OS ATRIBUTOS “PONTO DE TESTE”	72
TABELA 9 – RESULTADOS GERAIS OBTIDOS NAS METODOLOGIAS 01 E 02 PARA O CASO 02	93
TABELA 10 – RESULTADOS CONSIDERANDO APENAS OS ATRIBUTOS “PONTO DE TESTE”	93
TABELA 11 – RESULTADOS MÉDIOS GERAIS PARA O CASO 01 BASEADO NAS METODOLOGIAS .	106
TABELA 12 – RESULTADOS MÉDIOS GERAIS PARA O CASO 02 BASEADO NAS METODOLOGIAS .	106
TABELA 13 – RESULTADOS MÉDIOS GERAIS PARA O PT01	107
TABELA 14 – RESULTADOS MÉDIOS GERAIS PARA O PT02.....	108
TABELA 15 – RESULTADOS MÉDIOS GERAIS PARA O PT03.....	110
TABELA 16 – RESULTADOS MÉDIOS GERAIS PARA O PT04.....	111
TABELA 17 – RESULTADOS MÉDIOS GERAIS PARA O PT05.....	113

Lista de Abreviaturas

PT	Ponto de Teste
RNA	Rede Neural Artificial
MLP	<i>Multilayer Perceptron</i>
CA	Corrente Alternada
CC	Corrente Contínua
Li	Lítio
K	Potássio
Ca	Cálcio
Na	Sódio
La	Lantânio
Mg	Magnésio
Pu	Plutônio
Th	Tório
Be	Berílio
U	Urânio
Al	Alumínio
Ti	Titânio
Zr	Zircônio
Mn	Manganês
Nb	Nióbio
Cu	Cobre
O	Oxigênio
H	Hidrogênio
S	Enxofre
BEN	Balanco Energético Nacional

OIE	Oferta Interna de Energia
AM	Aprendizado de Máquina
ABNT	Associação Brasileira de Normas Técnicas
ISO	<i>International Organization for Standardization</i>
IBP	Instituto Brasileiro de Petróleo
NBR	Norma Brasileira

Capítulo 1 - Introdução

A constante busca de ferramentas para melhoria de processos industriais gera confiabilidade normalmente atrelada a uma diminuição do custo, tornando os resultados com mais clareza de interpretação, embasando possíveis ações estratégicas com decisões mais assertivas.

A melhora nos processos de manutenção e operação a partir dos dados que são obtidos das inspeções preventivas e preditivas por vezes são perdidos ou mal interpretados, pois a gestão deste volume de dados normalmente é realizada por profissionais que possuem várias outras atribuições e, alguns detalhes podem passar despercebidos até pelos olhos mais hábeis.

O tratamento do volume de dados de um determinado processo pode ser realizado de forma automática caso haja uma plataforma que trabalhe a informação disponível no banco de dados e passe os resultados automaticamente para o sistema de gerenciamento da manutenção que irá gerar as ordens corretivas necessárias ou não, bem como dar volume de informações e embasar as ações futuras.

Dentro deste contexto, o modal de transporte dutoviário de gás natural possui diversos processos que podem ser melhorados a partir das premissas apresentadas anteriormente.

O setor de transporte e distribuição de gás natural no Brasil passa por um processo de estagnação de investimentos e envelhecimento da malha existente, portanto, é natural o aumento nas incidências de manutenções corretivas, normalmente atreladas ao desgaste dos ativos. Outras circunstâncias como avarias por terceiros, furto e vandalismo também ocorrem, aumentando a incidência de correção.

O produto gás natural é uma mistura de hidrocarbonetos leves entre os quais se destaca o metano, que tem origem fóssil e procedente da decomposição da matéria orgânica espalhada entre os extratos rochosos. O gás natural é um produto incolor e inodoro, não tóxico e é mais leve do que o ar. Para a sua comercialização, o gás tem um odor adicionado a ele, por questões de segurança, pois se trata de um combustível. O gás natural é uma energia carente de enxofre e a sua combustão é completa, liberando como produtos o dióxido de carbono e vapor de água, o que faz do gás natural uma energia ecológica e não poluente [1].

De 2008 a 2018 o consumo de gás natural no Brasil cresceu 158% saindo de 2,89 para 4,57mil bilhões de m³/ano [2]. O BEN 2018 destaca o aumento da oferta do gás natural no Brasil que cresceu 6,7% em relação ao ano anterior, representando 13% da oferta interna de energia (OIE) [3].

No Brasil, o modal rodoviário é preponderante na matriz agregada de transporte, sendo responsável por movimentar 67,4% do total de volume de carga no país, contra apenas 3% do modal dutoviário (outras modalidades de transporte correspondem a 29,6% da matriz agregada). Entretanto, essa proporção não se mantém no setor de petróleo e gás, no qual o transporte dutoviário é considerado estratégico em diversas etapas da cadeia logística, por ser mais competitivo que os outros modais. Essa competitividade se estabelece à medida que existe a necessidade de transporte contínuo de grandes volumes de fluidos, de um ponto de oferta a um ponto de demanda, com os dutos entregando muita confiabilidade por não se sujeitarem a incertezas meteorológicas ou congestionamentos. Os grandes volumes transportados e os baixos custos variáveis característicos do modal compensam os altos investimentos necessários para a construção e implantação de uma malha de dutos, como equipamentos, mão de obra, desapropriações e direito de acesso às terras [4].

Para este trabalho, foi escolhido um processo de manutenção de empresas de transporte e distribuição de gás natural. Este processo consiste no combate a corrosão eletroquímica através da proteção catódica nos dutos de aço de transporte e distribuição de gás natural.

A Corrosão é a degradação de um material devido à reação com o meio ambiente, de natureza química ou eletroquímica, podendo ou não estar associado a esforços mecânicos.

A corrosão eletroquímica é a mais frequente na natureza e se caracteriza por:

- Presença de água no estado líquido;
- Temperaturas abaixo do ponto de orvalho da água;
- Formação de uma pilha ou célula de corrosão, com a circulação de elétrons pela superfície metálica.

A proteção catódica pode ser definida como abrandamento ou prevenção total de todas as formas de corrosão em estruturas metálicas imersas em um eletrólito [5]. A rotina de inspeção dos elementos de proteção catódica segue uma periodicidade de acordo com a importância destes para o sistema. Diversas medições são coletadas pelo inspetor de campo, e, a partir delas, ações são tomadas após a análise dos dados.

Eventualmente, dados faltantes da inspeção de campo ou do sistema de monitoramento tornam a análise de um determinado sistema de proteção catódica incompleta.

Diante deste problema e visando solucioná-lo, o aprendizado de máquina, mais especificamente o uso de redes neurais, mostra-se uma ferramenta poderosa para estimativa dos dados faltantes em um determinado período, tendo como base um sólido banco de dados, aumentando a confiabilidade do sistema e do setor de manutenção.

O aprendizado de Máquina (AM) é uma técnica que gera um modelo computacional da capacidade de aprender a partir da interação com seu ambiente visando melhorar o seu desempenho na execução de uma determinada tarefa [6]. Ainda que o AM seja uma ferramenta poderosa para a aquisição automática de conhecimento, deve ser observado que não existe um único algoritmo que apresente o melhor desempenho para todos os problemas. Portanto, é importante compreender o poder e a limitação dos diversos algoritmos de AM utilizando alguma metodologia que permita avaliar os conceitos induzidos por esses algoritmos em determinados problemas [7].

Uma forma de aprendizagem que pode ser aplicada em AM é a aprendizagem indutiva. A indução é a forma de inferência lógica que permite obter conclusões genéricas sobre um conjunto particular de exemplos, ou seja, um conceito é aprendido efetuando-se inferência indutiva sobre os exemplos apresentados, portanto, as hipóteses geradas através da inferência indutiva podem ou não preservar a verdade [7].

O aprendizado indutivo é efetuado a partir do raciocínio sobre exemplos fornecidos por um processo externo ao sistema de aprendizado. O aprendizado indutivo pode ser dividido em supervisionado e não supervisionado. No aprendizado supervisionado é fornecido ao algoritmo de aprendizado, ou indutor, um conjunto de exemplos de treinamento para os quais o rótulo da classe associada é conhecido, ou seja, devem existir dados de entrada com suas respectivas saídas, para serem apresentadas ao algoritmo de aprendizado utilizado durante o processo de treinamento [8]. O objetivo do algoritmo de indução é construir um tipo de classificador que possa determinar corretamente a classe de novos exemplos ainda não rotulados. Para rótulos de classe discretos, esse problema é conhecido como classificação e para valores contínuos como regressão.

1.1 Motivação

A necessidade das informações dos potenciais eletroquímicos hora não obtidos pelos métodos convencionais de inspeção do sistema de proteção catódica foi a principal motivação deste trabalho. Para uma interpretação completa do sistema de proteção catódica, todos os dados devem ser coletados, o que eventualmente não ocorre para determinados sistemas.

As estruturas metálicas protegidas catódicamente estão normalmente dispostas nos subsolos de grandes cidades, sendo assim, estão sujeitas a fatores não controlados, como o crescimento desordenado, ocupação irregular de espaços públicos, obras não informadas por outras concessionárias de serviços públicos que também utilizam o subsolo, falta de segurança pública, dentre diversos outros fatores que possivelmente irão dificultar a inspeção dos elementos de proteção catódica.

O investimento em ferramentas para coleta de dados remotos vem aumentando no mundo inteiro, inclusive no Brasil. Para o seguimento da proteção catódica, no Brasil, ainda não há uma empresa referência com uma solução remota eficiente. Ao importar equipamentos de países cuja cultura e clima são opostas à do Brasil, sem um estudo prévio, acaba tornando o investimento falho em relação ao objetivo inicial do projeto. Sem “tropicalizar” estes equipamentos levando em consideração a incidência elevada de raios, humidade relativa do ar e calor excessivo além da situação político-econômica do país cujo resultado culmina em ações tais como vandalismo, furto ou mesmo a baixa qualificação dos profissionais técnicos que trabalham em campo torna o avanço tecnológico em retrocesso, deste modo, informações hora obtidas são perdidas ou analisadas de forma inadequada gerando ações que não efetivamente resolvam o problema sinalizado ou detectado.

Ao entender que problemas relacionados à corrosão eletroquímica podem estar sendo mascarados devido à falta de informações de campo, o que pode estar colocando em risco a população local, meio ambiente e a imagem da empresa, é o primeiro passo para o desenvolvimento de ferramentas que auxiliem na interpretação dos dados recebidos e na tomada de decisão.

Existem diversas aplicações para o uso de RNA's na engenharia, entretanto, seu uso para estimar dados faltantes de um SPC ainda não foi explorado. Portanto, o uso de

dois sistemas distintos, ambos baseados em RNA's, demonstra que não há apenas uma maneira de fazer. Na verdade, existem diversas possibilidades de metodologias que podem ser exploradas com o uso de RNA's.

1.2 Objetivos

O objetivo principal desta dissertação é, a partir das bases de dados de dois sistemas da proteção catódica de gasodutos de distribuição de gás natural distintos, criar uma rede neural que, através de duas metodologias baseadas em RNA, possa estimar o dado faltante de um dos elementos do sistema em um determinado instante. Após a análise dos resultados obtidos, uma das metodologias será a escolhida a partir da comparação do erro associado. Parte das bases de dados não foram utilizadas no treinamento com o propósito de testar os resultados e embasar a escolha da metodologia mais adequada aos objetivos deste trabalho.

1.3 Objetivos Específicos do Trabalho

Os objetivos específicos deste trabalho são:

- Abordar o cenário contemporâneo do transporte e distribuição do gás natural;
- Discorrer sobre a problemática da corrosão e contextualizar com os métodos de combate e controle atualmente utilizados;
- Discorrer sobre o Aprendizado de Máquina, com ênfase em RNA;
- Abordar sobre o uso de RNA para estimativa de dados faltantes;
- Propor duas metodologias baseadas em RNA para atender ao objetivo principal desta dissertação;
- Demonstrar os resultados gerados no software MATLAB através das duas metodologias propostas;
- Apresentar os erros entre a saída real e a saída gerada para cada metodologia;
- Escolher uma das metodologias baseado no menor erro encontrado no passo anterior;
- Propor trabalhos futuros.

1.4 Publicação decorrente da pesquisa

No âmbito desta dissertação foi publicado e apresentado o seguinte artigo técnico nos anais do evento INTERCORR 2018

INTERCORR2018_231. Rede neural aplicada em sistemas de proteção catódica -Estudo de caso. Rodrigo M. S. Soares, Vitor H. Ferreira, Márcio Z. Fortes. In: INTERCORR, 2018, São Paulo.

1.5 Estrutura do Documento

A presente dissertação está dividida em 7 capítulos. O Capítulo 1, introdutório, foi apresentado o contexto em que o trabalho está envolvido, suas principais motivações e objetivos.

- O Capítulo 2, "CORROSÃO E MÉTODOS DE MITIGAÇÃO", apresenta os tipos de corrosão e os métodos de combate e controle, com ênfase nas tubulações metálicas enterradas.
- O Capítulo 3, "APRENDIZADO DE MÁQUINA", apresenta os conceitos, tipos de aprendizado com ênfase em redes neurais, *perceptrons* de múltiplas camadas e Inferência Bayesiana.
- O Capítulo 4, "METODOLOGIA", apresenta os critérios adotados, bases de dados e os métodos utilizados para a realização das simulações no ambiente MATLAB.
- O Capítulo 5, "RESULTADOS", apresenta os resultados obtidos a partir das simulações realizadas, com base nos dois bancos de dados e nas duas metodologias propostas.
- O capítulo 6 "CONCLUSÕES E TRABALHOS FUTUROS", apresenta as conclusões desta dissertação, com a escolha da metodologia que obteve maior aderência a proposta geral, evidenciando a importância e contribuição das análises aqui realizadas. Também será discutido sobre os trabalhos que possam ser realizados em complemento ao proposto neste trabalho.

Capítulo 2 - Corrosão e Métodos de Mitigação

2.1 Corrosão

A corrosão pode ser definida como a deterioração das propriedades físicas de um metal pela reação eletroquímica com o meio eletrolítico que o envolve, em outras palavras, a corrosão é a deterioração que ocorre quando um material reage com seu ambiente. Em uma perspectiva termodinâmica, a tendência de decréscimo energético é a principal força da corrosão metálica [9]. Existem alguns autores que englobam além do metal, materiais não metálicos como a madeira, o concreto e material polimérico como materiais que sofrem com o fenômeno da corrosão [10].

Quando um produto entra em processo de corrosão, a energia gasta para sua elaboração se perde, a matéria prima utilizada na sua fabricação, retorna ao meio, porém não de forma concentrada como encontrada inicialmente. A figura 1 exemplifica este processo, tomando como exemplo uma tubulação metálica.

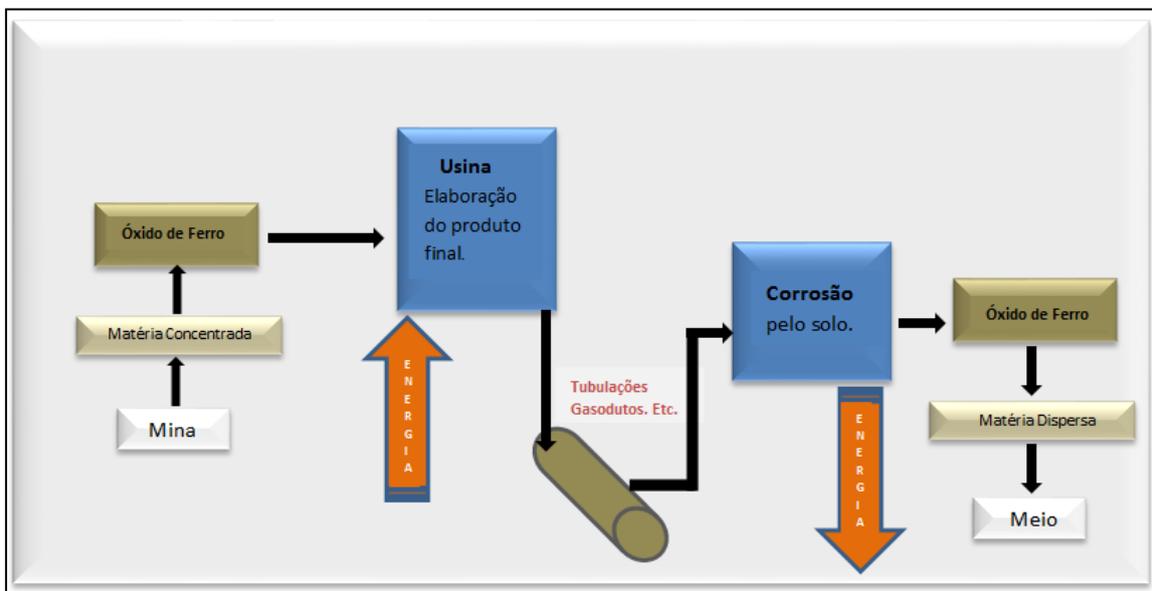


Figura 1– Ciclo da corrosão [10]

Uma reação é considerada eletroquímica se ela estiver associada a uma passagem de corrente elétrica através de uma distância finita, maior do que a distância interatômica. Esta passagem de corrente envolve o movimento de partículas carregadas: íons, elétrons

ou ambos. Dessa forma, na maioria das reações, que se manifestam em presença de uma superfície metálica, ocorre esta passagem de corrente através do metal, e a reação é eletroquímica em sua natureza. Como na corrosão de um metal a superfície está sempre presente, as reações básicas responsáveis pela corrosão são eletroquímicas[11].

A corrosão eletroquímica ocorre apenas na presença de um eletrólito. Este tipo de fenômeno tem como base duas reações parciais que são:

Reação anódica – É a reação de oxidação, na qual os elétrons são liberados e se destinam a outra parte do metal, ocasionando a sua dissolução e consequente corrosão.

Reação Catódica – É a parte do metal que recebe os elétrons liberados na reação de oxidação dando início a reação de redução. Este mecanismo pode ser exemplificado como uma pilha de corrosão e tem como principais componentes o anodo, o catodo, o meio eletrolítico e uma ligação metálica que une ambas as áreas anódicas e catódicas.

Em uma tubulação metálica enterrada, a região corroída é chamada de anodo, e a região protegida e chamada de catodo, tendo o solo como o meio eletrolítico onde as reações de oxidação e redução acontecem. Na região anódica, os átomos carregados positivamente deixam esta região e fluem, através do metal, até a região catódica (Figura 2) onde são utilizados na reação química da redução.

Pode ser tomado como exemplo um íon de hidrogênio próximo à região catódica que interage com um elétron que flui através do metal no sentido do anodo para o catodo, formando assim um átomo de hidrogênio $H^+ + e \rightarrow H$. Quando aparecem dois átomos de hidrogênio, estes podem interagir entre si, formando o gás hidrogênio H_2 . Já na área anódica com a perda do elétron o ferro de Fe^{3+} passa para o meio eletrolítico com Fe^{2+} e pode interagir com duas hidroxilas $2OH^-$, formando a reação de oxidação $Fe^{++} + 2OH^- \rightarrow Fe(OH)_2$ chamada de hidróxido ferroso, o qual em meio aerado, esse composto é oxidado a hidróxido de ferro $Fe(OH)_3$, que perde água e se transforma no óxido de ferro monohidratado $Fe_2O_3 \cdot H_2O$ que evolui até a formação da ferrugem $2Fe(OH)_2 + H_2O + \frac{1}{2} O_2 \rightarrow 2Fe(OH)_3$ [12][13].

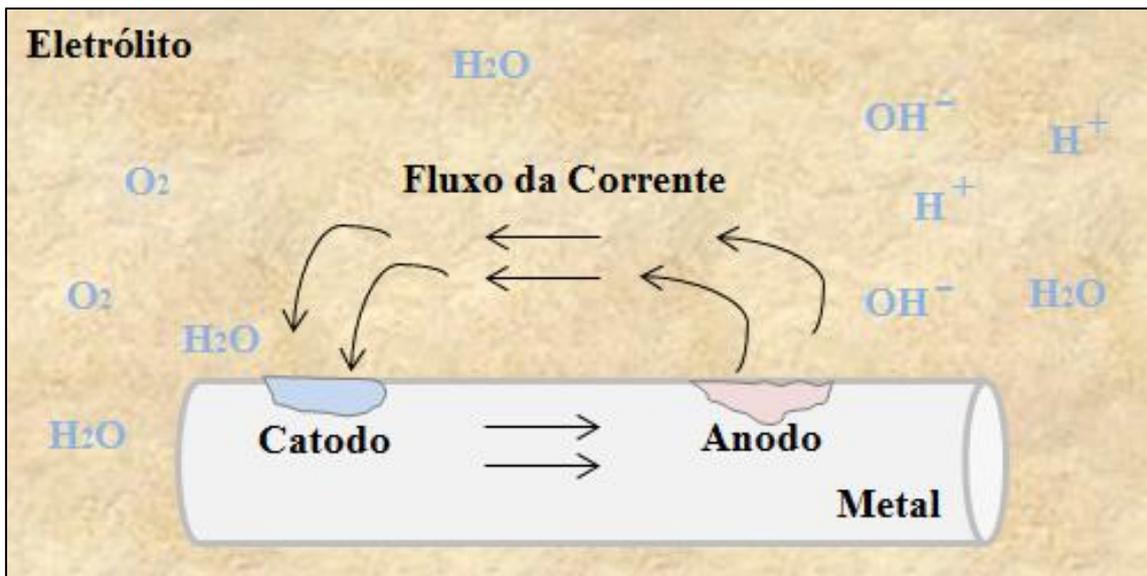


Figura 2– Célula e corrosão simples [10]

Existem duas principais vertentes para que ocorra a corrosão de um metal, a primeira está diretamente relacionada ao tipo do eletrólito (neste caso, o solo) e a segunda com as heterogeneidades do metal. Tomando como exemplo a Figura 2, o lado onde ocorre a oxidação (área anódica) pode ser caracterizado pela baixa resistividade do solo (eletrólito), se comparado com a área onde ocorre a redução (área catódica), assim caracterizando o primeiro caso. Já no segundo caso, falhas no revestimento do metal, ou mesmo heterogeneidades do metal podem fazer com que ocorra este fenômeno.

A corrosão de metais pelo solo pode ocorrer principalmente pelo fator físico-químico e biológico do solo, ou por fatores como pilhas de longo alcance, pilhas de concentração e correntes de fuga. Em grandes gasodutos, ocorrem de maneira mais comum as chamadas pilhas de longo alcance, pois estas estruturas normalmente têm grandes extensões, passando por solos de diversos níveis de resistividade. Já a pilha de concentração ocorre quando estas estruturas entram em contato com solos de diferentes teores de oxigênio, ou de diferentes teores de sais.

Alguns índices são determinantes no que se diz respeito à agressividade do solo. O teor de umidade, a composição química do solo e o pH são algumas das características mais comuns para definir este problema. Outros fatores também são utilizados como a permeabilidade do solo, a presença de bactérias e a presença de poluentes, porém o mais usual é determinar o índice de agressividade somente pela resistividade do solo em questão. Quanto maior o valor de sua resistividade menos agressivo é o solo, pois dificulta a circulação da corrente elétrica entre o anodo e o catodo.

A tabela 1 propõe uma avaliação da corrosividade do solo através da relação entre a resistividade elétrica no solo em que se situa e seu grau de agressividade. Observa-se que quanto menor a resistividade do solo, mais agressivo é o meio. Analogamente, quanto maior a resistividade do solo, menos agressivo é o meio [14].

Tabela 1 - Grau de corrosão do solo em função da resistividade elétrica.

Grau de Agressividade	Resistividade (Ω.cm)
Extremamente agressivo	< 1000
Fortemente agressivo	1000 a 2000
Moderadamente agressivo	2000 a 3500
Pouco agressivo	3500 a 5000
Ligeiramente agressivo	5000 a 10000
Não agressivo	> 10000

2.1.1 Polarização

A polarização é a variação do potencial causada por uma corrente elétrica em um eletrodo. A polarização pode ser anódica ou catódica. A polarização anódica é obtida através da variação de potencial no sentido positivo e, se esta variação for ao sentido negativo, temos a polarização catódica. A polarização pode ser obtida por três modos, polarização por ativação, polarização por concentração e polarização ôhmica. A primeira ocorre pelo retardamento das reações ou de fases das reações, a segunda ocorre pela concentração dos reagentes causando também retardamento das reações, já a última é consequência da resistência elétrica causada pela presença de uma película de algum tipo de produto localizado na superfície do eletrodo [13].

O fenômeno da polarização se faz muito importante para a proteção catódica, já que a essência desta consiste na polarização da estrutura a proteger.

2.1.2 Potencial do eletrodo

Quando um metal entra em contato com um eletrólito, uma diferença de potencial chamada de potencial do eletrodo aparece. Esta diferença de potencial pode ser positiva, negativa ou nula. Isto ocorre devido à presença de cargas elétricas de um sinal, no eletrólito (solo), e cargas negativas do sinal oposto no metal, e o valor desta diferença varia muito de acordo com o metal, fixando as variáveis ligadas ao eletrólito (Figura 3).

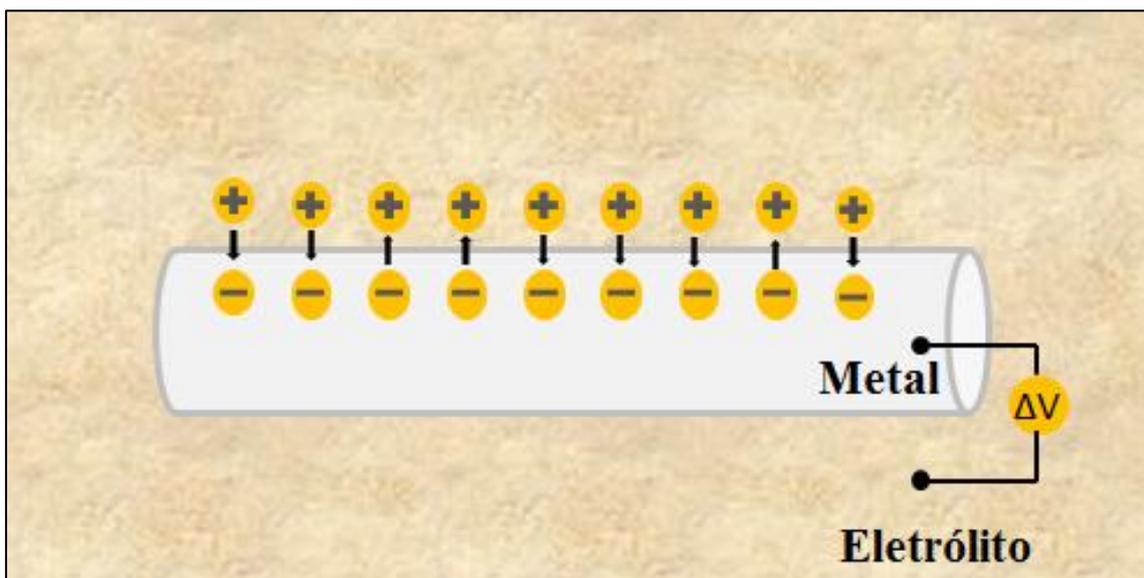


Figura 3 – Potencial do eletrodo

O conjunto metal eletrólito também é conhecido por meia pilha, meia célula, semicélula ou eletrodo.

Quando os metais reagem têm tendência a perder elétrons, sofrendo oxidação e, conseqüentemente corrosão. Os metais apresentam distintas tendências à oxidação e, portanto, interessante para a previsão de alguns problemas de corrosão ter em mãos uma tabela que indique a ordem preferencial de oxidação dos metais. A Tabela 2 apresenta os potenciais padrões de redução para alguns metais, ordenados em relação a um eletrodo de referência [16].

Tabela 2 - Potenciais padrão de redução [16].

Reação do Eletrodo	Potenciais de Eletrodo Padrão a 25°C (Volts)	Reação do Eletrodo	Potenciais de Eletrodo Padrão a 25°C (Volts)
$Li^+ + e^- \rightarrow Li$	- 3,045	$Zr^{2+} + 2e^- \rightarrow Zr$	- 0,763
$K^+ + e^- \rightarrow K$	- 2,925	$Cr^{3+} + 3e^- \rightarrow Cr$	- 0,740
$Ca^{2+} + 2e^- \rightarrow Ca$	- 2,870	$Fe^{2+} + 2e^- \rightarrow Fe$	- 0,440

$Na^+ + e^- \rightarrow Na$	- 2,714	$Cd^{2+} + 2e^- \rightarrow Cd$	- 0,403
$La^{3+} + 3e^- \rightarrow La$	- 2,520	$In^{3+} + 3e^- \rightarrow In$	- 0,342
$Mg^{2+} + 2e^- \rightarrow Mg$	- 2,370	$Co^{2+} + 2e^- \rightarrow Co$	- 0,277
$Pu^{3+} + 3e^- \rightarrow Pu$	-2,070	$Ni^{2+} + 2e^- \rightarrow Ni$	- 0,250
$Th^{4+} + 4e^- \rightarrow Th$	-1,090	$Mo^{3+} + 3e^- \rightarrow Mo$	- 0,200
$Be^{2+} + 2e^- \rightarrow Be$	-1,850	$Sn^{2+} + 2e^- \rightarrow Sn$	- 0,136
$U^{3+} + 3e^- \rightarrow U$	- 1,800	$Pb^{2+} + 2e^- \rightarrow Pb$	- 0,126
$Al^{3+} + 3e^- \rightarrow Al$	- 1,660	$2H^+ + 2e^- \rightarrow H_2$	0,000
$Ti^{2+} + 2e^- \rightarrow Ti$	- 1,630	$Cu^{2+} + 2e^- \rightarrow Cu$	0,337
$Zr^{4+} + 4e^- \rightarrow Zr$	- 1,530	$Ag^+ + e^- \rightarrow Ag$	0,799
$Mn^{3+} + 3e^- \rightarrow Mn$	- 1,180	$Pd^{2+} + 2e^- \rightarrow Pd$	0,987
$Nb^{3+} + 3e^- \rightarrow Nb$	- 1,100	$Au^{3+} + 3e^- \rightarrow Au$	1,500

2.1.3 Passividade

Na presença de um eletrólito, metais podem corroer ativamente, entretanto, quando estes metais estão com certo valor de potencial, passam a apresentar uma corrente anódica reduzida, que corresponde a uma corrosão pequena ou desprezível. Estes metais e ligas normalmente ativos perdem sua reatividade química e tornam-se extremamente inertes. Este fenômeno, conhecido como passivação, ocorre no cromo, ferro, níquel, titânio e muitas de suas ligas. Diz-se, então que o metal está passivo. Este comportamento pode ser observado na curva de polarização anódica. Este fenômeno da passivação pode ser explicado em termos da curva de potencial de polarização contra a densidade de corrente. A curva de polarização para um metal que passiva terá uma forma geral apresentada na Figura 4.

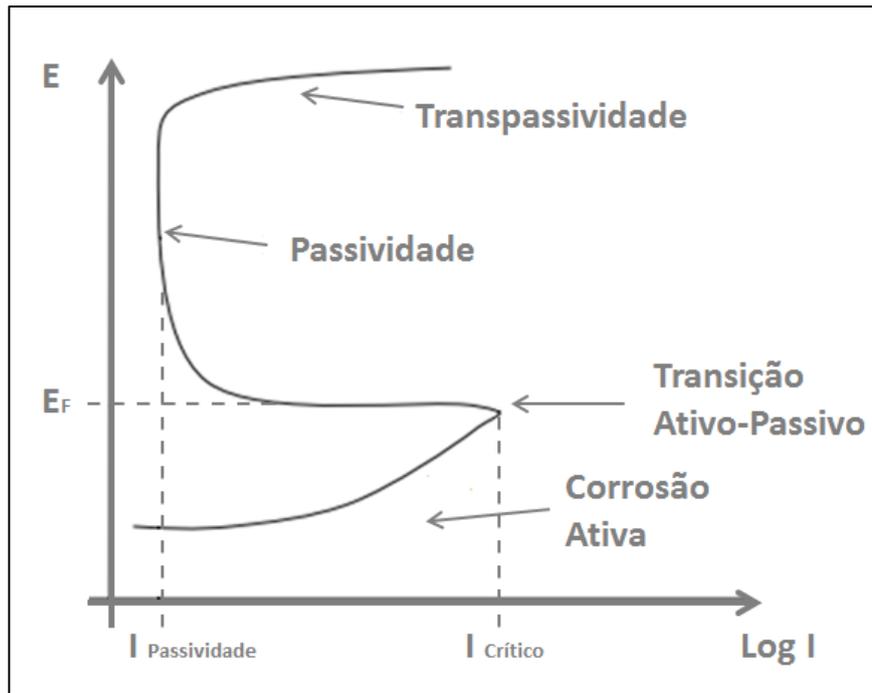


Figura 4 – Curva de polarização anódica [15]

Com o aumento do potencial, a densidade de corrente diminui subitamente para um valor muito baixo que permanece independentemente da variação do potencial. Esta região é denominada de região passiva. Finalmente, para altos valores de potencial, a densidade de corrente novamente aumenta de forma súbita caracterizando uma região denominada de transpassiva [16].

2.1.4 Diagrama de Pourbaix

A partir da variação da energia livre de Gibbs, podem ser calculados de forma teórica os potenciais de eletrodo dos metais envolvidos, a partir da expressão[17]:

$$\Delta G = -nFE \quad (1)$$

Onde:

- n = Número de elétrons envolvidos na reação;
- F = Constante de Faraday;
- E = Potencial de eletrodo (V).

Marcel Pourbaix descobriu a existência de relações entre o potencial do eletrodo e o pH das soluções, para os sistemas quando em equilíbrio. Destas relações surgiu o diagrama de Pourbaix[17] (Figura 5).

Tomando como base a água, suas características químicas podem ser separadas em H^+ e OH^- , ou seja, reduzidos evoluindo H_2 ou oxidados evoluindo O_2 , desta forma obtem-se duas retas, considerando as condições de equilíbrio:

- Reta “a”: Com a relação $H_2 \rightleftharpoons 2H^+ + 2e$, o potencial pode ser obtido de acordo com a equação de Nernst[17]:

$$E = E_0 + \left(\frac{RT}{2F}\right) \times 2,3 \log\left(\frac{[H^+]^2}{PH_2}\right) \quad (2)$$

Onde:

E = Potencial;

E_0 = Potencial padrão do H_2 que é zero;

T = Temperatura em K;

R = Constante dos gases perfeitos;

F = Constante de Faraday.

É obtida para concentrações diferentes a equação $E = 0 - 0,0591 \text{ pH}$, que representa a reta “a”.

- Reta “b”: Com a relação $H_2O + \frac{1}{2} O_2 + 2e \rightleftharpoons 2 OH^-$, pode ser obtido a equação (3)[17]:

$$E = +0,401 + \left(\frac{0,0591}{2}\right) \log\left(\frac{[PO]^{\frac{1}{2}}}{[OH^-]^2}\right) \quad (3)$$

Tendo a pressão do oxigênio igual a 1 e considerando que $-\log [OH^-] = 14$, a reta “b” pode ser definida por $E = 1,229 - 0,0591 \text{ pH}$.

Com este raciocínio Pourbaix conseguiu traçar curvas das condições de equilíbrio, definindo o domínio de estabilidade termodinâmica de vários elementos químicos, inclusive de metais. É importante enfatizar que para problemas práticos, o diagrama de Pourbaix não é confiável já que necessita de temperatura e pressão padrão.

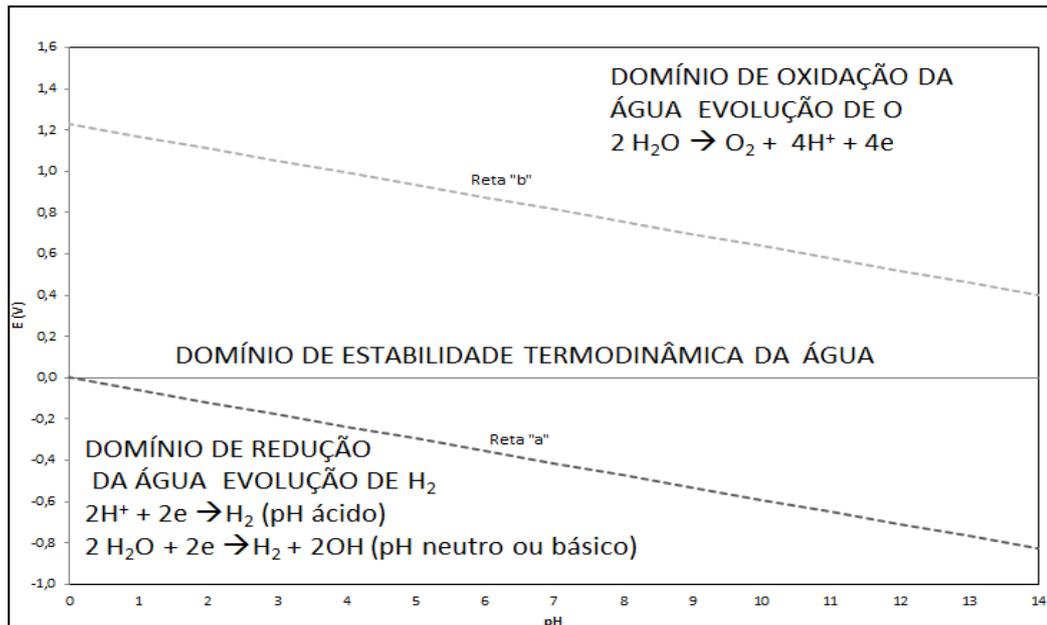


Figura 5 – Diagrama de pourbaix [17]

2.2 Métodos para inibição da corrosão

2.2.1 Critérios para combate e controle da corrosão

Para garantir que a proteção catódica esteja sendo aplicada de maneira correta em todas as áreas dos dutos envolvidos, deve ser construído ao longo deste, pontos de monitoração de potencial respeitando os seguintes critérios:

- 1) Para obter uma boa proteção, a taxa de corrosão do metal deverá ser inferior a 0,01 mm por ano.
- 2) É adotado como valor máximo de proteção, o potencial de -0,85 V com relação ao eletrodo de referência de Cu/CuSO₄, ou seja, $E_p = -0,85 \text{ V}$.
- 3) É adotado como valor mínimo de proteção, o potencial de -1,20 V com relação ao eletrodo de referência de Cu/CuSO₄, ou seja $E_c = -1,2 \text{ V}$.

O potencial metal-eletrólito (E) deverá ser menor ou igual ao potencial de proteção (E_p) e que por sua vez, deverá ser maior ou igual que o potencial crítico (E_c). Este último se deve ao fenômeno da fragilização por hidrogênio provocado por potenciais muito

baixos, tendo como ônus o descolamento do revestimento protetor do metal e/ou um alto pH na superfície do metal [17]. Valores entre o potencial de proteção e o potencial crítico são considerados ideais para a proteção das tubulações de gás enterradas [18].

$$E_c \leq E \leq E_p \quad (4)$$

O critério utilizado pela companhia de gás na adequação do valor do potencial de proteção tubo-solo é o mesmo utilizado na norma brasileira e americana [19] e [20]. É adotado como valor mínimo de proteção, o potencial de -0,85V com relação ao eletrodo de referência de Cu/CuSO₄, ou seja, $E_p = -0,85V$. Já o potencial crítico, de acordo com a norma, o potencial tubo-solo não poderá ser mais negativo do que $E_c = -1,2V$. Qualquer valor entre o potencial de proteção e o potencial crítico é considerado como ideal para proteção das tubulações de gás enterradas [18].

O critério E_c normalmente é superado (superproteção), com valores mais negativos ao critério apresentado, principalmente em locais próximos aos pontos de injeção de corrente de proteção catódica e pontos próximos a fontes de interferência elétrica, temas que serão abordados mais a frente nesta dissertação. A fragilização do aço pelo H₂ e descolamento do revestimento anti-corrosivo são os principais problemas relacionados à superproteção.

Para solos anaeróbicos, ou quando houver bactérias redutoras de sulfato, o potencial de proteção tubo-solo passa para $E_p = -0,95V$. Já para os dutos que operam em solos com a resistividade muito alta o valor do potencial de proteção poderá ser alterado. Para solos com a resistividade entre 100 e 1000 $\Omega.m$, o potencial de proteção poderá ser de $E_p = -0,75V$, acima de 1000 $\Omega.m$, o potencial de proteção então poderá ser de $E_p = -0,65V$.

Existe outra medida padrão que pode ser utilizada como alternativa dos potenciais vistos neste subtópico. Pode ser utilizada como critério uma polarização catódica mínima de 0,100 mV entre a superfície do duto e um eletrodo de referência em contato com o eletrólito. Este critério não deverá ser utilizado em temperatura alta de operação, em solos com bactérias redutoras de sulfato, ou com correntes de interferência, correntes de equalização e correntes telúricas[18].

2.2.2 Revestimento de dutos enterrados

Neste tópico serão comentados sobre os tipos de revestimentos utilizados em dutos de aço enterrados.

Se houvesse um revestimento totalmente perfeito, isto é, um ótimo isolante elétrico, que ao ser aplicado na tubulação sem conter quaisquer tipos de falhas e conseguisse continuar intacto durante toda a vida útil da estrutura, os problemas com a corrosão não existiriam. Na realidade o revestimento aplicado possui micro falhas ou até falhas visíveis contraídas com o deslocamento, instalação dos dutos e com a degradação natural.

O principal objetivo do revestimento é fazer uma barreira entre o metal e o eletrólito. Para a proteção de dutos enterrados é utilizada o revestimento tipo orgânico, porém existem outros dois tipos de revestimentos que são os inorgânicos e os metálicos.

Os revestimentos orgânicos utilizados no passado e mais utilizados atualmente para tubulações de gás enterradas são:

- Alcatrão de piche e asfalto: No início do século XX o uso de alcatrão de piche e asfalto para revestimento de tubulações era a opção de menor custo e maior disponibilidade no mercado, sendo assim, largamente utilizado nesta época sem a preocupação de que esta seria uma boa opção de revestimento para proteção anticorrosiva da tubulação.
- Revestimento com esmalte de alcatrão e hulha (coal-tar): Este tipo de revestimento está em desuso. Pode ser utilizado de duas formas, com o chamado revestimento simples ou revestimento duplo. O revestimento simples é aplicado com espessura de 3 a 5mm. Normalmente este tipo de revestimento é utilizado em locais de baixa ou média agressividade de solo. Já o revestimento duplo é aplicado de 6 a 8mm, e é normalmente utilizado em solos agressivos ou com correntes de interferência. Este tipo de revestimento é encontrado em várias tubulações, porém sua resistividade elétrica decai sensivelmente com o tempo, existe um alto índice de falhas decorrentes de manuseio e instalação dos dutos e pode causar sérios riscos ao meio ambiente devido ao desprendimento de gases tóxicos.

- Revestimento com polietileno extrudado (PE3L): Possui excelente resistência a corrosão, excelente resistência a danos mecânicos e é impermeável. É aplicado com espessura de 3 a 5 mm. Este tipo de revestimento é mais eficiente do que os revestimentos vistos anteriormente e é o mais utilizado em novas tubulações de gás enterradas desta companhia de gás.
- Revestimento com tinta epóxi em pó (*FusionBondedEpoxi*): Este tipo de revestimento possui grande eficiência anticorrosiva. Além de ser impermeável, possui uma película de 400 μm com elevada resistência, dando mais flexibilidade no manuseio e instalação.

Os principais requisitos de um revestimento são: baixa permeabilidade, resistência química ao meio agressivo, dilatação térmica compatível com o substrato, propriedades físicas adequadas aos abusos que receberá por abrasão, tráfego, impacto, flexão, etc. Suas características são: monolítico (sem emendas), remota ocorrência de trincas ou fissuras, não permite infiltrações, fácil e rápida aplicação, aceita reparos localizados, equipamentos de suporte simples e baixo custo[20].

2.2.3 Proteção Catódica

A proteção dos materiais metálicos pode ser realizada através da aplicação de corrente anódica ou impedindo a difusão de oxigênio através do processo catódico. Em 1954, Endeleanu sugeriu a possibilidade do emprego da proteção catódica que, de maneira simples, baseia-se na formação de uma película protetora, nos materiais metálicos, por aplicação de corrente anódica externa. Essa corrente ocasiona a polarização anódica, que possibilita a passivação do material metálico[21].

Como visto no item 2.1, o fenômeno da corrosão eletroquímica é o principal fator para que exista a proteção catódica. A corrosão eletroquímica sempre ocorre quando há uma estrutura metálica em contato com o eletrólito, fazendo com que pilhas de corrosão apareçam. Estas podem ocorrer pelas características do metal e/ou pelas características inerentes ao meio.

Um processo artificial pode fazer com que ambas as áreas anódicas e catódicas se transformem em áreas apenas catódicas, eliminando dali a reação de oxidação e consequente perda das propriedades físicas e químicas do metal. Os metais que apresentam a maior diferença de potencial com relação aos metais a serem protegidos e que apresentam características de polarização anódica baixa, são os escolhidos para se tornarem a parte anódica do sistema de proteção catódica.

A proteção catódica pode ocorrer de duas formas, ou através de um sistema galvânico, empregando anodos de sacrifício, ou através do sistema de corrente impressa, utilizando retificadores para ajudar na composição do sistema de proteção catódica.

2.2.3.1 Proteção catódica galvânica

Este tipo de proteção só pode ser considerado nos casos de dutos de pequeno diâmetro, ou em pequenos trechos de dutos com diâmetros maiores, principalmente em solos de baixa resistividade ou onde não houver possibilidade de alimentação elétrica. Este sistema se mostra ineficaz quanto existem correntes de fuga.

Neste tipo de proteção, a corrente elétrica surge através da força eletromotriz existente entre o metal a proteger e o metal que servirá como o anodo do sistema, esta corrente é o resultado de seus diferentes potenciais eletroquímicos. Quando ocorre a ligação entre o anodo e a estrutura, estando ambos inseridos em um eletrólito, a corrente resultante ocorre a partir da reação apresentada em (5)[13]:



Existem dois tipos de anodos mais utilizados neste sistema que são o anodo de liga de zinco e o anodo de liga de magnésio. Para ilustrar este procedimento, segue na Figura 6, o sistema de proteção catódica por anodo de sacrifício.

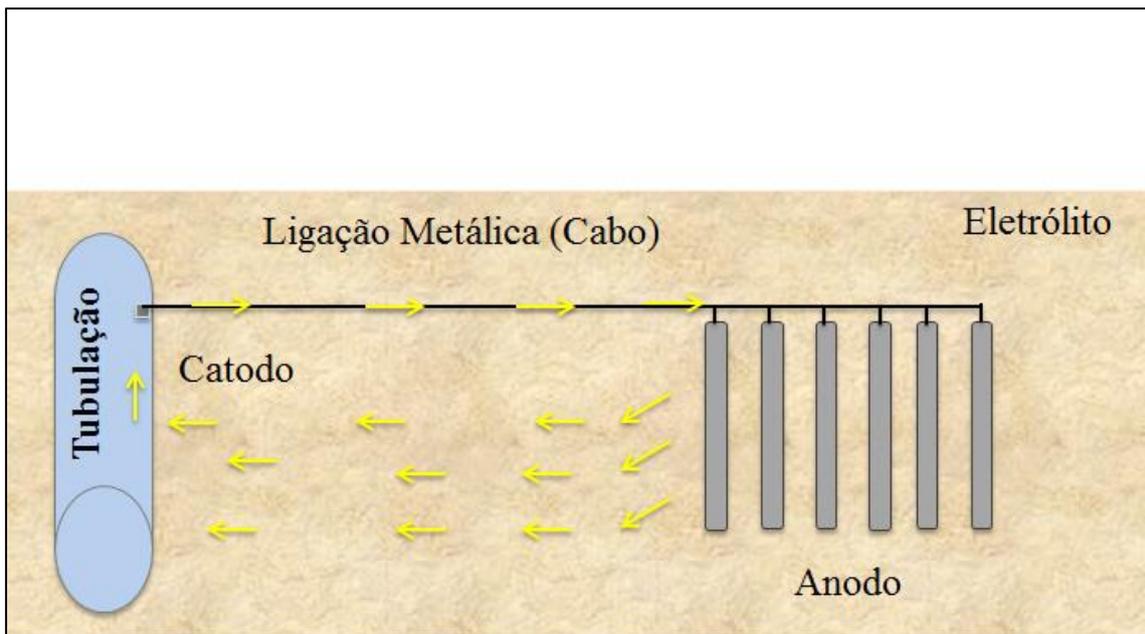


Figura 6 – Sistema de proteção catódica por anodo de sacrifício

2.2.3.2 Proteção catódica por corrente impressa

Para pensar em proteção catódica por corrente impressa, primeiro deverá ser feito um estudo levando em consideração a disponibilidade de alimentação elétrica no local, levantar a resistividade do solo na área onde ficará o leito de anodos, verificar se existirá bom acesso aos equipamentos de proteção, se existem dutos enterrados que poderão ser afetados, se existem fontes de correntes de interferência, entre outros, que deverão ser levados em consideração quando se diz respeito à proteção catódica por corrente impressa.

Neste sistema, a corrente elétrica é proveniente de uma fonte geradora de corrente contínua. É usual a utilização de retificadores específicos para este processo, que utiliza os anodos, denominados leito de anodos, como dispersor de corrente no eletrólito. Este sistema também é o mais indicado quando se têm grandes estruturas a proteger, em solos de baixa, média, alta ou altíssima resistividade. Como estes sistemas normalmente são alimentados pela concessionária de energia elétrica, estão mais sujeitos a interrupções no fornecimento.

Em geral, os anodos mais utilizados na proteção catódica por corrente impressa são o anodo de ferro-silício, de ferro-silício-cromo, e o de titânio tubular revestido com óxido de metais nobres e magnetita.

É importante ressaltar que o polo positivo da fonte deverá ser ligado ao leito de anodos e o polo negativo ligado à estrutura a se proteger conforme figura 7. Se estas polaridades forem invertidas a tubulação passará a ser o anodo do sistema, o que acelerará o processo corrosivo deste ativo.

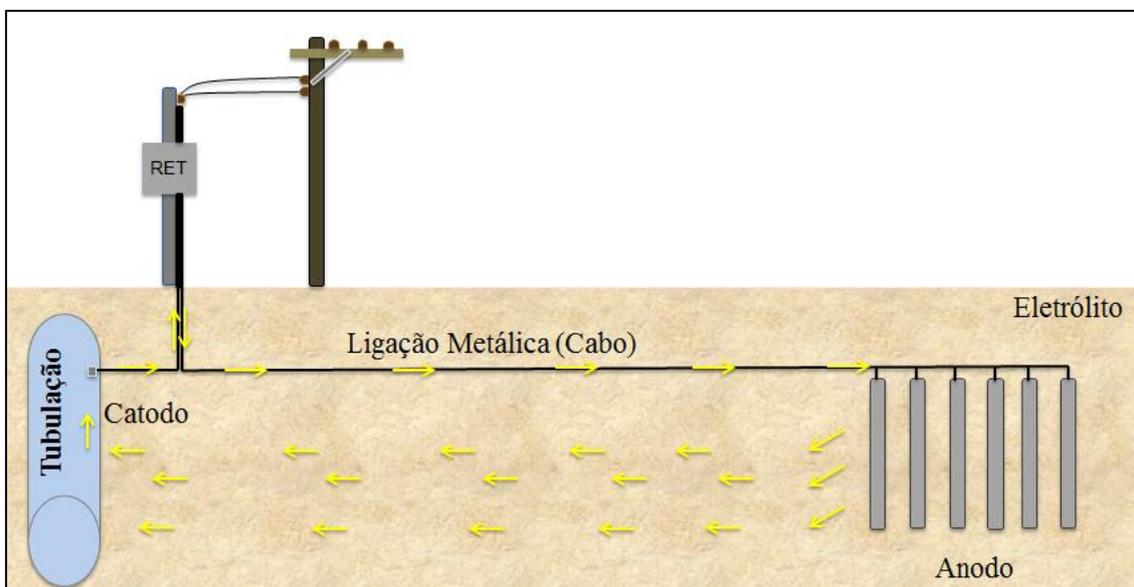


Figura 7 – Sistema de proteção catódica por anodo de sacrifício

2.2.4 Pontos de medição de potencial

Para garantir que a proteção catódica esteja sendo aplicada de maneira correta em todas as áreas dos dutos envolvidos, deve ser construído ao longo deste, pontos de monitoração de potencial respeitando os critérios mencionados no item 2.2.1. Os pontos de medição de potencial podem ser divididos de acordo com a necessidade de medição de potencial na tubulação. O número de pontos de teste em um duto varia de acordo com as características do local onde o mesmo for instalado. Na área urbana, é pedido para que se tenha no mínimo um ponto de teste por quilômetro, e na área rural um ponto de teste a cada três quilômetros. Para obter o potencial duto-eletrólito, é utilizado um eletrodo de referência (ER) de Cu/CuSO_4 posicionado na geratriz superior do duto, com sua base porosa em perfeito contato com o solo. Este eletrodo estará conectado ao polo negativo de um multímetro digital, tendo o lado positivo conectado a um cabo que está soldado na parede externa do duto.

Uma caixa é instalada de modo a acomodar o cabo soldado a tubulação e deve ser previsto um espaço dentro da caixa com a terra do local de modo que se possa encostar a ponta porosa do eletrodo de referência (ER) no eletrólito. Estes pontos ficam exatamente em cima da tubulação. Em locais de difícil acesso, é instalado um moirão de concreto de modo a sinalizar a existência da tubulação e do cabo elétrico soldado. Na Figura 8, pode ser observado o esquema de medição do potencial eletroquímico.

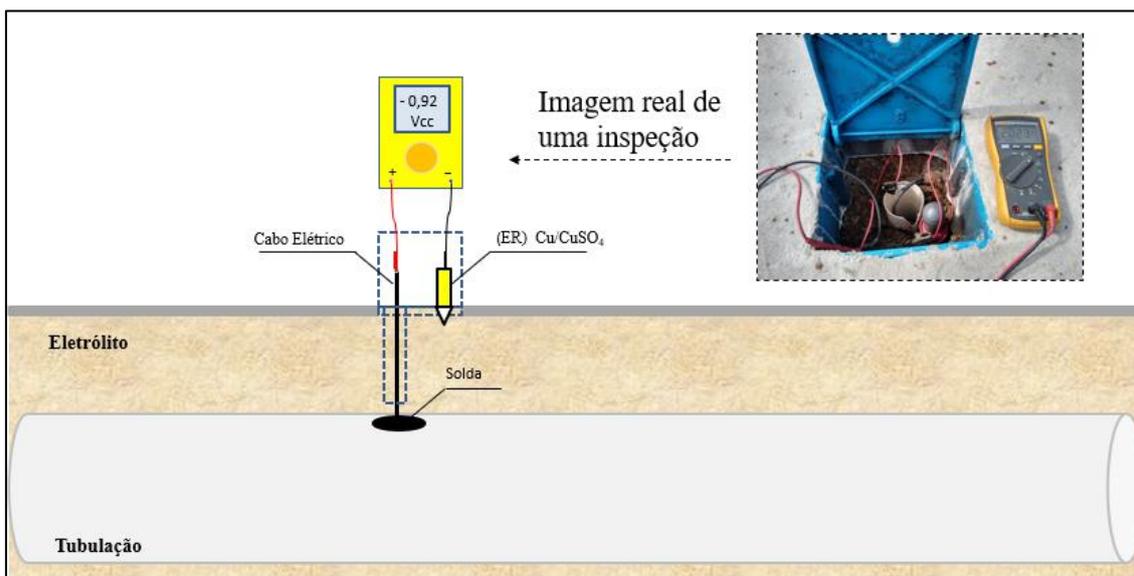


Figura 8 – Sistema de proteção catódica por corrente impressa

2.2.5 Interferência elétrica

A corrosão por correntes indesejáveis em dutos metálicos se difere das demais causas da corrosão pelo fato destas correntes serem provenientes de uma fonte externa ao duto e normalmente não existe uma continuidade elétrica. Esta corrente é captada pelo duto através do solo, e quando retorna a este circuito original, ocorre a corrosão. Estas correntes indesejáveis podem ser provenientes de uma fonte de corrente contínua, que é a mais comum, ou por corrente alternada.

As correntes de fuga provenientes de sistemas c.c. podem causar corrosão em estruturas metálicas enterradas. Esta corrosão, conhecida como eletrolítica, é particularmente crítica, uma vez que seus efeitos podem aparecer em um curto espaço de tempo. Por esta razão, é fundamental identificar o problema desde o projeto e, em

sistemas existentes, implantar medidas mitigadoras com rapidez, verificando regularmente a efetividade das mesmas [22].

Uma interferência de sistemas c.a. pode ocorrer em função de acoplamentos resistivo, indutivo ou capacitivo, em regime permanente ou durante um curto-circuito. São três os principais aspectos que devem ser observados nos casos de aproximações de linhas de transmissão com estruturas metálicas de terceiros: a segurança pessoal, a integridade física das estruturas e a corrosão [23].

As técnicas de mitigação das interferências elétricas estão disponíveis nas normas NBR 16563-1 - Mitigação de efeitos de interferências elétricas em sistemas dutoviários. Parte 1: Sistemas de corrente alternada [23] e NBR 16563-2 - Mitigação de efeitos de interferências elétricas em sistemas dutoviários. Parte 2: Sistemas de corrente contínua [24]. Para o estudo de caso deste trabalho, não será necessária a apresentação das técnicas de mitigação das interferências elétricas.

2.2.6 Inspeção do sistema de proteção catódica

O sistema de proteção catódica deverá ser inspecionado, regularmente, com intervalos definidos pela importância dos locais ou pelos equipamentos a serem inspecionados. A importância da inspeção se dá, pois quando os retificadores saem de operação ou estão fora de um ajuste ótimo, o processo de corrosão poderá iniciar.

O sistema de proteção galvânica normalmente não precisa de inspeção com muita frequência, pois se constituem de anodos soldados na estrutura a ser protegida, e a possibilidade de falha no funcionamento é bem menor que o sistema por corrente impressa, além do que, normalmente este tipo de técnica é utilizada em estruturas que não estão sujeitas a correntes de interferência.

O sistema de corrente impressa e o sistema de drenagem de corrente deverão ter uma periodicidade maior de inspeção, pois os dois se constituem de equipamentos eletroeletrônicos, que estão sujeitos a falhas de funcionamento. O primeiro sistema está sujeito a problemas causados por falhas nos seus componentes, a problemas na distribuição de energia elétrica por conta da concessionária, interrupção causada pelos dispositivos de proteção, falhas nas conexões dos anodos, etc. Já ao segundo, deverá ser atribuída uma atenção especial, pois estes equipamentos tem grande importância no sistema de proteção catódica quanto à correntes indesejáveis.

A periodicidade das inspeções varia de acordo com a normativa interna das empresas, com a ressalva da periodicidade mínima prevista na norma brasileira [19] que se encontra na tabela 3.

Tabela 3 - Periodicidade mínima de medições e verificações de rotina [18].

Item	Ação	Periodicidade
Estação de corrente impressa	Verificação visual do retificador e leitura da tensão e corrente de saída.	Um a três meses
Estação de corrente impressa	Ensaio funcionas abrangentes (por exemplo, verificação do retificador, controle do eletrodo permanente, determinação da resistência do leito de anodos, medição do sistema de aterramento, controle de instrumentos) e medição da tensão e corrente de saída.	Um a três anos
Estações de drenagem unidirecional	Verificação visual da drenagem unidirecional e leitura da instrumentação de proteção catódica integral	Um mês
Estações de drenagem unidirecional	Ensaio funcionais abrangentes (por exemplo, controle do eletrodo permanente, funcionalidade dos diodos e seus dispositivos de proteção, ajuste de resistores, controle de instrumentos) e medição do potencial e corrente do ponto de retorno de corrente.	Um a três anos
Conexões com estruturas externas (ligação resistiva ou direta).	Medição do fluxo de corrente	Um ano
Conexões com estruturas externas (ligação resistiva ou direta).	Ensaio funcionais abrangentes no dispositivo e medição do fluxo de corrente (magnitude e sentido) e do potencial	Um a três anos

Sistemas de aterramento conectados ao duto e relacionados ao sistema de proteção catódica ou sistema de mitigação de corrosão	Ensaio funcionais abrangentes dos dispositivos e determinação da resistência de aterramento e do fluxo de corrente através do sistema de mitigação (se houver).	Um a três anos
Eletrodo de referência permanente	Comparação com um eletrodo de referência cuja precisão pode ser rastreada até um eletrodo mestre	Um a seis anos ^b
Pontos de teste selecionados	Medição de potencial E_{ON}	Um ano
Todos os pontos de teste	Medição do potencial E_{OFF} ^a	Três anos
Estação de anodos galvânicos	Verificação visual das estações e medição do potencial tubo-eletrólito	Um ano
Estação de anodos galvânicos	Ensaio funcionais abrangentes da estação galvânica (por exemplo, ajuste do resistor, eficiência da ligação e medição do potencial tubo-eletrólito)	Três anos
<p>a Quando correntes de interferência influenciam nas medições de potencial E_{OFF}, podem ser consideradas técnicas de medição alternativas se forem comprovadamente precisas e eficazes.</p> <p>b O desempenho e a Repetibilidade do eletrodo de referencia permanente dependem do tipo de eletrodo de referencia e da frequência das medições. A periodicidade, portanto, pode variar de um a seis anos.</p>		

Capítulo 3 - Redes Neurais

Este capítulo apresenta a base teórica para o desenvolvimento de duas metodologias baseadas em RNA: Uma metodologia baseada em *perceptrons* de múltiplas camadas (MLP) e a outra metodologia baseada em redes bayesianas. Objetiva-se aplicar as metodologias no tratamento dos dados vinculados a dois sistemas de proteção catódica de gasodutos distintos, de modo a estimar, com a maior qualidade de predição, um dado faltante em um determinado momento.

Uma rede neural é um processador maciçamente e paralelamente distribuído constituído de unidades de processamento simples, que tem a capacidade natural para depositar conhecimento experimental e torná-lo disponível para aplicação. A semelhança com o cérebro pode ser verificada a partir de dois aspectos:

- 1) O conhecimento é adquirido pela rede a partir de seu ambiente através de um processo de aprendizagem;
- 2) Forças de conexão entre neurônios, conhecidas como pesos sinápticos, são utilizadas para armazenar o conhecimento adquirido.

O procedimento utilizado para realizar o processo de aprendizagem é chamado de algoritmo de aprendizagem, cuja função é modificar os pesos sinápticos da rede de uma forma ordenada para alcançar um objetivo de projeto desejado.

A modificação dos pesos sinápticos é o método tradicional para o projeto de redes neurais. Esta abordagem é bastante próxima da teoria dos filtros adaptativos lineares, que já está bem estabelecida e foi aplicada com sucesso em diversas áreas, entretanto, é possível também para uma rede neural modificar sua própria topologia, o que é motivado pelo fato de os neurônios no cérebro humano poderem morrer e que novas conexões sinápticas possam crescer [25].

A figura 9 demonstra o neurônio através de um diagrama esquemático. A saída c é dada pela equação 7 [27]:

$$c = \phi \left(\sum_{i=1}^n \omega_i x_i + b \right) \quad (6)$$

- c representa a saída do neurônio ;
- ω o conjunto de pesos sinápticos: $\underline{\omega} \in \mathbb{R}^n, \underline{\omega} = [\omega_1 \dots \omega_n]^t$;

- x o conjunto de entradas: $\underline{x} \in \mathbb{R}^n, \underline{x} = [x_1 \dots x_n]^t$;
- b o *bias* associado: $b \in \mathbb{R}$;
- ϕ função ativação: $\phi(\cdot) : \mathbb{R} \rightarrow \mathbb{R}$

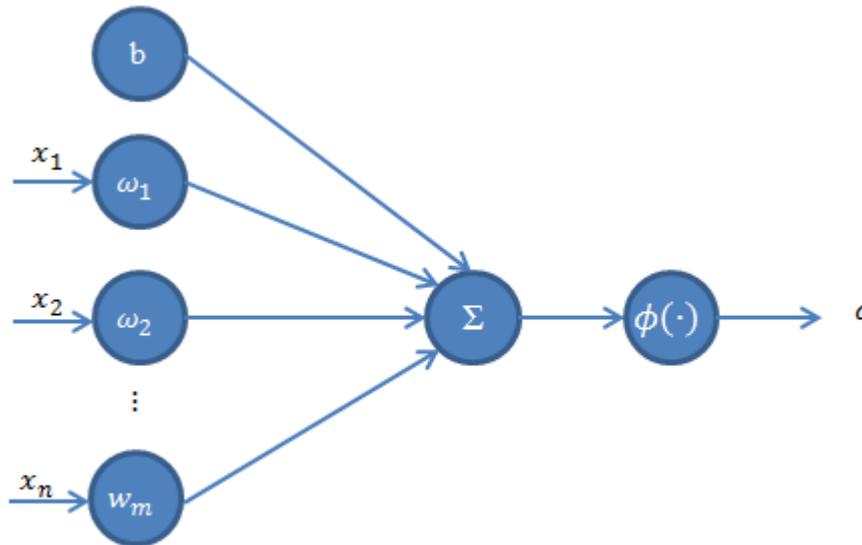


Figura 9 – Diagrama esquemático de um neurônio - adaptado de [27]

3.1 Redes *perceptrons* de Múltiplas Camadas (MLP)

Se um único neurônio existe em uma rede, ele é geralmente chamado de *perceptron*. Os *perceptrons* encontram um hiperplano de separação linear e a prova pode ser dada para mostrar que convergirá para uma solução, se ela existir [29].

RNA's com várias camadas são muito mais difíceis de treinar. Na verdade, esse foi um grande obstáculo no desenvolvimento de RNA até que um algoritmo eficiente de treinamento fosse desenvolvido. Esse algoritmo é conhecido como o algoritmo de retropropagação, assim chamado porque os erros que são a força motriz no treinamento (se não houver erro, não há necessidade de alterar os pesos do RNA) são propagados a partir da camada de saída, através de as camadas ocultas, para a camada de entrada. Esse algoritmo permite que a rede seja treinada de acordo com a necessidade. Avanços adicionais foram atingidos através de métodos de segunda ordem para treinamento, que alcançaram uma convergência mais rápida [29].

O sinal de erro na saída do neurônio j , na iteração n , é definido por [26]:

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

O valor instantâneo da energia do erro para o neurônio j é definido por:

$$\frac{1}{2} e_j^2(n) \quad (8)$$

O valor instantâneo total do erro $E(n)$ é obtido somando-se os termos da energia do erro de todos os neurônios da camada de saída. Desse modo, os sinais de erro podem ser calculados diretamente:

$$E(n) = \frac{1}{2} \sum_{j \in C} e_j^2(n) \quad (9)$$

Onde o conjunto C inclui todos os neurônios da camada de saída da rede. O número total de exemplos contidos no conjunto de treinamento será representado por N . O erro quadrado médio é obtido somando-se os $E(n)$ para todos os n e normalizado pelo tamanho do conjunto N :

$$E_{med} = \frac{1}{N} \sum_{n=1}^N E(n) \quad (10)$$

O objetivo do processo de aprendizagem é ajustar os parâmetros livres da rede para minimizar o E_{med} . A medida que os pesos são atualizados de padrão em padrão formará uma época. A época é uma representação completa do conjunto de treinamento inteiro que está sendo processado.

Para superar as limitações práticas do *perceptron*, analisa-se para uma estrutura de rede neural conhecida como *perceptron* multicamadas. Os três pontos seguintes destacam as características básicas dos MLPs:

- O modelo de cada neurônio na rede inclui uma função de ativação não linear que é diferenciável.
- A rede contém uma ou mais camadas ocultas dos nós de entrada e saída.
- A rede exibe um alto grau de conectividade, cuja extensão é determinada pelos pesos sinápticos da rede.

Essas mesmas características, no entanto, também são responsáveis pelas deficiências em nosso conhecimento sobre o comportamento da rede. Em primeiro lugar, a presença de uma forma distribuída de não linearidade e a alta conectividade da rede dificulta a análise teórica de um *perceptron* multicamadas. Em segundo lugar, o uso de neurônios ocultos dificulta a visualização do processo de aprendizado. Em um sentido

implícito, o processo de aprendizagem deve decidir quais características do padrão de entrada devem ser representadas pelos neurônios ocultos. O processo de aprendizagem torna-se, portanto, mais difícil, porque a busca deve ser conduzida em um espaço muito maior de funções possíveis, e uma escolha deve ser feita entre representações alternativas do padrão de entrada [28].

Análogo a biologia, os neurônios constituintes das RNA's são dispostos em camadas, e a arquitetura com que as camadas são interligadas define o modelo. De forma geral, existem duas arquiteturas: As redes alimentadas adiante (*feedforward*) e as redes recorrentes. A Figura 16 mostra o grafo arquitetural de um modelo de redes “alimentadas adiante” comumente conhecidas como *feedforward* com duas camadas ocultas e uma camada de saída. Observa-se que o *perceptron* de múltiplas camadas na sua forma geral, é totalmente conectado. Um neurônio em qualquer camada da rede está conectado a todos os nós/neurônios da camada anterior [26]. As camadas são conectadas de forma consecutiva e adjacente, com o sinal fluindo da entrada para a saída em sentido único da esquerda para a direita e de camada em camada. As redes recorrentes apresentam um ou mais laços de realimentação, baseado na mesma estrutura da figura 10.

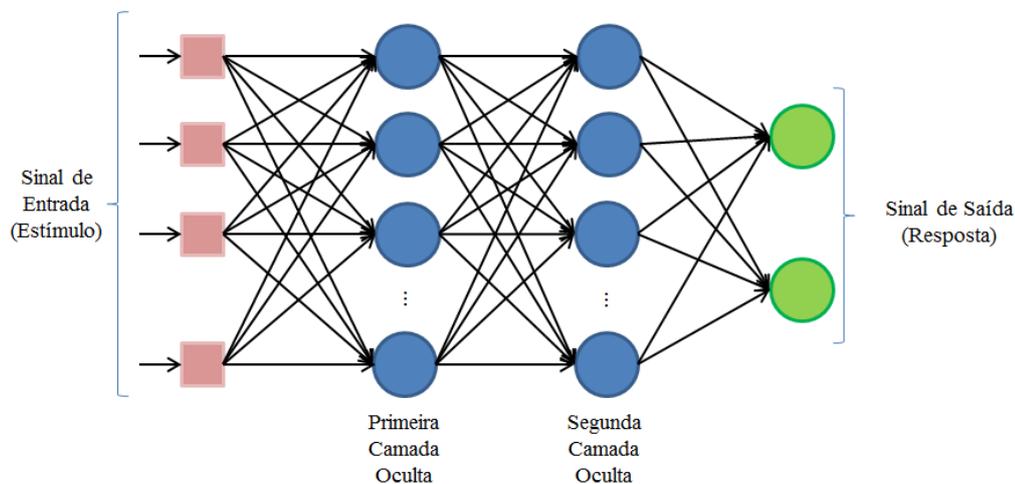


Figura 10 – Grafo arquitetural de um perceptron de múltiplas camadas com duas camadas ocultas - adaptado de [26]

Para este tipo de rede, dois sinais são identificados (Figura 11):

Sinais de função: sinal de entrada (estímulo) que chega na entrada, propaga-se para frente (neurônio por neurônio) através da rede e emerge na saída da rede como um sinal de saída. Também conhecido como sinais de entrada [30].

Sinais de erro: origina em um neurônio de saída da rede e propaga-se de volta (camada por camada) através da rede [30].

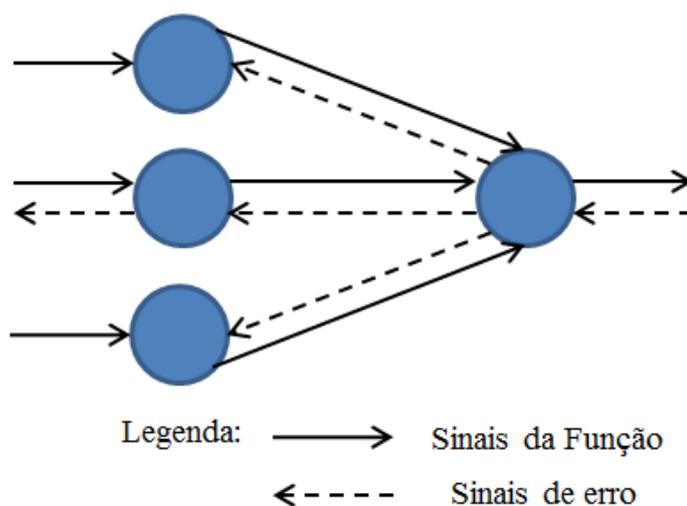


Figura 11 – Direções dos fluxos de sinal básico de um perceptron de múltiplas camadas - adaptado de [26]

A primeira camada oculta é nutrida pelo sinal de entrada, através das unidades sensoriais (nós de fonte). As saídas resultantes da primeira camada oculta são aplicadas à próxima camada oculta e assim por diante para o resto da rede. Cada neurônio escondido ou de saída de um *perceptron* de múltiplas camadas realiza duas computações:

- Sinal de função que aparece na saída de um neurônio, expresso como uma função não-linear contínua do sinal de entrada e pesos sinápticos associados.
- Estimativa do vetor gradiente (isto é, os gradientes da superfície de erro com respeito aos pesos conectados às entradas de um neurônio), necessária para o passo para trás através da rede (retropropagação) [30].

Segundo o teorema da aproximação universal, modelos de retropropagação podem aproximar com precisão arbitrária qualquer função contínua $F(\underline{x}): \mathbb{R}^n \rightarrow \mathbb{R}$. Deste modo, sua estrutura deve apresentar ao menos uma camada oculta contendo neurônios com função de ativação contínua, não-constante, limitada, e uma saída linear, representando a aproximação de $F(\underline{x})$ gerada pelo modelo. Portanto, estes modelos com uma única camada escondida contendo um número suficiente de neurônios com função de ativação com as características apresentadas anteriormente podem aproximar qualquer função contínua [27].

A aproximação universal do teorema é importante do ponto de vista teórico porque fornece a ferramenta matemática necessária para a viabilidade de redes *feedforward* com uma única camada oculta como uma classe de soluções aproximadas. Sem esse teorema, poderíamos estar à procura de uma solução que não pode existir.

Apesar da vantajosa característica da aproximação universal, o objetivo do desenvolvimento de uma máquina de aprendizagem não habita na representação exata do conjunto de dados disponíveis, mas sim na obtenção de um modelo estatístico do processo gerador de tais dados [31]. Portanto, é desejado que o modelo apresente resultados satisfatórios tanto para os dados disponíveis quanto para novos dados a serem apresentados.

A saída y deste tipo de modelo, contendo m neurônios na camada oculta, é dado por:

$$\mathbf{y} = \sum_{j=1}^m \omega_j \phi \left(\sum_{j=1}^n \omega_j x_j + b_j \right) + b \quad (11)$$

Os pesos que ligam os neurônios da camada oculta ao neurônio linear de saída são representados por $\underline{\omega}_s \in \mathbb{R}^m$, $\underline{\omega}_s = [\omega_1, \omega_2 \dots \omega_m]^t$ e $\underline{\omega}_j \in \mathbb{R}^n$, $\underline{\omega}_j = [\omega_1, \omega_2 \dots \omega_n]^t$. Estes são formados pelos pesos que unem as entradas ao j -ésimo neurônio da camada escondida. O bias do neurônio da camada escondida é dado por $b_j \in \mathbb{R}$ e o bias do neurônio de saída é dado por $b \in \mathbb{R}$ [27].

Dentre as principais vantagens dos modelos neurais, podem ser citadas as seguintes [28]:

- Não-linearidade: para o caso em que a função de ativação dos neurônios da camada oculta é não-linear, o modelo neural resultante da interconexão destas unidades mais simples apresenta considerável grau de não-linearidade. Essa característica pode ser prejudicial no caso de dados ruidosos. Evitar o ajuste excessivo dos dados de treinamento, também conhecido como *overfitting*, constitui uma das principais tarefas no desenvolvimento de modelos neurais.
- Mapeamento entrada-saída: a partir de um conjunto de pares entrada-saída, as redes neurais realizam um mapeamento destes dados, sem a necessidade do desenvolvimento de modelos matemáticos específicos.
- Adaptabilidade: estes modelos apresentam elevada capacidade de adaptação em virtude de mudanças das premissas do ambiente para o qual

o modelo foi treinado. Deste modo, basta treinar novamente a rede, incluindo o novo conjunto de dados.

- Implementação simples: os modelos *feedforward* podem ser vistos como grafor orientados e a implementação destas estruturas é extremamente simples, quando comparada com demais modelos que podem ser gerados.

O nível de não-linearidade disponibilizado pelo modelo da rede neural está propriamente relacionado com as suas entradas. Deste modo, o uso de variáveis irrelevantes que não possuem nenhum grau de interdependência com a saída pode comprometer a precisão das previsões. Assim, em conjunto com o controle de complexidade, é necessária a seleção adequada do espaço de entrada de modo a obter estruturas com elevada capacidade de generalização [28].

3.1.1 Normalização dos dados de entrada

Um dos pontos mais importantes na estrutura da rede neural são suas entradas. Os MLPs comumente utilizam em suas camadas ocultas funções de ativação conhecidas como sigmoidais. Trata-se de uma função de ativação para limitar a amplitude da saída do neurônio. A ativação em uma faixa limitada de funções é geralmente chamada de funções de esmagamento [32]. Como exemplos de funções de ativação sigmoideal podem ser citados a função logística e a função tangente hiperbólica.

A função tangente hiperbólica ϕ_{hiper} está limitada ao intervalo $[-1,1]$, ou seja, $\phi_{hiper}(x): \mathbb{R} \rightarrow [-1,1]$, cuja expressão matemática é dada pela equação (12):

$$\phi_{hiper}(x) = \tanh(ax) \quad (12)$$

Essa função é especialmente vantajosa para uso em redes neurais treinadas por algoritmos de propagação reversa. Porque é fácil distinguir, e isso pode minimizar a capacidade de computação para treinamento. O termo sigmóide significa "em forma de S". A Figura 12 apresenta o gráfico de $\phi_{hiper}(x)$ para diferentes valores do ganho a . Usualmente este ganho a é feito igual a 1, permitindo que a inclinação da região linear de $\phi_{hiper}(x)$ seja calculada ao longo do processo de treinamento do modelo por meio do cálculo dos pesos sinápticos e bias.

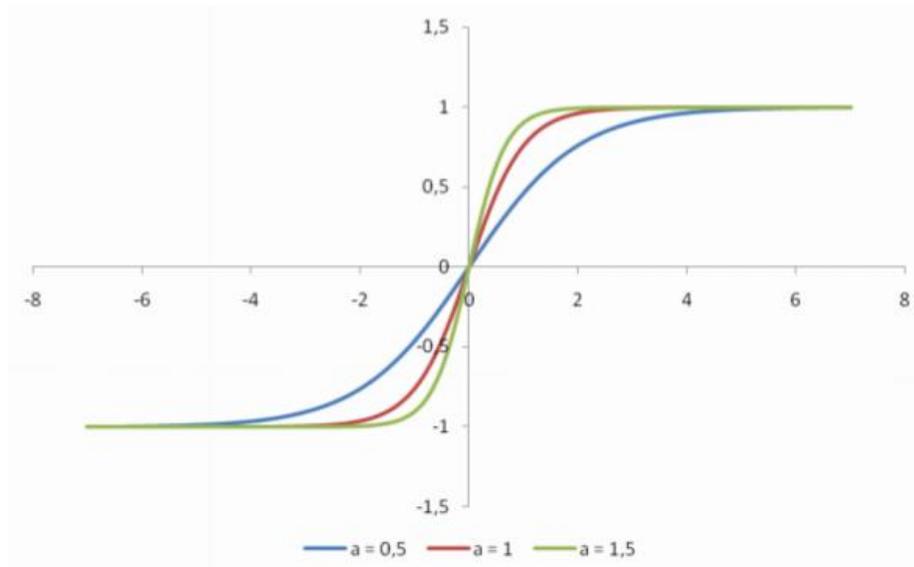


Figura 12 – Função Tangente Hiperbólica [33]

A função logística $\phi_{log}(x)$ está limitada ao intervalo $[0,1]$, ou seja, $\phi_{log}(x): \mathbb{R} \rightarrow [0,1]$, com sua expressão matemática sendo dada pela equação (13):

$$\phi_{log}(x) = \frac{1}{1 + e^{-ax}} \quad (13)$$

A Figura 13 apresenta o gráfico de $\phi_{log}(x)$ da função logística para diferentes valores do ganho a . De forma análoga ao descrito anteriormente para a função tangente hiperbólica, usualmente este ganho a é feito igual a 1 permitindo que a inclinação da região linear de $\phi_{log}(x)$ seja calculada ao longo do processo de treinamento do modelo por meio do cálculo dos pesos sinápticos e bias.

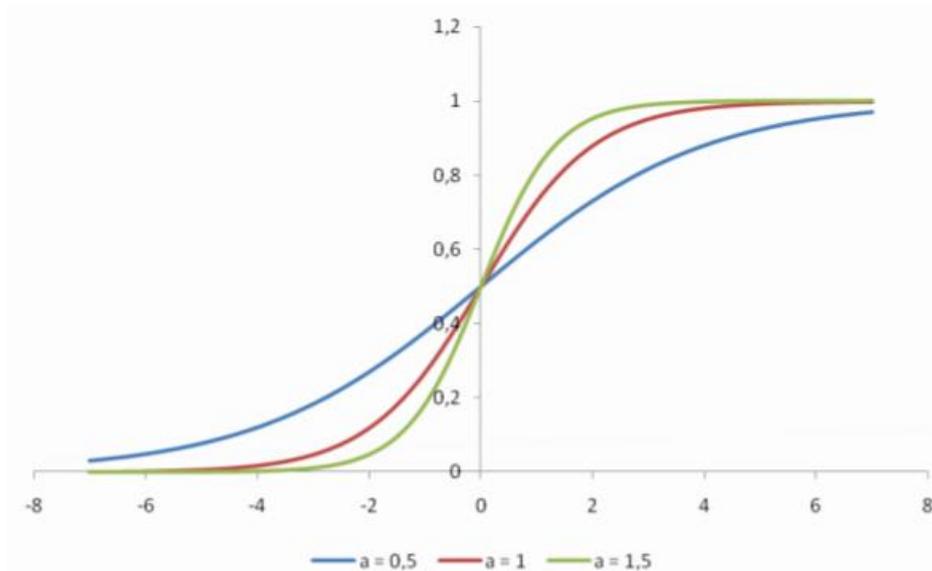


Figura 13 – Função Logística [33]

As funções de ativação da camada de saída de MLPs geralmente são definidas a partir da natureza do problema a ser abordado. Em problemas de aproximação funcional ou regressão, onde o objetivo do desenvolvimento de uma máquina de aprendizagem reside na estimação de um modelo para representação da superfície geradora dos dados, uma vez que a saída da rede neural pode assumir valores para quaisquer pontos dessa superfície, usualmente é utilizada nos neurônios de saída uma função de ativação linear, ou seja, $\phi_{saída}(x) = x$.

3.2 Rede *Bayesiana*

A estrutura Bayesiana para comparação e regularização de modelos é demonstrada por estudos de interpolação e problemas de classificação modelados com modelos lineares e não-lineares. Quando aplicado a 'redes neurais', a estrutura *Bayesiana* possibilita [34]:

- Comparação objetiva de soluções usando arquiteturas de redes alternativas;
- Regras objetivas de parada para poda de rede ou procedimentos de crescimento;

- Escolha objetiva do tipo de termos de decaimento do peso (ou regularizadores);
- Técnicas on-line para otimizar a magnitude da queda de peso (ou constante de regularização);
- Uma medida do número efetivo de parâmetros bem determinados em um modelo;
- Estimativas quantificadas das barras de erro nos parâmetros da rede e na saída da rede.

Como o foco deste trabalho é na utilização da regressão para estimativa de dados faltantes, será apresentado um subitem dedicado a este tema. Problemas relacionados a classificação poderão ser verificados em [27], [34] e [35].

3.2.1 Regressão

Definida a estrutura a ser utilizada, ou seja, número de camadas ocultas, número de neurônios por camada e tipo de função de ativação de cada neurônio, dado o conjunto $D = \{X, Y\}$, sendo $X = \{\underline{x}_1, \dots, \underline{x}_n\}$ e $Y = \{\underline{d}_1, \dots, \underline{d}_n\}$, $\underline{x} \in \mathbb{R}^n$, $\underline{x} = [x_1, \dots, x_n]^t$ e $\underline{d} \in \mathbb{R}$, onde $\underline{d} = F(\underline{x}) + \zeta$, o objetivo desse treinamento é determinar o vetor de parâmetros $\underline{w} \in \mathbb{R}^n$ que maximize a probabilidade a posteriori $p(\underline{w}|Y, X)$, dada pela equação 14:

$$p(\underline{w}|Y, X) = \frac{p(Y|\underline{w}, X)p(\underline{w})}{p(Y|X)} \quad (14)$$

Assim na equação 14, $p(Y|X) = \int p(Y|\underline{w}, X)p(\underline{w}|X) d\underline{w}$ é um fator de normalização, que garante que $\int p(\underline{w}|Y, X)d\underline{w} = 1$ e visto que os MLPs não modelam a distribuição de probabilidade $p(\underline{x})$ geradoras dos padrões de entrada e o conjunto X aparece como variável condicionante com todas as probabilidades envolvidas na equação 14.

Na ausência de conhecimento prévio sobre a solução, conforme é o caso do treinamento de MLPs, a escolha da distribuição $p(\underline{w})$ deve refletir tal falta de conhecimento. Visto que modelos apresentando componentes de \underline{w} com pequena magnitude reproduzem mapeamentos suaves [33], uma escolha razoável para a

distribuição $p(\underline{w})$ reside na distribuição gaussiana com vetor média nulo e matriz de covariância $\alpha^{-1}I$, $I \in \mathbb{R}^M \times \mathbb{R}^M$, igual à matriz identidade, dada pela equação 15:

$$p(\underline{w}) = \frac{1}{Z_{\underline{w}}(\alpha)} e^{-\left(\frac{\alpha}{2}\|\underline{w}\|^2\right)} \quad (15)$$

Onde,

$$Z_{\underline{w}}(\alpha) = \left(\frac{2\pi}{\alpha}\right)^{\frac{M}{2}}$$

Na equação 15, $\alpha \in \mathbb{R}^+$ é o chamado hiperparâmetro, cuja estimativa será apresentada mais a frente, e que, para este estágio da apresentação do algoritmo, é admitido como uma constante de valor conhecido.

Além do pressuposto de reprodução de mapeamentos suaves, a escolha da distribuição de probabilidade a $p(\underline{w})$ dada pela equação 15 simplifica as análises subsequentes, dando origem a um algoritmo analítico e iterativo para o cálculo de \underline{w} . Distribuições não-informativas, refletindo completo desconhecimento sobre o comportamento de \underline{w} , também podem ser utilizadas [36]. Neste contexto, a busca pelo valor mais provável de \underline{w} à luz dos dados é substituída pela estimativa da distribuição de probabilidade a posteriori $p(\underline{d}_{N+1}|\underline{d}_{N+1}, D, X)$ da saída a ser prevista \underline{d}_{N+1} dado o novo padrão de entrada \underline{x}_{N+1} e o conjunto de dados $\{X, D\}$. Esta distribuição é obtida pela marginalização de $p(\underline{d}_{N+1}|\underline{x}_{N+1}, \underline{\theta})$ sobre todo o espaço $\underline{\theta}$ de parâmetros desconhecidos (incluindo \underline{w}), através da integral dada pela equação 16:

$$p(\underline{d}_{N+1}|\underline{d}_{N+1}, D, X) = \int p(\underline{d}_{N+1}|\underline{x}_{N+1}, \underline{\theta})p(\underline{\theta}|D, X)d\underline{\theta} \quad (16)$$

Partindo-se da mesma analogia das equações: 15 e 16, resta agora especificar a distribuição de probabilidade do ruído ζ existente na saída desejada.

Supondo que a função $F(\underline{x})$ apresente certo grau de suavidade e que o ruído ζ possua distribuição gaussiana com média nula e variância β^{-1} , a probabilidade de ocorrência de uma saída específica d_k , dado o vetor de entrada \underline{x}_k e o vetor de parâmetros \underline{w} , é dada pela equação 17:

$$p(\underline{d}_k | \underline{x}_k, \underline{w}) = \frac{e^{\{\frac{\beta}{2}[d_k - f(x_k, \underline{w})]^2\}}}{\int e^{\{-\frac{\beta}{2}[d_k - f(x_k, \underline{w})]^2\}} d_k} \quad (17)$$

Partindo do pressuposto que os padrões do conjunto de treinamento foram obtidos de maneira independente a partir desta distribuição, pode-se obter uma expressão para verossimilhança $p(Y|\underline{w})$, dada pela equação 18:

$$p(Y|\underline{w}) = \frac{1}{Z_Y(\beta)} e^{-\left(\frac{\beta}{2} \sum_{k=1}^N [d_k - f(x_k, \underline{w})]^2\right)} \quad (18)$$

Onde,

$$Z_Y(\beta) = \left(\frac{2\pi}{\beta}\right)^{\frac{N}{2}}$$

Na equação 24, $\beta \in \mathbb{R}^+$ é outro hiperparâmetro. De posse das equações 15 e 18, calcula-se a probabilidade a posteriori de \underline{w} dado o conjunto de saídas desejadas Y , através da aplicação da regra de *Bayes* da equação 14, resultando na equação 19:

$$p(\underline{w}|Y) = \frac{1}{Z_S} e^{-[S(\underline{w})]} \quad (19)$$

Onde,

$$Z_S = \int e^{-S(\underline{w})} d\underline{w}$$

$$S(\underline{w}) = \frac{\beta}{2} \sum_{k=1}^N [d_k - f(\underline{x}_k, \underline{w})]^2 + \frac{\alpha}{2} \sum_{k=1}^N w_j^2$$

Analogamente, visto que a busca é pelo ponto ótimo \underline{w}^* que minimiza $S(\underline{w})$, da equação 19, $\lambda = \alpha/\beta$. Desta forma, o cálculo do hiperparâmetro α e β fornece uma estimativa do parâmetro de regularização λ .

Consequentemente, fazendo \underline{w} , α e β desconhecidos, a probabilidade a posteriori de \underline{w} , $p(\underline{w}|Y)$ passa a ser conhecido pela equação 20:

$$p(\underline{w}|Y) = \iint p(\underline{w}, \alpha, \beta|Y) d\alpha d\beta = \iint p(\underline{w}|\alpha, \beta, Y) p(\alpha, \beta|Y) d\alpha d\beta \quad (20)$$

Deste modo, utilizando a abordagem de Mackay [35] que parte do princípio que a probabilidade a posteriori $p(\alpha, \beta|Y)$ apresenta pouca dispersão em torno dos valores mais prováveis α^* e β^* , a equação 26 pode ser simplificada, têm-se a equação 21:

$$p(\underline{w}|Y) \simeq p(\underline{w}|\alpha^*, \beta^*, Y) \iint p(\alpha, \beta|Y) d\alpha d\beta = p(\underline{w}|\alpha^*, \beta^*, Y) \quad (21)$$

Para a determinação dos valores mais prováveis para os hiperparâmetros, a probabilidade a posteriori destes, $p(\alpha, \beta|Y)$, deve ser maximizada. Aplicando a regra de *Bayes*, é obtida a equação (22) para esta probabilidade:

$$p(\alpha, \beta|Y) = \frac{p(Y|\alpha, \beta)p(\alpha|\beta)}{p(Y)} \quad (22)$$

A maximização da evidência pode ser utilizada para estimativa dos hiperparâmetros mais prováveis α^* e β^* . A apresentação detalhada desse desenvolvimento pode ser encontrada em [28].

O algoritmo de treinamento bayesiano de MLPs para problemas de regressão pode ser resumido da forma que segue [28]:

1. Faça $l = 0$.
2. Inicialize o vetor de parâmetros $\underline{w}(l)$ e os hiperparâmetros $\alpha(l)$ e $\beta(l)$.
3. Utilizando alguma técnica de otimização, atualize o vetor de parâmetros $\underline{w}(l + 1)$ através da minimização do funcional $S(\underline{w})$ dado pela equação 23:

$$S(\underline{w}) = \frac{\beta}{2} \sum_{k=1}^N [d_k - f(\underline{x}_k, \underline{w})]^2 + \frac{1}{2} \sum_{i=1}^g (\alpha_i \sum_{j=1}^{M_i} w_{ij}^2) \quad (23)$$

4. Se o critério de parada for atendido, encerre o algoritmo. Do contrário, vá para o passo 5.
5. Calcule a matriz hessiana $\underline{H}(\underline{w})|_{\underline{w}=\underline{w}(l)}$.
6. Atualize os hiperparâmetros $\alpha_i(l + 1)$ e $\beta(l + 1)$, utilizando as equações 24 a 27 para $\alpha_i(l + 1)$ e a equação 28 para $\beta(l + 1)$, respectivamente.

$$\underline{\underline{B}}_i(l) = [\underline{\underline{A}}(\underline{\underline{w}})]_{\underline{\underline{w}}=\underline{\underline{w}}(l)}^{-1} \underline{\underline{I}}_i \quad (24)$$

$$\gamma_i(l+1) = M_i - \text{trace}\{\underline{\underline{B}}_i(l)\} \quad (25)$$

$$\alpha_i(l+1) = \frac{\gamma_i(l+1)}{\|\underline{\underline{w}}_i(l)\|^2} \quad (26)$$

$$\gamma(l+1) = \sum_{i=1}^g \gamma_i(l+1) \quad (27)$$

$$\beta \sum_{k=1}^N [d_k - f(\underline{\underline{x}}_k, \underline{\underline{w}}^*)]^2 = N - \gamma \quad (28)$$

7. Faça $l = l + 1$ e retorne ao passo 3.

Segue na equação 29 o cálculo da variância de saída desejada:

$$\sigma_t^2 = \frac{1}{\beta} + g^T H^{-1} g \quad (29)$$

Onde σ_t^2 representa a variância da saída da rede neural, β^{-1} a variância da distribuição a priori, hiperparâmetro calculado pela equação 34, g o gradiente de $S(\underline{\underline{w}})$ e $\underline{\underline{H}}$ a matriz hessiana de $S(\underline{\underline{w}})$, todos calculados para o valor estimado de $\underline{\underline{w}}$. O primeiro termo corresponde à contribuição do ruído intrínseco sobre os dados de saída e o segundo termo corresponde à contribuição da largura da distribuição à posteriori dos pesos da rede neural.

Considerando um nível de confiança especificado, o intervalo de confiança para o valor previsto deve ser calculado através da equação 30:

$$[y_t^{min}, y_t^{max}] = [y_t - \alpha \sigma_t, y_t + \alpha \sigma_t] \quad (30)$$

Onde α representa o número de múltiplos do desvio padrão associados com o nível de confiança especificado. Por exemplo, conforme tabela 4, considerando um nível

de confiança de 50,00%, α é igual a 0,674 e seguindo a tabela, têm-se que para 99,73% de confiança, α é igual a três.

Tabela 4 – Níveis de confiança para valores de α

Nível de confiança	Valor de α
50,00%	0,674
68,27%	1
80,00%	1,282
90,00%	1,645
95,45%	2
99,00%	2,576
99,73%	3

3.2.2 Método Automático de Treinamento de MLPs utilizando rede *Bayesiana*

Tomando como base um grupo específico dos parâmetros que definem o MLP, a relação entre α_i e a magnitude de w_i pode ser utilizada para mensuração da relevância de cada entrada. Conhecido como determinação automática de relevância (DAR) [37], esta metodologia para a seleção de entradas divide o vetor de parâmetros w em $n + m + 2$ grupos. Os n primeiros grupos representam os pesos que ligam cada uma das entradas à camada oculta (n grupos com cada grupo contendo m pesos). O restante dos conjuntos está atrelado aos pesos que ligam cada neurônio da camada oculta à camada de saída (m grupos com cada grupo contendo C pesos), outro grupo com os *bias* dos neurônios da camada oculta (m componentes) e por fim, um grupo com os *bias* dos neurônios da camada de saída (C componentes).

A escolha específica de agrupamento dos pesos permite considerar a relevância de cada sinal através da análise dos n hiperparâmetros α_i . As entradas com valores elevados de α_i ao final do processo de treinamento são ligadas ao modelo através de pesos com pequena magnitude, deste modo, contribuindo menos para o cálculo da saída. Logo, quanto maior α_i , menor a importância da respectiva entrada no cálculo da saída. Utilizando este agrupamento específico de pesos, a análise dos hiperparâmetros α obtidos fornece uma estrutura para mensuração da relevância de cada entrada. Mesmo sendo

capaz de ordenar estas variáveis segundo a importância de cada uma no cálculo da saída, este método não apresenta ferramentas para detecção de entradas irrelevantes, sendo necessária a definição de um limiar de relevância α_0 para identificação das variáveis irrelevantes. O método baseado na inserção de variáveis aleatórias de prova [38], [39] é utilizado para definição empírica do limiar de relevância α_0 .

A regra de *Bayes* foi aplicada para estimação dos parâmetros do modelo, supondo que os hiperparâmetros que regem as probabilidades a priori envolvidas eram conhecidos e sendo definida uma estrutura para o modelo. Como os hiperparâmetros também são desconhecidos, o fator de normalização $p(X, D)$ nada mais é que a evidência para os hiperparâmetros, utilizada na regra de *Bayes*. Neste ponto, é suposto que o modelo é conhecido, o que não corresponde à realidade. Desta maneira, a parcela $p(X, D)$ é a evidência para os modelos, utilizada para seleção da estrutura, conforme indicada na equação 31:

$$\ln p(Y|H_h) = -S(\underline{w}) - \frac{1}{2} \ln |\underline{A}(\underline{w})| + \frac{1}{2} \sum_{i=1}^g M_i \alpha_i + 2 \ln m + \ln m! + \frac{1}{2} \sum_{i=1}^g \ln \left(\frac{2}{\gamma_i} \right) \quad (31)$$

Na expressão (31), $S(\underline{w})$ e a matriz $\underline{A}(\underline{w})$ são avaliados para \underline{w}^* e $\underline{\alpha}^*$, sendo também utilizados os respectivos números efetivos de parâmetros γ_i^* estimados. Utilizando DAR, o número de grupos g é igual a $n + m + 2$. Deste modo, a evidência para os modelos poderá ser utilizada para determinação do número de neurônios na camada oculta de MLPs através da escolha do modelo que apresentar maior $\ln p(Y|H_h)$, ou seja, aquele mais provável com base nos dados. O algoritmo *bayesiano* para desenvolvimento de modelos neurais autônomos pode ser resumido da maneira que segue [33]:

1. Escolha o número mínimo m_{min} e máximo m_{max} de neurônios na camada oculta.
2. Para $i = m_{min}, \dots, m_{max}$, faça:
 - 2.1. Insira variáveis de prova no conjunto de entradas original seguindo as distribuições de probabilidade apresentadas pela equação 32 e equação 33:

$$p(x_{PC}) = \begin{cases} \frac{1}{b-a}, & \text{se } a \leq x_{PC} < b \\ 0, & \text{se } x_{PC} < a \text{ ou } x_{PC} > b \end{cases} \quad (32)$$

$$p(x_{PD}) = \begin{cases} \frac{1}{k}, & \text{se } x_{PC} \in C \\ 0, & \text{se } x_{PC} \notin C \end{cases} \quad (33)$$

Se as variáveis de entrada forem somente contínuas, insira somente o sinal de prova. Caso contrário, insira também a variável de prova discreta.

2.2. Estime o vetor de parâmetros \underline{w} e o conjunto de hiperparâmetros $\underline{\alpha}$ e $\underline{\beta}$ utilizando os algoritmos listados nas seções 3.2.1.

2.3. Separe os hiperparâmetros α_i associados a entradas contínuas e discretas em vetores distintos e ordene de forma decrescente estes vetores.

2.4. Em cada vetor, selecione as variáveis de entrada situadas abaixo do respectivo limiar de relevância, ou seja, $\alpha_i < \alpha_0$, com $\alpha_0 = \alpha_C$ para variáveis contínuas e $\alpha_0 = \alpha_D$ para as discretas.

2.5. Utilizando somente as variáveis relevantes selecionadas no passo 2.4, estime novamente o vetor de parâmetros \underline{w} e o conjunto de hiperparâmetros $\underline{\alpha}$ utilizando o algoritmo de treinamento *bayesino*.

2.6. Calcule o logaritmo natural da evidência do modelo, $\ln p(Y|H_h)$, dado pela equação 31.

3. Escolha o modelo com maior evidência.

Visando à estimação do vetor de parâmetros, o algoritmo descrito acima requer somente a definição do número mínimo e máximo de neurônios na camada escondida, m_{min} e m_{max} . Os hiperparâmetros dos algoritmos de retropropagação são definidos de forma geral, geralmente utilizando valores padrão independentes da aplicação. Por outro lado, m_{min} e m_{max} estão relacionados basicamente com o esforço computacional, uma vez que devem ser avaliados $m_{max} - m_{min} + 1$ modelos.

Destaca-se que, dado um conjunto inicial de entradas, são determinadas automaticamente as variáveis mais relevantes para cada tipo de modelo. Visto que o nível de não-linearidade da estrutura neural está diretamente associado à representação do espaço de entrada [40], é esperado que o conjunto de entradas relevantes varie entre as diversas estruturas. Desta maneira, utilizando algum método de seleção do conjunto inicial de entradas, a intervenção do usuário ocorre somente para a definição dos limites para a busca pelo número de neurônios na camada escondida, mostrando o considerável grau de automação e adaptabilidade do método [33].

Capítulo 4 - Aplicação de RNAs na estimação de dados faltantes de proteção catódica

Propostas de utilização de RNAs demonstram resultados promissores em comparação às metodologias tradicionais, em diversos segmentos da engenharia. Neste capítulo, serão apresentados os dados de dois sistemas de proteção catódica que atuam em tubulações distintas de aço enterrado de distribuição de gás natural, presentes na região sudeste do Brasil. Através da ferramenta MATLAB em conjunto com uma sólida base de dados, é possível estimar dados faltantes de um determinado elemento com precisão aceitável através de RNA.

Serão analisados dois sistemas de proteção catódica, definidos como caso 01 e caso 02. Pretende-se simular os dois casos com duas metodologias distintas. Uma metodologia determinística (metodologia 01) através da regressão não linear e outra metodologia probabilística (metodologia 02) através de rede bayesiana (Figura 14). Em ambos os casos serão realizadas simulações variando a saída das bases de dados dos sistemas caso 01 e caso 02 com o objetivo de verificar qual dos atributos tem maior aderência ao propósito deste trabalho. Pretende-se então, demonstrar que com a utilização da RNA, é possível prever um dado faltante de inspeção do sistema de proteção catódica, com precisão aceitável, de modo a dar melhores subsídios ao setor de manutenção na tomada de decisão e análise de dados.

É importante lembrar que cada sistema de proteção catódica possui condições específicas de acordo com fatores construtivos, tempo de vida dos anodos, condições de manutenção e de solo que varia com relação à região e profundidade. Portanto, a aplicação da metodologia para estimação de um dado faltante somente poderá ser realizada para o próprio sistema de proteção catódica, cuja base foi gerada e aplicada no treinamento da RNA.

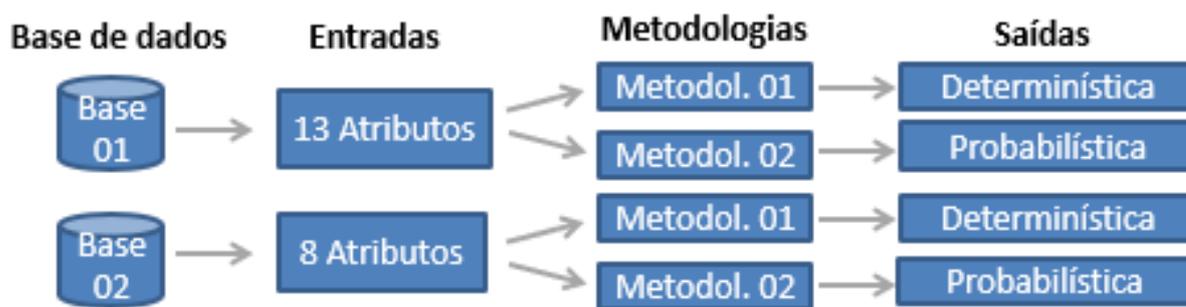


Figura 14 – Estrutura para treinamento

4.1 Estrutura dos sistemas de proteção catódica

Neste item serão apresentados os dois sistemas de proteção catódica, caso 01 e caso 02. As grandezas consideradas para elaboração de ambos bancos de dados dos casos descritos foram:

- Em retificadores: Considerado a corrente de saída.
- Em pontos de teste: Considerado o potencial tubo-solo.

A grandeza mais importante de um sistema de proteção catódica é o potencial tubo-solo e o objetivo do setor de manutenção é manter esta grandeza dentro da faixa de proteção, conforme indicado no capítulo 2 desta dissertação.

Os retificadores possuem tensão fixa em operação normal, podendo ser variada manualmente de acordo com a necessidade de obter mais saída de corrente e consequente proteção do sistema. A demanda de corrente pode aumentar ou diminuir com o tempo devido ao aparecimento/reparo de falhas no revestimento e/ou regime de chuvas, que altera a resistividade do solo. Para manter o sistema polarizado e dentro dos critérios, é necessária a injeção de corrente dos anodos para a tubulação. Como a tensão é fixa, dentro da faixa determinada, a saída de corrente passa a ser o alvo e será considerada como atributo da base de dados dos sistemas de proteção catódica analisados.

Para que o objetivo desta dissertação seja cumprido, o banco de dados deve percorrer diversas situações de modo a dar subsídios para que a saída da RNA possua o menor erro possível.

4.1.1 Estrutura caso 01

O caso 01 possui banco de dados coletado durante três anos com resumos diários através de um sistema remoto de medição. A expectativa inicial da análise deste banco de dados era obter 1095 medições (linhas), entretanto, após análise prévia, foram excluídos dias com dados inconsistentes. Portanto, foram considerados 801 dados em 13 elementos que compõe o sistema de proteção catódica desta região. A Figura 15 apresenta a imagem do Google Earth com os elementos do Caso 01 espacialmente dispostos. A seguir, apresenta-se pequeno resumo dos elementos considerados:

- N° de retificadores: 4;
- N° de pontos de teste: 9
- Banco de Dados: 13 x 801 = 10.413 dados
- km de rede: ≈ 40

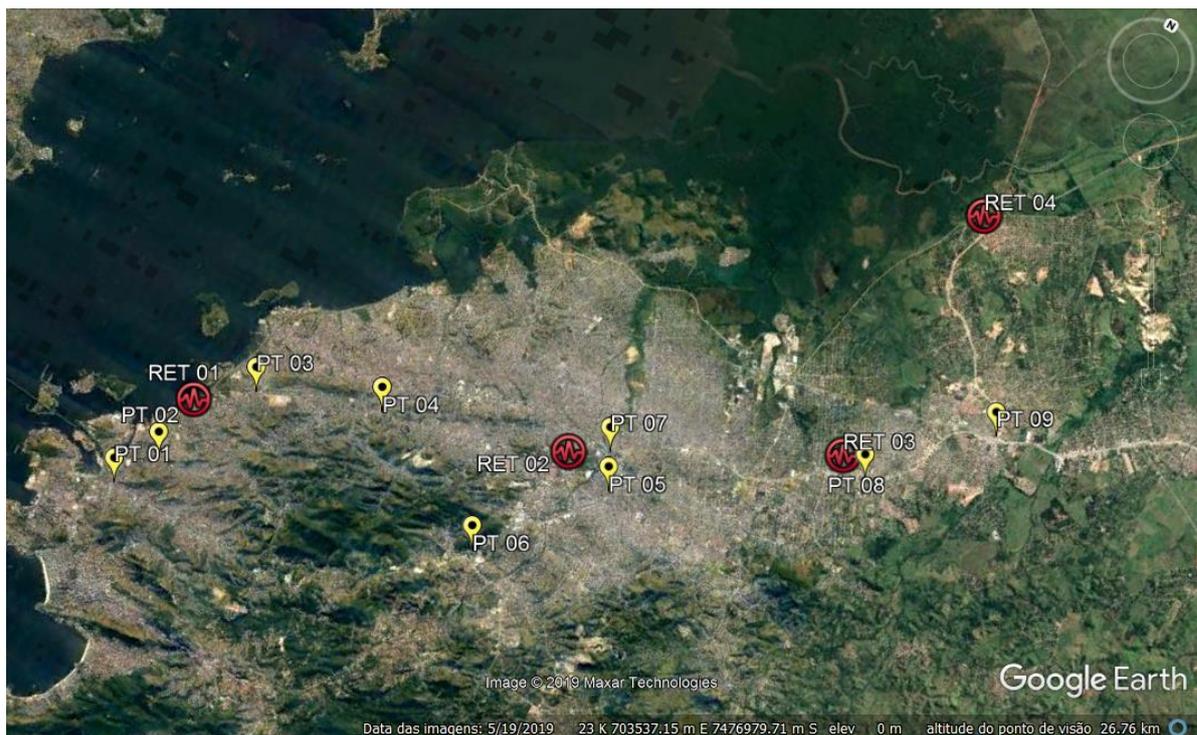


Figura 15– Imagem dos elementos de proteção catódica do caso 1 dispostos no mapa

Segue na Figura 16, os dados dispostos na mesma base de tempo. É possível observar algumas semelhanças entre dados, principalmente entre pontos de teste e retificadores próximos.

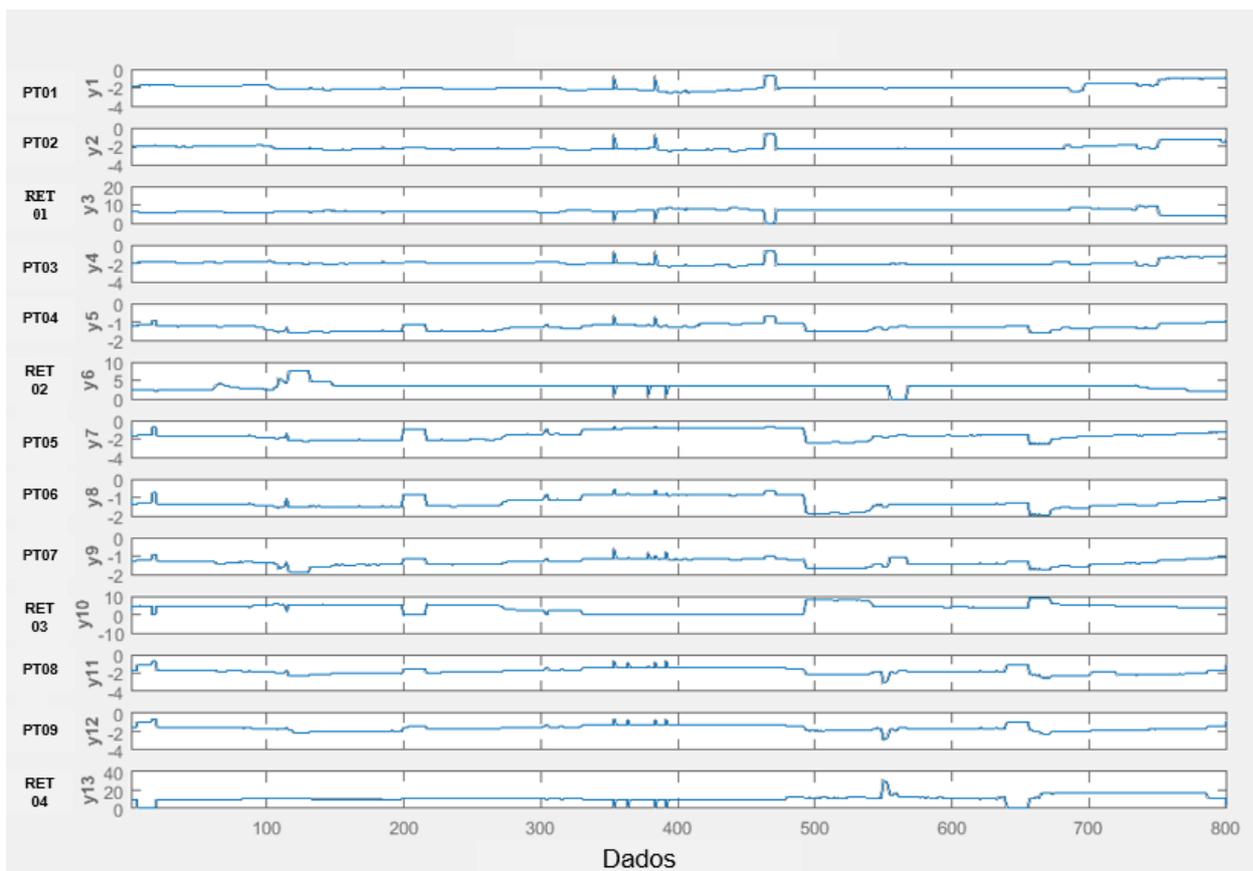


Figura 16 – Gráfico com os dados do caso 01.

4.1.2 Estrutura caso 02

O caso 02 possui banco de dados coletado durante 1 ano, com informações a cada 2 minutos simultaneamente através de sistema remoto de medição. A expectativa inicial da análise deste banco de dados era obter 262.800 medições (linhas), entretanto, após análise prévia foram excluídas medições com dados inconsistentes. Portanto, foram considerados 75.167 dados em 8 elementos que compõe o sistema de proteção catódica desta região. A Figura 17 apresenta a imagem do Google Earth com os elementos do Caso 02 espacialmente dispostos. A seguir, apresenta-se pequeno resumo dos elementos considerados:

- N° de retificadores: 3;
- N° de pontos de teste: 5
- Banco de Dados: $8 \times 75167 = 601.336$ dados
- km de rede: ≈ 43

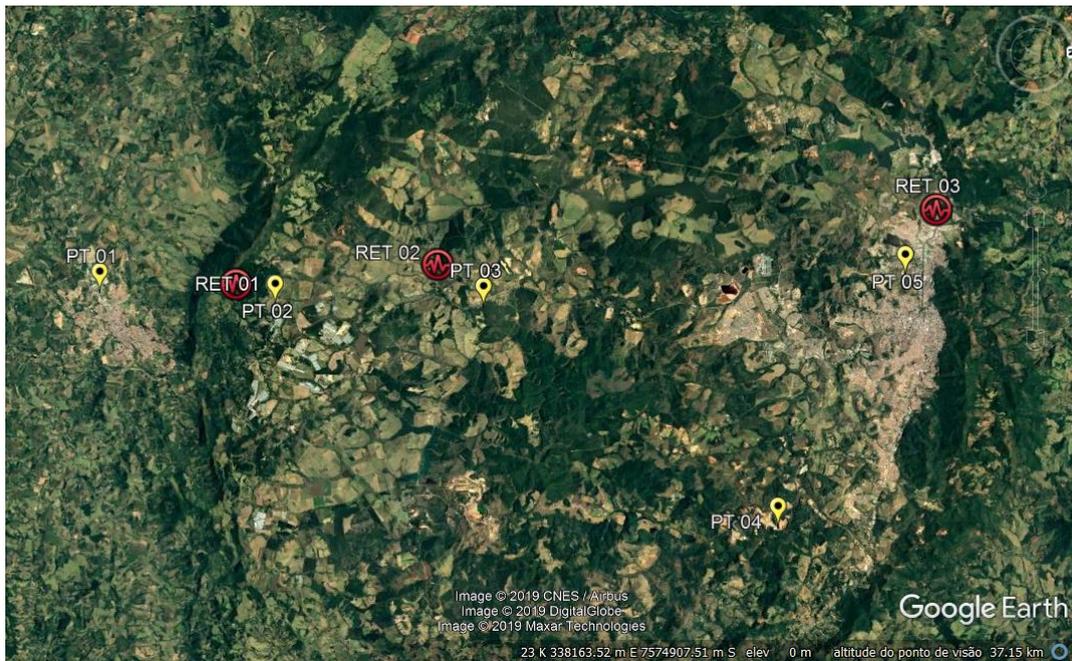


Figura 17 – Imagem dos elementos de proteção catódica do caso 2 dispostos no mapa

Segue na Figura 18, os dados dispostos ao mesmo tempo. É possível observar algumas semelhanças entre dados, principalmente entre pontos de teste e retificadores próximos, entretanto, observa-se menor relação entre estes dados em comparação aos dados do caso 1.

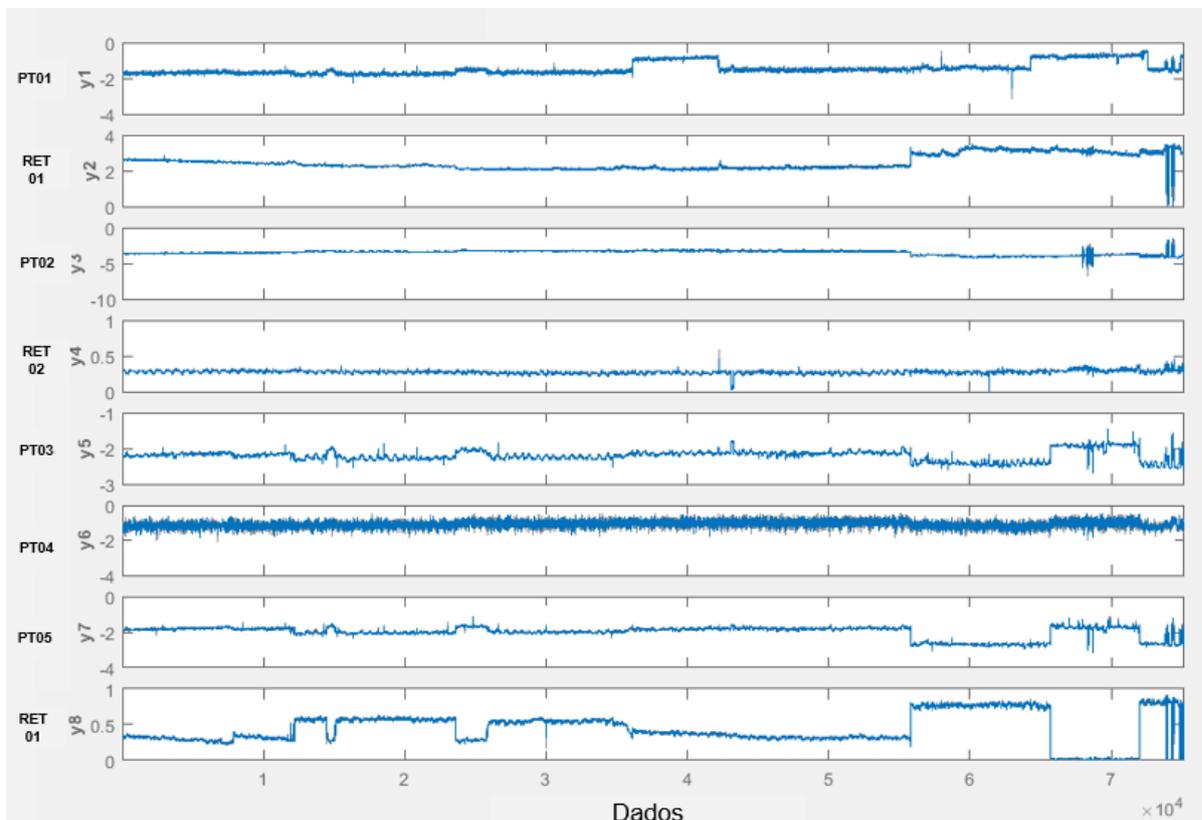


Figura 18 – Gráfico com os dados do caso 02.

4.2 Desenvolvimento da RNA

Através das bases de dados obtidas ao longo de um determinado período e tratadas pelo especialista, foi possível elaborar duas metodologias com o uso de RNA's. A primeira trata-se de regressão não linear, elaborada através do uso de MLP's e a segunda baseada em rede bayesiana, adaptada de [28]. As RNA's desenvolvidas realizam os procedimentos de treinamento, validação e teste para as duas bases de dados. A metodologia que apresentar os melhores resultados para os casos estudados, será a escolhida.

A base de dados 1 e 2 foram divididas da seguinte maneira para o treinamento e validação:

Caso 1: 10.413 atributos de dados onde 2/3 (6.942 dados) foram escolhidos para o treinamento e validação e 1/3 (3.471 dados) para a teste.

Caso 2: 601.336 atributos de dados onde 2/3 (405.104) foram escolhidos para o treinamento e validação e 1/3 (196.232) para teste.

Com o objetivo de obter o melhor resultado possível, percorrendo todas as possibilidades de situações da base de dados, para os dois casos, foi feita a separação entre o treinamento, validação e teste da seguinte forma:

- A cada 3 (três) atributos, os 2 (dois) primeiros foram separados para o treinamento e validação e o último para o teste e assim em diante.

Caso a separação dos dados fosse realizada por batelada, separando por exemplo 2/3 do total inicial e 1/3 do fim, algumas situações não seriam mapeadas o que aumentaria o erro do treinamento independente das metodologias propostas.

Definindo as entradas e a saída esperada, será dado um *input* da base de dados no sistema. A fim de evitar problemas relativos à estagnação do treinamento, as entradas e saída serão normalizadas em uma faixa, neste caso definida entre [-1 1] pois os dados têm valores e características diferentes. Este passo foi aplicado em ambas metodologias.

4.2.1 Metodologia 01

A metodologia 01, cuja base teórica é apresentada no item 3.1 desta dissertação, baseia-se na utilização do aprendizado supervisionado. A não linearidade, particularmente, é fruto da geração do sinal de entrada que é inteiramente não-linear.

Após normalização dos dados, será definido o número de neurônios que serão utilizados. Segundo o teorema da aproximação universal [41], uma camada oculta é o suficiente, tendo apenas que definir o número de neurônios que serão utilizados neste sistema, que será definido de acordo com a melhor aproximação ou menor erro da saída. Após testes na faixa entre 1 a 20 neurônios, foram definidas configurações de acordo com a variação dos dados de entrada e saída, portanto, para a base de dados 1 foram 13 resultados e para a base de dados 2 foram 8 resultados sistêmicos diferentes. Segue na Tabela 5 os resultados obtidos relacionados ao número de neurônios definidos pela metodologia aplicada. Na Figura 19 é apresentada a estrutura da RNA. Para os dados do caso 1 serão 12 entradas e para os dados do caso 2 serão 7 entradas e, em ambos os casos, uma única saída alvo.

Tabela 5 – Resultados e variação do número de neurónios aplicados na metodologia 01

Base 1	Saída (y)	Número de Neurônios	Base 2	Saída (y)	Número de Neurônios
1	PT01	16	1	PT01	2
2	PT02	17	2	RET01	17
3	RET 01	13	3	PT02	1
4	PT03	11	4	RET02	1
5	PT04	6	5	PT03	1
6	RET 02	15	6	PT04	13
7	PT05	14	7	PT05	1
8	PT06	16	8	RET03	16
9	PT07	15			
10	RET 03	1			
11	PT08	20			
12	PT09	20			
13	RET 04	1			

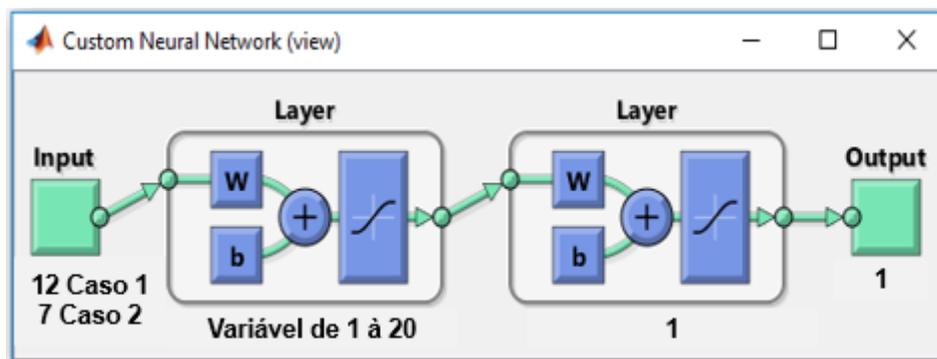


Figura 19 – Estrutura da rede neural artificial.

Além de 1/3 de dados para o teste final, foram utilizados 20% dos 2/3 dos dados para validação e reforço da metodologia.

A medida do número de vezes que todos os vetores de treinamento são usados para atualizar os pesos é chamada de *epoch*. Nesta rede neural foram utilizados 1000 *epochs*. Todas as amostras de treinamento passam pelo algoritmo de aprendizado simultaneamente em uma época antes que os pesos sejam atualizados. Neste primeiro momento, a saída do sistema é conhecida para definição do modelo.

4.2.2 Metodologia 02

A metodologia 02, cuja base teórica é apresentada no item 3.2 desta dissertação, baseia-se na utilização da rede bayesiana, adaptada de [28].

Após normalização dos dados, será definido o número de neurônios que serão utilizados. Segundo o teorema de aproximação universal [41], uma camada oculta é o suficiente, tendo apenas que definir o número de neurônios que serão utilizados neste sistema, que será definido de acordo com a melhor aproximação ou menor erro da saída. Após testes na faixa entre 1 a 25 neurônios, foram definidas configurações de acordo com a variação dos dados de entrada e saída, portanto, para a base de dados 1 foram 13 resultados e para a base de dados 2 foram 8 resultados sistêmicos diferentes. Segue na Tabela 6 os resultados obtidos relacionados ao número de neurônios definidos pela metodologia aplicada.

Tabela 6 – Resultados e variação do número de neurónios aplicados na metodologia 02

Base 1	Saída (y)	Número de Neurônios	Base 2	Saída (y)	Número de Neurônios
1	PT01	8	1	PT01	22
2	PT02	17	2	RET01	19
3	RET 01	22	3	PT02	25
4	PT03	17	4	RET02	24
5	PT04	19	5	PT03	25
6	RET 02	24	6	PT04	13
7	PT05	24	7	PT05	25
8	PT06	21	8	RET03	24
9	PT07	19			
10	RET 03	24			

11	PT08	21
12	PT09	24
13	RET 04	21

4.3 Metodologia para comparação entre as saídas e erro associado.

Conforme mencionado no item 4.2, 1/3 das bases de dados não foram utilizadas com o objetivo de testar os resultados obtidos por atributo que será analisado. Para tanto, será definido a magnitude do erro que pode ser considerada como aceitável.

Uma tubulação de aço enterrada, sem proteção catódica e sem interferência elétrica de outros sistemas, possui potencial eletroquímico natural entre o duto e o solo em torno de $-0,54 V_{cc}$ [13]. O potencial mínimo de proteção tubo-solo, é $-0,85V_{cc}$ [13]. Considerando o erro de $\pm 0,1V_{cc}$ como aceitável, poderíamos obter um potencial equivocado de $-0,75V_{cc}$ (utilizando a referência do potencial mínimo de proteção), entretanto, este potencial não é um problema a curto prazo e, caso venha a se repetir durante os meses subsequentes, o especialista deverá enviar a equipe de inspeção para comprovação e prova real.

Apesar do banco de dados possuir atributos de corrente de saída dos retificadores, nos concentraremos nos resultados dos potenciais eletroquímicos dos pontos de teste analisados nos dois bancos de dados propostos e nas duas metodologias apresentadas.

Para a escolha entre as metodologias, será feita a comparação entre as saídas baseado no erro absoluto (Ea), erro percentual (Ep), coeficiente de determinação (R^2) e desvio padrão (σ).

O erro absoluto é a diferença entre o valor alvo (y) e o valor retornado pelo sistema (y'), conforme apresentado em (34), ou seja:

$$Ea = |y - y'| \quad (34)$$

O erro percentual (Ep) é a representação percentual do erro relativo (Er). O erro relativo é a razão entre o Ea e o valor alvo, representado pela equação (35), ou seja:

$$Er = \frac{Ea}{y} \quad (35)$$

$$Ep = Er100\% \quad (36)$$

O coeficiente de determinação (R^2), é uma função que devolve o coeficiente de correlação de dois intervalos de células. O objetivo é utilizar o coeficiente de correlação para determinar a relação entre duas propriedades [42]. Neste caso, será relacionado a coluna com o valor alvo versus a coluna com o resultado das metodologias propostas.

A equação para o coeficiente de correlação é [42]:

$$R^2(X, Y) = \frac{\sum(x - \bar{x})(y - \bar{y})}{\sqrt{\sum(x - \bar{x})^2 \sum(y - \bar{y})^2}} \quad (37)$$

Onde,

\bar{x} e \bar{y} são as médias das amostras.

O desvio padrão (σ) é uma medida do grau de dispersão dos valores com relação ao valor médio.

$$\sigma = \sqrt{\frac{\sum(x - \bar{x})^2}{(n - 1)}} \quad (38)$$

Onde,

n é o tamanho da amostra.

Capítulo 5 - Resultados

Neste capítulo serão apresentados os resultados das simulações das metodologias 1 e 2 a partir das bases de dados dos casos 1 e 2, ambos definidos no capítulo 4. Com objetivo de verificar e analisar os resultados das simulações, serão comparadas as metodologias por atributo inclusive para definir qual destes obteve melhor aderência e confiabilidade nos resultados obtidos. Por fim, uma das metodologias será escolhida, baseada no menor erro encontrado.

5.1 Resultados Caso 01

Para o caso 01, foram utilizados 10.413 dados de 13 atributos (4 retificadores e 9 pontos de teste).

Foram realizadas 26 simulações, 13 utilizando a metodologia 01 e 13 utilizando a metodologia 02. A variável de saída foi sendo modificada ao passo de cada simulação realizada. Segue na Tabela 7 os resultados com as comparações dos erros médios, a partir da base teste, por metodologia.

Tabela 7 – Resultados gerais obtidos nas metodologias 01 e 02 para o caso 01

Caso 01								
Saída	Metodologia 01				Metodologia 02			
	Ea	Ep	R ²	σ	Ea	Ep	R ²	σ
PT01	0,0221	1,10%	0,9263	0,0965	0,023	1,20%	0,973	0,0568
PT02	0,012	0,62%	0,9798	0,0436	0,0141	0,72%	0,9876	0,0342
RET 01	0,0084	0,20%	0,9979	0,0534	0,0227	0,60%	0,9952	0,0819
PT03	0,0047	0,26%	0,9906	0,025	0,0066	0,36%	0,9932	0,0213
PT04	0,0077	0,63%	0,988	0,0186	0,0093	0,77%	0,9829	0,0222
RET 02	0,0085	0,94%	0,9925	0,0794	0,0104	0,79%	0,9988	0,0311
PT05	0,0164	0,94%	0,9921	0,0481	0,015	1,04%	0,9965	0,0286
PT06	0,0002	0,02%	1	0,0019	0,0009	0,08%	1	0,0018
PT07	0,0017	0,19%	0,9928	0,0164	0,0009	0,08%	0,9999	0,0022
RET 03	0,0217	0,57%	0,9995	0,0512	0,0099	0,26%	0,9999	0,0262
PT08	0,0138	0,92%	0,9823	0,0483	0,0193	1,08%	0,9893	0,0357
PT09	0,0184	1,25%	0,9539	0,0681	0,0168	1,02%	0,9821	0,0421
RET 04	0,5666	5,76%	0,9461	0,8537	0,0993	1,47%	0,9921	0,3257
Média	0,054	1,03%	0,9801	0,108	0,0191	0,73%	0,9916	0,0546

Como os resultados possuem saídas com magnitudes diferentes, excluiremos da análise os resultados da saída de corrente dos retificadores. Vale lembrar que a corrente de saída é um atributo importante para o treinamento, visto que a polarização ocorre através dos elétrons que fluem dos anodos para a tubulação, entretanto, o foco desta dissertação é na saída dos potenciais eletroquímicos. Desta forma, segue na Tabela 8 os resultados considerando apenas os pontos de teste:

Tabela 8 – Resultados considerando apenas os atributos “ponto de teste”

Caso 01 – Somente ponto de teste								
Saída	Metodologia 01				Metodologia 02			
	Ea	Ep	R ²	σ	Ea	Ep	R ²	σ
PT01	0,0221	1,10%	0,9263	0,0965	0,023	1,20%	0,973	0,0568
PT02	0,012	0,62%	0,9798	0,0436	0,0141	0,72%	0,9876	0,0342
PT03	0,0047	0,26%	0,9906	0,025	0,0066	0,36%	0,9932	0,0213
PT04	0,0077	0,63%	0,988	0,0186	0,0093	0,77%	0,9829	0,0222
PT05	0,0164	0,94%	0,9921	0,0481	0,015	1,04%	0,9965	0,0286
PT06	0,0002	0,02%	1	0,0019	0,0009	0,08%	1	0,0018
PT07	0,0017	0,19%	0,9928	0,0164	0,0009	0,08%	0,9999	0,0022
PT08	0,0138	0,92%	0,9823	0,0483	0,0193	1,08%	0,9893	0,0357
PT09	0,0184	1,25%	0,9539	0,0681	0,0168	1,02%	0,9821	0,0421
Média	0,0108	0,66%	0,9784	0,0407	0,0118	0,70%	0,9894	0,0272

5.1.1 Caso1 – Resultado PT01

A Figura 20 apresenta o gráfico com a comparação entre os resultados das duas metodologias e o valor real (alvo) do potencial eletroquímico do PT01 *versus* o tempo (número de linhas dos conjuntos para o teste).

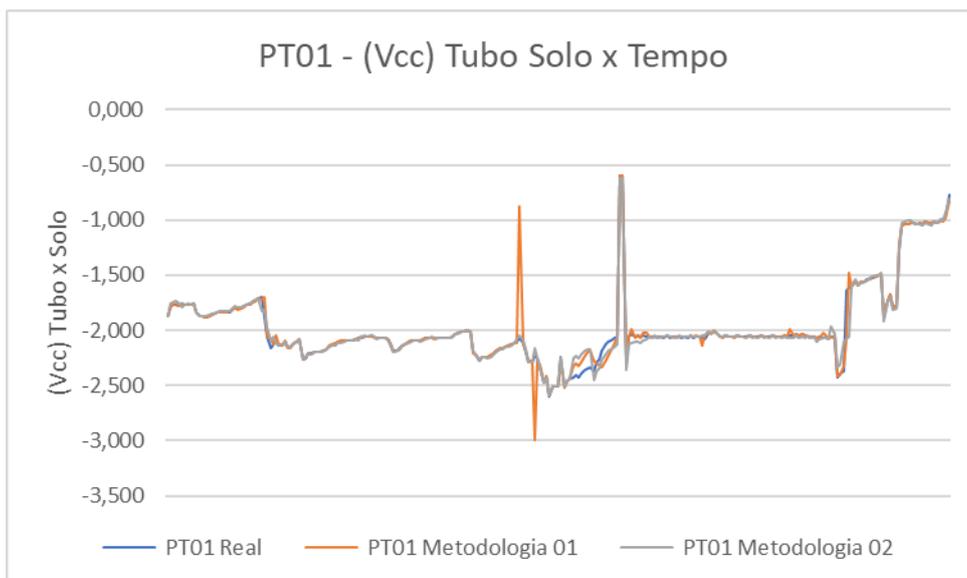


Figura 20 – Gráfico Tubo Solo x Tempo: PT01, Caso 01

A Figura 21 apresenta o gráfico com a dispersão dos resultados das metodologias propostas.

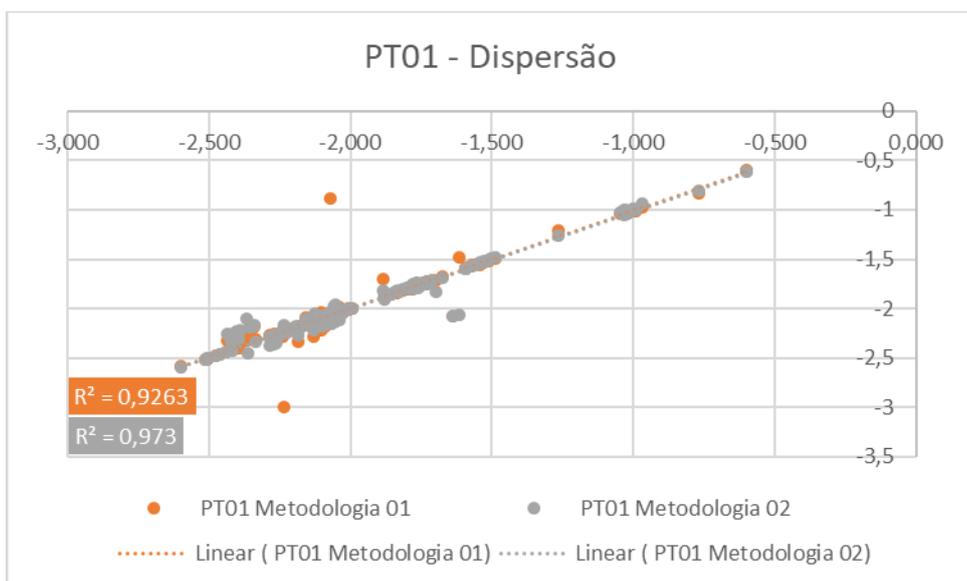


Figura 21 – Gráfico Dispersão: PT01, Caso 01

As Figuras 22 e 23 apresentam os gráficos com o histograma do erro relativo, para cada metodologia proposta.

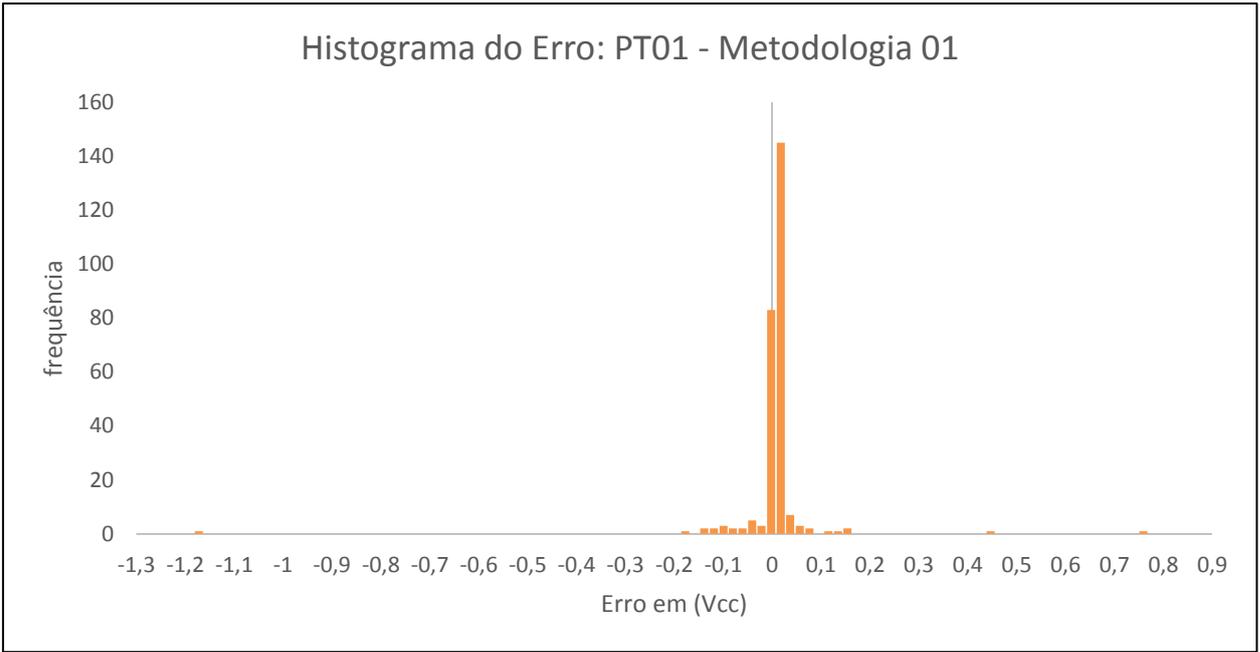


Figura 22 – Gráfico Histograma: PT01, Caso 01, Metodologia 01

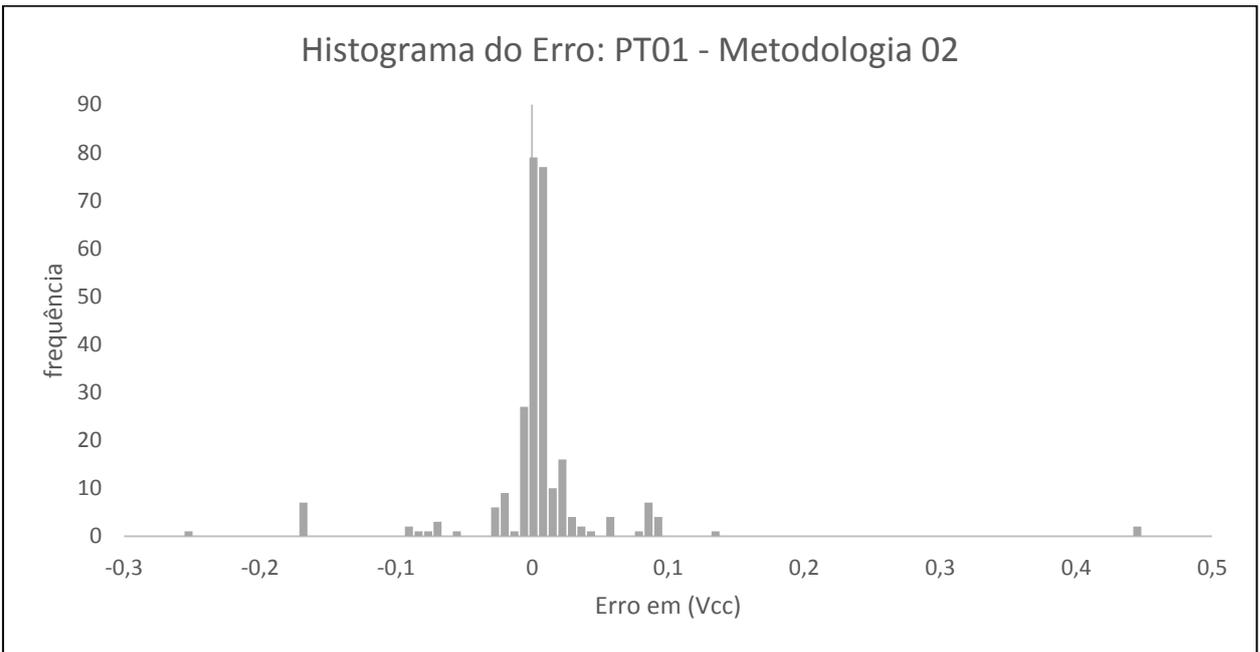


Figura 23 – Gráfico Histograma: PT01, Caso 01, Metodologia 02

5.1.2 Caso1 – Resultado PT02

A Figura 24 apresenta o gráfico com a comparação entre os resultados das duas metodologias e o valor real (alvo) do potencial eletroquímico do PT02 versus o tempo (número de linhas dos conjuntos para o teste).

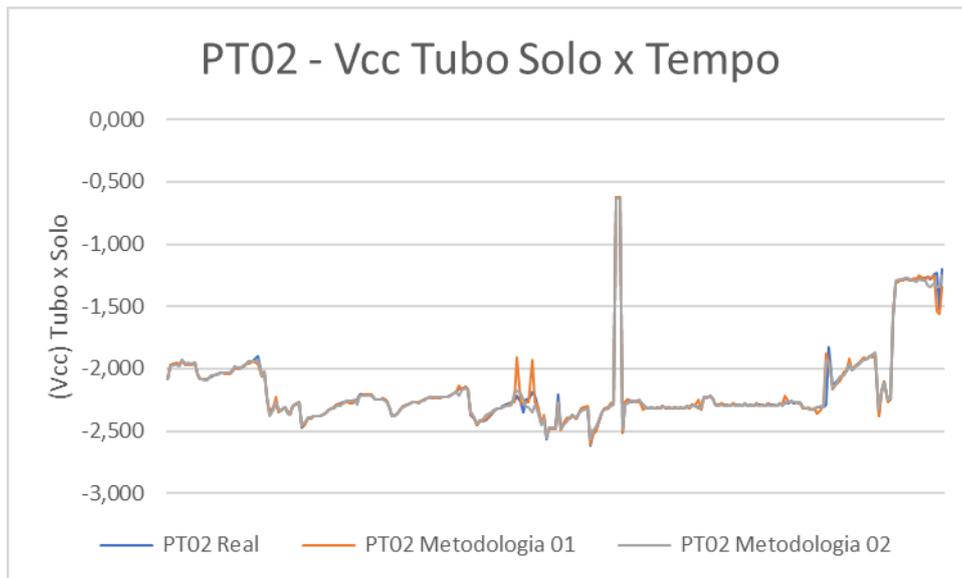


Figura 24 – Gráfico Tubo Solo x Tempo: PT02, Caso 01

A Figura 25 apresenta o gráfico com a dispersão dos resultados das metodologias propostas.

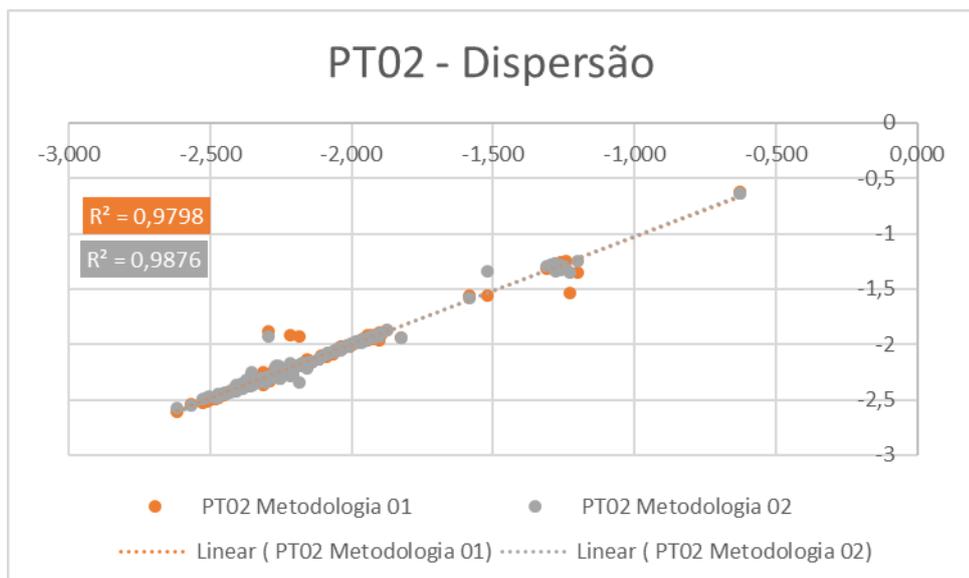


Figura 25 – Gráfico Dispersão: PT02, Caso 01

As Figuras 26 e 27 apresentam os gráficos com o histograma do erro relativo, para cada metodologia proposta.

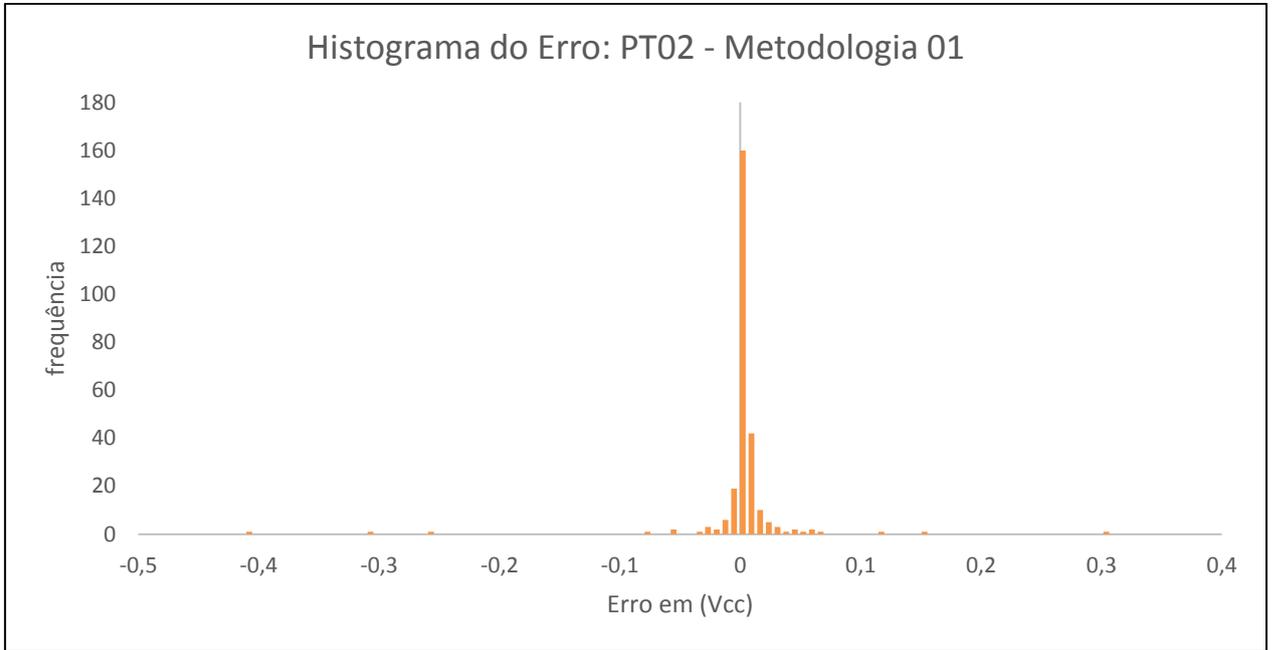


Figura 26 – Gráfico Histograma: PT02, Caso 01, Metodologia 01

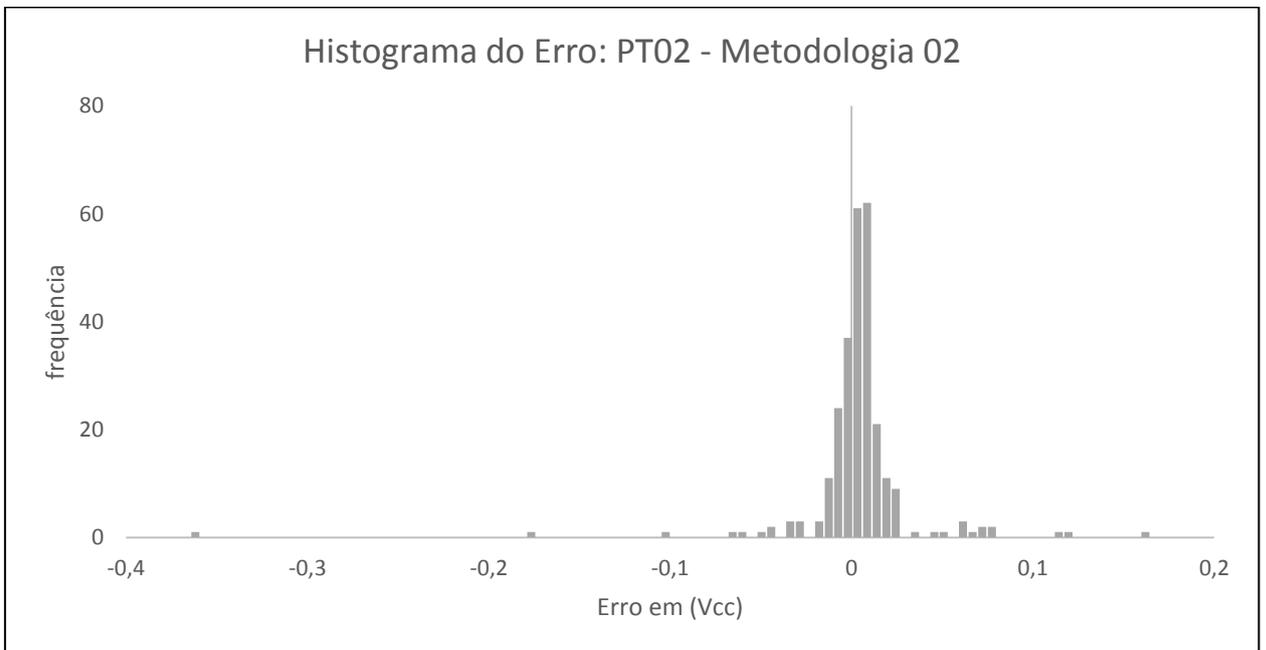


Figura 27 – Gráfico Histograma: PT02, Caso 01, Metodologia 02

5.1.3 Caso1 – Resultado PT03

A Figura 28 apresenta o gráfico com a comparação entre os resultados das duas metodologias e o valor real (alvo) do potencial eletroquímico do PT03 versus o tempo (número de linhas dos conjuntos para o teste).

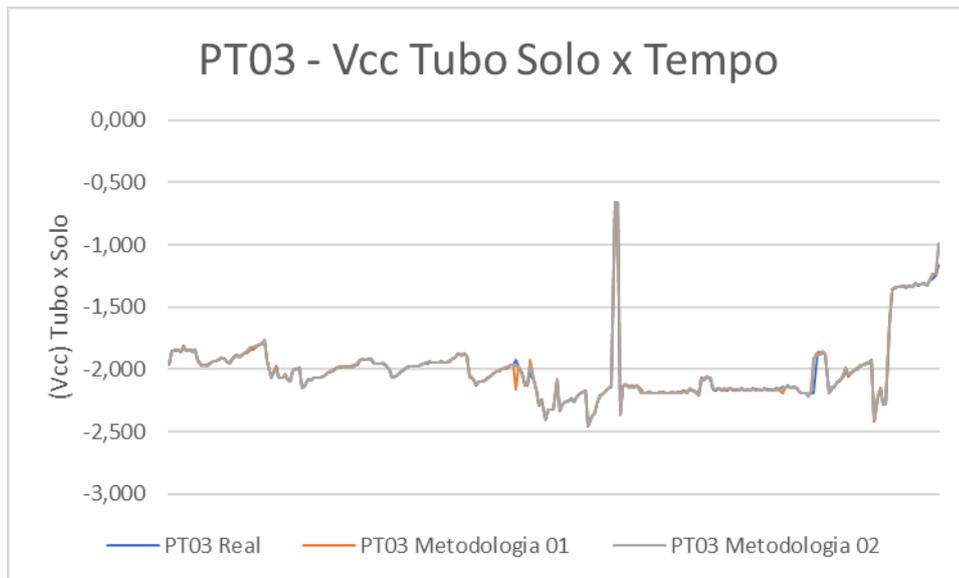


Figura 28 – Gráfico Tubo Solo x Tempo: PT03, Caso 01

A Figura 29 apresenta o gráfico com a dispersão dos resultados das metodologias propostas.

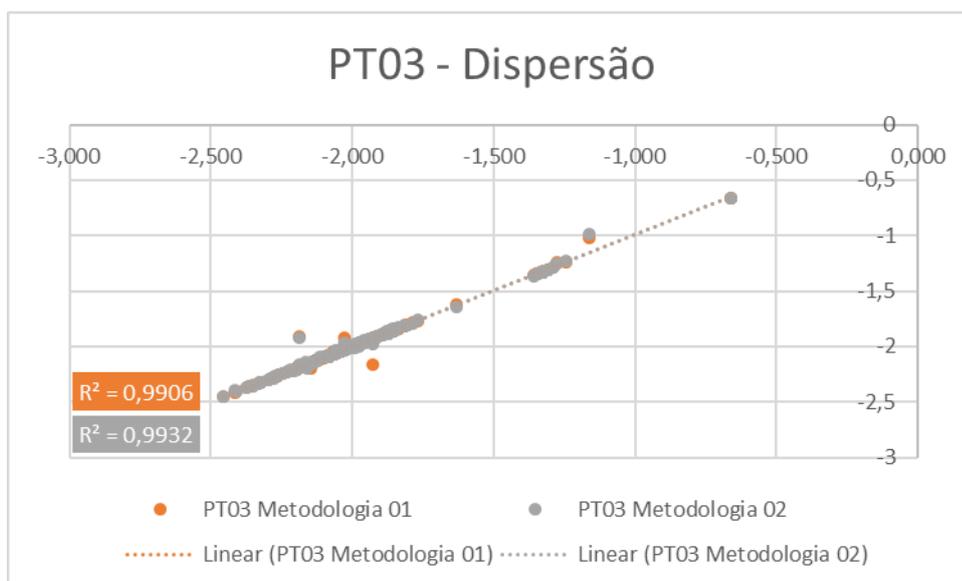


Figura 29 – Gráfico Dispersão: PT03, Caso 01

As Figuras 30 e 31 apresentam os gráficos com o histograma do erro relativo, para cada metodologia proposta.

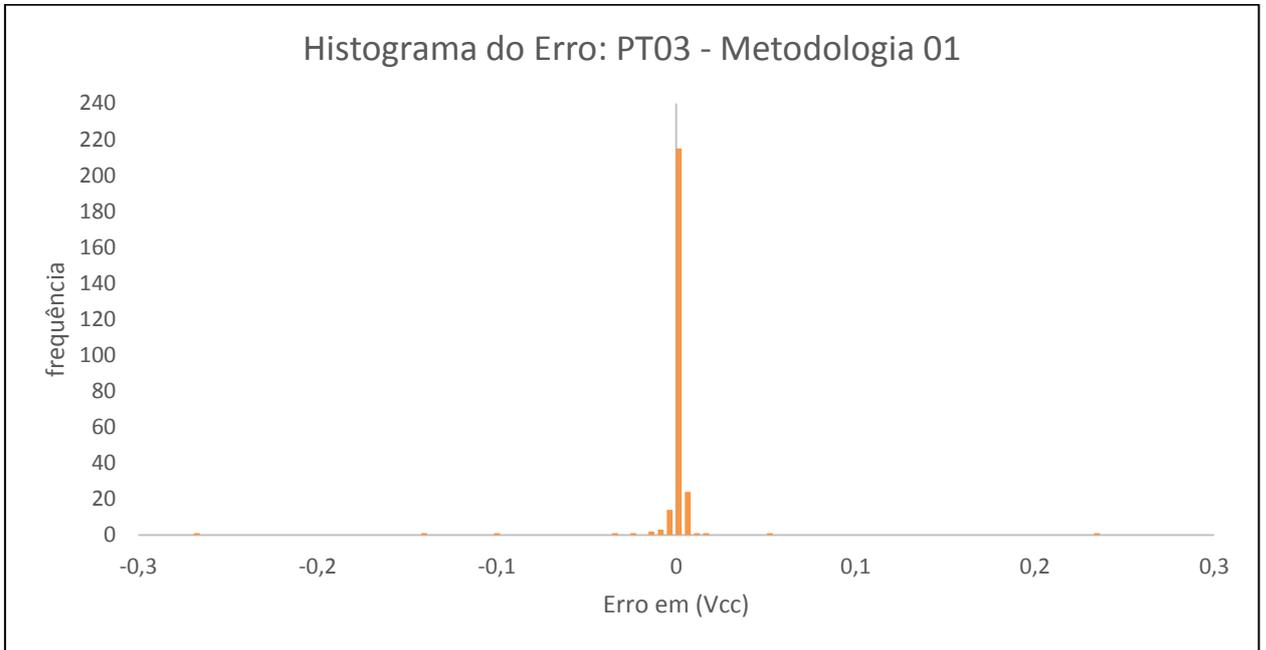


Figura 30 – Gráfico Histograma: PT03, Caso 01, Metodologia 01

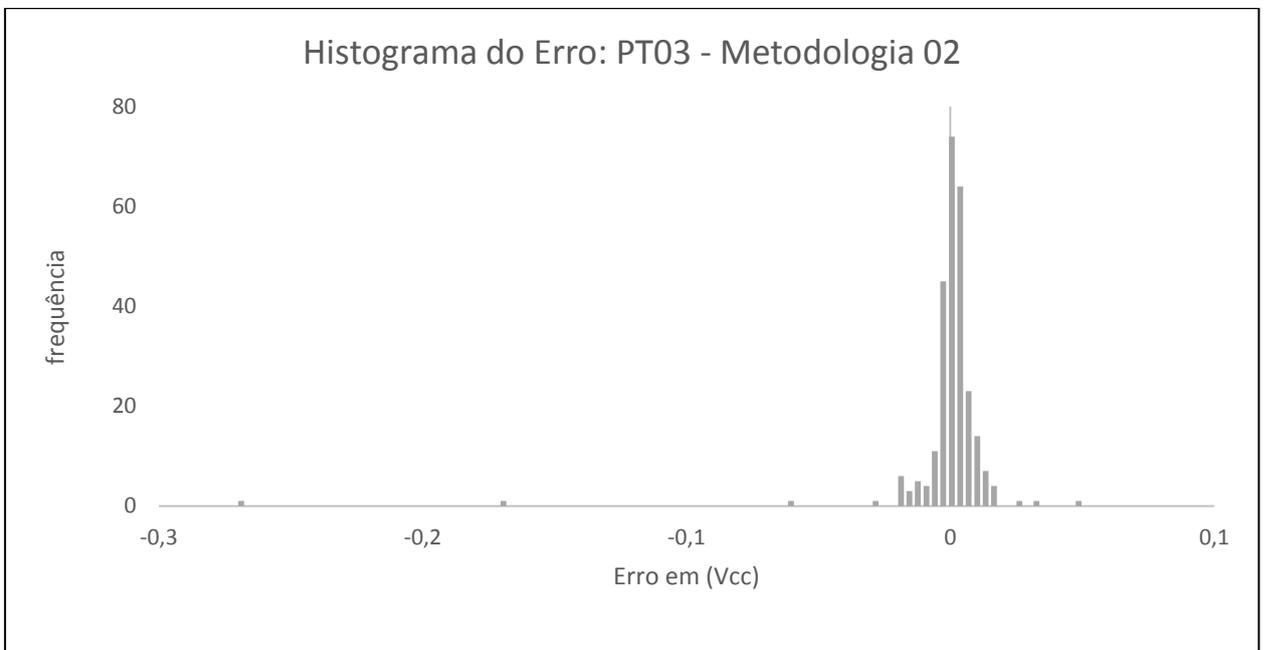


Figura 31 – Gráfico Histograma: PT03, Caso 01, Metodologia 02

5.1.4 Caso1 – Resultado PT04

A Figura 32 apresenta o gráfico com a comparação entre os resultados das duas metodologias e o valor real (alvo) do potencial eletroquímico do PT04 versus o tempo (número de linhas dos conjuntos para o teste).

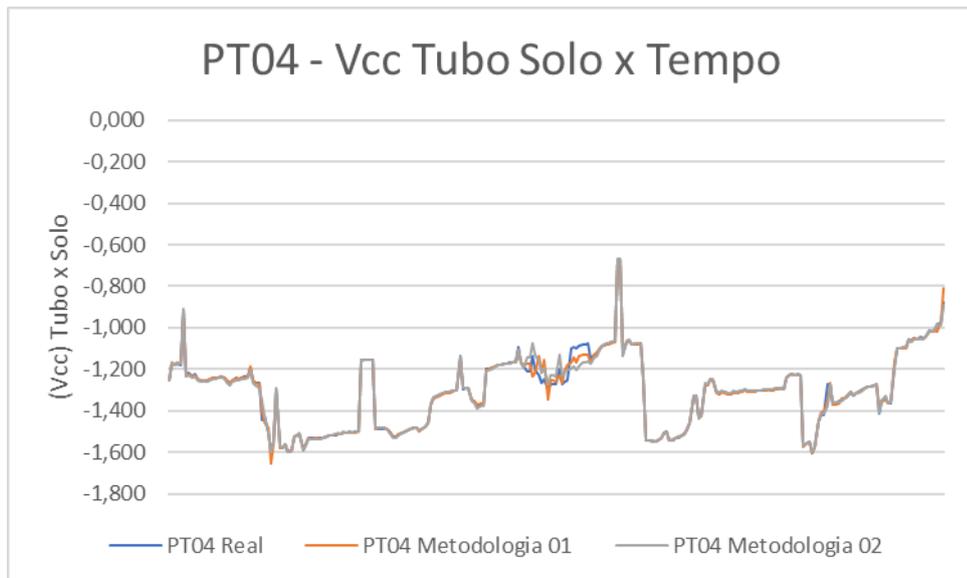


Figura 32 – Gráfico Tubo Solo x Tempo: PT04, Caso 01

A Figura 33 apresenta o gráfico com a dispersão dos resultados das metodologias propostas.

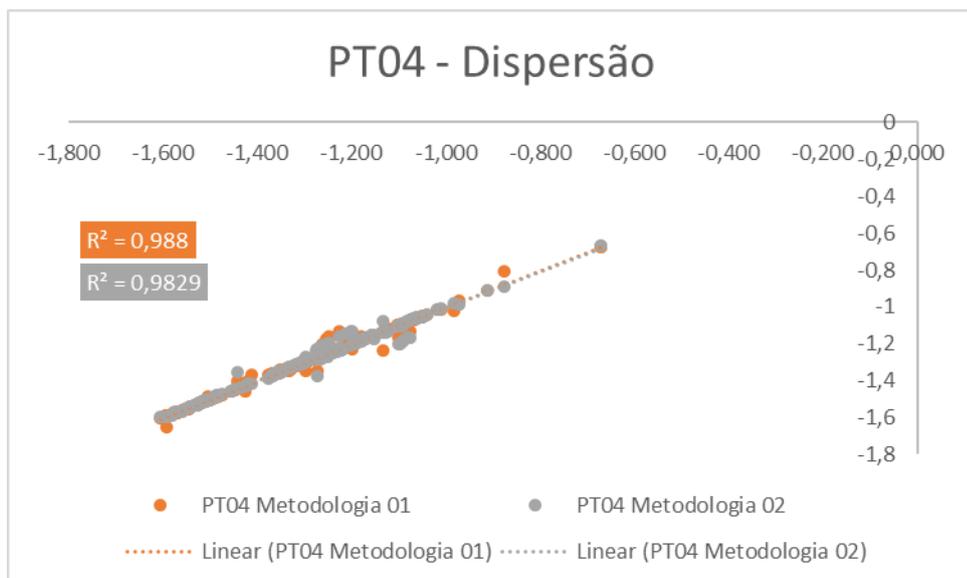


Figura 33 – Gráfico Dispersão: PT04, Caso 01

As Figuras 34 e 35 apresentam os gráficos com o histograma do erro relativo, para cada metodologia proposta.

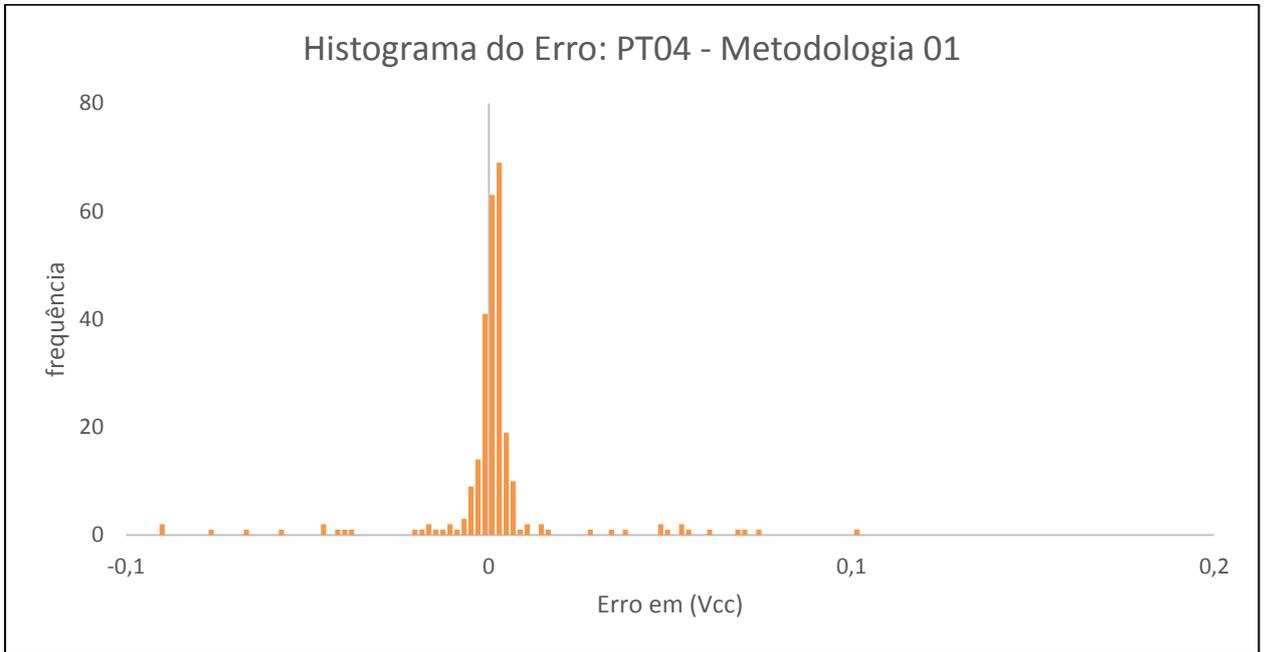


Figura 34 – Gráfico Histograma: PT04, Caso 01, Metodologia 01

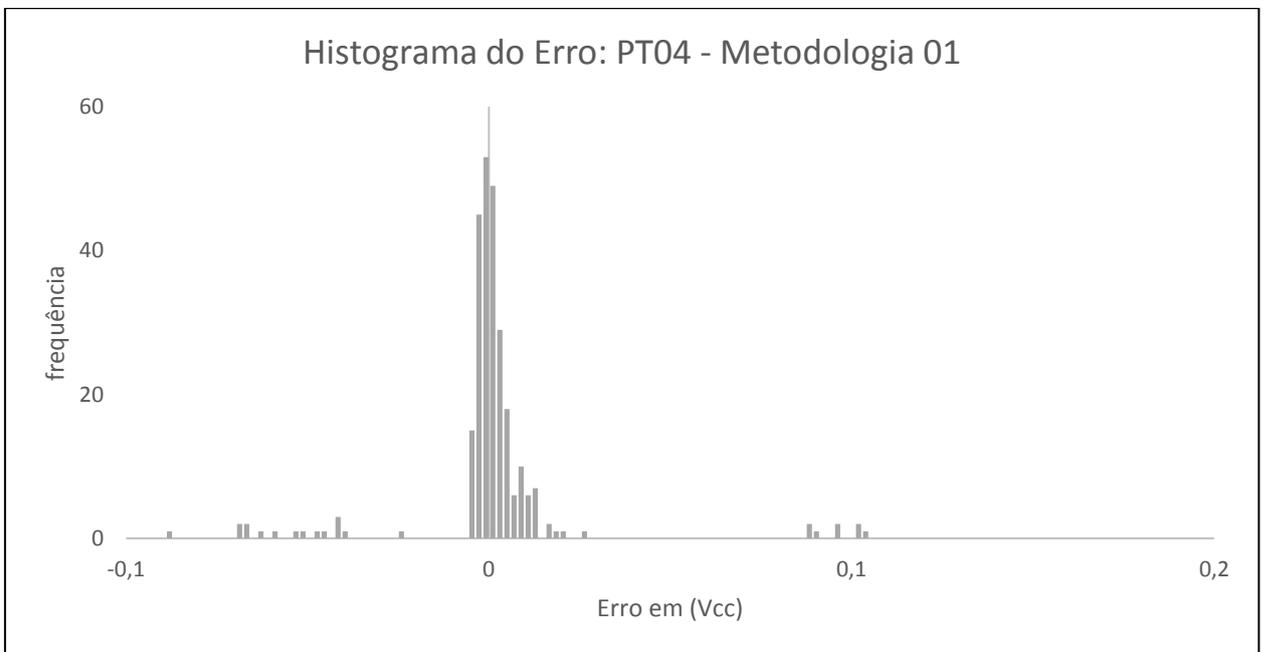


Figura 35 – Gráfico Histograma: PT04, Caso 01, Metodologia 02

5.1.5 Caso1 – Resultado PT05

A Figura 36 apresenta o gráfico com a comparação entre os resultados das duas metodologias e o valor real (alvo) do potencial eletroquímico do PT05 versus o tempo (número de linhas dos conjuntos para o teste).

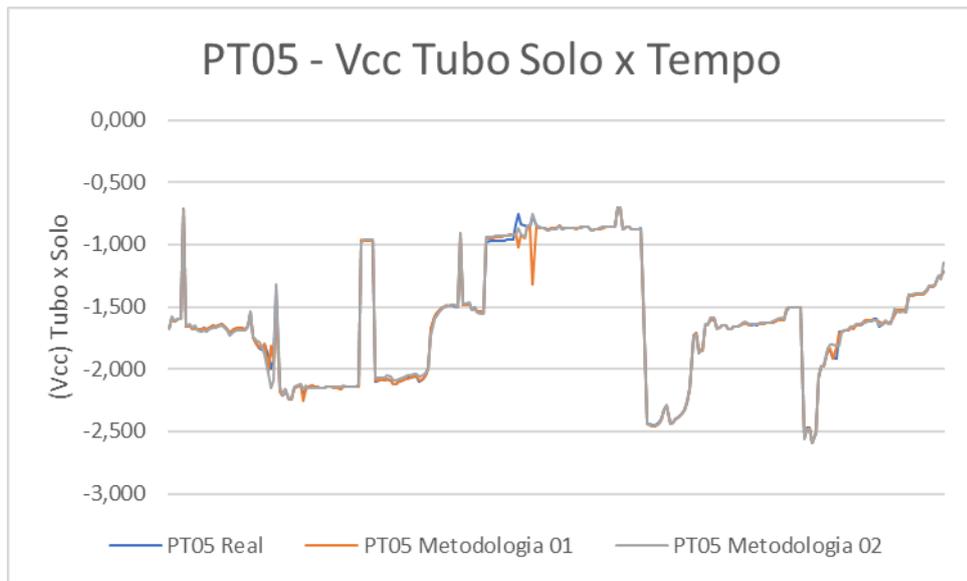


Figura 36 – Gráfico Tubo Solo x Tempo: PT05, Caso 01

A Figura 37 apresenta o gráfico com a dispersão dos resultados das metodologias propostas.

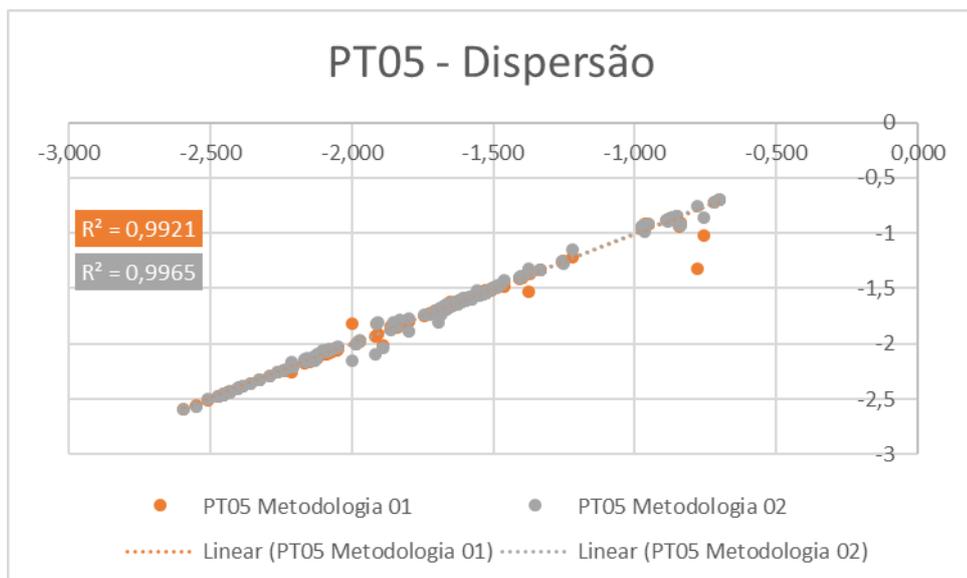


Figura 37 – Gráfico Dispersão: PT05, Caso 01

As Figuras 38 e 39 apresentam os gráficos com o histograma do erro relativo, para cada metodologia proposta.

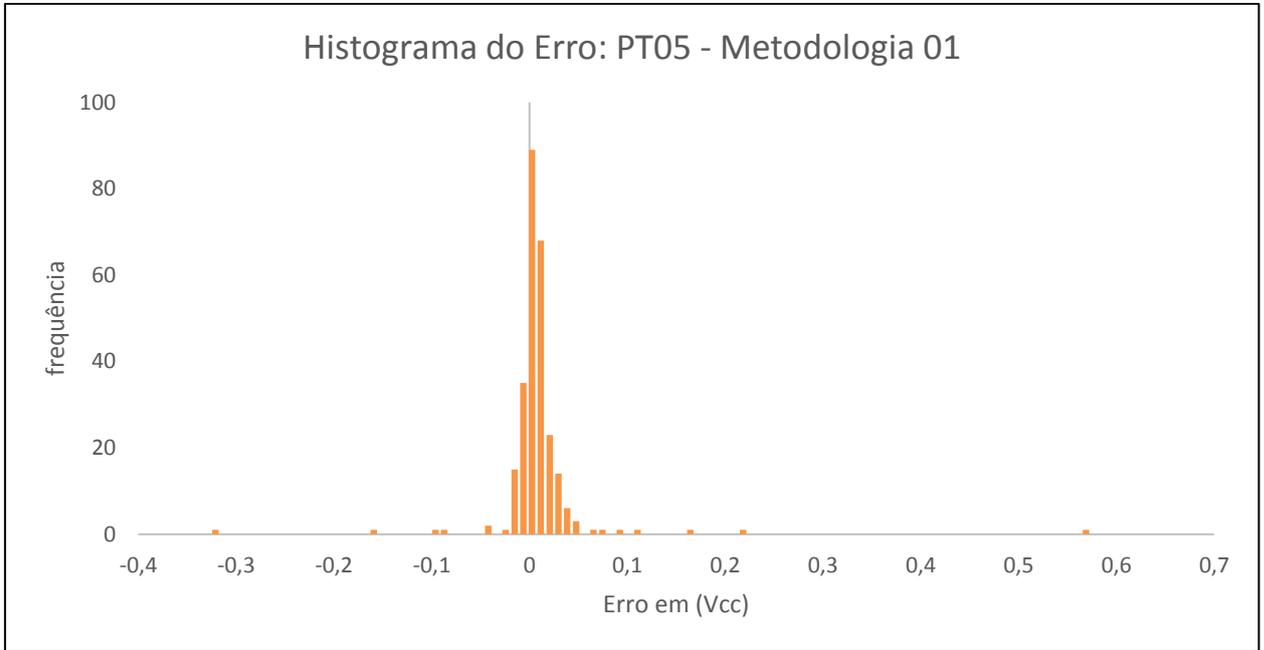


Figura 38 – Gráfico Histograma: PT05, Caso 01, Metodologia 01

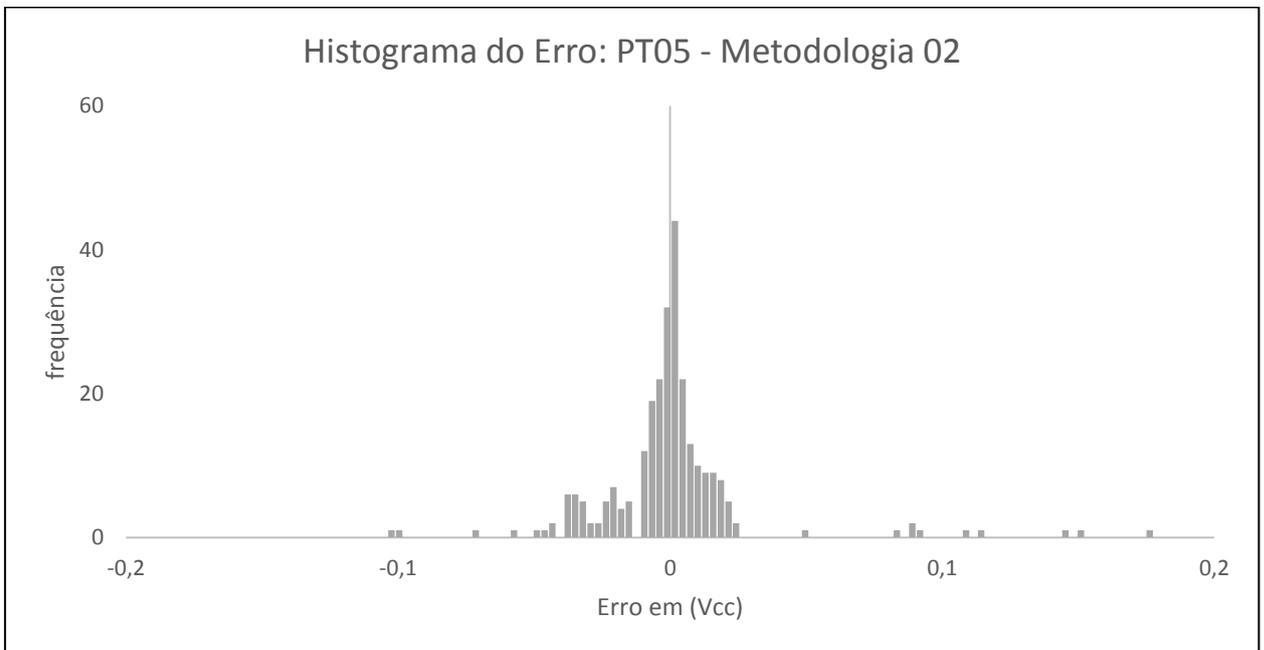


Figura 39 – Gráfico Histograma: PT05, Caso 01, Metodologia 02

5.1.6 Caso1 – Resultado PT06

A Figura 40 apresenta o gráfico com a comparação entre os resultados das duas metodologias e o valor real (alvo) do potencial eletroquímico do PT06 versus o tempo (número de linhas dos conjuntos para o teste).

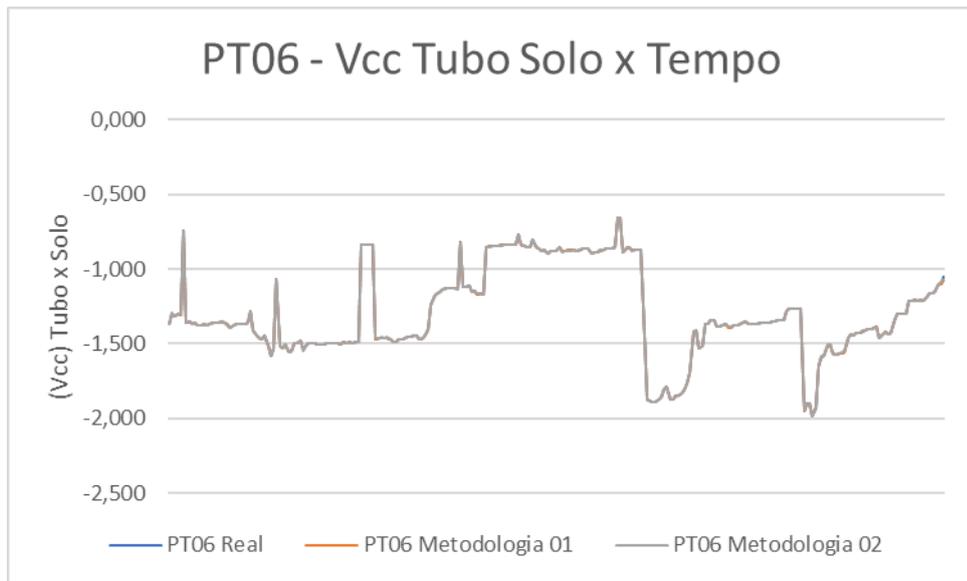


Figura 40 – Gráfico Tubo Solo x Tempo: PT06, Caso 01

A Figura 41 apresenta o gráfico com a dispersão dos resultados das metodologias propostas.

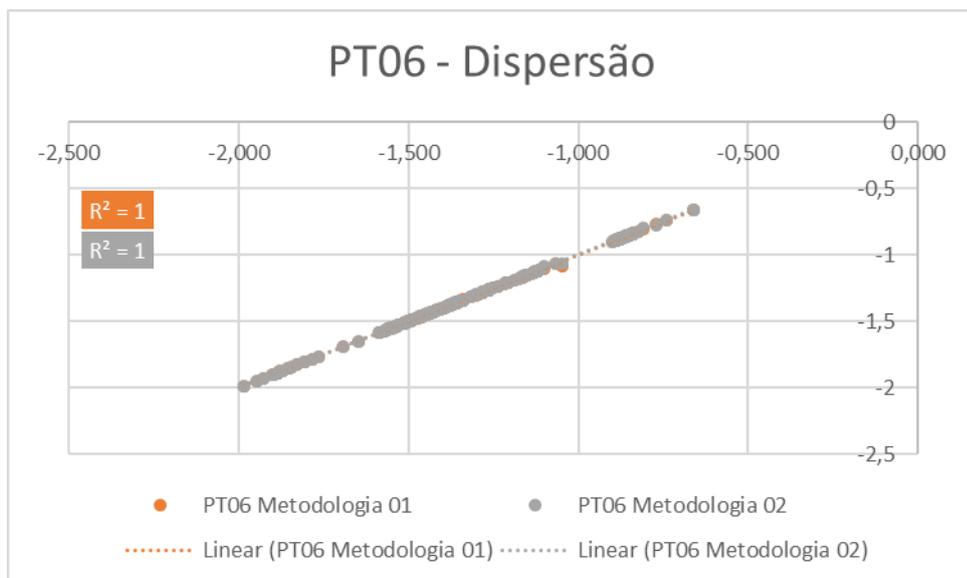


Figura 41 – Gráfico Dispersão: PT06, Caso 01

As Figuras 42 e 43 apresentam os gráficos com o histograma do erro relativo, para cada metodologia proposta.

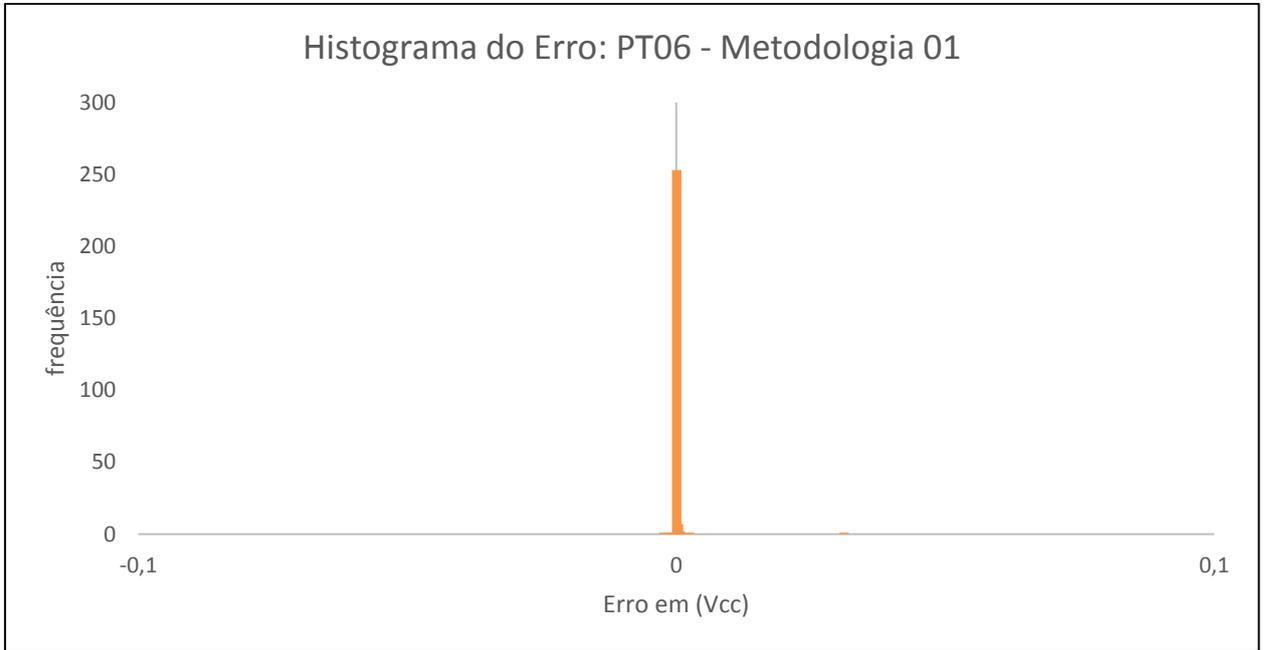


Figura 42 – Gráfico Histograma: PT06, Caso 01, Metodologia 01

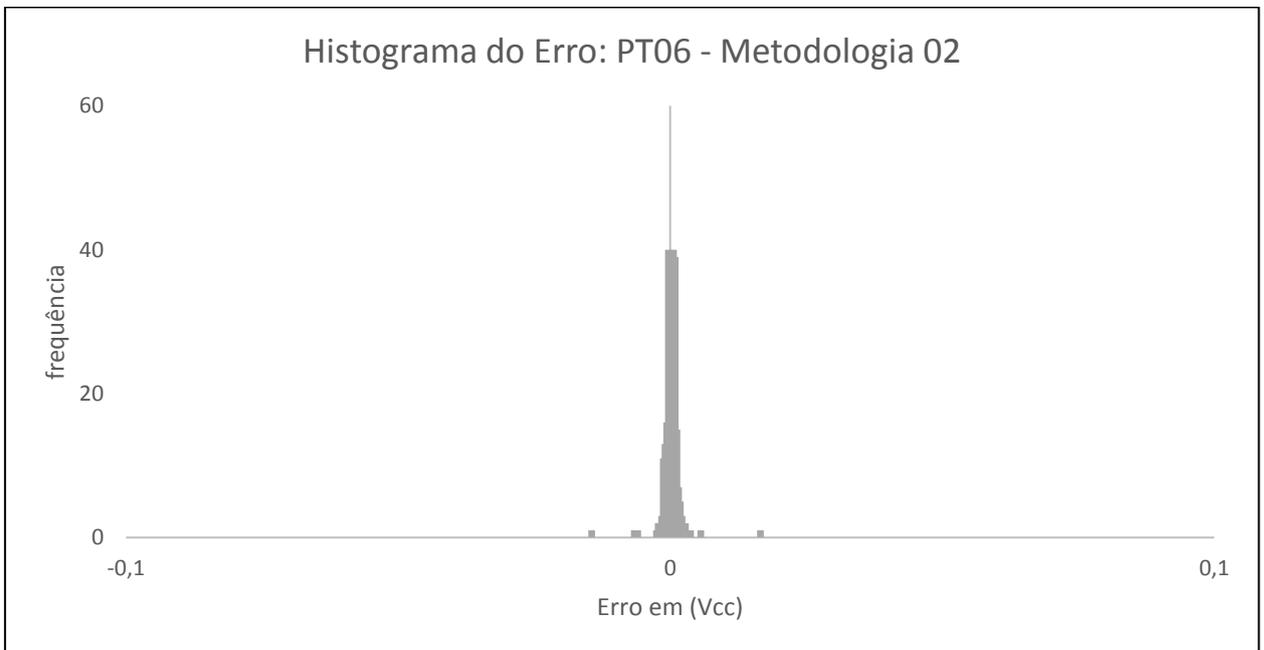


Figura 43 – Gráfico Histograma: PT06, Caso 01, Metodologia 02

5.1.7 Caso1 – Resultado PT07

A Figura 44 apresenta o gráfico com a comparação entre os resultados das duas metodologias e o valor real (alvo) do potencial eletroquímico do PT07 versus o tempo (número de linhas dos conjuntos para o teste).

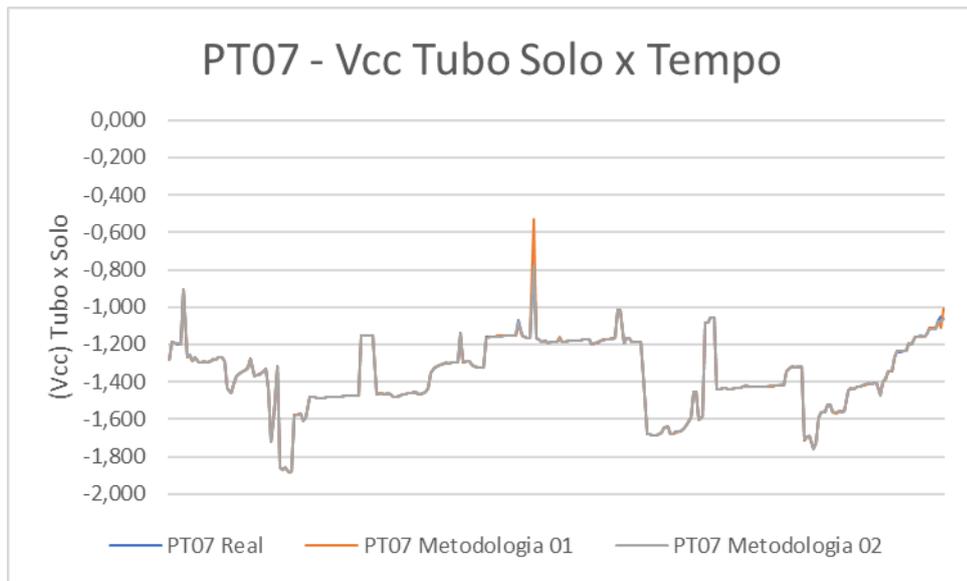


Figura 44 – Gráfico Tubo Solo x Tempo: PT07, Caso 01

A Figura 45 apresenta o gráfico com a dispersão dos resultados das metodologias propostas.

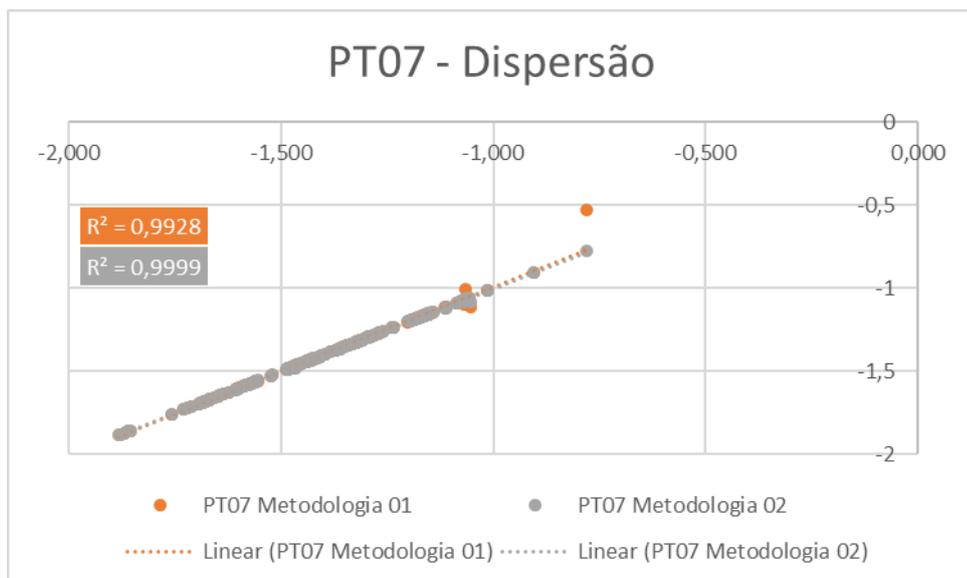


Figura 45 – Gráfico Dispersão: PT07, Caso 01

As Figuras 46 e 47 apresentam os gráficos com o histograma do erro relativo, para cada metodologia proposta.

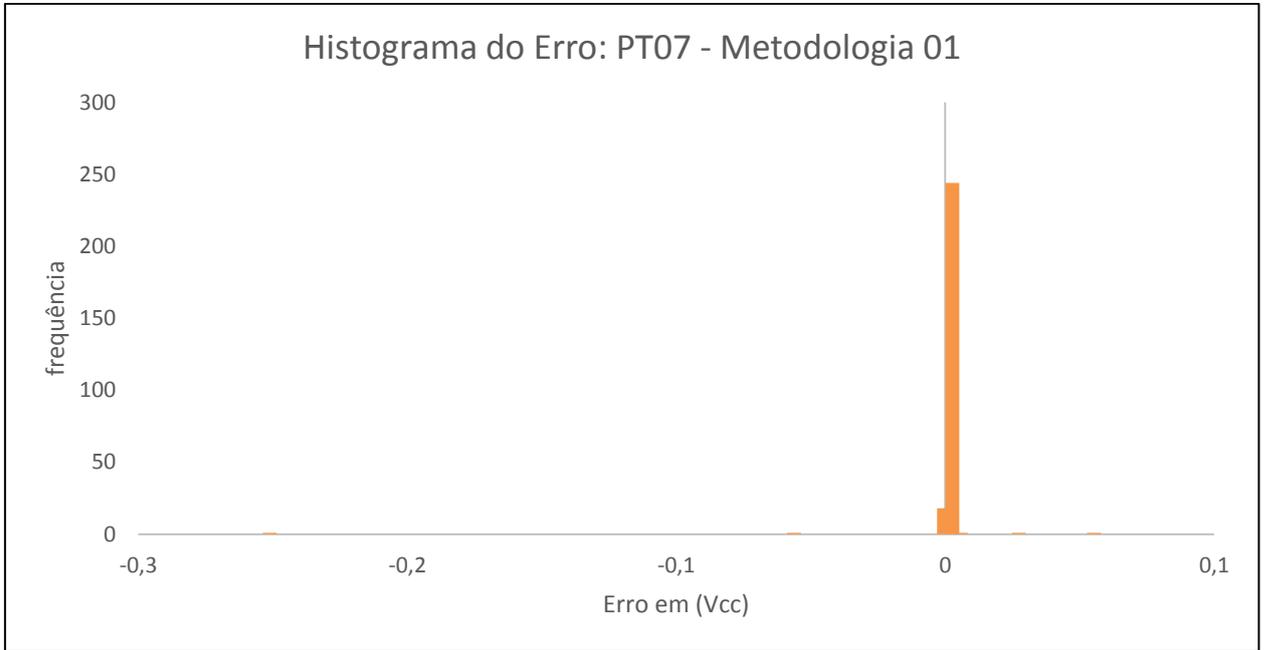


Figura 46 – Gráfico Histograma: PT07, Caso 01, Metodologia 01

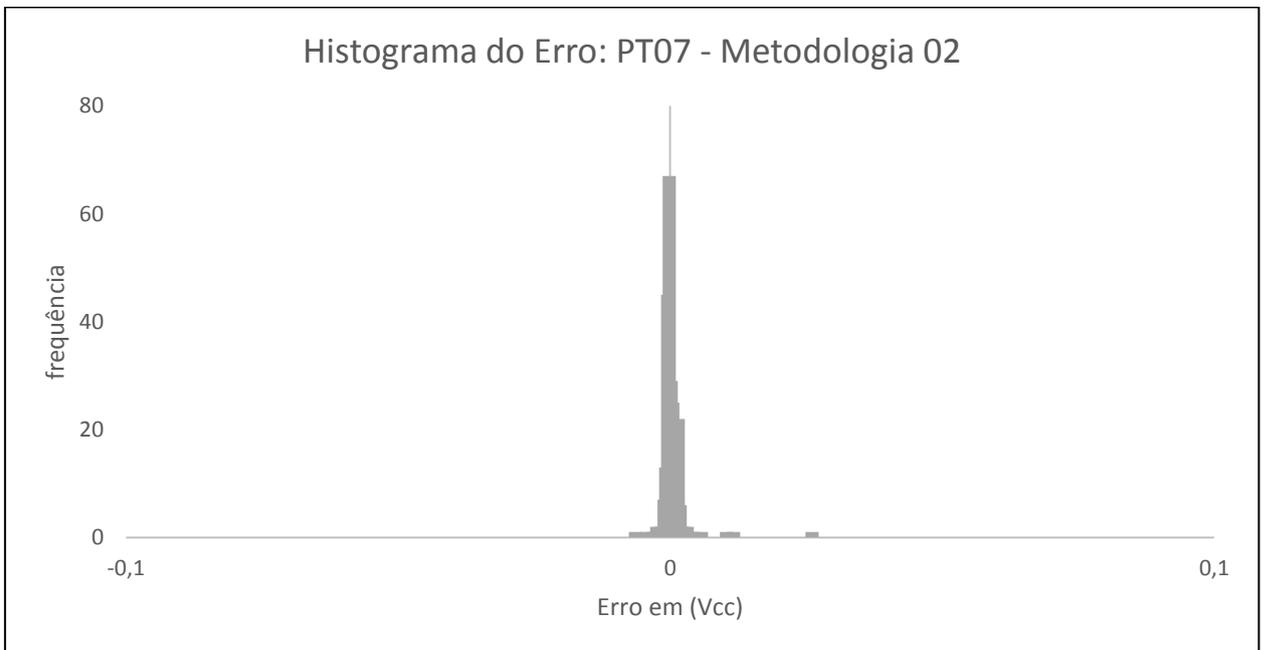


Figura 47 – Gráfico Histograma: PT07, Caso 01, Metodologia 02

5.1.8 Caso1 – Resultado PT08

A Figura 48 apresenta o gráfico com a comparação entre os resultados das duas metodologias e o valor real (alvo) do potencial eletroquímico do PT08 versus o tempo (número de linhas dos conjuntos para o teste).

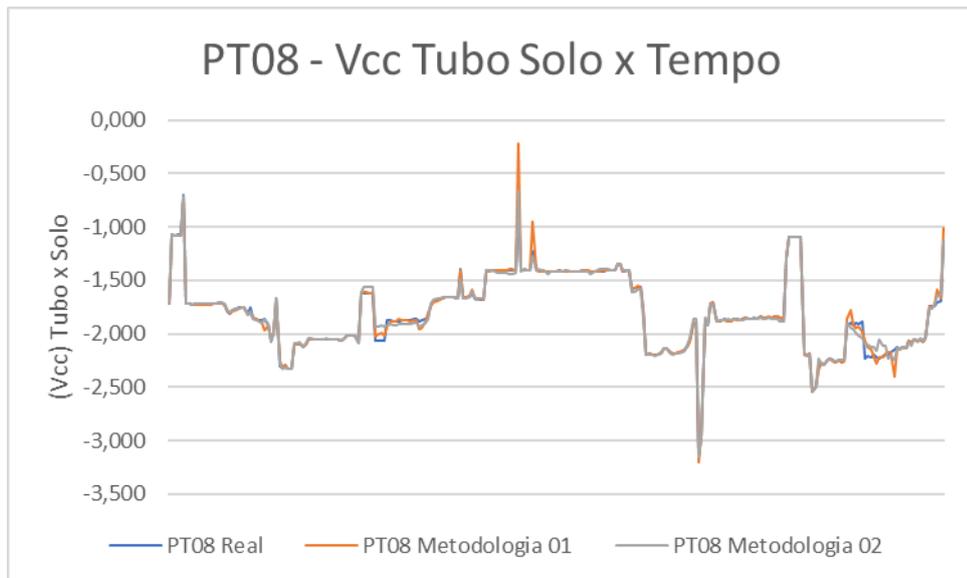


Figura 48 – Gráfico Tubo Solo x Tempo: PT08, Caso 01

A Figura 49 apresenta o gráfico com a dispersão dos resultados das metodologias propostas.

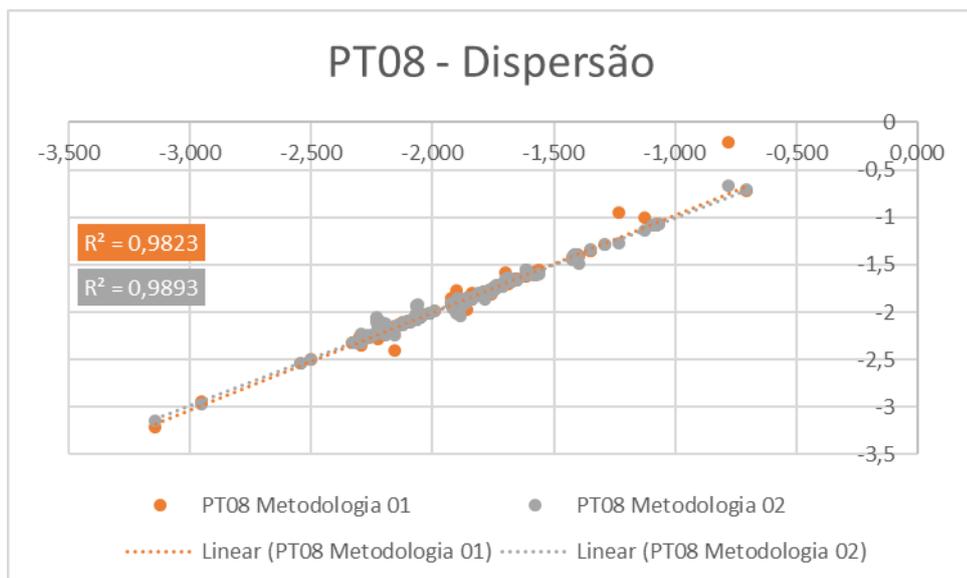


Figura 49 – Gráfico Dispersão: PT08, Caso 01

As Figuras 50 e 51 apresentam os gráfico com o histograma do erro relativo, para cada metodologia proposta.

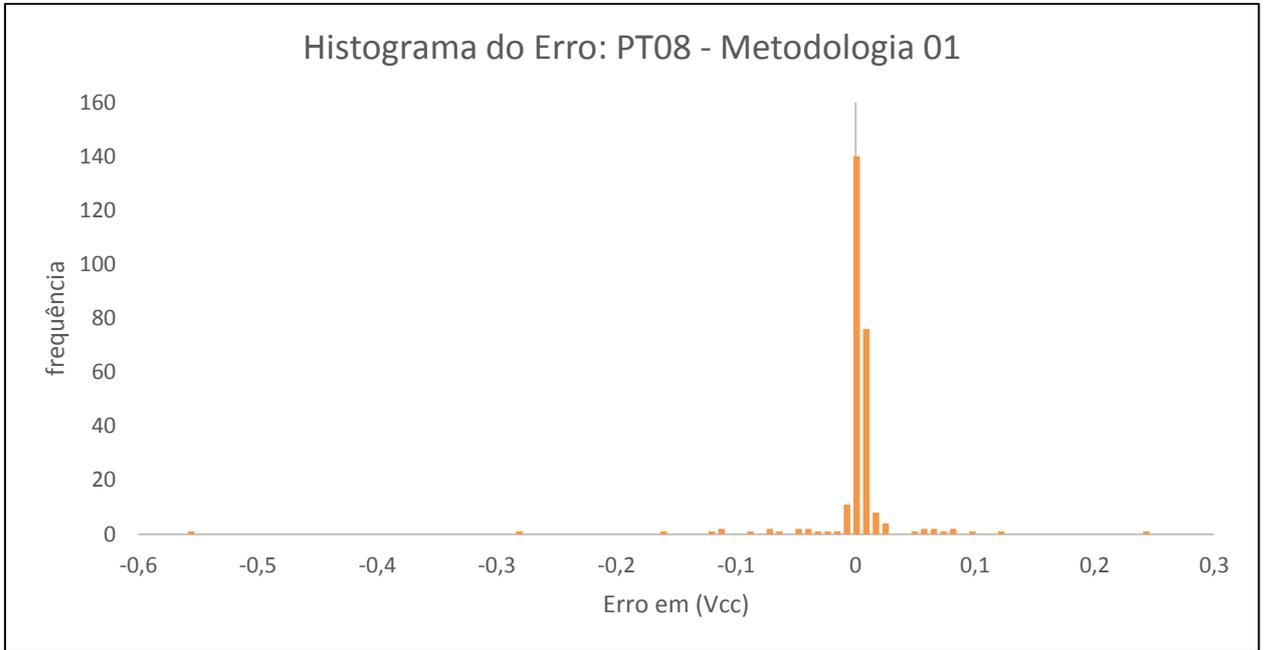


Figura 50 – Gráfico Histograma: PT08, Caso 01, Metodologia 01

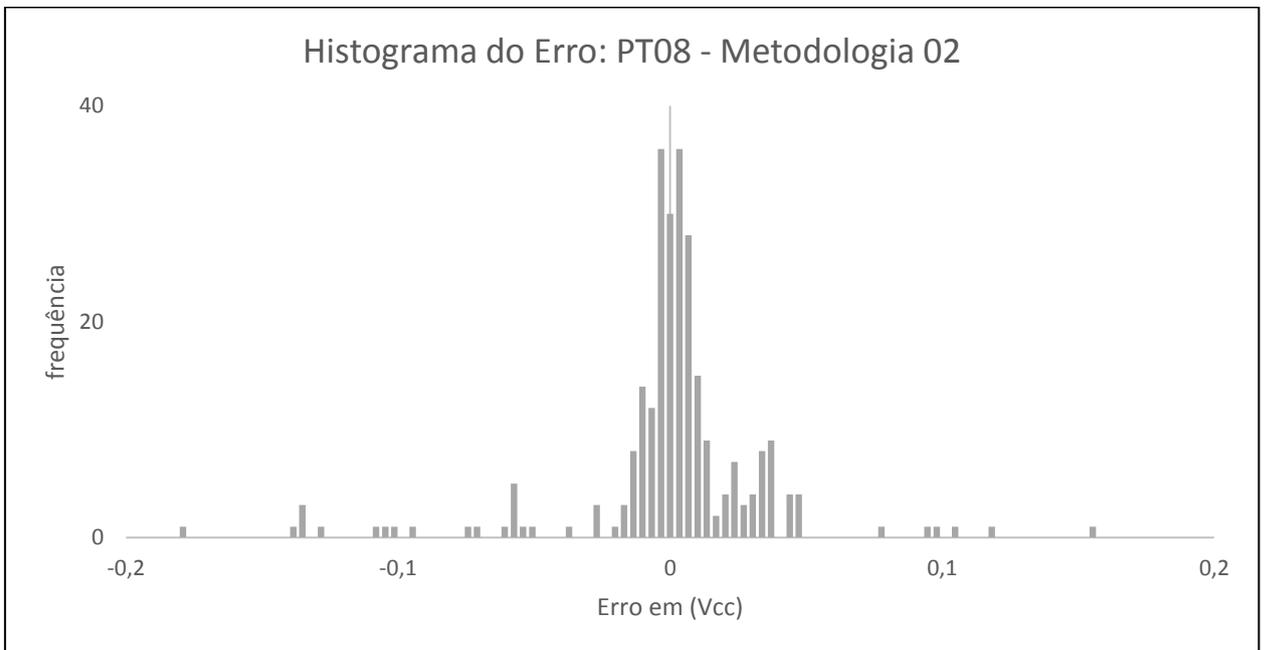


Figura 51 – Gráfico Histograma: PT08, Caso 01, Metodologia 02

5.1.9 Caso1 – Resultado PT09

A Figura 52 apresenta o gráfico com a comparação entre os resultados das duas metodologias e o valor real (alvo) do potencial eletroquímico do PT09 versus o tempo (número de linhas dos conjuntos para o teste).

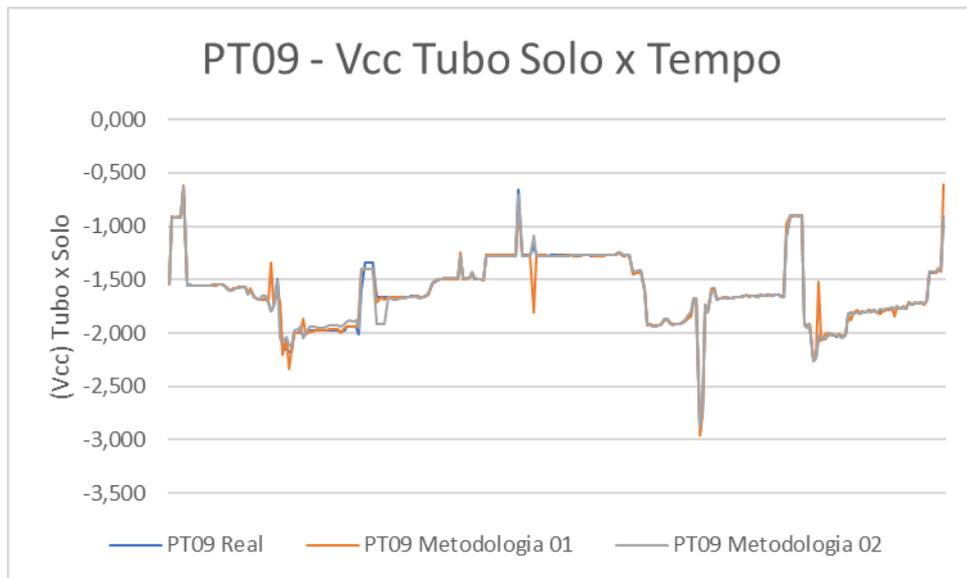


Figura 52 – Gráfico Tubo Solo x Tempo: PT09, Caso 01

A Figura 53 apresenta o gráfico com a dispersão dos resultados das metodologias propostas.

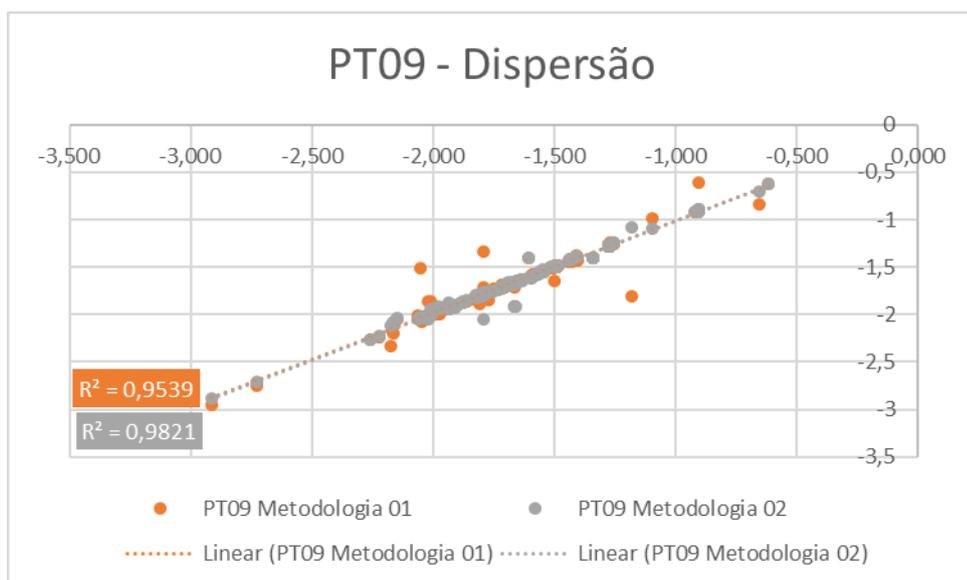


Figura 53 – Gráfico Dispersão: PT09, Caso 01

As Figuras 54 e 55 apresentam os gráficos com o histograma do erro relativo, para cada metodologia proposta.

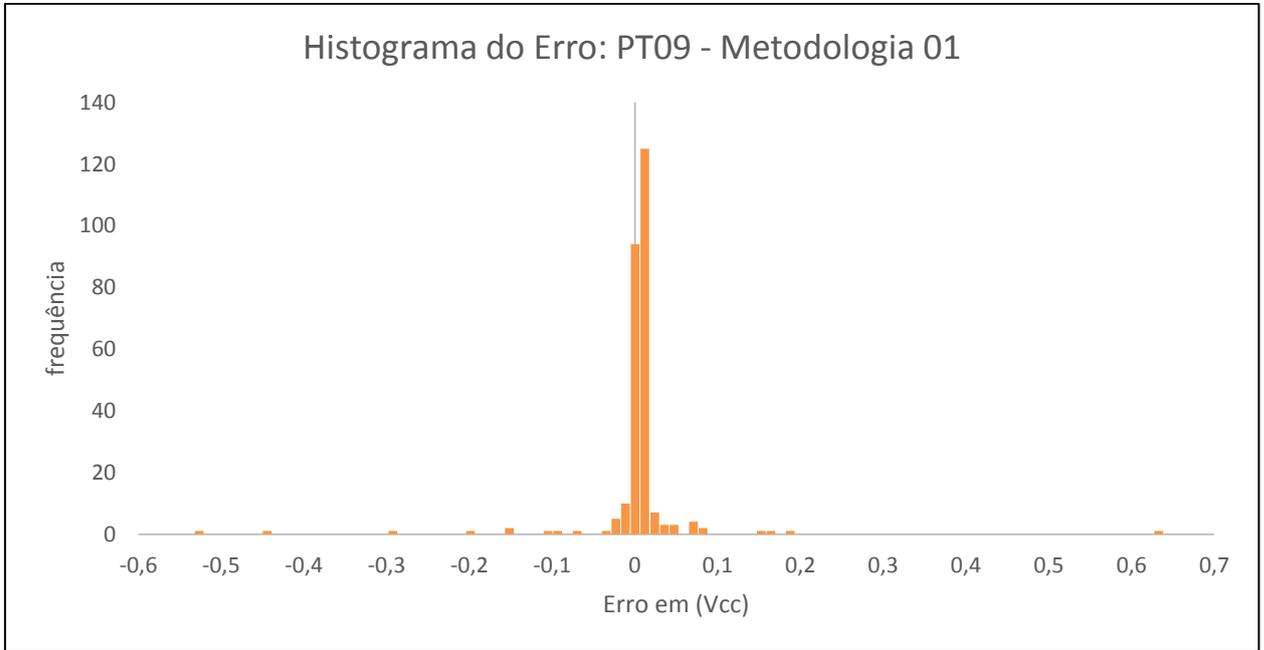


Figura 54 – Gráfico Histograma: PT09, Caso 01, Metodologia 01

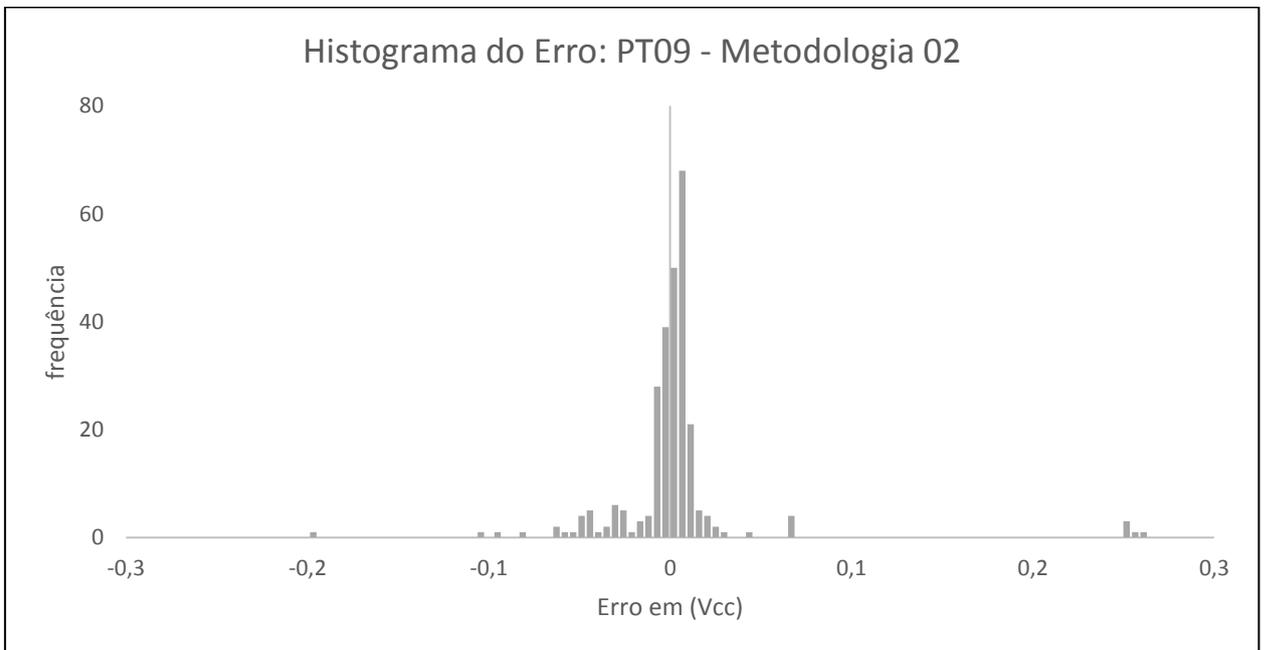


Figura 55 – Gráfico Histograma: PT09, Caso 01, Metodologia 02

5.1.10 Resultado Geral – Caso 01

A Figura 56 apresenta o gráfico com a comparação entre as duas metodologias propostas através do erro absoluto (Ea) médio para cada atributo:

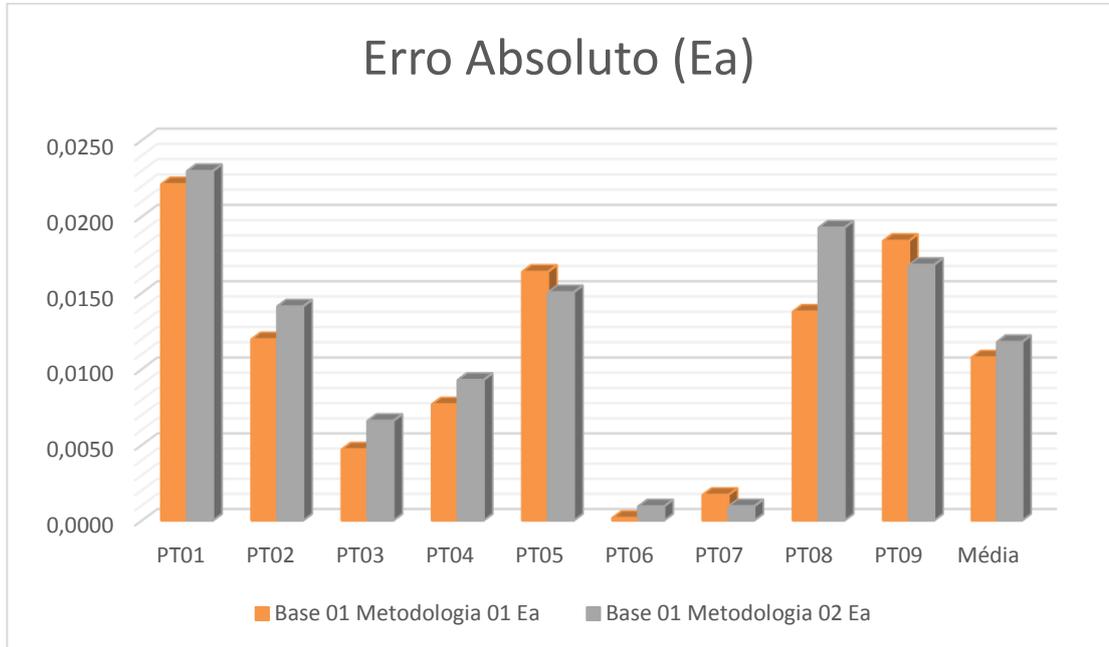


Figura 56 – Erro Absoluto Médio para os resultados do Caso 1

A Figura 57 apresenta o gráfico com a comparação entre as duas metodologias propostas através do erro percentual (Ep) médio para cada atributo:

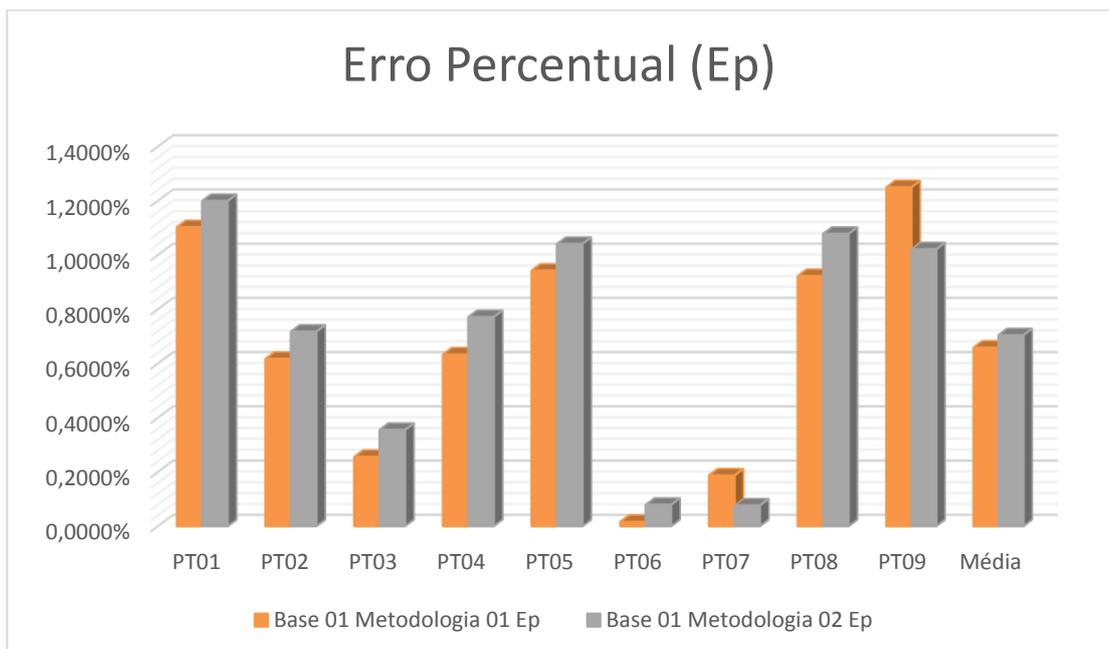


Figura 57 – Erro Percentual Médio para os resultados do Caso 1

A Figura 58 apresenta a função R^2 ao quadrado, comparando os resultados das duas metodologias propostas para cada atributo:

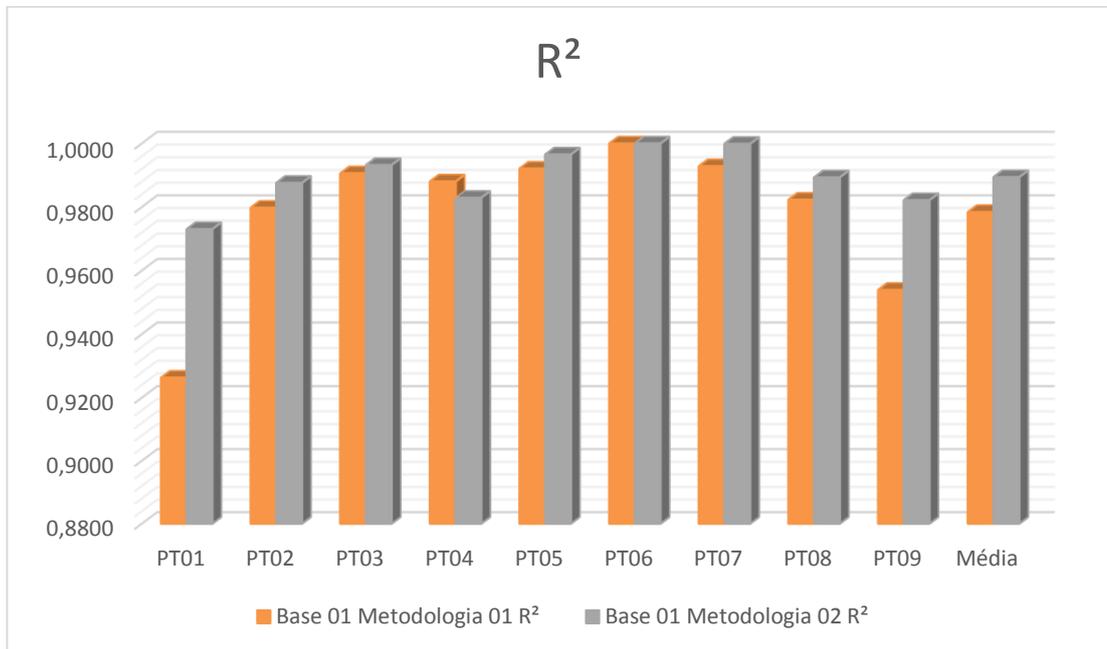


Figura 58 – Comparação da função R^2 para os resultados do Caso 1

A Figura 59 apresenta o desvio padrão (σ), comparando os resultados das duas metodologias propostas para cada atributo:

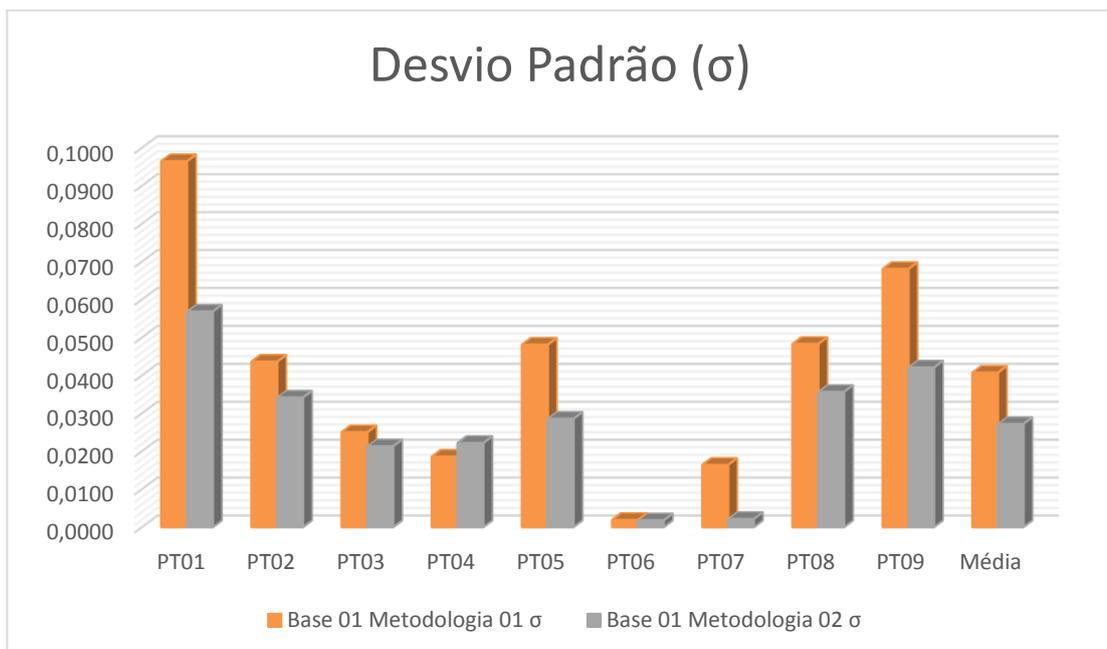


Figura 59 – Comparação do Desvio Padrão para os resultados do Caso 1

5.2 Resultados Caso 02

Para o caso 02, foram utilizados 601.336 dados de 8 atributos (3 retificadores e 5 pontos de teste).

Foram realizadas 16 simulações, 8 utilizando a metodologia 01 e 8 utilizando a metodologia 02. A variável de saída foi sendo modificada ao passo de cada simulação realizada. Segue na Tabela 9 os resultados com as comparações dos erros médios, a partir da base teste, por metodologia.

Tabela 9 – Resultados gerais obtidos nas metodologias 01 e 02 para o caso 02

Caso 02								
Saída	Metodologia 01				Metodologia 02			
	Ea	Ep	R ²	σ	Ea	Ep	R ²	σ
PT01	0,1613	14,41%	0,5125	0,2372	0,0774	6,71%	0,86	0,1271
RET01	0,3182	14,86%	0,381	0,3456	0,0546	2,49%	0,9687	0,0716
PT02	0,0588	1,68%	0,9025	0,0953	0,0518	1,47%	0,9186	0,087
RET02	0,0288	11,45%	0,0832	0,0352	0,0112	4,52%	0,6248	0,015
PT03	0,0436	2,01%	0,9076	0,0448	0,0204	0,93%	0,9638	0,0281
PT04	0,1	9,83%	0,2028	0,1335	0,0873	8,37%	0,2901	0,1256
PT05	0,098	4,83%	0,9487	0,0751	0,0286	1,47%	0,9848	0,0404
RET03	0,0727	38,55%	0,8594	0,0897	0,0173	9,52%	0,9838	0,0275
Média	0,1102	12,20%	0,5997	0,132	0,0436	4,44%	0,8243	0,0653

Como os resultados possuem saídas com magnitudes diferentes, excluiremos da análise os resultados da saída de corrente dos retificadores. Desta forma, segue na Tabela 10 os resultados considerando apenas os pontos de teste:

Tabela 10 – Resultados considerando apenas os atributos “ponto de teste”

Caso 02 – Somente ponto de teste								
Saída	Metodologia 01				Metodologia 02			
	Ea	Ep	R ²	σ	Ea	Ep	R ²	σ
PT01	0,1613	14,41%	0,5125	0,2372	0,0774	6,71%	0,86	0,1271
PT02	0,0588	1,68%	0,9025	0,0953	0,0518	1,47%	0,9186	0,087
PT03	0,0436	2,01%	0,9076	0,0448	0,0204	0,93%	0,9638	0,0281
PT04	0,1	9,83%	0,2028	0,1335	0,0873	8,37%	0,2901	0,1256
PT05	0,098	4,83%	0,9487	0,0751	0,0286	1,47%	0,9848	0,0404
Média	0,0923	6,55%	0,6948	0,1172	0,0531	3,79%	0,8034	0,0816

5.2.1 Caso2 – Resultado PT01

A Figura 60 apresenta o gráfico com a comparação entre os resultados das duas metodologias e o valor real (alvo) do potencial eletroquímico do PT01 versus o tempo (número de linhas dos conjuntos para o teste).

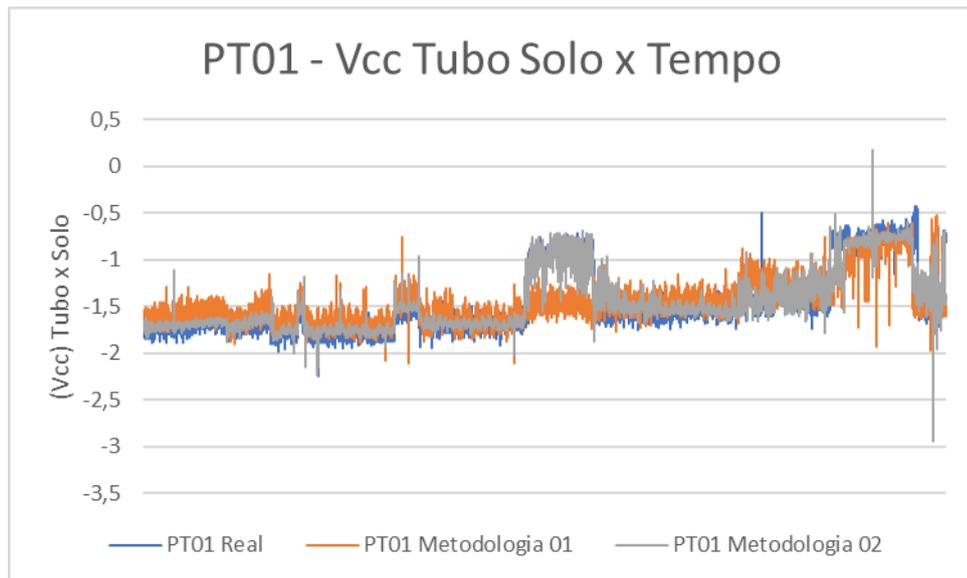


Figura 60 – Gráfico Tubo Solo x Tempo: PT01, Caso 02

A Figura 61 apresenta o gráfico com a dispersão dos resultados das metodologias propostas.

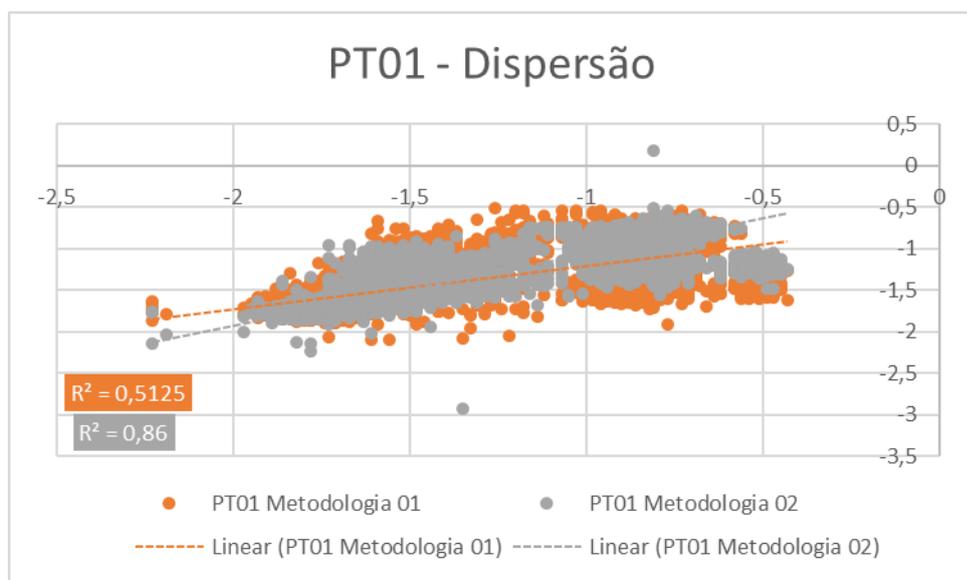


Figura 61 – Gráfico Dispersão: PT01, Caso 02

As Figuras 62 e 63 apresentam os gráficos com o histograma do erro relativo, para cada metodologia proposta.

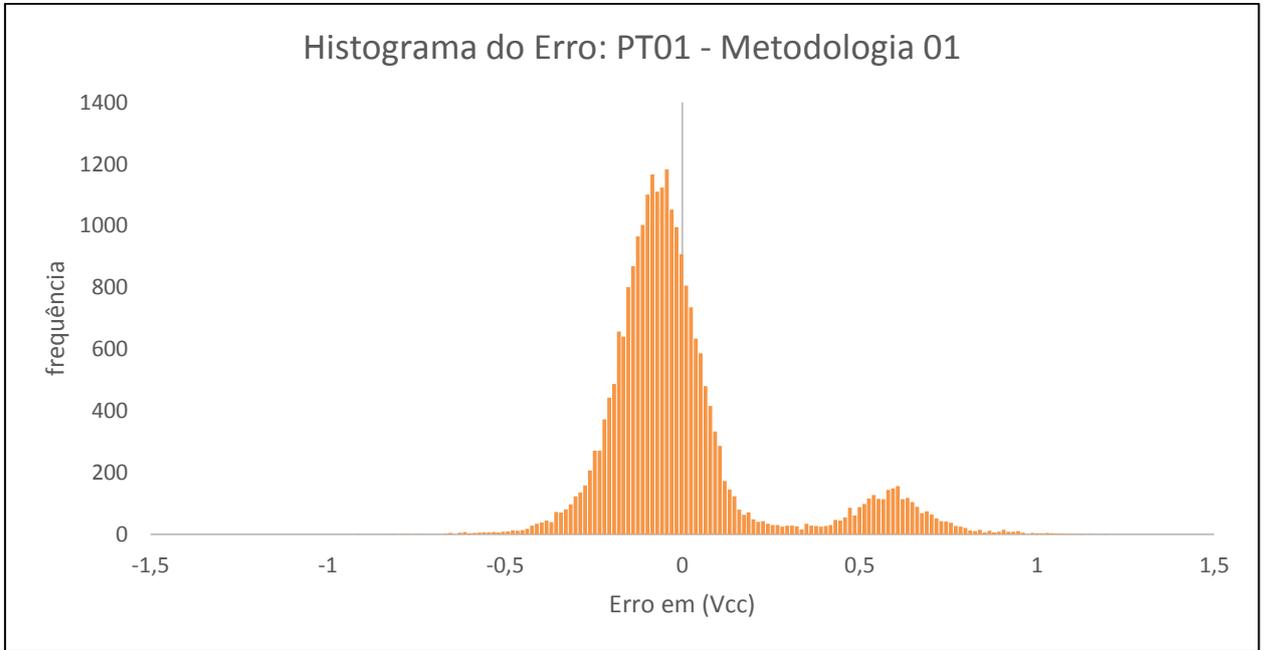


Figura 62 – Gráfico Histograma: PT01, Caso 02, Metodologia 01

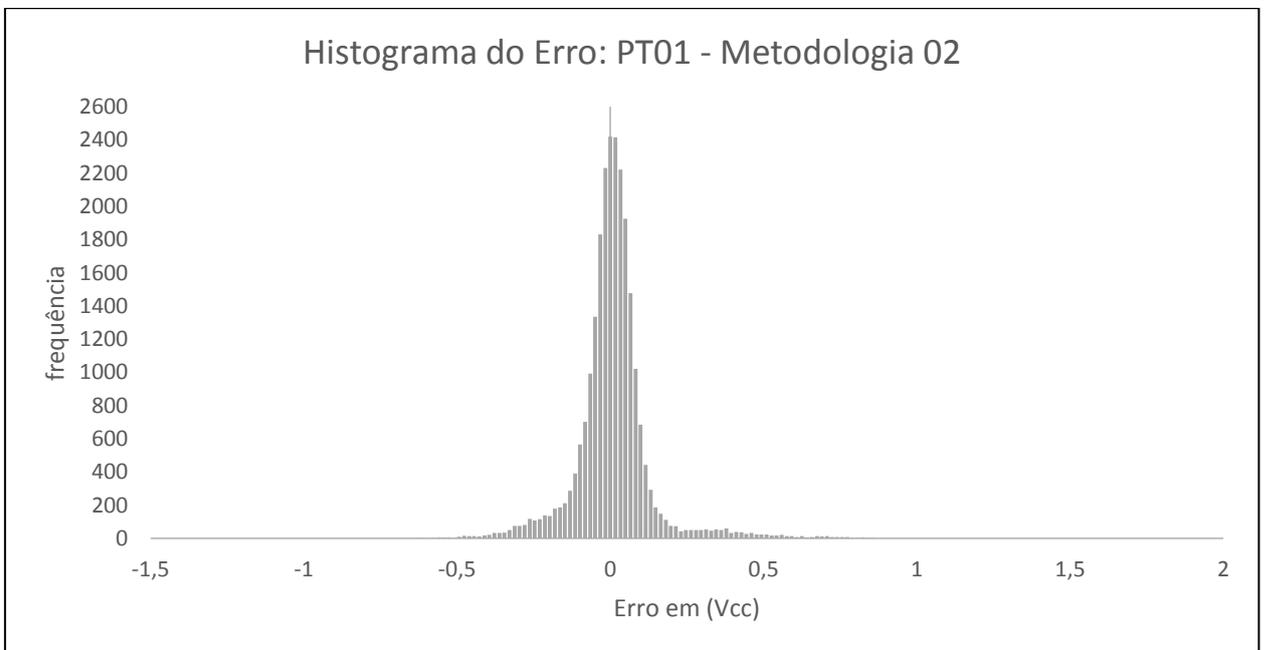


Figura 63 – Gráfico Histograma: PT01, Caso 02, Metodologia 02

5.2.2 Caso2 – Resultado PT02

A Figura 64 apresenta o gráfico com a comparação entre os resultados das duas metodologias e o valor real (alvo) do potencial eletroquímico do PT02 versus o tempo (número de linhas dos conjuntos para o teste).

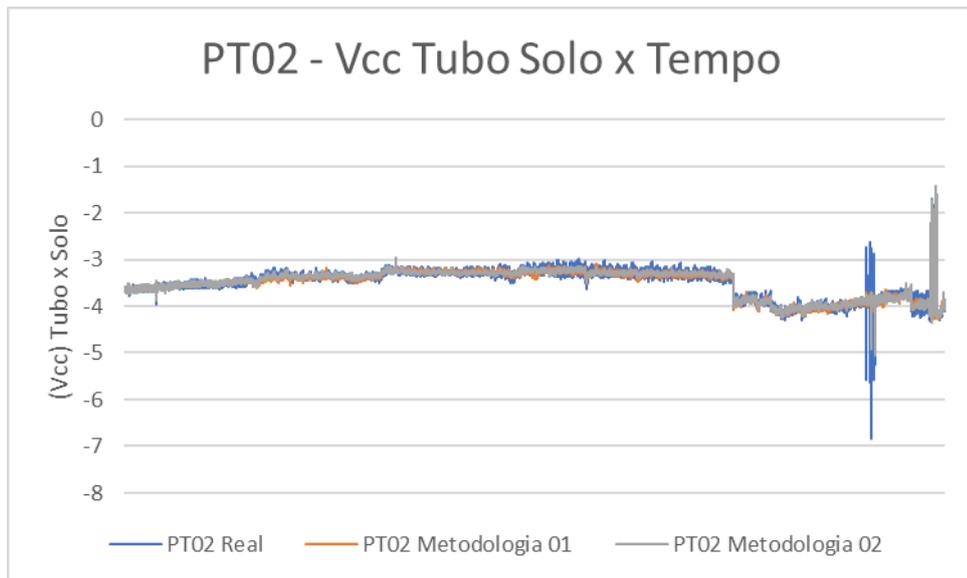


Figura 64 – Gráfico Tubo Solo x Tempo: PT02, Caso 02

A Figura 65 apresenta o gráfico com a dispersão dos resultados das metodologias propostas.

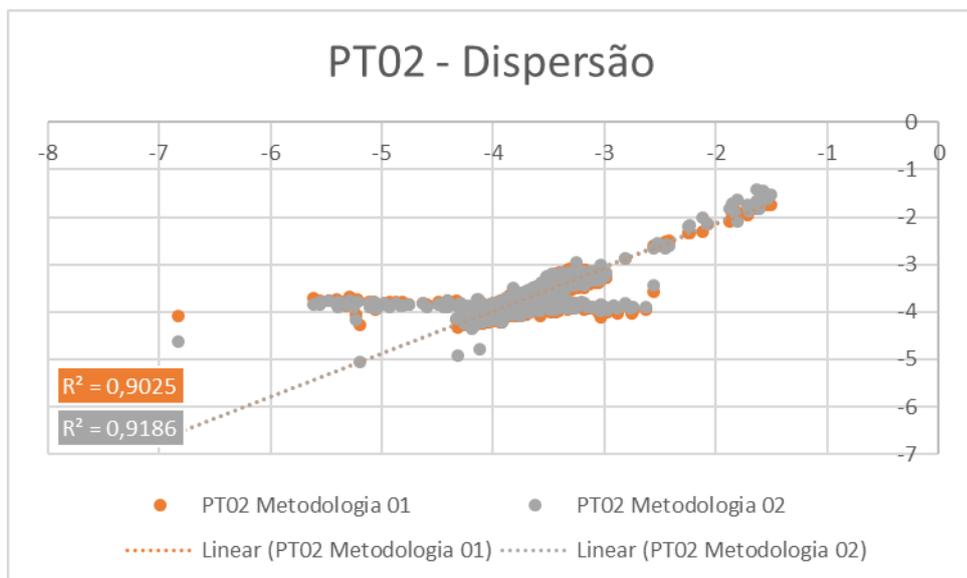


Figura 65 – Gráfico Dispersão: PT02, Caso 02

As Figuras 66 e 67 apresentam os gráficos com o histograma do erro relativo, para cada metodologia proposta.

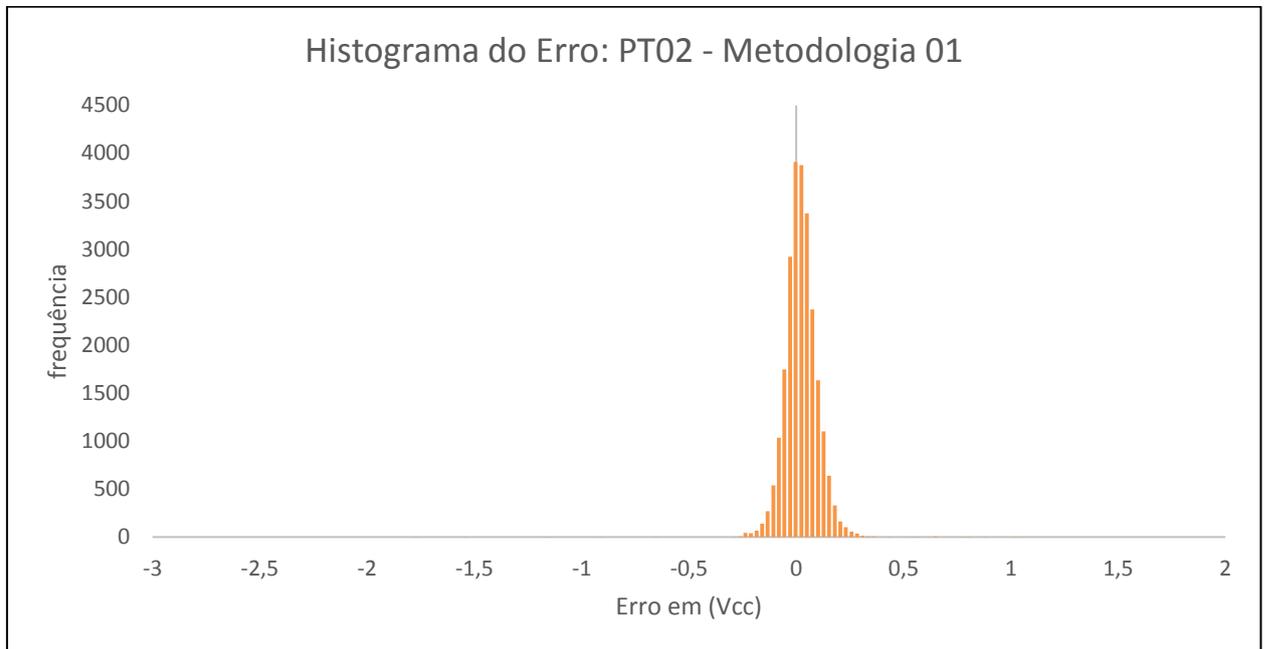


Figura 66 – Gráfico Histograma: PT02, Caso 02, Metodologia 01

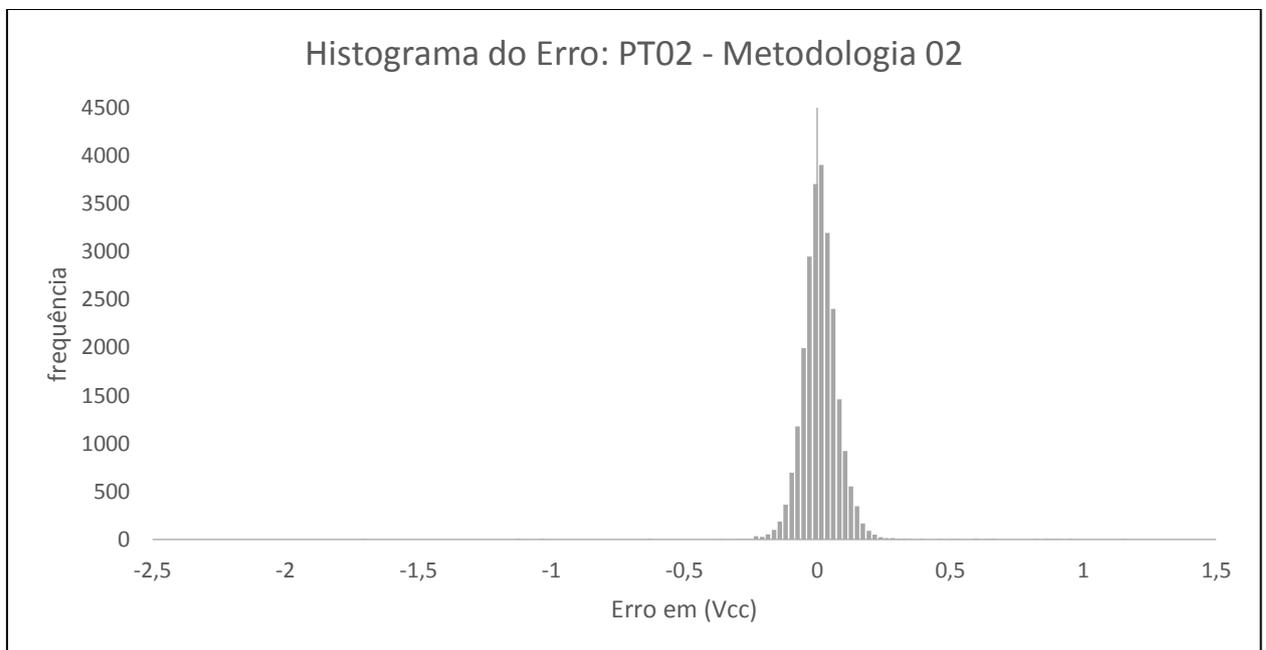


Figura 67 – Gráfico Histograma: PT02, Caso 02, Metodologia 02

5.2.3 Caso2 – Resultado PT03

A Figura 68 apresenta o gráfico com a comparação entre os resultados das duas metodologias e o valor real (alvo) do potencial eletroquímico do PT03 versus o tempo (número de linhas dos conjuntos para o teste).

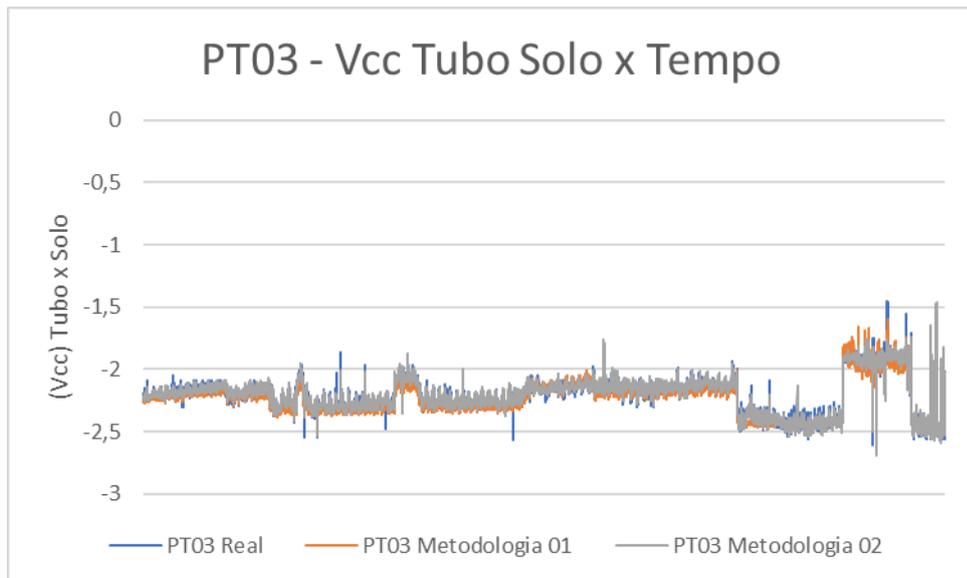


Figura 68 – Gráfico Tubo Solo x Tempo: PT03, Caso 02

A Figura 69 apresenta o gráfico com a dispersão dos resultados das metodologias propostas.

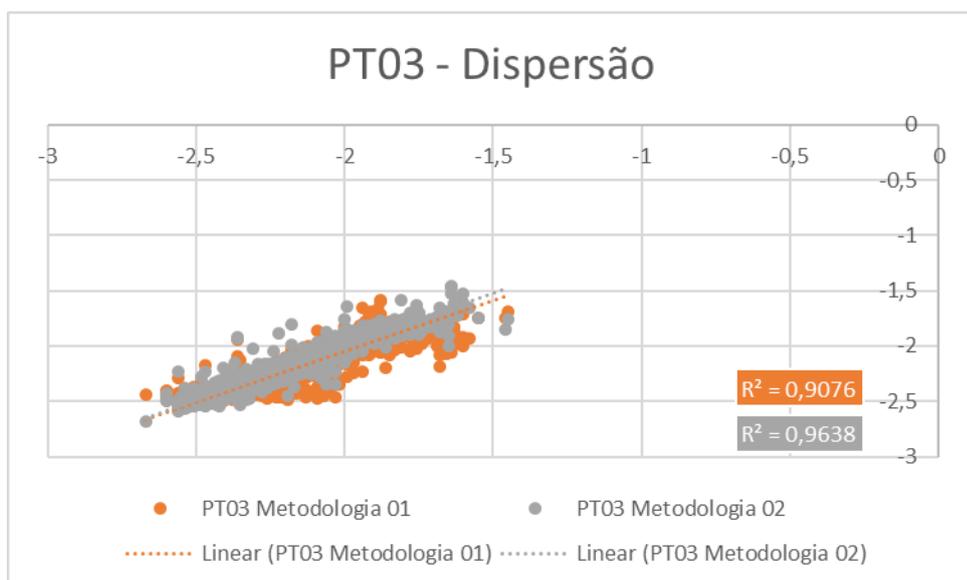


Figura 69 – Gráfico Dispersão: PT03, Caso 02

As Figuras 70 e 71 apresentam os gráficos com o histograma do erro relativo, para cada metodologia proposta.

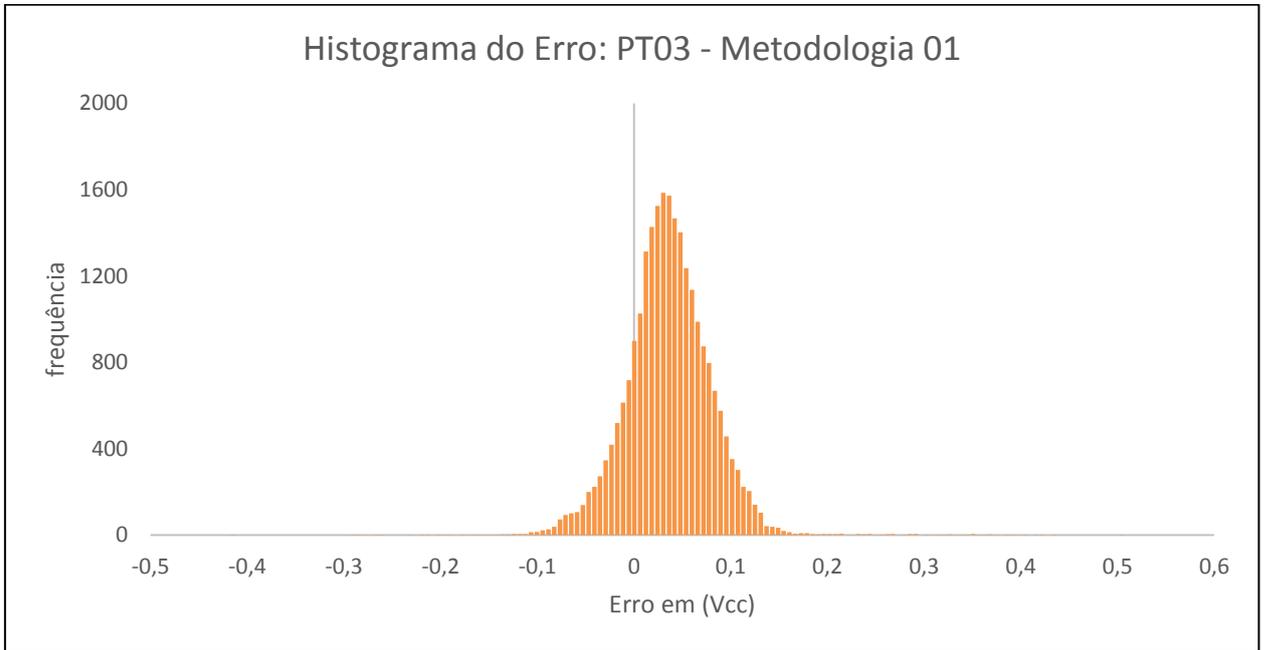


Figura 70 – Gráfico Histograma: PT03, Caso 02, Metodologia 01

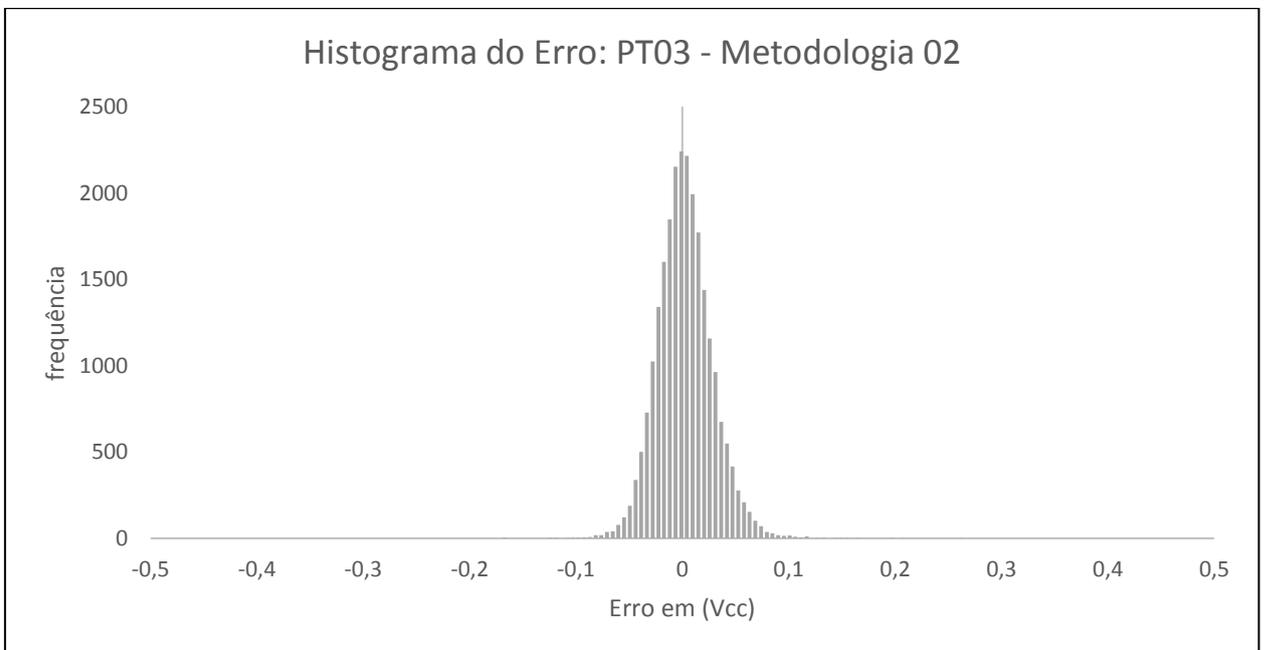


Figura 71 – Gráfico Histograma: PT03, Caso 02, Metodologia 02

5.2.4 Caso2 – Resultado PT04

A Figura 72 apresenta o gráfico com a comparação entre os resultados das duas metodologias e o valor real (alvo) do potencial eletroquímico do PT04 versus o tempo (número de linhas dos conjuntos para o teste).

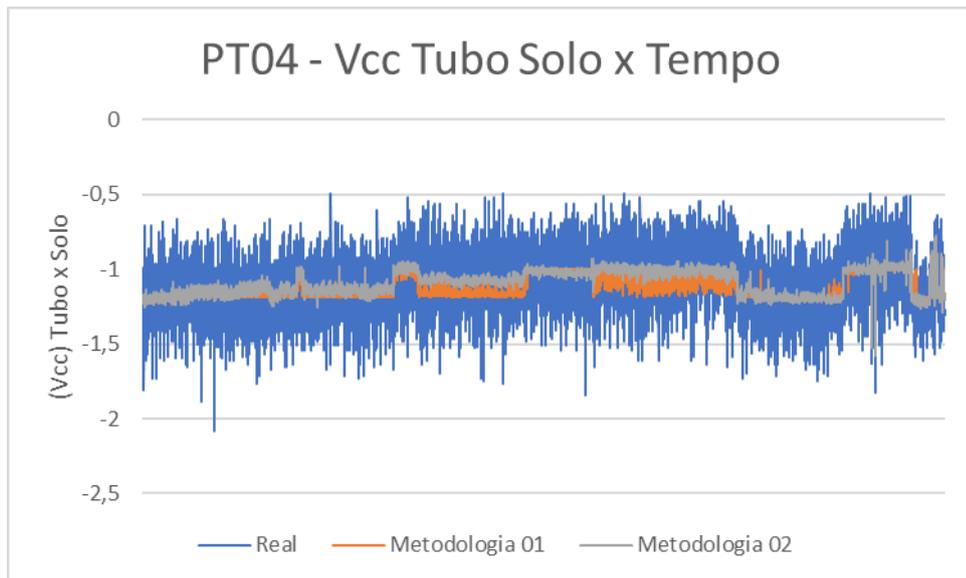


Figura 72 – Gráfico Tubo Solo x Tempo: PT04, Caso 02

A Figura 73 o gráfico com a dispersão dos resultados das metodologias propostas.

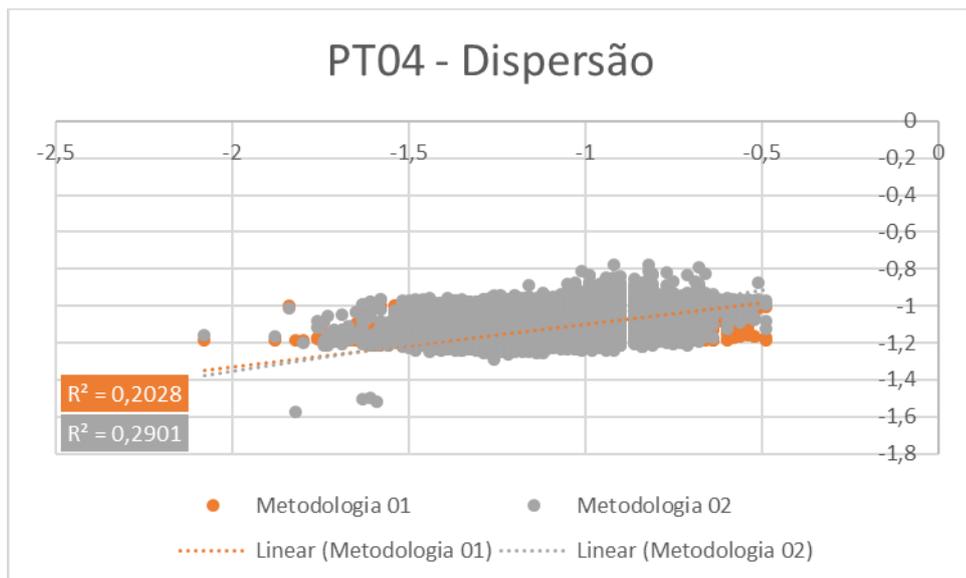


Figura 73 – Gráfico Dispersão: PT04, Caso 02

As Figuras 74 e 75 apresentam os gráficos com o histograma do erro relativo, para cada metodologia proposta.

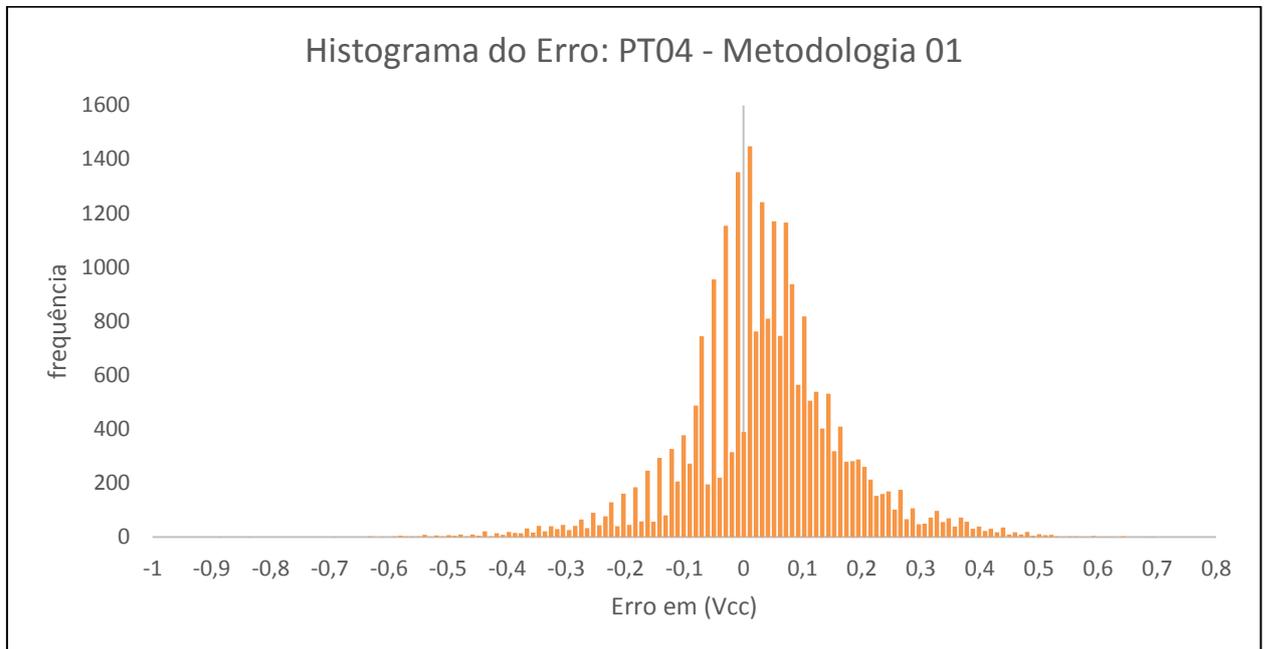


Figura 74 – Gráfico Histograma: PT04, Caso 02, Metodologia 01

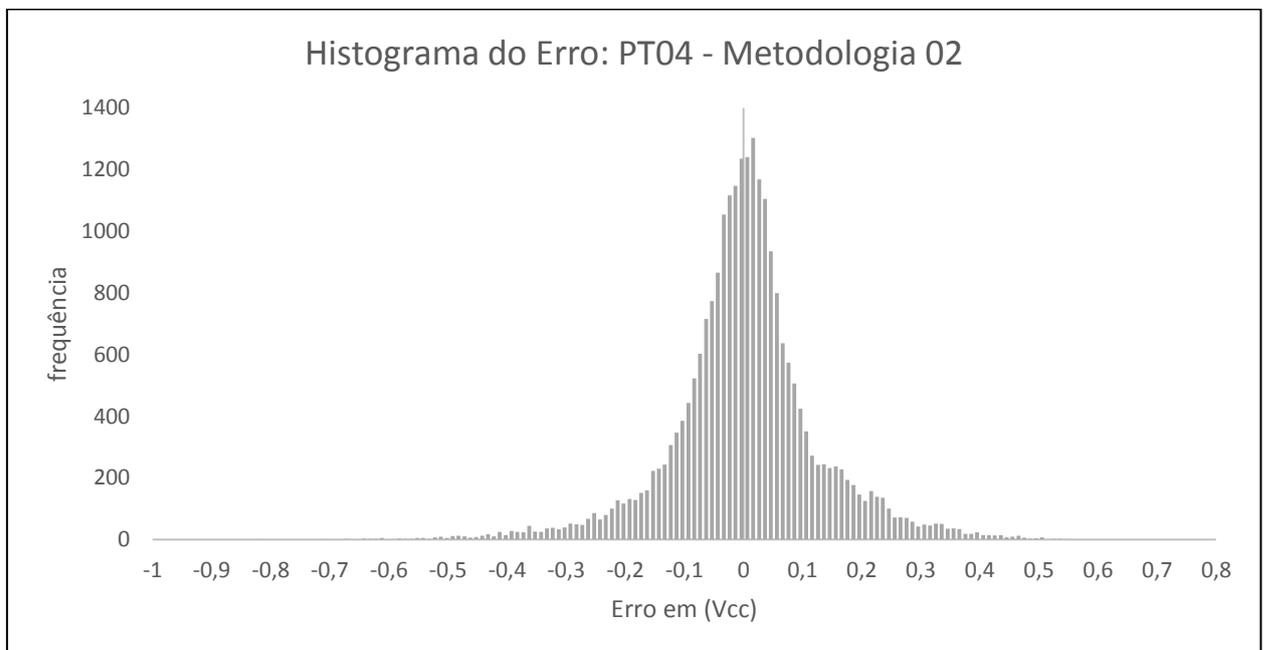


Figura 75 – Gráfico Histograma: PT04, Caso 02, Metodologia 02

5.2.5 Caso2 – Resultado PT05

A Figura 76 apresenta o gráfico com a comparação entre os resultados das duas metodologias e o valor real (alvo) do potencial eletroquímico do PT05 versus o tempo (número de linhas dos conjuntos para o teste).

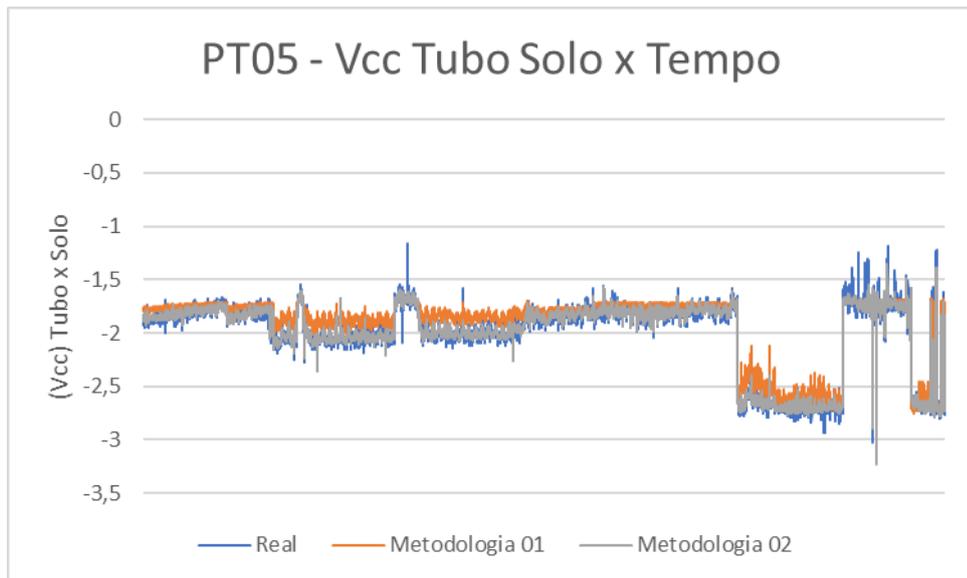


Figura 76 – Gráfico Tubo Solo x Tempo: PT05, Caso 02

A Figura 77 apresenta o gráfico com a dispersão dos resultados das metodologias propostas.

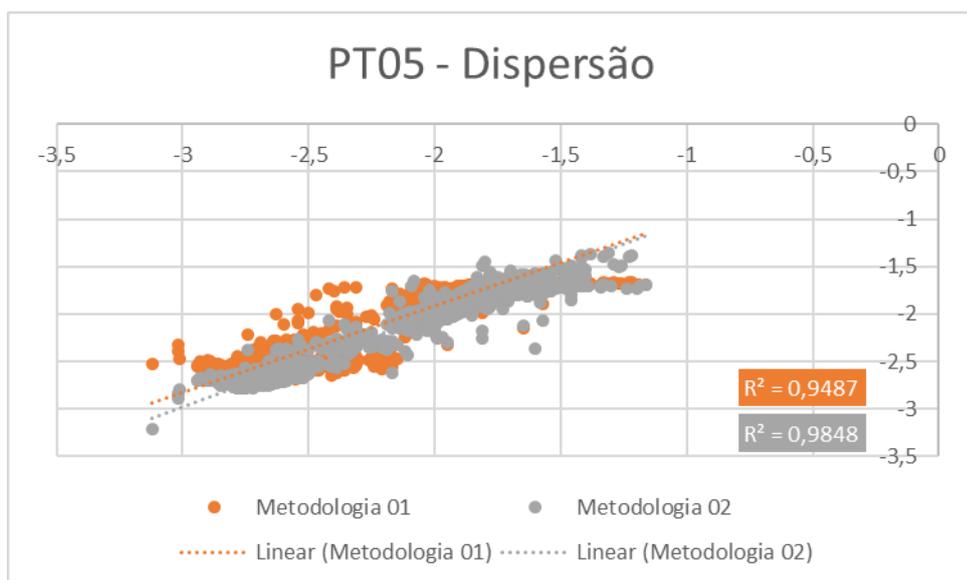


Figura 77– Gráfico Dispersão: PT05, Caso 02

As Figuras 78 e 79 apresentam os gráficos com o histograma do erro relativo, para cada metodologia proposta.

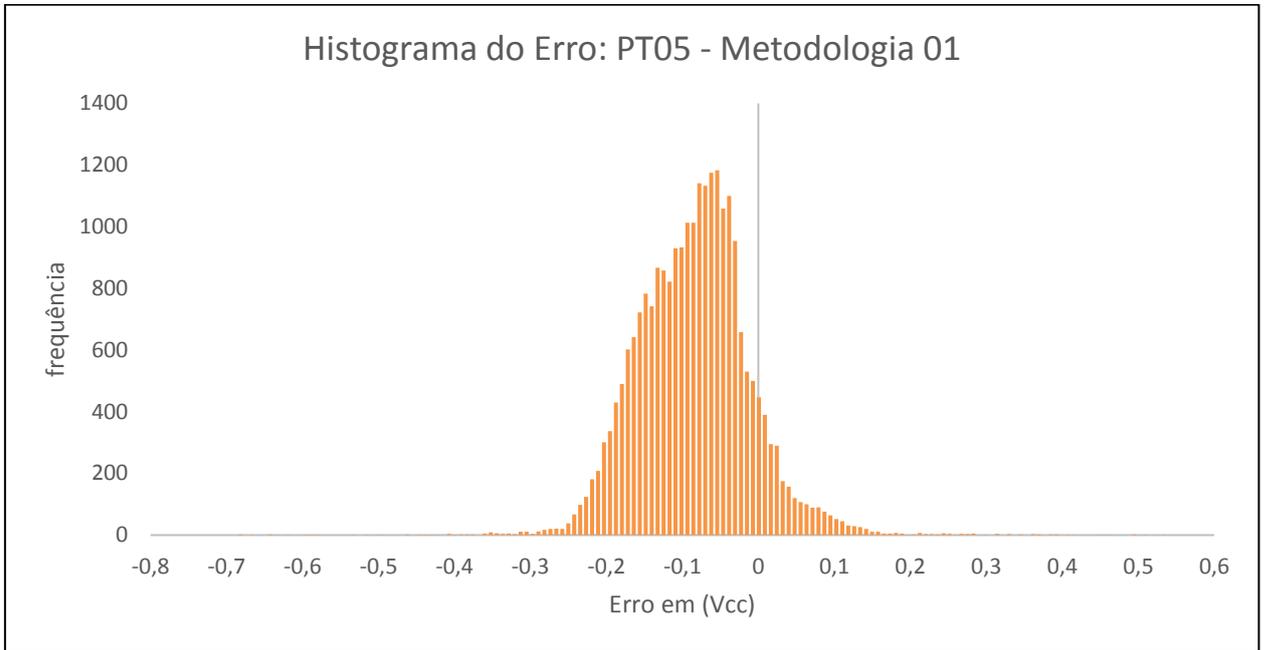


Figura 78 – Gráfico Histograma: PT05, Caso 02, Metodologia 01

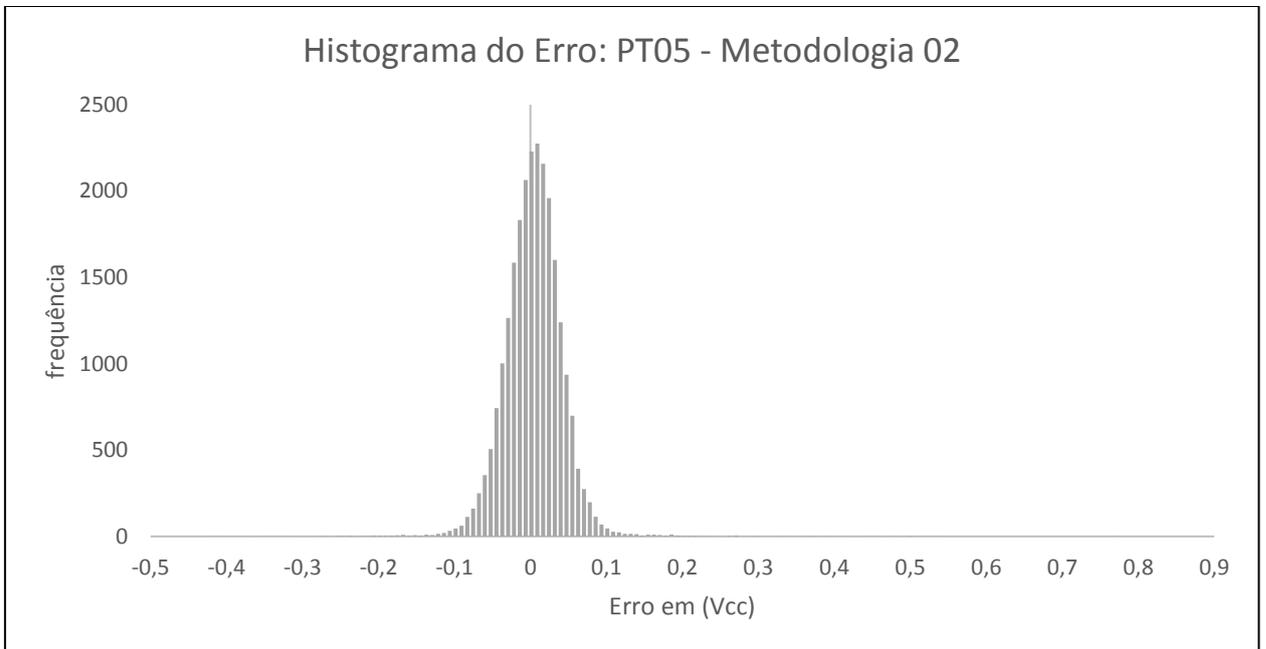


Figura 79 – Gráfico Histograma: PT05, Caso 02, Metodologia 02

5.2.6 Resultado Geral – Caso 02

A Figura 80 apresenta o gráfico com a comparação entre as duas metodologias propostas através do erro absoluto (Ea) médio para cada atributo:

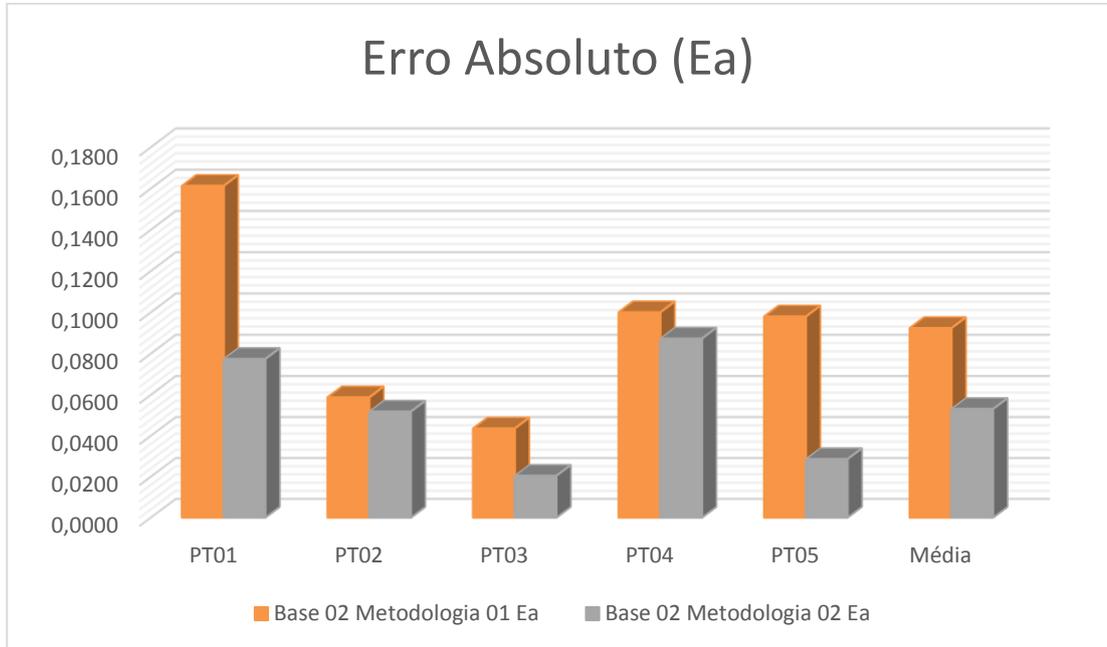


Figura 80 – Erro Absoluto Médio para os resultados do Caso 2

A Figura 81 apresenta o gráfico com a comparação entre as duas metodologias propostas através do erro percentual (Ep) médio para cada atributo:

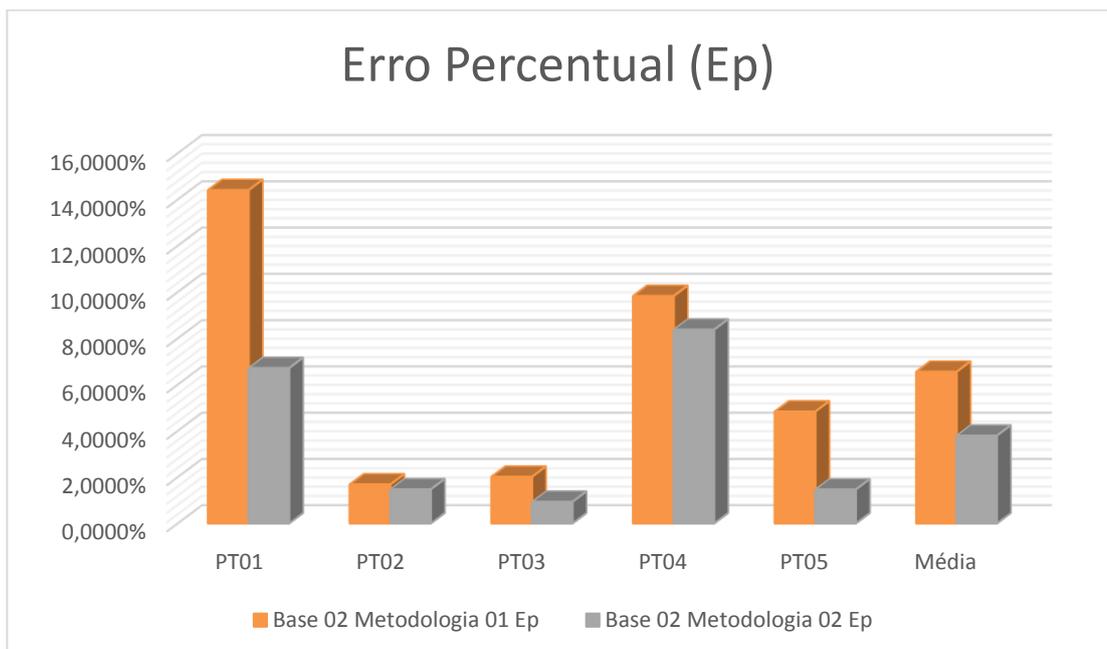


Figura 81 – Erro Percentual Médio para os resultados do Caso 2

A Figura 82 apresenta a função R^2 , comparando os resultados das duas metodologias propostas para cada atributo:

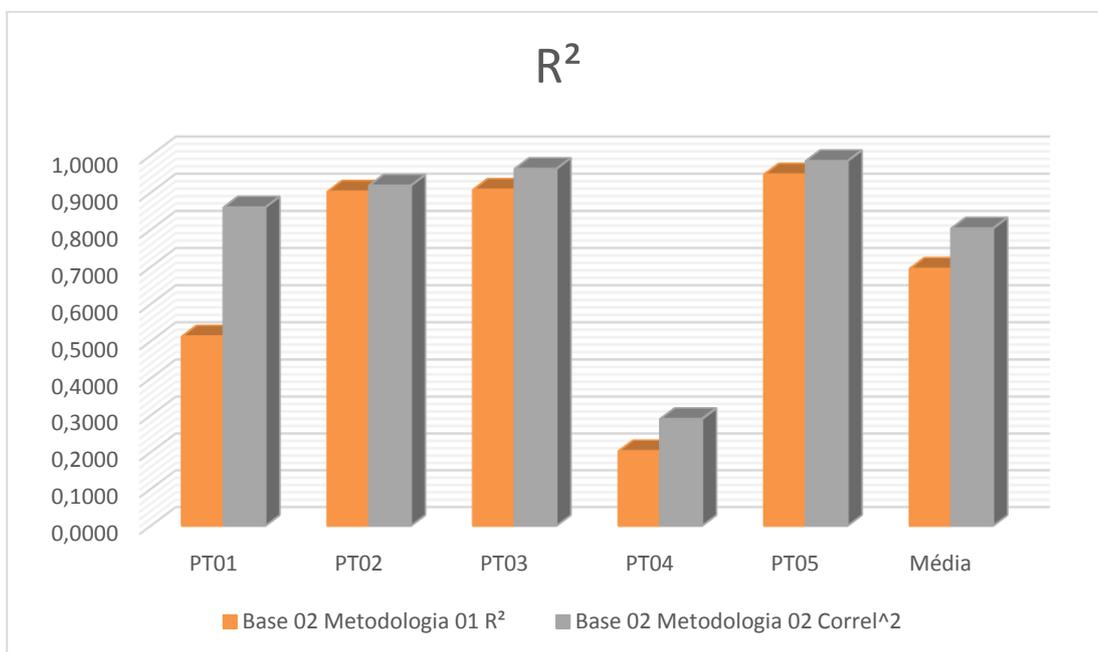


Figura 82 – Comparação da função R^2 para os resultados do Caso 2

A Figura 83 apresenta o desvio padrão (σ), comparando os resultados das duas metodologias propostas para cada atributo:

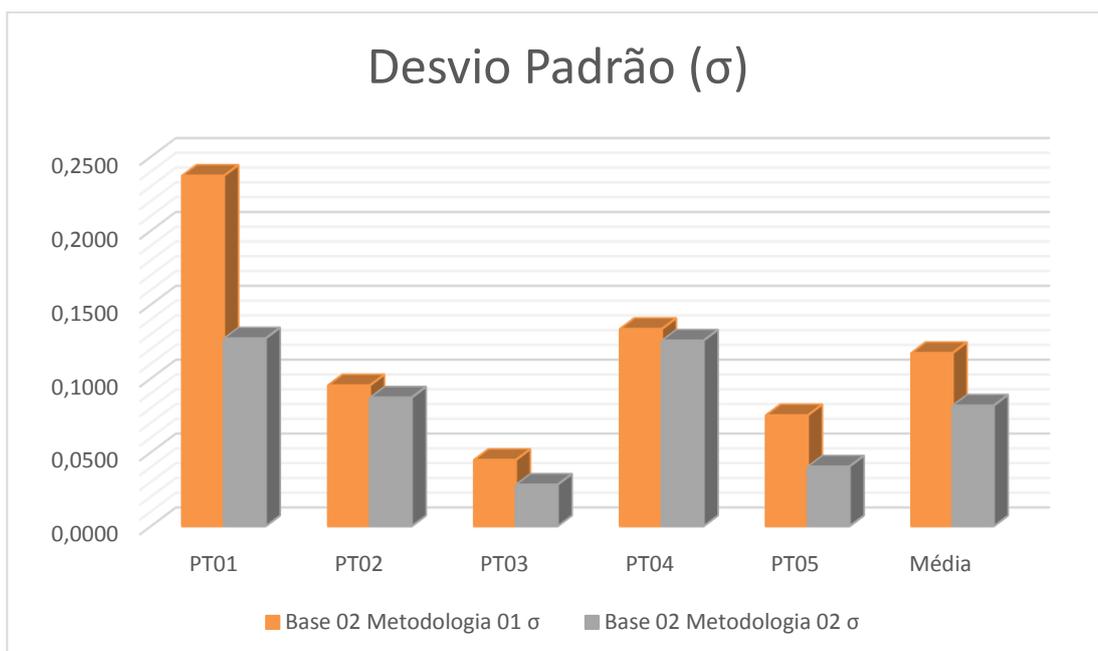


Figura 83 – Comparação do Desvio Padrão para os resultados do Caso 2

5.3 Considerações gerais

Dentre os erros, correlação e desvio padrão, de maneira geral, temos a situação descrita na Tabela 11:

Tabela 11 – Resultados médios gerais para o caso 01 baseado nas metodologias

Caso 01 – Somente ponto de teste – Resultado Médio								
Saída	Metodologia 01				Metodologia 02			
	Ea	Ep	R ²	σ	Ea	Ep	R ²	σ
Média	0,0108	0,66%	0,9784	0,0407	0,0118	0,70%	0,9894	0,0272

Observa-se que a metodologia 01 obteve melhores resultados no erro absoluto e percentual e que a metodologia 02 obteve melhores resultados de correlação e desvio padrão. Além disto, todos os elementos obtiveram êxito superior a 90% na taxa de assertividade relacionada ao erro absoluto de 0,1Vcc. A escolha entre uma das metodologias foi definida a partir da análise dos dados do caso 02.

Para o caso 02, dentre os erros, correlação e desvio padrão, de maneira geral, temos a situação descrita na Tabela 12:

Tabela 12 – Resultados médios gerais para o caso 02 baseado nas metodologias

Caso 02 – Somente ponto de teste – Resultado Médio								
Saída	Metodologia 01				Metodologia 02			
	Ea	Ep	R ²	σ	Ea	Ep	R ²	σ
Média	0,0923	6,55%	0,6948	0,1172	0,0531	3,79%	0,8034	0,0816

Observa-se que a metodologia 02 obteve melhores resultados médios em todos os quesitos de comparação.

Para definir qual dos elementos do caso 02 possui aderência adequada a proposta desta dissertação, serão comentados os resultados, com mais detalhes, baseado nos gráficos do histograma do erro em conjunto com a curva da distribuição normal bem como o intervalo de confiança associado.

5.3.1 PT01 – Caso 02

Tendo o elemento PT01 como alvo, foram obtidos resultados médios dos erros, correlação e desvio padrão conforme pode ser apreciado na Tabela 13:

Tabela 13 – Resultados médios gerais para o PT01

Caso 02 – PT01								
Saída	Metodologia 01				Metodologia 02			
	Ea	Ep	R ²	σ	Ea	Ep	R ²	σ
PT01	0,1613	14,41%	0,5125	0,2372	0,0774	6,71%	0,86	0,1271

A Figura 84 apresenta os gráficos do histograma do erro em conjunto com a curva da distribuição normal para a metodologia 01 e 02.

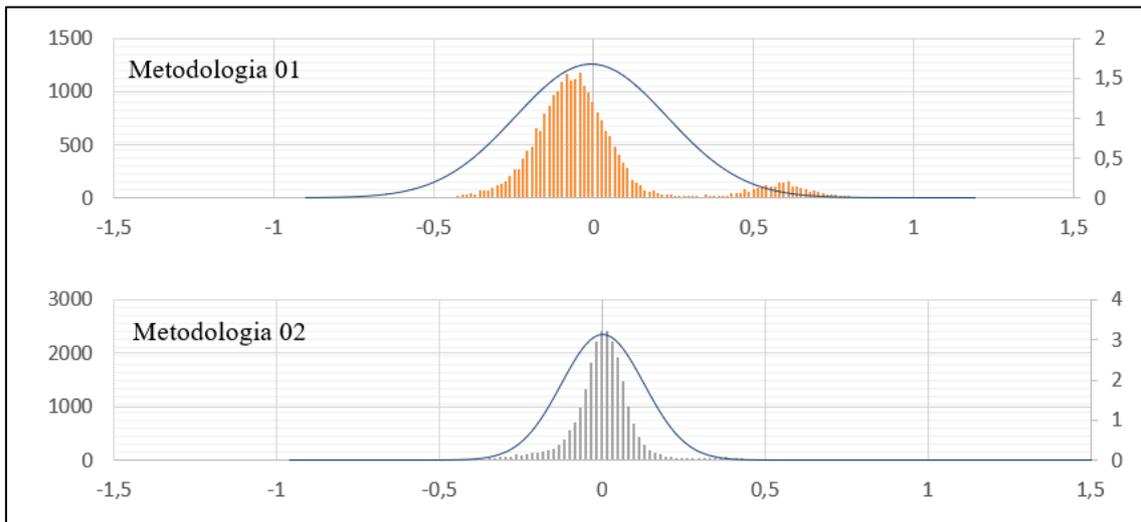


Figura 84 – histograma do erro em conjunto com a curva da distribuição normal – PT01

O histograma do erro da metodologia 01 apresenta uma pequena “ilha” isolada e uma distribuição normal fora do eixo. Já a metodologia 02 possui histograma simétrico, tipo distribuição normal, onde a frequência é mais alta no centro e decresce gradualmente para as “caudas” de maneira simétrica em forma de sino.

Observa-se que todos os resultados foram melhores para a metodologia 02. Além disto, a taxa de assertividade relacionada ao erro absoluto de 0,1Vcc ficou em 48,4% para a metodologia 01 e em 78,9% para a metodologia 02.

A Figura 85 apresenta a variância estimada gerada a partir da metodologia 02 para o PT01, cuja teoria está descrita no item 3.2.1 desta dissertação.

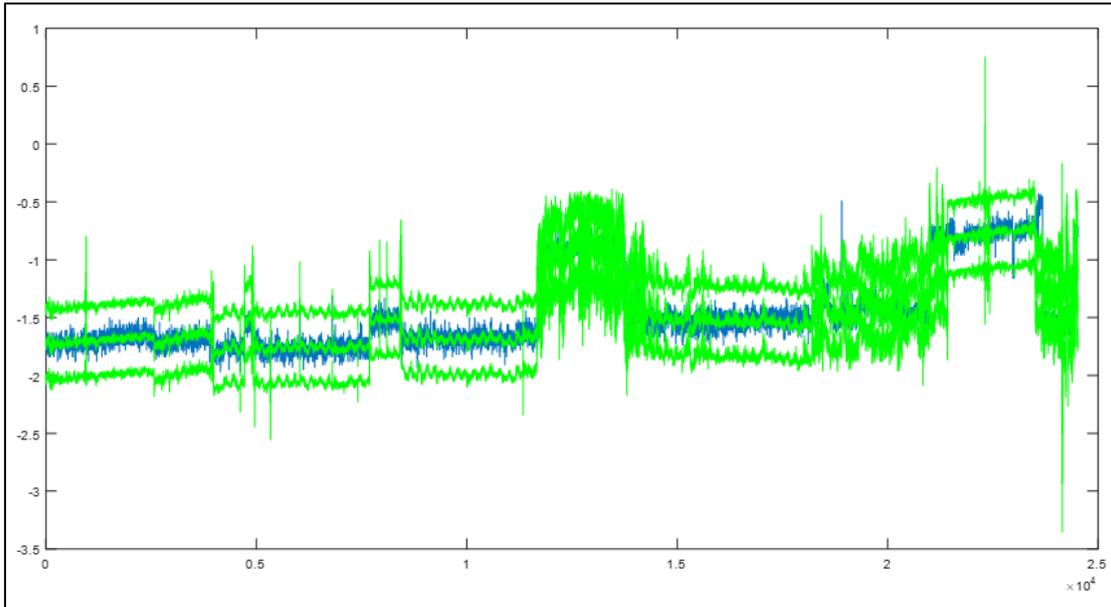


Figura 85 – Variância Estimada para o PT01, Caso 02

Conseqüentemente, além da previsão pontual, é possível demonstrar através deste modelo o intervalo de confiança do dado e verificar se ele se encontra dentro da faixa admissível para os valores previstos para o PT01 com base na metodologia 02.

5.3.2 PT02 – Caso 02

Tendo o elemento PT02 como alvo, foram obtidos resultados médios dos erros, correlação e desvio padrão conforme pode ser apreciado na Tabela 14:

Tabela 14 – Resultados médios gerais para o PT02

Caso 02 – PT02								
Saída	Metodologia 01				Metodologia 02			
	Ea	Ep	R ²	σ	Ea	Ep	R ²	σ
PT02	0,0588	1,68%	0,9025	0,0953	0,0518	1,47%	0,9186	0,087

Segue na Figura 86 os gráficos do histograma do erro em conjunto com a curva da distribuição normal para a metodologia 01 e 02.

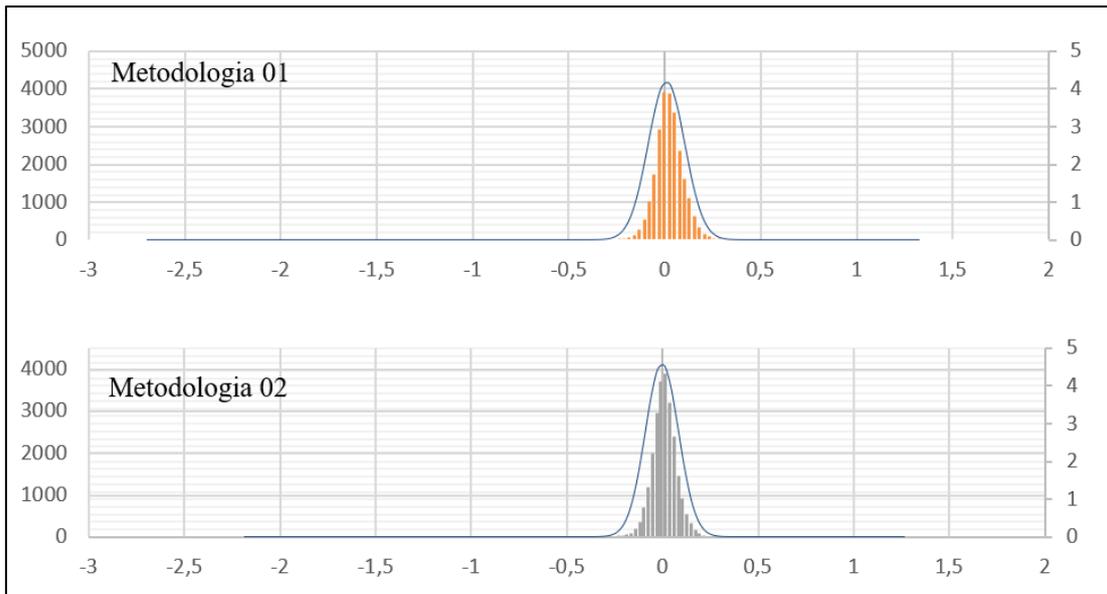


Figura 86 – histograma do erro em conjunto com a curva da distribuição normal – PT02

O histograma do erro da metodologia 01 e da metodologia 02 apresentam histograma simétrico, tipo distribuição normal, onde a frequência é mais alta no centro e decresce gradualmente para as “caudas” de maneira proporcional em forma de sino.

Observa-se que todos os resultados foram melhores para a metodologia 02. Além disto, a taxa de assertividade relacionada ao erro absoluto de $0,1V_{cc}$ ficou em 83,9% para a metodologia 01 e em 88,4% para a metodologia 02.

A Figura 87 apresenta a variância estimada gerada a partir da metodologia 02 para o PT02, cuja teoria está descrita no item 3.2.1 desta dissertação.

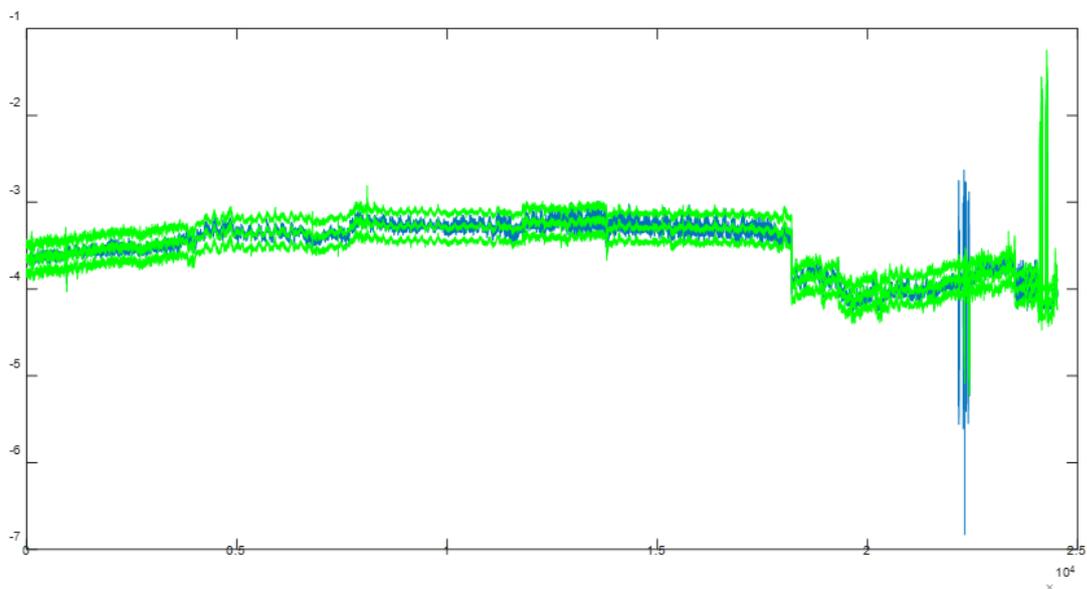


Figura 87 – Variância Estimada para o PT02, Caso 02

Consequentemente, além da previsão pontual, é possível demonstrar através deste modelo o intervalo de confiança do dado e verificar se ele se encontra dentro da faixa admissível para os valores previstos para o PT02 com base na metodologia 02.

5.3.3 PT03 – Caso 02

Tendo o elemento PT03 como alvo, foram obtidos resultados médios dos erros, correlação e desvio padrão conforme pode ser apreciado na Tabela 15 :

Tabela 15 – Resultados médios gerais para o PT03

Caso 02 – PT03								
Saída	Metodologia 01				Metodologia 02			
	Ea	Ep	R ²	σ	Ea	Ep	R ²	σ
PT03	0,0436	2,01%	0,9076	0,0448	0,0204	0,93%	0,9638	0,0281

Segue na Figura 88 os gráficos do histograma do erro em conjunto com a curva da distribuição normal para a metodologia 01 e 02.

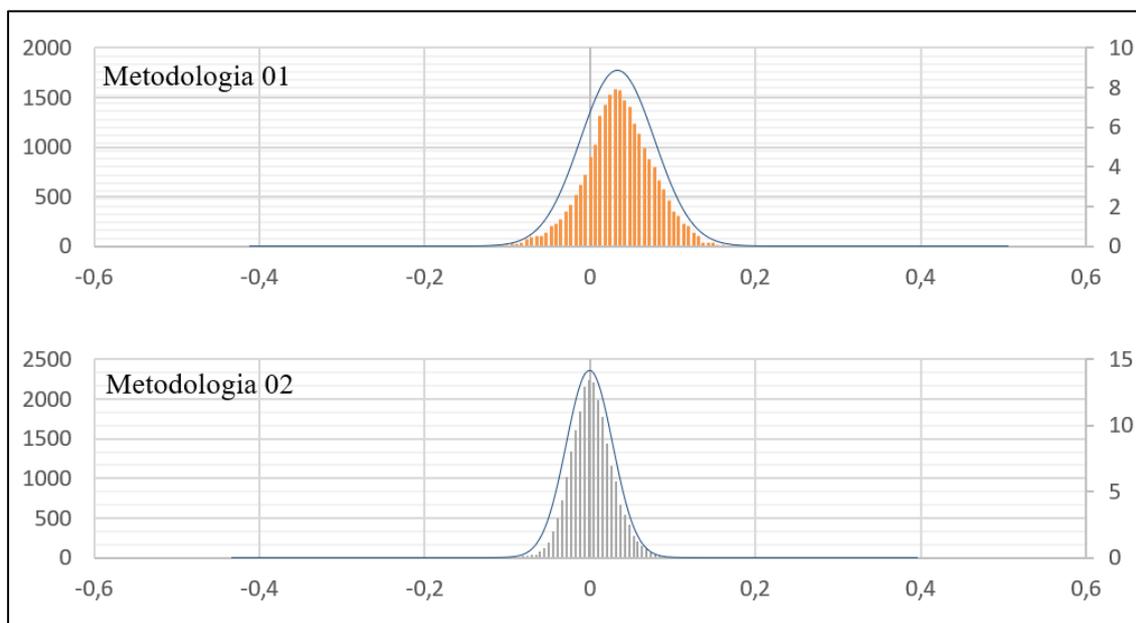


Figura 88 – histograma do erro em conjunto com a curva da distribuição normal – PT03

O histograma do erro da metodologia 01 apesar de apresentar simetria, o centro está deslocado do “zero” da abscissa. Já a metodologia 02 apresenta histograma simétrico,

tipo distribuição normal, onde a frequência é mais alta no centro e decresce gradualmente para as “caudas” de maneira proporcional em forma de sino.

Observa-se que todos os resultados foram melhores para a metodologia 02. Além disto, a taxa de assertividade relacionada ao erro absoluto de 0,1Vcc ficou em 94,4% para a metodologia 01 e em 99,6% para a metodologia 02.

A Figura 89 apresenta a variância estimada gerada a partir da metodologia 02 para o PT03, cuja teoria está descrita no item 3.2.1 desta dissertação.

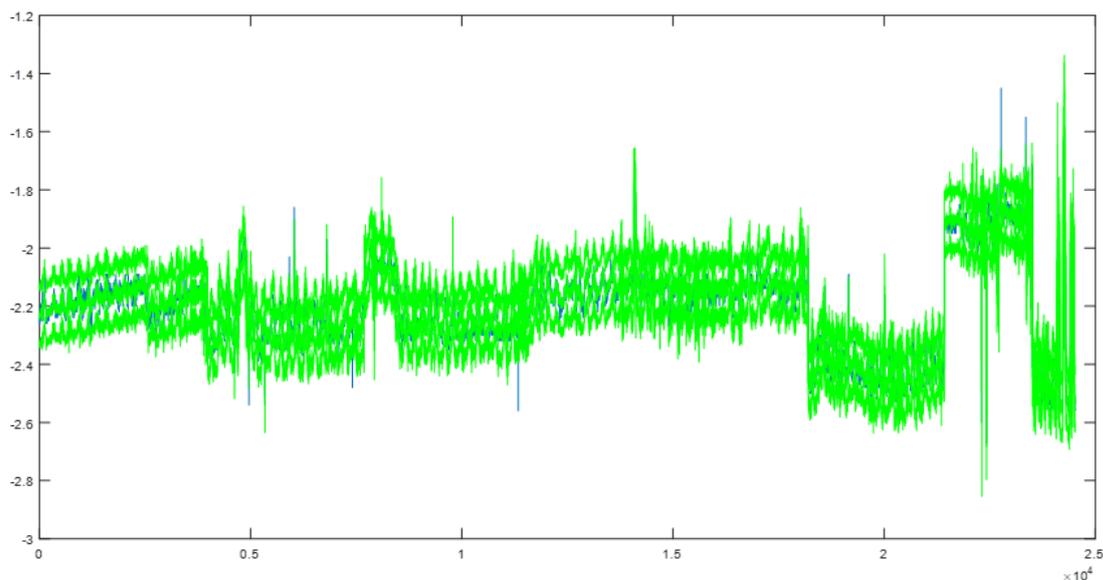


Figura 89 – Variância Estimada para o PT02, Caso 02

Consequentemente, além da previsão pontual, é possível demonstrar através deste modelo o intervalo de confiança do dado e verificar se ele se encontra dentro da faixa admissível para os valores previstos para o PT03 com base na metodologia 02.

5.3.4 PT04 – Caso 02

Tendo o elemento PT04 como alvo, foram obtidos resultados médios dos erros, correlação e desvio padrão conforme pode ser apreciado na Tabela 16 :

Tabela 16 – Resultados médios gerais para o PT04

Caso 02 – PT04								
Saída	Metodologia 01				Metodologia 02			
	Ea	Ep	R ²	σ	Ea	Ep	R ²	σ
PT04	0,1	9,83%	0,2028	0,1335	0,0873	8,37%	0,2901	0,1256

Segue na Figura 90 os gráficos do histograma do erro em conjunto com a curva da distribuição normal para a metodologia 01 e 02.

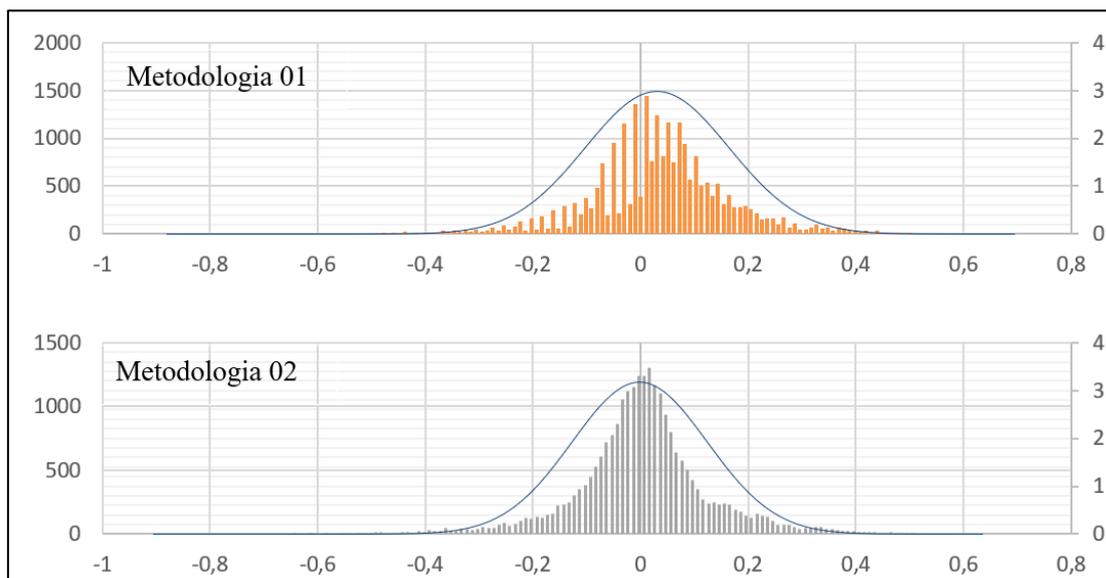


Figura 90 – histograma do erro em conjunto com a curva da distribuição normal – PT04

O histograma do erro da metodologia 01 apresenta vários pequenos picos, sem simetria, e com o centro um pouco deslocado do “zero” da abscissa. Já a metodologia 02 apresenta histograma simétrico, tipo distribuição normal, onde a frequência é mais alta no centro e decresce gradualmente para as “caudas” de maneira proporcional..

Observa-se que todos os resultados foram melhores para a metodologia 02. Além disto, a taxa de assertividade relacionada ao erro absoluto de $0,1V_{cc}$ ficou em 63,7% para a metodologia 01 e em 69,9% para a metodologia 02.

A Figura 91 apresenta a variância estimada gerada a partir da metodologia 02 para o PT04, cuja teoria está descrita no item 3.2.1 desta dissertação.

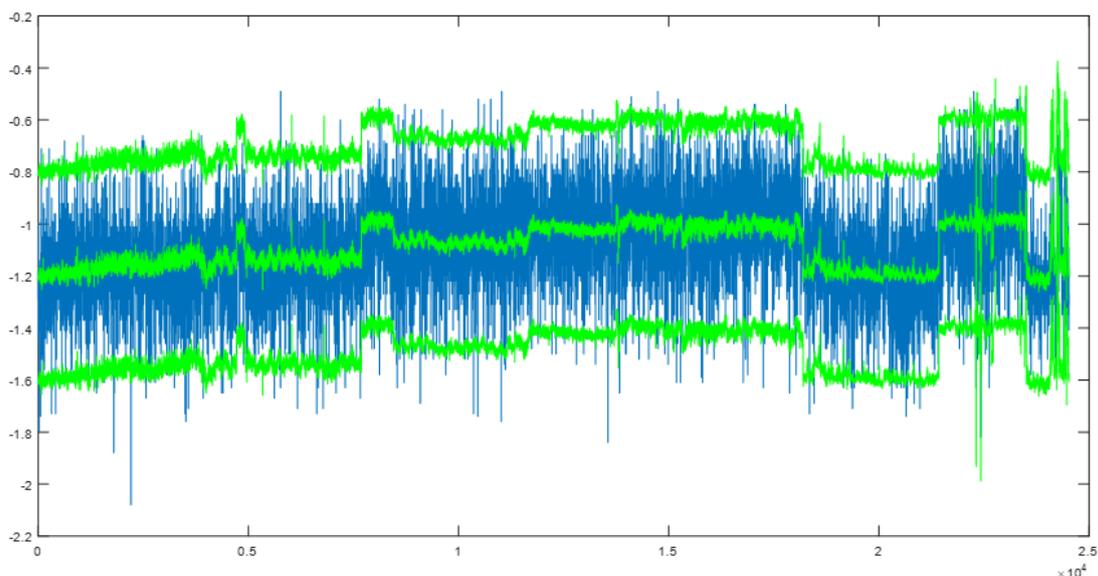


Figura 91 – Variância Estimada para o PT02, Caso 02

Consequentemente, além da previsão pontual, é possível demonstrar através deste modelo o intervalo de confiança do dado e verificar se ele se encontra dentro da faixa admissível para os valores previstos para o PT04 com base na metodologia 02.

5.3.5 PT05 – Caso 02

Tendo o elemento PT05 como alvo, foram obtidos resultados médios dos erros, correlação e desvio padrão conforme pode ser apreciado na Tabela 17 :

Tabela 17 – Resultados médios gerais para o PT05

Caso 02 – PT05								
Saída	Metodologia 01				Metodologia 02			
	Ea	Ep	R ²	σ	Ea	Ep	R ²	σ
PT05	0,098	4,83%	0,9487	0,0751	0,0286	1,47%	0,9848	0,0404

Segue na Figura 92 os gráficos do histograma do erro em conjunto com a curva da distribuição normal para a metodologia 01 e 02.

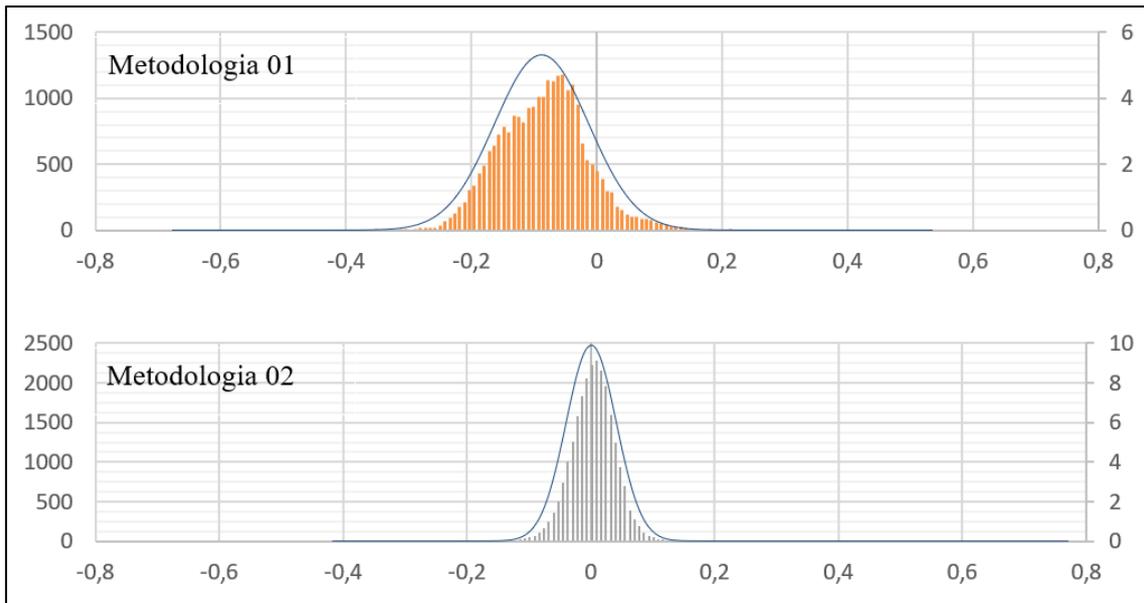


Figura 92 – histograma do erro em conjunto com a curva da distribuição normal – PT05

O histograma do erro da metodologia 01 não apresenta simetria e o centro está deslocado do “zero” da abscissa. Já a metodologia 02 apresenta histograma simétrico, tipo distribuição normal, onde a frequência é mais alta no centro e decresce gradualmente para as “caudas” de maneira proporcional em forma de sino.

Observa-se que todos os resultados foram melhores para a metodologia 02. Além disto, a taxa de assertividade relacionada ao erro absoluto de $0,1V_{cc}$ ficou em 56,0% para a metodologia 01 e em 98,4% para a metodologia 02.

A Figura 93 apresenta a variância estimada gerada a partir da metodologia 02 para o PT05, cuja teoria está descrita no item 3.2.1 desta dissertação.

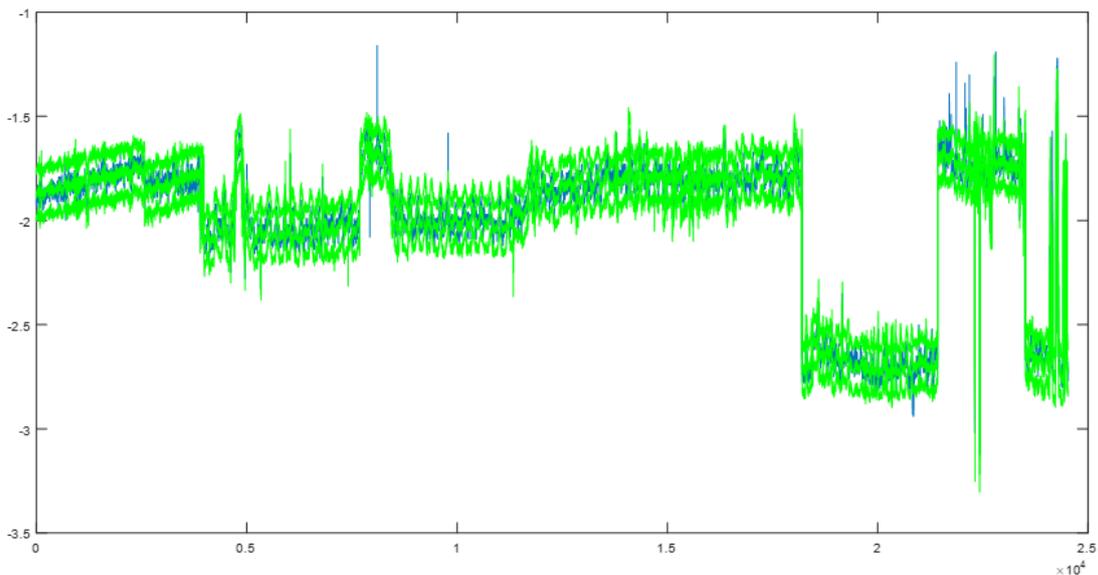


Figura 93 – Variância Estimada para o PT02, Caso 02

Consequentemente, além da previsão pontual, é possível demonstrar através deste modelo o intervalo de confiança do dado e verificar se ele se encontra dentro da faixa admissível para os valores previstos para o PT05 com base na metodologia 02.

Capítulo 6 - Conclusões

Esta dissertação teve por objetivo a elaboração de um sistema inteligente para estimação de potencial de proteção catódica baseado em RNA's. Foram apresentadas duas bases de dados e duas metodologias para que o objetivo principal fosse cumprido.

Inicialmente realiza-se uma contextualização das motivações para elaboração desta dissertação. A manutenção do sistema de proteção catódica do modal dutoviário no transporte e distribuição de gás natural no Brasil requer avanços dos quais são atingíveis com investimentos em tecnologia e aplicação de metodologias de RNA's já existentes a nichos de atividades e setores mais específicos da atividade industrial.

Em seguida foram expostos os cenários da proteção catódica e do aprendizado de máquina no Brasil e no mundo. Foram apresentados os principais conceitos destes seguimentos o que deu base teórica para elaboração desta dissertação.

Utilizaram-se então dois sistemas de proteção catódica de dutos de distribuição de gás natural instaladas na região sudeste do Brasil para estudo. Um duto com 13 elementos e outro com 8 elementos que geram dados com características distintas devido as particularidades dos locais onde os dutos foram instalados, revestimento, tempo de vida do ativo bem como o tipo de sistema remoto instalado o que difere no tempo da aquisição dos dados.

Posteriormente desenvolveram-se duas metodologias para treinamento, validação e teste dos dados dos sistemas descritos no parágrafo anterior. A metodologia 01, cuja base teórica é apresentada no item 3.1 desta dissertação, baseia-se na utilização de redes MLP. A não linearidade, particularmente, é fruto da geração do sinal de entrada que é inteiramente não-linear. A metodologia 02, cuja base teórica é apresentada no item 3.2 desta dissertação, baseia-se na utilização da rede bayesiana, adaptada de [28].

Verifica-se que a base de dados tem papel fundamental para o que se propõe nesta dissertação. Sinais incomuns, fora da magnitude normal e elementos com pouca relação geram resultados menos satisfatórios. Ambas bases de dados foram “tratadas”, excluindo-se os valores com magnitude anormal. Entende-se que a distância entre os elementos também influencia na relação entre eles, o que diminui a taxa de assertividade das metodologias propostas para o caso 02.

A base de dados do caso 01 percorre situações com pouca dinâmica e variação dos potenciais eletroquímicos e da saída de corrente dos retificadores. Possui também forte relação entre as grandezas elétricas entre os elementos, o que aumenta a taxa de assertividade dos dois sistemas propostos. Levando em consideração apenas o caso 01, seria difícil definir a melhor metodologia já que a taxa de erro foi bem pequena e, em poucos casos, os potenciais obtiveram variação maior que $\pm 0,1V_{cc}$, ou seja, todos os elementos possuem forte aderência as duas metodologias propostas.

A base de dados do caso 02 percorre situações com variação dos potenciais eletroquímicos e da saída de corrente dos retificadores com maior frequência e pouca relação entre as grandezas elétricas entre os elementos. Levando em consideração os dados obtidos no caso 02, fica mais claro que os melhores resultados foram obtidos através da metodologia 02, apesar do PT04 possuir baixa aderência as duas metodologias propostas para esta dissertação. Neste caso, avaliando a variação e inconsistência dos dados para este elemento, é possível afirmar que as informações recebidas pelo servidor através deste equipamento de monitoramento remoto estão equivocadas e uma verificação em campo se faz necessária. Apesar disto, ainda assim, houve uma assertividade considerável baseado na metodologia 02

Portanto, a melhor metodologia obtida para a aplicação proposta é a metodologia 02, rede Bayesiana. Esta metodologia, além de prever o resultado, informa o grau de confiança da medida, o que pode poupar a ida de uma equipe a campo de acordo com o grau de confiança e intervalo obtidos. Um outro diferencial desta metodologia é que não há o ajuste das entradas o que evita problemas de super ajuste dos parâmetros.

6.1 Trabalhos Futuros

Propõe-se para análises futuras, reavaliar as duas metodologias retirando do treinamento a base de dados do elemento PT04, do caso 02. Deve-se verificar se haverá melhora dos dados, com conseqüente diminuição do erro associado.

Propõe-se para análises futuras, tubulações com interferência elétrica de corrente contínua proveniente de sistemas de trens eletrificados. Deve-se verificar se a aderência da metodologia proposta pode ser aplicada em redes com este tipo de interferência.

Propõe-se realizar simulação do sistema de proteção catódica de uma tubulação de aço no PSCAD, definindo parâmetros hora desconhecidos tais como resistência do

circuito por trecho de ponto de medição e comparar a saída do sistema proposto nesta dissertação com o valor calculado pelo circuito do PSCAD para um dado de potencial desconhecido.

Sugere-se que seja analisado se o regime de chuvas interfere no potencial analisado, cruzando as informações de precipitação com o potencial da tubulação nos trechos analisados.

Propõe-se o uso de RNA para identificar a taxa de uso dos anodos de um determinado sistema e informar o período aproximado do qual o sistema deverá ser renovado.

Propõe-se um sistema autônomo, baseado em RNA, alimentado com informações *online* de sistemas de monitoramento remoto dos principais equipamentos do sistema de proteção catódica estrategicamente distribuídos, e com retificadores automáticos com possibilidade de comando remoto. Este tipo de sistema poderá gerar um menor consumo de energia elétrica (melhoria da eficiência energética), ações *on line* e *on time* frente a flutuação dos potenciais eletroquímicos, minimizando a taxa de corrosão, aumentando a vida útil do ativo. Avisos de parada do funcionamento de retificadores poderão ser identificados não só pelo envio dos dados remotos destes equipamentos como também pelo potencial eletroquímico dos pontos de teste do entorno. A parada poderá ser compensada pelos demais retificadores que influenciam na região afetada.

Com o tempo, o sistema proposto nesta pesquisa poderá definir trechos com necessidade de inspeção no revestimento devido ao aumento da demanda de corrente de proteção catódica o que indica incidência de falhas no revestimento.

Capítulo 7 - Bibliografia

- [1] B. S. B. Andrade, “O Setor de Distribuição de Gás Liquefeito e Natural no Brasil: Panorama e Perspectivas Frente aos Desinvestimentos da Petrobras” Universidade Federal do Rio de Janeiro, Setembro de 2018.
- [2] "Produção Nacional de Gás Natural" [Online]. Available: <http://www.anp.gov.br/dados-estatisticos> [Acesso em 10 05 2019].
- [3] "Balanço Energético Nacional 2018" [Online]. Available: <http://www.epe.gov.br/pt/publicacoes-dados-abertos/publicacoes/balanco-energetico-nacional-2018> [Acesso em 10 05 2019].
- [4] "Transporte Dutoviário de Combustíveis no Brasil: Desafios e Oportunidades" [Online]. Available: <https://www.ilos.com.br/web/transporte-dutoviario-de-combustiveis-no-brasil-desafios-e-oportunidade/> [Acesso em 10 05 2019].
- [5] L. Lazzari, P. Pedferri, “Cathodic Protection” 1st edition, Polipress Milano, 2006.
- [6] R. S. Sutton e A.G. Barto, “Reinforcement Learning: An Introduction”. Cambridge, Massachusetts, London, England, 2012.
- [7] M. C. Monart e J. A. Baranauskas, “Conceitos sobre Aprendizado de Máquina” Capítulo 4, 2003.
- [8] F. Borsato e M. Figueiredo, “Aprendizagem por Reforço em Redes Neurais Multicamadas Aplicadas em Controle Autônomo”. Departamento de Computação, Universidade Federal de São Carlos, São Paulo, 2007.
- [9] "Fundamentos da Corrosão" [Online]. Available: <https://docplayer.com.br/15411556-Fundamentos-da-corrosao.html> [Acesso em 12 05 2019].
- [10] E. T. Serra, "Corrosão e proteção anticorrosiva dos metais no solo". Centro de Pesquisas de Energia Elétrica – CEPEL 2006.
- [11] S. Wolynech, "Técnicas Eletroquímicas em corrosão". São Paulo: Editora da Universidade de São Paulo, 2003.
- [12] A. C. Dutra e L. P. Nunes, "Proteção catódica, técnica de combate a corrosão". 5. Ed., 2011. p.16.
- [13] "Tipos de Corrosão" [Online]. Available: <https://brasilecola.uol.com.br/quimica/tipos-corrosao.htm> [Acesso em 16 11 2019].

- [14] M. Neveux, "La Corrosion des Conduites d'Eau et de Gas – Causes et Remèdes. Éd. Eyrolles", Paris p. 169-170, 1968.
- [15] "Passividade em Metais" [Online]. Available: http://www.ufrgs.br/lapec/wa_files/passividade.pdf [Acesso em 15 05 2019].
- [16] "Eletroquímica Aplicada e Corrosão" [Online]. Available: <http://www.gea.ufpr.br/arquivos/lea/material/Fundamentos%20da%20Corrosao.pdf> [Acesso em 12 05 2019].
- [17] "Predição da composição do clínquer industrial utilizando minimização da energia livre de Gibbs" [Online]. Available: http://www.scielo.br/scielo.php?pid=S0366-69132015000100005&script=sci_arttext [Acesso em 14 05 2019].
- [18] A. C. Dutra e L. P. Nunes, "Proteção catódica, técnica de combate a corrosão". 5. Ed., 2011.
- [19] ABNT NBR ISSO 15589-1:2016 – Indústria do Petróleo e Gás Natural – Proteção Catódica para Sistemas de Transporte de Dutos Parte1: Dutos Terrestres.
- [20] ISO 15589-1:2015 - Petroleum, petrochemical and natural gas industries - Cathodic protection of pipeline systems - Part 1: On-land pipelines
- [21] C. F. Santos, M. A. Albuquerque, M. C. C. Oliveira; A. Echevarria, " A Corrosão e os Agentes Anticorrosivos" Rev. Virtual Quim., 2014.
- [22] V. Gentil, Corrosão. LTC – Livros Técnicos e farmacológico e usos clínicos. Química Nova Científicos Editora: Rio de Janeiro, 2007
- [23] NBR 16563-1 - Mitigação de efeitos de interferências elétricas em sistemas dutoviários. Parte 1: Sistemas de corrente alternada.
- [24] NBR 16563-2 - Mitigação de efeitos de interferências elétricas em sistemas dutoviários. Parte 2: Sistemas de corrente contínua.
- [25] L. P. Daister "Estratégias para Desenvolvimento de Sistemas de Múltiplos Classificadores em Aprendizado Supervisionado" COPPE/UFRJ 2007.
- [26] V. N. Vapnik, "Statistical Learning Theory", 1998.
- [27] M. I. Jordan e D. E. Rumelhart "Supervised Learning With a Distal Teacher" 1992.
- [28] V. H. Ferreira, "Desenvolvimento de Modelos Neurais Autônomos para Previsão de Carga Elétrica", Rio de Janeiro, COPPE/UFRJ, 2008.

- [29] S. Haykin, "Neural Networks and Learning Machines", Third Edition, 2008.
- [30] O. Bousquet, U. V. Luxburg e G. Ratsch, "Advanced Lectures on Machine Learning", ML Summer Schools, 2003.
- [31] BISHOP, C.M.; *Neural Networks for Pattern Recognition*, Oxford, New York.
- [32] J. L. Yuan, T. L. Fine, "Neural-Network Design for Small Training Sets of High Dimension", IEEE Transactions on Neural Networks, v.9, n.2, pp. 266-280, Mar. 1998.
- [33] D. S. Lopes, "Modelos Neurais Autônomos para Classificação e Localização de Defeitos em Linhas de Transmissão", Niterói, UFF, 2017.
- [34] J. L. G. Rosa, "Perceptron Multicamadas" Cap. 5, USP, 2011
- [35] D. J. C. Mackay, "Bayesian Methods for Adaptive Models", Ph.D. dissertation, California Institute of Technology, Pasadena, California, USA, 1992.
- [36] C. M. BISHOP, "Neural Networks for Pattern Recognition", Oxford, New York: Oxford University Press, 1995.
- [37] H. K. H. Lee, "A Noninformative Prior for Neural Networks", Machine Learning, vol. 50, n.1-2, pp. 197-212, Jan. 2003
- [38] V.H Ferreira, J. C. S. Souza e M. B. C. Filho, " Inferência Bayesiana Aplicada ao Desenvolvimento de Modelos Neurais para Tratamento de Alarmes em Subestações"
- [39] V.H. Ferreira, A.P. Alves da Silva, "Toward Estimating Autonomous Neural Network Based Electric Load Forecasters", IEEE Transactions on Power Systems, v.22, n.4, 1554-1562, Nov. 2007.
- [40] I V.H. Ferreira, A. Lazzaretti, H.V. Neto, R. Riella, J. Omeri, "Classificação de Eventos em Redes de Distribuição de Energia utilizando Transformada Wavelet e Modelos Neurais Autônomos", Learning and Nonlinear Models – Journal of the Brazilian Society on Neural Networks, v.8, n.2, 93-99, 2010]
- [41] G. Cybenko, "Approximation by Superpositions of a Sigmoidal Function," Mathematics of Control Signals Systems (1989), pp. 303-314, 1989.
- [42] "Função Correl" [Online]. Available: <https://support.office.com/pt-pt/article/correl-fun%C3%A7%C3%A3o-correl-995dcef7-0c0a-4bed-a3fb-239d7b68ca92> [Acesso em 20 08 2019].

Anexo 01



INTERCORR2018_231

Copyright 2018, ABRACO

Trabalho apresentado durante o INTERCORR 2018, em São Paulo, no mês de maio de 2018.

As informações e opiniões contidas neste trabalho são de exclusiva responsabilidade do(s) autor(es).

Rede Neural Aplicada em Sistemas de Proteção Catódica – Estudo de Caso Rodrigo M.S. Soares^a, Vitor H. Ferreira^b, Márcio Z. Fortes^c.

Abstract

Cathodic protection is a technique used for the preservation of metallic structures immersed in an electrolyte, whose function is to minimize corrosion to derisory values. The control to verify the good functioning of these systems is given through the field inspection or through monitoring via datalogger. In both cases, there is a periodicity of verification in the field, whose interval between inspections will be done according to the importance that each one has for a given system. The cathodically protected metal structures are usually arranged in the subsoil of large cities, and thus are subject to various events that are likely to hamper their inspection. Having a good database of previous inspections, aligned with a regression-supervised learning algorithm, missing measurements can be estimated, leading to greater system reliability by improving maintenance, safety, and environmental indices. The objective of this work is, from the cathodic protection database of a natural gas distribution pipeline, to create a neural network that, through the regression, can estimate the missing data of one of the measurement points at a given instant.

Keywords: Machine learning, supervised learning, regression, cathodic protection.

Resumo

A proteção catódica é uma técnica utilizada para preservação de estruturas metálicas imersas em um eletrólito, cuja função é minimizar a corrosão a valores irrisórios. O controle para verificação do bom funcionamento deste sistema se dá através da inspeção em campo ou por meio de monitoração via *datalogger*. Em ambos os casos há uma periodicidade de verificação em campo, cujo intervalo entre inspeções será feito de acordo com a importância que cada um tem para um determinado sistema. As estruturas metálicas protegidas catódicamente estão normalmente dispostas nos subsolos de grandes cidades, sendo assim, estão sujeitas a diversos acontecimentos que possivelmente irão dificultar sua inspeção. Tendo uma boa base de dados de inspeções anteriores, alinhada com um algoritmo de aprendizado supervisionado por regressão, as medições faltantes poderão ser estimadas, levando a uma maior confiabilidade ao sistema melhorando os índices da manutenção, segurança e meio ambiente. O objetivo deste trabalho é, a partir da base de dados da proteção catódica de um gasoduto de distribuição de gás

^a Mestrando, Engenheiro Eletricista – GÁS NATURAL FENOSA.

^b Doutor, Engenheiro Eletricista – UFF.

^c Doutor, Engenheiro Eletricista – UFF.

natural, criar uma rede neural que, através da regressão, possa estimar o dado faltante de um dos pontos de medição em um determinado instante.

Palavras-chave: Aprendizado de máquina, aprendizado supervisionado, regressão, proteção catódica.

Introdução

O aprendizado de Máquina (AM) é uma técnica que gera um modelo computacional da capacidade de aprender a partir da interação com seu ambiente visando melhorar o seu desempenho na execução de uma determinada tarefa (1). Ainda que o AM seja uma ferramenta poderosa para a aquisição automática de conhecimento, deve ser observado que não existe um único algoritmo que apresente o melhor desempenho para todos os problemas. Portanto, é importante compreender o poder e a limitação dos diversos algoritmos de AM utilizando alguma metodologia que permita avaliar os conceitos induzidos por esses algoritmos em determinados problemas (2).

Uma forma de aprendizagem que pode ser aplicada em AM é a aprendizagem indutiva. A indução é a forma de inferência lógica que permite obter conclusões genéricas sobre um conjunto particular de exemplos, ou seja, um conceito é aprendido efetuando-se inferência indutiva sobre os exemplos apresentados, portanto, as hipóteses geradas através da inferência indutiva podem ou não preservar a verdade (2).

O aprendizado indutivo é efetuado a partir do raciocínio sobre exemplos fornecidos por um processo externo ao sistema de aprendizado. O aprendizado indutivo pode ser dividido em supervisionado e não supervisionado. No aprendizado supervisionado é fornecido ao algoritmo de aprendizado, ou indutor, um conjunto de exemplos de treinamento para os quais o rótulo da classe associada é conhecido, ou seja, devem existir dados de entrada com suas respectivas saídas, para serem apresentadas ao algoritmo de aprendizado utilizado durante o processo de treinamento (3). O objetivo do algoritmo de indução é construir um tipo de classificador que possa determinar corretamente a classe de novos exemplos ainda não rotulados. Para rótulos de classe discretos, esse problema é conhecido como classificação e para valores contínuos como regressão.

A proteção catódica pode ser definida como abrandamento ou prevenção total de todas as formas de corrosão em estruturas metálicas imersas em um eletrólito (4). A rotina de inspeção dos elementos de proteção catódica segue uma periodicidade de acordo com a importância destes para o sistema. Diversas medições são coletadas pelo inspetor de campo, e, a partir delas, ações são tomadas após a análise dos dados.

Eventualmente, dados faltantes da inspeção de campo ou do sistema de monitoramento tornam a análise de um determinado sistema de proteção catódica incompleta.

Diante deste problema e visando solucioná-lo, o aprendizado de máquina, mais especificamente o aprendizado supervisionado, mostra-se uma ferramenta poderosa para estimativa dos dados faltantes em um determinado período, com base em um sólido banco de dados, aumentando a confiabilidade do sistema e do setor de manutenção.

Metodologia

Para garantir que a proteção catódica esteja sendo aplicada de maneira correta em todas as áreas dos dutos envolvidos, deve ser construído ao longo deste, pontos de monitoração de potencial respeitando os seguintes critérios:

- 1) Para obter uma boa proteção, a taxa de corrosão do metal deverá ser inferior a 0,01 mm por ano.
- 2) É adotado como valor máximo de proteção, o potencial de -0,85 V com relação ao eletrodo de referência de Cu/CuSO₄, ou seja, $E_p = -0,85$ V.
- 3) É adotado como valor mínimo de proteção, o potencial de -1,20 V com relação ao eletrodo de referência de Cu/CuSO₄, ou seja $E_c = -1,2$ V.

O potencial metal-eletrólito (E) deverá ser menor ou igual ao potencial de proteção (E_p) e que por sua vez, deverá ser maior ou igual que o potencial crítico (E_c). Este último se deve ao fenômeno da fragilização por hidrogênio provocado por potenciais muito baixos, tendo como ônus o descolamento do revestimento protetor do metal e/ou um alto pH na superfície do metal (5). Valores entre o potencial de proteção e o potencial crítico são considerados ideais para a proteção das tubulações de gás enterradas (6) e (7).

Foi utilizado como base para este problema um gasoduto de aço carbono API 5L grau B, de aproximadamente 40 km de extensão, 350 mm de diâmetro, 12 kgf/cm² de pressão, de uma companhia de distribuição de gás natural. Esta linha possui 9 pontos de medição de potencial e 4 retificadores de proteção catódica, monitorados via *datalogger* diariamente.

Segue na Figura 1 imagem via satélite dos equipamentos posicionados no *Google Earth*. A linha vermelha representa o gasoduto, os 4 pontos vermelhos representam os retificadores de proteção catódica, os pontos em amarelo representam os pontos de medição de potencial.

Das 12 entradas no programa, 4 são valores de corrente de saída (retificadores) e 8 de potencial tubo-solo (entre o gasoduto e uma semi-célula de cobre/sulfato de cobre - Cu/CuSO₄) e uma saída de potencial tubo-solo. Dentro da configuração deste sistema, o ponto que tem potencial mais próximo ao limite do potencial de proteção (E_p) é o ponto PT06, portanto é o ponto que deverá ter atenção especial e será a saída de dados deste estudo. Desta maneira, a base de dados

possui uma tabela com 13 colunas e 773 linhas com medições diárias coletadas entre o período de 21/10/2014 à 01/12/2016. Os valores coletados foram do resumo diário do sistema por elemento. Neste resumo há valores mínimos, médios, máximos e predominantes. O *datalogger* faz medições a cada segundo durante 24 horas, ou seja, são 86.400 medidas que irão gerar o resumo e conseqüente envio do mesmo para o sistema. Para este trabalho foram utilizados os valores da predominância diária por elemento



Figura 1 – Equipamentos posicionados no *Google Earth*

Definindo as entradas e a saída esperada, será dado um *input* da base de dados no sistema. A fim de evitar problemas relativos à estagnação do treinamento, as entradas e saída serão normalizadas em uma faixa, neste caso definida entre $[-1 \ 1]$ pois os dados têm valores e características diferentes.

Após normalização dos dados, será definido o número de neurônios que serão utilizados. Segundo o teorema de aproximação universal, uma camada oculta é o suficiente, tendo apenas que definir o número de neurônios que serão utilizados neste sistema, que será definido de acordo com a melhor aproximação ou menor erro da saída. Após testes na faixa entre 30 e 50 neurônios, foi definida a configuração com 36 neurônios, como sendo o melhor modelo para este sistema. Segue na Figura 02 o diagrama de blocos do sistema.

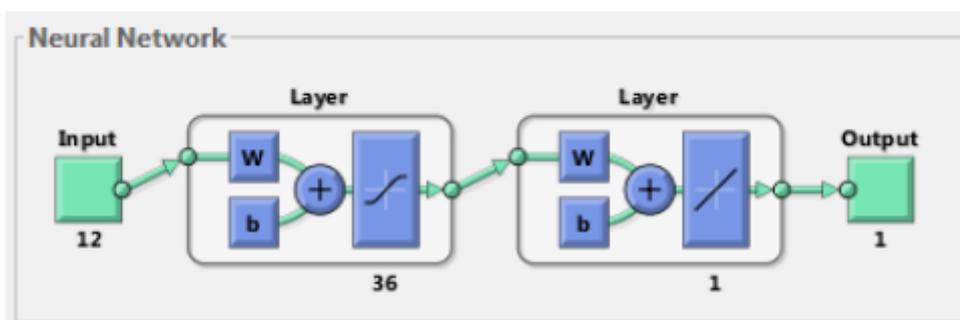


Figura 2 – Diagrama de blocos da rede neural (MATLAB)

Foram utilizados 20% dos dados para validação e escolha do modelo. A medida do número de vezes que todos os vetores de treinamento são usados para atualizar os pesos é chamada de *epoch*. Nesta rede neural foram utilizados 1000 *epochs*. Todas as amostras de treinamento passam pelo algoritmo de aprendizado simultaneamente em uma época antes que os pesos sejam atualizados. Neste primeiro momento, a saída do sistema é conhecida para definição do modelo.

Após definição dos parâmetros de funcionamento da rede neural, o usuário poderá imputar ao sistema 12 novas entradas, que serão normalizadas e verificadas pela rede. Após as iterações, uma saída será dada pelo sistema. Os dados imputados para obtenção desta nova saída não foram utilizados no treinamento do sistema, sendo assim, possível verificar a confiabilidade deste resultado.

Resultados e discussão

Na aprendizagem supervisionada a saída desejável é conhecida, portanto, é possível mensurar o erro cometido pelo modelo. O sinal de erro aciona um mecanismo de controle responsável pelo ajuste dos parâmetros visando otimizar uma determinada medida de desempenho (8). Segue na Figura 3 o diagrama de funcionamento do aprendizado supervisionado.

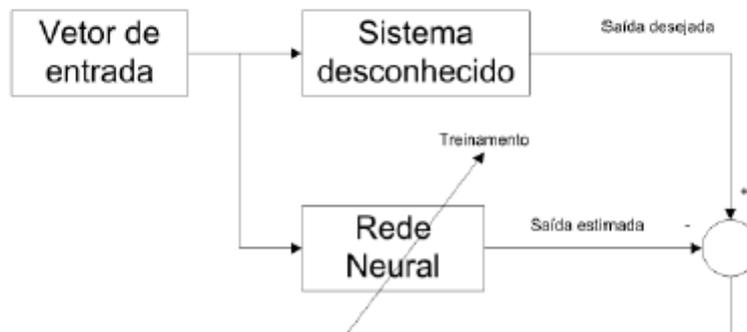


Figura 3 – Sistema aprendido supervisionado

A fim de validar a rede neural criada, serão imputados ao sistema 5 conjuntos de entradas com as respectivas saídas conhecidas que serão comparadas com a saída da rede. Para melhor visualização do problema, segue na Figura 4 as variáveis alocadas em um modelo linear.

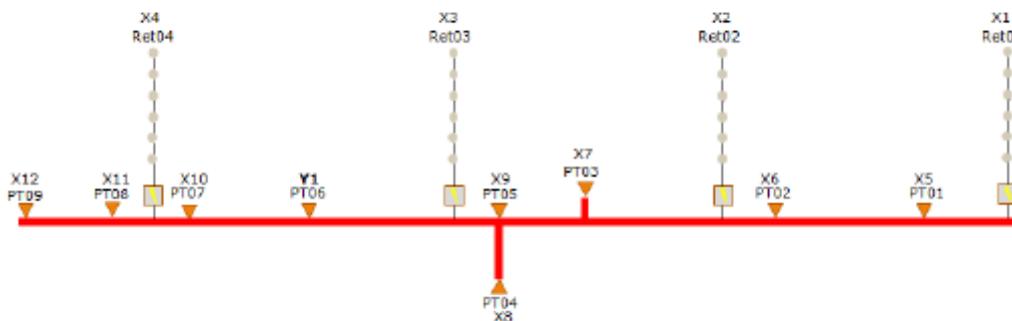


Figura 4 – Variáveis alocadas em um modelo linear

Na Tabela 1 verifica-se o conjunto X1 a X12 de entrada, a saída esperada Y1, a saída da rede neural e o erro da rede para cada conjunto.

Tabela 1 – Conjunto de entradas e saídas da rede neural

X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	Saída Esperada	Saída Rede	Erro (%)
7,3	4,9	9,1	11,1	-2,4	-3,3	-1,9	-1,8	-1,6	-1,6	-1,8	-1,8	-1,8	-1,798	0,094
7,4	5,1	9,2	11,1	-2,2	-2,9	-1,7	-1,8	-1,5	-1,6	-1,6	-1,5	-1,9	-1,889	0,558
5,8	5,1	9,1	11,1	-1,8	-2,2	-1,5	-1,6	-1,4	-1,3	-1,5	-1,4	-1,3	-1,300	0,000
6,0	4,3	3,9	10,2	-1,9	-2,4	-1,5	-2,0	-1,4	-1,4	-1,5	-1,6	-0,9	-0,911	1,178
6,0	4,3	2,7	10,2	-1,8	-2,3	-1,5	-1,9	-1,4	-1,4	-1,5	-1,6	-0,8	-0,803	0,375
													Erro Médio	0,441

Segue na Figura 04 o gráfico representando o erro quadrático médio pela *epoch*, cujo melhor desempenho de treinamento foi na *epoch* 1000, com um valor de 0,00041437.

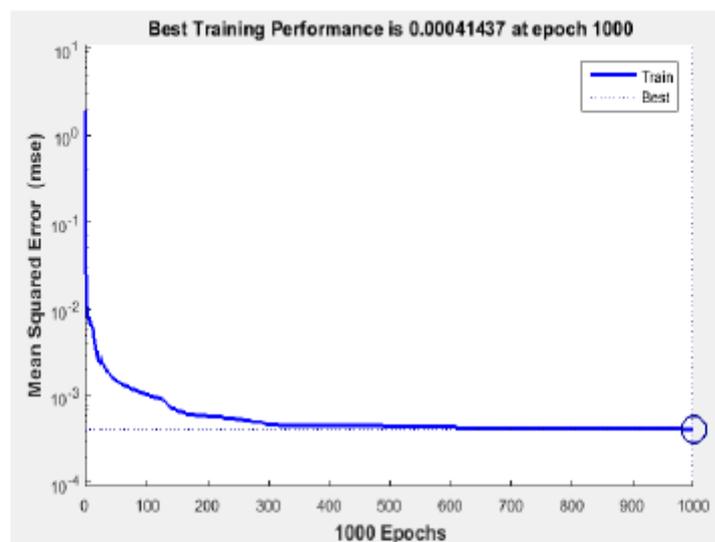


Figura 4 – Gráfico erro quadrático médio x *epoch*

O valor R na Figura 5 é uma indicação da relação entre as saídas do sistema e a saída esperada. Se $R = 1$, isso indica que há uma relação linear exata entre elas, no entanto, se R é próximo de zero, não há relação linear. No treinamento R atingiu o valor de $R=0,99844$, próximo de 1, indicando relação linear entre a saída do sistema e a saída esperada.

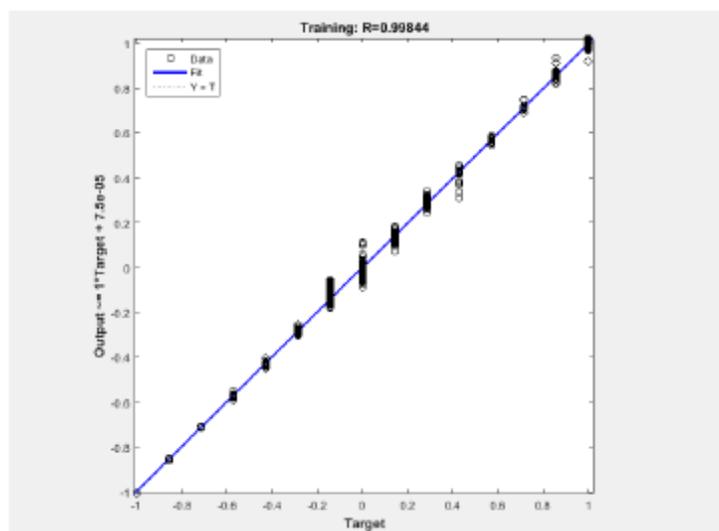


Figura 5 – Relação entre a saída e a saída esperada

Conclusões

A rede neural criada para este sistema de proteção catódica, com base no histórico que possuía, mostrou-se efetiva nos testes realizados, com erros consideráveis dentro de uma margem aceitável para este problema, aumentando a confiabilidade do sistema, dando uma maior segurança na tomada de decisões do setor de manutenção.

O sistema analisado possui pouca variação histórica de potenciais e não possui problemas de interferência elétrica, o que facilita na hora do treinamento, pois não há variáveis desconhecidas ou fora de um padrão normal. Caso fosse uma linha com interferência elétrica, poderia inviabilizar esta análise ou aumentar o erro de forma considerável.

Referências bibliográficas

- (1) R.S.SUTTON E A.G. BARTO, "Reinforcement Learning: An Introduction". 2012
- (2) M.C.MONART E J.A.BARANAUSKAS, "Conceitos sobre Aprendizado de Máquina" Capítulo 4.
- (3) F.BORSATO E M.FIGUEIREDO, "Aprendizagem por Reforço em Redes Neurais Multicamadas Aplicadas em Controle Autônomo".
- (4) L. LAZZARI, P.PEDEFERRI, "Cathodic Protection," 1st edition, Polipress Milano, 2006.
- (5) A.C.DUTRA E L.P. NUNES, "Proteção Catódica – Técnica de Combate à corrosão" 5ª Edição – 2011.
- (6) ABNT NBR ISSO 15589-1:2016 – Indústria do Petróleo e Gás Natural – Proteção Catódica para Sistemas de Transporte de Dutos Parte1: Dutos Terrestres.
- (7) ISO 15589-1:2015 - Petroleum, petrochemical and natural gas industries - Cathodic protection of pipeline systems - Part 1: On-land pipelines
- (8) GEDSON FARIA E ROSELI F. ROMERO, "Explorando o Potencial de Algoritmos de Aprendizado com Reforço em Robôs". 1999