

UNIVERSIDADE FEDERAL DO PARÁ
INSTITUTO DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

**SISTEMA INTELIGENTE DE SUPORTE A ANÁLISE DA QUALIDADE DA
ENERGIA ELÉTRICA EM REGIME PERMANENTE**

GABRIEL VIANNA SOARES ROCHA

DM: 10 / 2017

UFPA / ITEC / PPGEE
Campus Universitário do Guamá
Belém-Pará-Brasil
2017

UNIVERSIDADE FEDERAL DO PARÁ
INSTITUTO DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

GABRIEL VIANNA SOARES ROCHA

**SISTEMA INTELIGENTE DE SUPORTE A ANÁLISE DA QUALIDADE DA
ENERGIA ELÉTRICA EM REGIME PERMANENTE**

DM: 10 / 2017

UFPA / ITEC / PPGEE
Campus Universitário do Guamá
Belém-Pará-Brasil
2017

UNIVERSIDADE FEDERAL DO PARÁ
INSTITUTO DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

GABRIEL VIANNA SOARES ROCHA

**SISTEMA INTELIGENTE DE SUPORTE A ANÁLISE DA QUALIDADE DA
ENERGIA ELÉTRICA EM REGIME PERMANENTE**

Dissertação de Mestrado submetida à Banca Examinadora do Programa de Pós-Graduação em Engenharia Elétrica da UFPA para a obtenção do Grau de Mestre em Engenharia Elétrica na área de Sistemas de Energia Elétrica.

UFPA / ITEC / PPGEE
Campus Universitário do Guamá
Belém-Pará-Brasil
2017

Dados Internacionais de Catalogação-na-Publicação (CIP)
Sistema de Bibliotecas da UFPA

Rocha, Gabriel Vianna Soares, 1992-

Sistema inteligente de suporte a análise da qualidade da energia elétrica em regime permanente / Gabriel Vianna Soares Rocha. - 2017.

Orientadora: Maria Emília de Lima Tostes

Dissertação (mestrado) - Universidade Federal do Pará, Instituto de Tecnologia, Programa de Pós-Graduação em Engenharia Elétrica, Belém, 2017.

1. Energia elétrica - distribuição - qualidade. 2. Sistemas de energia elétrica - modelos matemáticos. 3. Harmônicos (ondas elétricas). I. Título.

CDD 23. ed. 621.319

UNIVERSIDADE FEDERAL DO PARÁ
INSTITUTO DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

**"SISTEMA INTELIGENTE DE SUPORTE A ANÁLISE DA QUALIDADE DA
ENERGIA ELÉTRICA EM REGIME PERMANENTE"**

AUTOR: GABRIEL VIANNA SOARES ROCHA

DISSERTAÇÃO DE MESTRADO SUBMETIDA À AVALIAÇÃO DA BANCA EXAMINADORA APROVADA PELO COLEGIADO DO PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA DA UNIVERSIDADE FEDERAL DO PARÁ E JULGADA ADEQUADA PARA OBTENÇÃO DO GRAU DE MESTRE EM ENGENHARIA ELÉTRICA NA ÁREA DE SISTEMA ELÉTRICO DE POTÊNCIA.

APROVADA EM: 21/02/2017

BANCA EXAMINADORA:

Prof.^a Dr.^a Maria Emília de Lima Tostes
(ORIENTADORA – PPGEE/UFPA)

Prof. Dr. João Paulo Abreu Vieira
(Avaliador Interno – PPGEE/UFPA)

Prof. Dr. Ubiratan Holanda Bezerra
(Avaliador Interno – PPGEE/UFPA)

Prof. Dr. Herminio Simões Gomes
(Avaliador Externo ao Programa – FACMAT/ICEN/UFPA)

VISTO:

Prof. Dr. Evaldo Gonçalves Pelaes
(COORDENADOR DO PPGEE/ITEC/UFPA)

DEDICATÓRIA

"À minha família"

AGRADECIMENTOS

Agradeço primeiramente a Deus por ter me dado a dádiva da vida e por ter colocado ao longo do meu caminho, pessoas que se tornaram especiais para mim e que me ajudaram a superar os obstáculos nos momentos difíceis.

A minha orientadora e professora Dra. Maria Emília de Lima Tostes que acreditou em mim e aceitou esta missão me guiando no decorrer de toda esta jornada.

Aos professores do CEAMAZON, agradeço pelos ensinamentos passados ao longo dos anos em que convivemos e pelo incentivo que sempre me deram. Agradeço em especial aos professores: Dr. João Paulo Abreu Vieira, Dr. Ubiratan Holanda Bezerra e Dr. Edson Ortiz de Matos.

Aos meus amigos Flávia Monteiro, Diorge Lima e André Carvalho que me deram todo o suporte e auxílio que precisei para cumprir esta etapa de minha vida. Em especial ao Allan Manito que se dispôs a me ajudar de perto.

Aos amigos que adquiri durante o período que estive na UFPA, em especial aos do LINC e do CEAMAZON, pela amizade e diversos momentos que compartilhamos.

A todos os amigos que fiz durante a vida, que sempre acreditaram em mim e me deram forças pra chegar até aqui. Á Aldemir Junior, Ana Laura, Anderson Dias, Camila Alves, Iago Cavalcante, Igor Araújo, Jamilly Azevedo, Patrick Monteiro e Wellington Lobato os quais considero como parte de minha família e que levarei comigo até o final de meus dias.

EPÍGRAFO

“Portanto, eu digo: Tudo o que vocês pedirem em oração, creiam que
já o receberam, e assim sucederá.”

Marcos 11:24

SUMÁRIO

LISTA DE FIGURAS	XI
LISTA DE TABELAS	XIII
LISTA DE ABREVEATURAS E SIGLAS	XIV
RESUMO	XV
ABSTRACT	XVI
1 INTRODUÇÃO	1
1.1 CONTEXTUALIZAÇÃO	1
1.2 MOTIVAÇÃO	2
1.3 OBJETIVOS	4
1.3.1 Gerais.....	4
1.3.2 Específicos.....	4
1.4 REVISÃO BIBLIOGRÁFICA	5
1.5 PRINCIPAIS RESULTADOS.....	7
1.6 ORGANIZAÇÃO DO TEXTO	8
2 FUNDAMENTAÇÃO TEÓRICA.....	9
2.1 CONSIDERAÇÕES INICIAIS	9
2.2 NORMAS ESTABELECIDAS PELO PRODIST	9
2.2.1 Tensão em Regime Permanente	9
2.2.2 Fator de Potência	12
2.2.3 Desequilíbrio	12
2.3 CARACTERIZAÇÃO DE HARMÔNICOS.....	14
2.4 INTELIGÊNCIA COMPUTACIONAL.....	19
2.4.1 Aprendizado de Máquina.....	19
2.4.2 Regressão Linear	23
2.4.3 Análise de Regressão.....	28
2.4.4 Árvores de Decisão.....	31
2.4.5 Redes Neurais Artificiais.....	34
2.5 CONSIDERAÇÕES FINAIS	38
3 DESENVOLVIMENTO DE SOFTWARE.....	39
3.1 CONSIDERAÇÕES INICIAIS	39
3.2 BASE DE DADOS	39
3.3 ARQUITETURA DE SOFTWARE	44
3.4 INTERFACE GRÁFICA	49
3.4.1 Módulo de Banco de Dados	51
3.4.2 Módulo de Análise de Medições	53

3.5	CONSIDERAÇÕES FINAIS	57
4	MÓDULOS DE INTELIGÊNCIA COMPUTACIONAL.....	58
4.1	CONSIDERAÇÕES INICIAIS	58
4.2	PATAMAR DE CARGA.....	58
4.3	MÓDULO DE REGRESSÃO LINEAR.....	62
4.4	MÓDULO DE REDES NEURAIS ARTIFICIAIS	69
4.5	MÓDULO DE ÁRVORES DE REGRESSÃO	80
4.6	CONSIDERAÇÕES FINAIS	82
5	ESTUDOS DE CASO	84
5.1	CONSIDERAÇÕES INICIAIS	84
5.2	ANÁLISE DE QEE	85
5.3	REGRESSÃO LINEAR	90
5.4	REDES NEURAIS ARTIFICIAIS	92
5.5	ÁRVORES DE REGRESSÃO	96
5.6	CONSIDERAÇÕES FINAIS	102
6	CONSIDERAÇÕES FINAIS	104
6.1	CONCLUSÕES	104
6.2	PROPOSTAS DE TRABALHOS FUTUROS	105
6.3	RESUMO DAS PUBLICAÇÕES	106
	REFERÊNCIAS BIBLIOGRÁFICAS	107
	APÊNDICES	111
	APÊNDICE A – Organização de Arquivos e Diretórios.....	111
	APÊNDICE B – Regras das Árvores de Regressão	113

LISTA DE FIGURAS

Figura 1. Representação das componentes simétricas.....	13
Figura 2. Composição harmônica de um sinal distorcido.	15
Figura 3. Características de uma carga linear.....	16
Figura 4. Características de uma carga não linear.	16
Figura 5. Hierarquia de aprendizagem.	21
Figura 6. Variável aleatória associada com o lançamento de uma moeda.	24
Figura 7. Distribuição binomial com $n = 20$ e $p = 0,25$	27
Figura 8. Três distribuições normais.	28
Figura 9. Modelo de regressão linear simples.	30
Figura 10. Modelo de árvore de decisão para diagnóstico de pacientes.....	32
Figura 11. Estrutura básica de um neurônio artificial.	34
Figura 12. Funções de ativação que podem ser usadas por um neurônio artificial.	35
Figura 13. Validação Cruzada.	37
Figura 14. Arquitetura de uma rede MLP de duas camadas escondidas.	38
Figura 15. Diagrama Entidade Relacionamento.....	40
Figura 16. Configuração HIOKI para porcentagem da fundamental.	41
Figura 17. Arquivo de saída do HIOKI para harmônicos em porcentagem da fundamental. ..	42
Figura 18. Arquivo csv dos harmônicos em valores nominais.....	42
Figura 19. Arquivo de saída do RMS para potência.....	43
Figura 20. Arquivo de saída do RMS para harmônicos.	43
Figura 21. Caracterização do problema de identificação da contribuição de múltiplas fontes harmônicas na distorção harmônica da tensão em um ponto de interesse da rede elétrica.	44
Figura 22. Estruturação da Metodologia de Desenvolvimento do Projeto.....	45
Figura 23. Arquitetura MVC.	47
Figura 24. Estrutura do Software.....	48
Figura 25. Tela Inicial do SISQEE.....	50
Figura 26. Tela para criação de uma nova campanha de medição.	51
Figura 27. Tela de importação de uma base de dados.	51
Figura 28. Tela de Importação de Dados.....	52
Figura 29. Tela para excluir uma base de dados.....	53
Figura 30. Demonstração do processo de geração de gráficos pelo módulo de Análise de Medições.....	55
Figura 31. Gráfico de Corrente gerado pelo módulo de Análise de Medições.	56
Figura 32. Demonstração do ajuste do eixo x no gráfico gerado.	56
Figura 33. Demonstração alternativa do ajuste do eixo x no gráfico gerado	57
Figura 34. Legenda do Gráfico.....	57
Figura 35. Exportar Gráfico.....	57
Figura 36. Aba Patamar de Carga.....	59
Figura 37. Aba Arquivo Externo.	59
Figura 38. Aba Construir Modelo.	60
Figura 39. Tela Principal do Módulo de Regressão Linear.	65
Figura 40. Resultados Regressão Linear: Aba Modelo.	65
Figura 41. Resultados Regressão Linear: Aba Estatística.	66
Figura 42. Resultados Regressão Linear: Aba Background - Tabela de Métricas.....	66
Figura 43. Resultados Regressão Linear: Aba Background - Tabela de Impacto.....	67
Figura 44. Resultados Regressão Linear: Gráfico do Modelo de Regressão Gerado.....	67
Figura 45. Resultados Regressão Linear: Gráfico de Impactos Pontuais.....	68

Figura 46. Resultados Regressão Linear: Gráfico de Impactos Pontuais <i>HTML</i> .	68
Figura 47. Esquema geral do modelo proposto para a estimação da tensão harmônica em uma determinada barra de interesse.	69
Figura 48. Tela de Configuração de Parâmetros de RNA.	75
Figura 49. Topologia da RNA.	75
Figura 50. Tela Principal do Módulo de Redes Neurais Artificiais.	76
Figura 51. Resultados RNA: Aba Modelo.	76
Figura 52. Resultados RNA: Aba Background - Tabela de MAE.	77
Figura 53. Resultados RNA: Aba Vector Norm - Impacto.	77
Figura 54. Gráfico das séries temporais das tensões estimadas pelo modelo de RNA.	78
Figura 55. Gráfico dos fatores de impacto do modelo de RNA.	78
Figura 56. Gráfico das séries temporais das tensões estimadas pelo modelo de RNA em html.	79
Figura 57. Gráfico das séries temporais das tensões estimadas pelo modelo de RNA em html.	79
Figura 58. Tela de Configuração de Parâmetros de AR.	81
Figura 59. Exemplo contendo parte do ramo direito da AR.	81
Figura 60. Tela Principal do Módulo de Árvores de Regressão.	82
Figura 61. Diagrama Unifilar da Subestação Distrito Industrial.	85
Figura 62. Diagrama Unifilar da Campanha de Medição 6.	86
Figura 63. Tensão RMS medida no lado de baixa tensão do transformador DITF4-04 da SE DI.	87
Figura 64. Taxa de distorção harmônica total de tensão medida do transformador DITF4-04 da SE DI.	88
Figura 65. Corrente RMS medida no lado de baixa tensão do transformador DITF4-04 da SE DI.	89
Figura 66. Taxa de distorção harmônica total de corrente medida no transformador DITF4-04 da SE DI.	90
Figura 67. Modelo de RL do 3º harmônico para as três fases da fonte DIAL2-16 em relação a transformador DITF4-04 da SE DI.	91
Figura 68. Fator de Impacto Pontual do 3º harmônico para as três fases da fonte DIAL2-16 em relação a transformador DITF4-04 da SE DI.	91
Figura 69. Topologia das RNAs do Estudo de Caso.	93
Figura 70. Tensões de saída da RNA de 3º harmônico para as três fases na barra de 13,8 kV do transformador DITF4-04.	94
Figura 71. Impactos de 3º harmônico do modelo de RNA para as três fases na barra de 13,8 kV do transformador DITF4-04.	95
Figura 72. Tensões de saída do modelo de AR de 3º harmônico para as três fases na barra de 13,8 kV do transformador DITF4-04.	98
Figura 73. Impactos de 3º harmônico do modelo de AR para as três fases na barra de 13,8 kV do transformador DITF4-04.	99
Figura 74. Árvores de regressão das três fases.	101
Figura A1. Estrutura dos diretórios quando uma base de dados é importada.	111
Figura A2. Estrutura de diretórios dos gráficos.	111
Figura A3. Arquivos html criados no processo de geração de gráficos pelo Módulo de Análise de Medições.	112
Figura A4. Conjunto de Arquivos CSV's.	112
Figura A5. Resultados Regressão Linear: Diretório de Resultados.	112

LISTA DE TABELAS

Tabela 1 - Pontos de Conexão em Tensão Nominal igual a 220 V.....	11
Tabela 2 - Pontos de Conexão em Tensão Nominal igual a 380 V.....	11
Tabela 3 - Pontos de Conexão em Tensão Nominal igual a 440 V.....	11
Tabela 4 - Pontos de Conexão em Tensão Nominal superior a 1 kV e inferior a 69 kV.	11
Tabela 5 - Pontos de Conexão em Tensão Nominal igual ou superior a 69 kV e inferior a 230 kV.....	11
Tabela 6 - Pontos de Conexão em Tensão Nominal igual ou superior a 230 kV.....	11
Tabela 7 - Valores de Referência globais das distorções harmônicas totais (em porcentagem da tensão fundamental).....	18
Tabela 8 - TR armazenadas no banco de dados.....	41
Tabela 9 - Recursos Computacionais.....	46
Tabela 10. Faixas de Classificação do R ² Quadrado.....	64
Tabela 11. Descrição da Campanha de Medição.....	84
Tabela 12. Valores máximos, mínimos e médios de tensão do transformador DITF4-04 da SE DI.....	87
Tabela 13. DRP transformador DITF4-04 da SE DI.....	87
Tabela 14. DRC transformador DITF4-04 da SE DI.....	87
Tabela 15. Valores máximos, mínimos, médios e percentis de DHT _v medidos no lado de baixa tensão do transformador DITF4-04 da SE DI.....	88
Tabela 16. Valores máximos, mínimos e médios de corrente medidos no lado de baixa tensão do transformador DITF4-04 da SE DI.....	89
Tabela 17. Valores máximos, mínimos, médios e percentis de DHT _i medidos no lado de baixa tensão do transformador DITF4-04 da SE DI.....	90
Tabela 18. Comparação dos impactos harmônicos na distorção de tensão devido às fontes de corrente harmônica e de background usando modelo de RL construído para o estudo de caso.....	92
Tabela 19. Fator de Impacto (%) com Background do Estudo de Caso.....	93
Tabela 20. Métricas de avaliação do modelo de RNA do Estudo de Caso.....	96
Tabela 21. Fator de Impacto (%) com Background do Estudo de Caso.....	97
Tabela 22. Métricas de avaliação do modelo de AR do Estudo de Caso.....	100
Tabela B1. Relações causa-efeito da árvore de regressão que levam ao menor valor de distorção harmônica de tensão de 3 ^a ordem da fase A.....	113
Tabela B2. Relações causa-efeito da árvore de regressão que levam ao menor valor de distorção harmônica de tensão de 3 ^a ordem da fase B.....	114
Tabela B3. Relações causa-efeito da árvore de regressão que levam ao menor valor de distorção harmônica de tensão de 3 ^a ordem da fase C.....	115

LISTA DE ABREVEATURAS E SIGLAS

AM	Aprendizado de Máquina
ANEEL	Agência Nacional de Energia Elétrica
AR	Árvore de Regressão
CART	Classification and Regression Trees
DHT	Distorção Harmônica Total
DRC	Duração Relativa da Transgressão de Tensão Crítica
DRCM	Duração Relativa da Transgressão Máxima de Tensão Crítica
DRP	Duração Relativa da Transgressão de Tensão Precária
DRPM	Duração Relativa da Transgressão Máxima de Tensão Precária
IC	Inteligência Computacional
IF	<i>Impact Factor - Fator de impacto</i>
MAE	<i>Mean Absolut Error</i>
MLP	<i>Multi-Layer Perceptron</i>
MSE	<i>Mean Square Error</i>
MVC	<i>Model-View-Controller</i>
ONS	Operador Nacional do Sistema Elétrico
PAC	Ponto de Acoplamento Comum
PRODIST	Procedimentos de Distribuição
QEE	Qualidade da Energia Elétrica
RL	Regressão Linear Simples
RMSE	<i>Root Mean Saquare Error</i>
RNA	Redes Neurais Artificiais
SE	Sub Estação
SISQEE	Sistema Computacional de Suporte à Avaliação da Qualidade de Energia em Regime Permanente em Redes de Distribuição
TL	Tensão de Leitura
TR	Tensão de Referência
UC	Unidade Consumidora
VBG	Tensão de Background
VE	Tensão Estimada

RESUMO

Este trabalho apresenta o desenvolvimento de um software que fornece um ambiente computacional para que o usuário possa realizar análises de Qualidade de Energia Elétrica (QEE) seguindo as normas estipuladas pela Agência Nacional de Energia Elétrica (ANEEL) através dos Procedimentos de Distribuição (PRODIST). Também é possível realizar análises de contribuição de impactos causados pela injeção de harmônicos na rede elétrica, apontando as possíveis parcelas percentuais de contribuição de cada Unidade Consumidora (UC), através de módulos que realizam um estudo de sensibilidade por meio de técnicas de Inteligência Computacional (IC) como Redes Neurais Artificiais (RNA), Árvores de Regressão (AR) e Regressão Linear Simples (RL). Para construção dos modelos são utilizados dados de séries temporais de harmônicos de corrente e de tensão que foram obtidos através de campanhas de medição.

PALAVRAS-CHAVES: Qualidade de Energia Elétrica, Harmônicos, Software, Regressão Linear, Árvore de Regressão, Redes Neurais Artificiais.

ABSTRACT

This work presents the development of software that provides a computational environment so that the user can carry out analyzes of Quality of Electrical Energy (QEE) following the rules stipulated by the National Electrical Power Agency (ANEEL) through the Distribution Procedures (PRODIST). Also it is possible to do analysis of contribution of impacts caused by the injection of harmonics in the electric network, pointing out the possible percentage shares of contribution of each Consumer Unit (UC), through modules that perform a sensitivity study through Computational Intelligence (CI) techniques as Artificial Neural Network (ANN) Regression Trees (RT) and Simple Linear Regression (LR). For the construction of the models are used time series data of current and voltage harmonics that were obtained through measurement campaigns.

KEYWORDS: Quality of Electrical Energy, Harmonics, Software, Linear Regression, Regression Tree, Artificial Neural Network.

1 INTRODUÇÃO

1.1 CONTEXTUALIZAÇÃO

Com o avanço da tecnologia, torna-se cada vez maior o número de equipamentos que injetam harmônicos na rede elétrica, devido a grande mudança do perfil destes equipamentos gerada pela inclusão de eletrônica de potência nos mesmos, os quais eram em sua maioria eletromecânicos, e passaram a possuir um conjunto de componentes de eletrônica de potência como: diodos, capacitores, resistores entre outros. Estes componentes modificam o comportamento na forma de onda de tensão causando distorções na rede elétrica, produzindo assim uma energia elétrica de menor qualidade.

O aumento contínuo dos níveis de distorção harmônica nas redes de distribuição causa preocupação às concessionárias de energia elétrica, pois a presença dos harmônicos na rede de distribuição provoca, entre outros, aumento de ocorrências de ressonâncias harmônicas, queima de células capacitivas em bancos de capacitores, aumento das perdas elétricas, introdução de torque em oposição à rotação de máquinas elétricas e surgimento de zumbidos em transformadores, este último efeito é conhecido como magnetostricção (ESTEVEES, 2012).

Hoje ainda não existe uma legislação definida destinada a punição de indústrias por injetarem harmônicos na rede elétrica, ou a punição de concessionárias de energia elétrica por fornecerem uma forma de onda de tensão distorcida para um determinado cliente. No entanto, está sendo discutida a criação de normas técnicas para atender a este propósito, que resultarão em multas para as indústrias e concessionários pelo não cumprimento destas.

A partir do módulo 8 do manual do Procedimentos de Distribuição (PRODIST) estipulado pela Agência Nacional de Energia Elétrica (ANEEL) que estabelece os procedimentos relativos à qualidade da energia elétrica (QEE), abordando a qualidade do produto e a qualidade do serviço prestado, serão criadas as normas que estabelecerão a qualidade do serviço prestado pelas distribuidoras aos consumidores, utilizando os indicadores das faixas de classificação de tensões, tornando este estudo de caso de suma importância, uma vez que não existem trabalhos com esta ênfase sendo desenvolvidos nas concessionárias e nem na literatura.

Diante do cenário exposto, esta dissertação apresenta o desenvolvimento de um sistema computacional para apoio na tomada de decisão na busca pela Unidade Consumidora (UC) que

estiver gerando maior impacto, causado pela injeção de harmônicos na rede elétrica. Para isso, são usadas técnicas de Inteligência Computacional (IC) como Redes Neurais Artificiais (RNA), Árvores de Regressão (AR) e Regressão Linear Simples (RL). Para a construção dos modelos do sistema elétrico, são usados dados de harmônicos de corrente e de tensão extraídos de campanhas de medições simultâneas realizadas em consumidores industriais. Cada técnica utiliza uma série temporal de tensão harmônica como saída e séries de corrente como entrada, então a partir de uma análise de sensibilidade é extraído a porcentagem de impacto de cada entrada. Além disso, o software fornece um ambiente computacional para que o usuário possa realizar análises de QEE seguindo as normas estipuladas pelo PRODIST.

1.2 MOTIVAÇÃO

Cada vez mais o crescente uso de dispositivos de eletrônica de potência e cargas elétricas não lineares em geral nas instalações residenciais, comerciais e industriais, tem contribuído de forma significativa, para o aumento das distorções harmônicas de tensão, como observado amplamente nas redes de distribuição de energia.

Dando a devida atenção a este cenário, no Brasil, a ANEEL desenvolveu o PRODIST, que é um conjunto de documentos composto de nove módulos voltados para a normatização e padronização das atividades técnicas relacionadas ao funcionamento e desempenho dos sistemas de distribuição de energia elétrica.

Estas normas específicas para a distorção harmônica, que de forma similar a outras normas internacionais como a IEEE-519 (SHOJAIE; MOKHTARI, 2014) e IEC-61000-3 (UNSAR et al., 2014), estabelecem procedimentos a serem seguidos pelas concessionárias de energia elétrica, a respeito da QEE. No módulo 8 do PRODIST são estabelecidos valores de referência para indicadores de qualidade, como por exemplo, os índices de distorção harmônica total e individual, de tensão e corrente, os quais limitam os níveis de distorção harmônica na rede elétrica em patamares que não representem degradação da qualidade da energia para a operação do sistema elétrico, e que também não afetem o ciclo de vida dos dispositivos e equipamentos elétricos e eletrônicos.

Diante desse contexto, as concessionárias de energia elétrica estão cada vez mais preocupadas com o problema da distorção harmônica, uma vez que as mesmas podem sofrer sanções caso suas redes apresentem violações dos limites de distorção harmônica estabelecidos

pelo PRODIST, tornando-se de fundamental importância o monitoramento constante da rede elétrica, através da implantação de um sistema de monitoração ou a realização de campanhas de medição em suas redes, de tal forma, a acompanhar periodicamente a evolução dos níveis de distorção das mesmas, visando detectar e identificar cargas suspeitas que possam estar impactando significativamente as distorções harmônicas de tensão, observadas em pontos específicos de interesse da rede elétrica, para que então possa ser realizada a mitigação do problema antecipadamente à avaliação da ANEEL.

Dessa forma, as concessionárias precisam dispor de procedimentos que possam identificar a contribuição das cargas não lineares dos consumidores poluidores, nas distorções harmônicas de tensão observadas na rede, e poder encaminhar, com esses consumidores, um tratamento diferenciado visando à mitigação do problema.

Atualmente, a forma mais comum de mitigação do problema da distorção harmônica se baseia na aplicação de filtros passivos, porém um aspecto importante que deve ser levado em consideração é identificar as principais fontes de harmônicos dentro de um sistema elétrico e conhecer a porcentagem de contribuição de cada uma de tal forma a mitigar com maior eficiência o problema.

Atualmente, não há metodologias consolidadas que sejam capazes de estimar, com segurança, a contribuição harmônica das cargas geradoras de harmônicas nos sistemas elétricos e, portanto, que sejam capazes de atribuir a parcela de responsabilidade a essas cargas pela distorção.

Sob este ponto de vista, o presente trabalho apresenta um novo software de avaliação do impacto de múltiplas cargas geradoras de harmônicos no sistema elétrico, capaz de estimar a porcentagem de contribuição de múltiplas fontes harmônicas na distorção de um sistema elétrico qualquer.

Esta dissertação de mestrado teve origem em um projeto de Pesquisa e Desenvolvimento (P&D) financiado pela *Eletrobras Amazonas Energia* em parceria com a *Universidade Federal do Pará* (UFPA) e o *Instituto de Tecnologia e Educação Galileo da Amazônia* (ITEGAM). O título do projeto é *Desenvolvimento De Um Sistema Inteligente Para Determinação Dos Impactos Harmônicos De Múltiplos Consumidores Industriais Nas Redes De Distribuição De Energia Elétrica*.

1.3 OBJETIVOS

1.3.1 Gerais

O principal objetivo deste trabalho é apresentar uma ferramenta que seja capaz de determinar e quantificar a distorção harmônica individual de consumidores comerciais e industriais nas redes de distribuição de energia elétrica, utilizando medições simultâneas de tensões e correntes. Para isso são utilizadas técnicas de estatística não paramétrica e de inteligência computacional como: Redes Neurais Artificiais (RNA), Árvores de Regressão (AR) e Regressão Linear Simples (RL).

Além disso, a ferramenta possui um módulo para auxiliar a realizar a análise de QEE em regime permanente e análise de harmônicos em um determinado período de tempo contido dentro das campanhas de medições realizadas.

1.3.2 Específicos

- Criar um servidor de um banco de dados relacional. Para esta finalidade foi utilizado o PostgreSQL como Sistema Gerenciador de Banco de Dados e para a alimentação das bases de dados foram utilizados os dados obtidos através de campanhas de medições simultâneas de clientes comerciais e industriais, realizadas por equipamentos analisadores de QEE de duas marcas distintas, o primeiro *MARH-32* desenvolvido pela *RMS Sistemas Eletrônicos* e o segundo *PW3198* desenvolvido pela *HIOKI E.E. CORPORATION*, ambos devidamente instalados nas concessionárias e indústrias analisadas.
- Normalizar o banco de dados, permitindo que possa receber dados de diferentes analisadores de QEE, além dos analisadores da *HIOKI* e da *RMS*.
- Efetuar o pré-processamento nos arquivos de saída gerados pelos analisadores de QEE, o que inclui a criação de arquivos padronizados e normalizados, filtragem e tratamento dos dados.
- Criar um software em Java com interface gráfica amigável para o usuário.
- Calcular indicadores de conformidade do nível de tensão em regime permanente descritos no módulo 8 do manual do Prodist: O Índice de Duração Relativa da transgressão de tensão Precária (DRP) e o Índice de duração Relativa da transgressão de tensão crítica (DRC), os quais expressam o percentual do tempo no qual a unidade consumidora permaneceu com tensão precária e com tensão crítica.

- Calcular o *percentil*, índice matemático, o qual é utilizado para avaliar a Distorção Harmônica Total (DHT).
- Utilizar técnicas de estatística e modelos de inteligência computacional, bem como métodos de correlação, RL, RNA e AR, para que possa ser realizado um diagnóstico apontando as possíveis indústrias que possuem uma maior contribuição para a distorção harmônica na rede elétrica.
- Gerar relatórios contendo os gráficos e tabelas pertinentes aos módulos de análise de QEE e análise de impactos.

1.4 REVISÃO BIBLIOGRÁFICA

Atualmente, há poucos trabalhos voltados para a determinação da contribuição de cargas geradoras de harmônicos na distorção da forma de onda da tensão de um sistema elétricos. A seguir, são abordados brevemente e de forma cronológica esses trabalhos.

Em Xu e Liu (1999) foi proposto um método para quantificar as responsabilidades do consumidor e da rede quanto às violações dos limites de distorção estabelecidos pelas normas de qualidade da energia causada tanto por variações da fonte harmônica quanto por variações da impedância harmônica. As ideias centrais deste método são: o uso de uma impedância de referência e a conversão das variações da impedância em variações da corrente da fonte harmônica equivalente.

De maneira semelhante, em Xu e Liu (2000), foi proposto um método voltado para determinar a contribuição harmônica do consumidor e da concessionária de energia no ponto de acoplamento comum. Esse método é capaz de separar a corrente e tensão harmônicas em duas componentes: uma devido ao consumidor e a outra devido à rede elétrica.

Em Xu, Liu e Liu (2003), foi apresentado um estudo com objetivo de investigar a validade do método da direção da potência para a localização de fontes harmônicas, no qual mostrou que este método, que está baseado na direção da potência ativa, está teoricamente incorreto, e, portanto pode levar a resultado incorreto, uma vez que a potência ativa é principalmente influenciada pelos ângulos de fases das fontes harmônicas. Além disso, este trabalho sugere que a potência reativa possa representar uma melhor opção para avaliar a contribuição das fontes harmônicas.

Em Li *et al* (2004), foi proposto um método para determinar se o consumidor ou a concessionária apresenta maior contribuição harmônica no ponto de acoplamento comum, inspirado no fato de que a potência reativa, em detrimento da potência ativa, constitui um indicador mais confiável para identificar fontes harmônicas dominantes.

Em Nino e Xu (2007), foi apresentado um método de medição que pode determinar a fonte e a impedância harmônica para sistemas residenciais e comerciais supridos por um transformador monofásico de tal forma que seja possível determinar a contribuição da fonte supridora de energia elétrica e do consumidor no ponto de acoplamento comum.

Em Dán (2009), foi proposta o um método de identificação de fontes harmônicas com base em medições “on-site”. A partir das medições de tensão e corrente harmônica, esta método propõe seja calculada a impedância da rede e da carga, e seja separada a contribuição da rede e da carga de tal forma a quantificar essa contribuição.

Em Xu *et al* (2009), foi proposta uma metodologia, baseada no método de correlação de dados, para determinar a contribuição do impacto harmônico de múltiplas cargas geradoras de harmônicos distribuídas num sistema elétrico. Nessa metodologia, foi aplicado o método de regressão linear simples nos dados de corrente das fontes harmônicas e de tensão de uma barra do sistema (XU et al., 2009).

Em Mazin, Xu e Huang (2011), foi proposto um indicador quantitativo para mensurar o impacto harmônico de múltiplas cargas geradoras de harmônico num sistema de transmissão ou distribuição. Este indicador foi obtido a partir das medições da tensão e da corrente harmônica com base da teoria de inferência estatística e da regressão linear múltipla.

Soares (2012) implementou uma metodologia para determinar o impacto harmônico de tensão a partir das correntes medidas provenientes das cargas não lineares que compõe um determinado sistema. A metodologia proposta consiste na criação de modelos desenvolvidos com base no método de regressão polinomial local de Kernel, sendo neste trabalho considerado todo o período de medição, a análise realizada ainda considera cada carga por vez na construção do modelo estatístico.

Manito *et al* (2014) propôs o desenvolvimento de um modelo baseado em redes neurais artificiais para avaliar a contribuição harmônica de múltiplas cargas não lineares em determinadas barras do sistema elétrico. Tal modelo foi baseado em campanhas de medições e considerou a interação de todas as cargas na análise, considerando todo o período de medição.

Em Matos *et al* (2016), foi apresentado um método prático para analisar a influência de múltiplas cargas não lineares sobre a distorção harmônica medida em pontos de interesse na rede elétrica. A metodologia foi baseada exclusivamente em medições simultâneas de magnitudes de tensão e corrente, e na aplicação de modelos de regressão linear e regressão não paramétrica para expressar a correlação entre a tensão medida na rede e a corrente de carga não linear correspondente na instalação do cliente.

Em Lima (2016) foi apresentado um ambiente computacional que permitiu auxiliar os estudos de propagação de harmônicos nos sistemas elétricos realizando diagnósticos quanto aos níveis de distorções harmônicas. Este ambiente tem como base o software ATP (Alternative Transient Program) para a modelagem do sistema elétrico. Para isso, utilizou-se de medidas de correntes harmônicas obtidas por meio de analisadores de qualidade de energia. Por fim, são realizadas simulações computacionais da rede elétrica observando o comportamento harmônico em diversos pontos da rede para diferentes topologias. Com isso, foi possível avaliar o cancelamento harmônico quando diferentes tipos de cargas são conectadas no mesmo sistema elétrico.

1.5 PRINCIPAIS RESULTADOS

Baseados nos objetivos descritos no item 1.3, algumas contribuições técnicas para o setor elétrico foram realizadas.

O produto gerado por este trabalho corresponde ao sistema computacional desenvolvido para a obtenção de estimativas para os impactos que as múltiplas cargas não lineares provocam na distorção harmônica de tensão, em pontos de interesse na rede elétrica, utilizando as metodologias apresentadas em Soares (2012) e Manito *et al* (2014). O software criado foi denominado SISQEE – Sistema Computacional de Suporte à Avaliação da Qualidade de Energia em Regime Permanente em Redes de Distribuição.

Vale ressaltar que todos os analisadores de QEE disponíveis no mercado, oferecem seus próprios módulos de software para o tratamento dos dados das campanhas de medição, e apresentação de resultados em forma gráfica e em tabelas das variáveis elétricas envolvidas nessas campanhas, para auxiliar nas análises desses resultados e nos cálculos dos principais indicadores de QEE. No entanto, esses softwares são proprietários, sem ter portabilidade para tratar com dados adquiridos por analisadores de outras marcas.

Além do mais, nenhum desses softwares disponíveis realiza análises de correlação para identificar os impactos das múltiplas fontes harmônicas nas distorções harmônicas da tensão nos pontos de interesse da rede elétrica, o que é uma característica única do SISQEE, até o momento. Além do mais, o SISQEE pode processar dados de campanhas de medição de diferentes marcas de analisadores de QEE disponíveis no mercado, e da mesma forma que os softwares desses analisadores, apresenta os resultados das campanhas de medição via interface amigável usuário-sistema, para as análises e cálculos dos indicadores de QEE segundo o PRODIST. Considerando que ainda não tem disponível nas Empresas do setor elétrico brasileiro ferramenta similar, pode-se afirmar que o produto disponibilizado pelo projeto é inovador e original.

1.6 ORGANIZAÇÃO DO TEXTO

O restante deste trabalho está organizado da seguinte maneira:

Cap. 2 – Fundamentação Teórica: Neste capítulo são apresentados os conceitos básicos pertinentes às áreas relacionadas a este trabalho. Mais especificamente, disserta-se acerca dos procedimentos relativos à QEE contidos no módulo 8 do PRODIST e sobre técnicas de inteligência computacional, apresentando os conceitos e técnicas de aprendizado de máquina.

Cap. 3 – Software: Neste capítulo é dissertado sobre a construção do sistema computacional, apresentando sua arquitetura geral e interface gráfica dos módulos criados para o gerenciamento de banco de dados e para a análise de medições.

Cap. 4 – Módulos de Inteligência Computacional: Neste capítulo são apresentados as interfaces gráficas e o funcionamento dos módulos que utilizam técnicas de IC para estimação de impacto harmônico.

Cap. 5 – Estudo de Caso: Neste capítulo é feita a descrição da campanha de medição utilizada para realizar os estudos utilizando os módulos desenvolvidos no software, apresentando os resultados obtidos por cada módulo.

Cap. 6 – Conclusões: Neste capítulo resumem-se os desafios de pesquisa e desenvolvimento enfrentados nesta dissertação, e por fim, são apresentadas as considerações finais sobre os resultados alcançados.

2 FUNDAMENTAÇÃO TEÓRICA

2.1 CONSIDERAÇÕES INICIAIS

O contexto de QEE engloba qualquer assunto relacionado com problemas manifestados na tensão, corrente ou desvio de frequência, que resulta em falha ou má operação de equipamentos dos consumidores (DUGAN et al., 2002). A medida da QEE considera a continuidade de serviço e a conformidade do sinal elétrico, avaliada por meio dos seguintes parâmetros: distorções harmônicas; flutuações de tensão; variações de tensão de curta duração; desequilíbrio de sistemas trifásicos; transitórios rápidos (SOTO; CANELOI, 2015).

Ainda não existe consenso para o termo "Qualidade de Energia Elétrica", podendo haver significados distintos. Uma definição aceita é: "Energia elétrica de boa qualidade é aquela que garante o funcionamento contínuo, adequado e seguro dos equipamentos elétricos e processos associados, sem afetar o meio ambiente e o bem estar das pessoas".

Dentre os aspectos da qualidade do produto em regime permanente e transitório, são considerados: Tensão em Regime Permanente, Fator de Potência, Harmônicos e Desequilíbrio de Tensão, Flutuação de Tensão, Variações de Tensão de Curta Duração e Variação de Frequência (ANEEL, 2010). Neste capítulo são apresentados alguns destes aspectos, além de alguns conceitos importantes para o entendimento deste trabalho.

2.2 NORMAS ESTABELECIDAS PELO PRODIST

2.2.1 Tensão em Regime Permanente

Permite analisar a conformidade da tensão elétrica, dividindo a tensão eficaz em faixas categorizadas como: adequada, precária e crítica. Estas faixas levam aos indicadores de conformidade do nível de tensão em regime permanente, os quais são descritos no módulo 8 do PRODIST.

Segundo o PRODIST, para cada unidade consumidora, a tensão deve ser medida ao longo de sete dias completos com intervalo de 10 minutos entre cada amostra, então são apurados os indicadores DRP (Duração Relativa da Transgressão de Tensão Precária) e DRC (Duração Relativa da Transgressão de Tensão Crítica), os quais expressam em termos percentuais o tempo

em que a unidade consumidora permaneceu com os níveis de tensão precária e crítica (ANEEL, 2010).

A ANEEL estipula limites para estes indicadores de tensão em regime permanente. O valor da Duração Relativa da Transgressão Máxima de Tensão Precária (DRPM) fica estabelecido em 3% temporal. O valor da Duração Relativa da Transgressão Máxima de Tensão Crítica (DRCM) fica estabelecido em 0,5% temporal.

Se houver transgressão destes limites, a concessionária distribuidora de energia elétrica tem um prazo de 90 dias para regularizar a tensão do consumidor, no caso do DRP, e de 15 dias no caso do DRC. Não havendo regularização dentro do prazo estipulado, a concessionária deve realizar uma compensação financeira ao consumidor.

O cálculo dos indicadores DRP e DRC é realizado de acordo com (1) e (2):

$$DRP = \frac{nlp}{1008} \cdot 100\% \quad (1)$$

$$DRC = \frac{nlc}{1008} \cdot 100\% \quad (2)$$

Onde nlp e nlc representam o maior valor entre as fases do número de leituras situadas nas faixas precária e crítica respectivamente.

Em outras palavras, se considerarmos o período de uma semana e dividirmos em intervalos de 10 minutos têm-se a quantia total de 1008 amostras que é o recomendado pelo PRODIST. Dentro destas 1008 amostras é conferido para as fases A B e C quantas vezes a tensão ficou fora do limite adequado, nlp e nlc representam a quantidade de intervalos pertencentes às faixas precária e crítica respectivamente. Para cada fase é conferido um nlp e um nlc , obtendo-se então nlc_A , nlc_B e nlc_C para tensão crítica e nlp_A , nlp_B e nlp_C para tensão precária, sendo nlp e nlc definidos como sendo o maior valor obtido dentre as fases.

$$nlp = \text{máximo}(nlp_A, nlp_B, nlp_C)$$

$$nlc = \text{máximo}(nlc_A, nlc_B, nlc_C)$$

O PRODIST estipula os limites que compõe as faixas adequada, precária e crítica em formato de tabelas. Os limites da Tensão de Leitura (TL) variam de acordo com a tensão nominal também conhecida como Tensão de Referência (TR). A seguir, são apresentadas as Tabelas 1-6 contidas no módulo 8 do PRODIST (ANEEL, 2010).

Tabela 1 - Pontos de Conexão em Tensão Nominal igual a 220 V.

Tensão de Atendimento	Faixa de Variação da TL (Volts)
Adequada	$201 \leq TL \leq 231$
Precária	$189 \leq TL < 201$ $231 < TL \leq 233$
Crítica	$TL < 189$ ou $TL > 233$

Tabela 2 - Pontos de Conexão em Tensão Nominal igual a 380 V.

Tensão de Atendimento	Faixa de Variação da TL (Volts)
Adequada	$348 \leq TL \leq 396$
Precária	$327 \leq TL < 348$ $396 < TL \leq 403$
Crítica	$TL < 327$ ou $TL > 403$

Tabela 3 - Pontos de Conexão em Tensão Nominal igual a 440 V.

Tensão de Atendimento	Faixa de Variação da TL (Volts)
Adequada	$402 \leq TL \leq 458$
Precária	$380 \leq TL < 402$ ou $458 < TL \leq 466$
Crítica	$TL < 380$ ou $TL > 466$

Tabela 4 - Pontos de Conexão em Tensão Nominal superior a 1 kV e inferior a 69 kV.

Tensão de Atendimento	Faixa de Variação da TL em Relação à TR
Adequada	$0,93 TR \leq TL \leq 1,05 TR$
Precária	$0,90 TR \leq TL < 0,93 TR$
Crítica	$TL < 0,90 TR$ ou $TL > 1,05 TR$

Tabela 5 - Pontos de Conexão em Tensão Nominal igual ou superior a 69 kV e inferior a 230 kV.

Tensão de Atendimento	Faixa de Variação da TL em Relação à TR
Adequada	$0,95 TR \leq TL \leq 1,05 TR$
Precária	$0,90 TR \leq TL < 0,95 TR$ $1,05 TR < TL \leq 1,07 TR$
Crítica	$TL < 0,90 TR$ ou $TL > 1,07 TR$

Tabela 6 - Pontos de Conexão em Tensão Nominal igual ou superior a 230 kV.

Tensão de Atendimento	Faixa de Variação da TL em Relação à TR
Adequada	$0,95 TR \leq TL \leq 1,05 TR$
Precária	$0,93 TR \leq TL < 0,95 TR$ $1,05 TR < TL \leq 1,07 TR$
Crítica	$TL < 0,93 TR$ ou $TL > 1,07 TR$

Alguns dos valores de tensão contidos nas tabelas extraídas diretamente do PRODIST estão em *volts*, porém outros estão em *pu* (Sistema Por Unidade). Sistema por unidade é uma forma de expressar as grandezas elétricas em um circuito de forma normalizada, para isso é definido um valor base, com base em valores pré-determinados. Para converter um valor de tensão medido de *volts* para *pu*, divide-se esta tensão por uma tensão nominal escolhida como tensão base, conforme a equação (3).

$$V_{pu} = \frac{V_{medido}}{V_{base}} \quad (3)$$

2.2.2 Fator de Potência

O Fator de Potência está relacionado com a razão entre a potência ativa e a potência aparente total e pode ser calculado através da Equação (4). Para uma unidade consumidora ou conexão entre distribuidoras com tensão inferior a 230 kV, o fator de potência no ponto de conexão deve estar compreendido entre 0,92 (noventa e dois centésimos) e 1,00 (um) tanto indutivo quanto capacitivo, de acordo com regulamentação vigente (PRODIST).

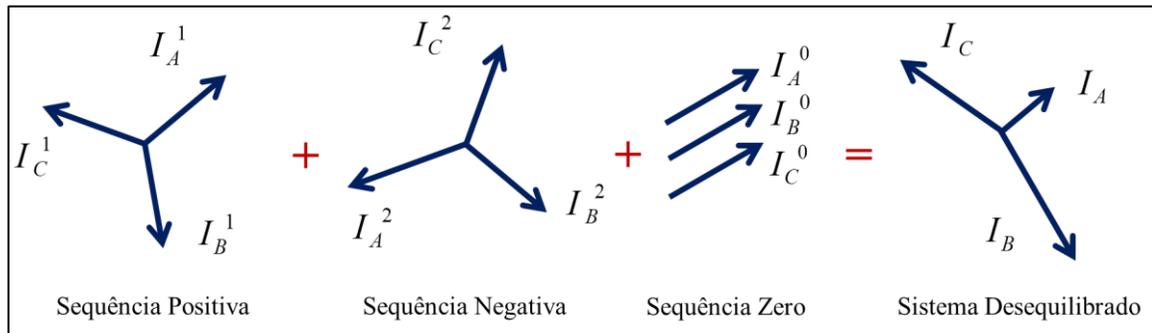
$$fp = \frac{P}{\sqrt{P^2 + Q^2}} \quad (4)$$

2.2.3 Desequilíbrio

O desequilíbrio de tensão é o fenômeno associado a alterações dos padrões trifásicos do sistema de distribuição. As tensões trifásicas no terminal do gerador são iguais em magnitude e com deslocamento de 120° entre as fases. Este sistema vetorial é, dessa forma, dito simétrico. Se os componentes do sistema elétrico são lineares e simétricos, e as fases individuais estão igualmente carregadas, as tensões medidas nas barras de carga continuam sendo simétricas.

As principais fontes de desequilíbrios permanentes são as cargas, principalmente as cargas monofásicas que estão presentes principalmente nos sistemas de baixa e média tensão. Uma técnica amplamente utilizada para tratar sistemas desequilibrados, é a técnica de decomposição em componentes simétricas, de tal forma que um conjunto assimétrico de fasores (tensão e/ou corrente) pode ser transformado em três conjuntos simétricos ditos: componentes de sequência positiva, sequência negativa e sequência zero. Portanto, um sistema trifásico composto por três fasores é decomposto por componentes de sequência, conforme apresentado na Figura (1).

Figura 1. Representação das componentes simétricas.



Fonte: LIMA, 2016.

O indicador responsável para avaliar o desequilíbrio de tensão no barramento da rede básica é o Fator de Desequilíbrio de Tensão. O valor de referência nos barramentos do sistema de distribuição, com exceção das redes de baixa tensão, deve ser igual ou inferior a 2%. Esse valor serve para referência do planejamento elétrico em termos de QEE e que, regulatoriamente, será estabelecido em resolução específica, após período experimental de coleta de dados (ANEEL, 2010). Este fator é obtido como a relação percentual da componente de sequência negativa (e/ou zero) pela componente de sequência positiva, de qualquer fase. O Fator de Desequilíbrio de Tensão é calculado pela equação (5).

$$FD = \frac{V_-}{V_+} \cdot 100\% \quad (5)$$

Onde V_- corresponde à tensão da sequência negativa e V_+ a tensão da sequência positiva. O desequilíbrio pode causar efeitos indesejáveis na rede elétrica, pois a circulação de correntes de sequência negativa e zero nos sistemas elétricos resultam em:

- Perdas adicionais de potência e energia.
- Aquecimento adicional, que como consequência pode limitar a capacidade de transporte das linhas para sequência positiva.
- Desequilíbrio das tensões nos nós elétricos.

O desequilíbrio de tensão afeta de forma adversa a operação de muitas cargas. Os motores assíncronos, geradores síncronos e retificadores são as cargas mais sensíveis a este aspecto.

Os motores assíncronos normalmente têm seus enrolamentos conectados em delta ou estrela com um ponto central aterrado. Dessa forma a operação do motor é somente afetada pelas componentes de sequência positiva e negativa. A corrente de sequência negativa cria um fluxo

que gira na direção oposta ao movimento do rotor. Este fluxo frenador pode causar (PIRES, 2010):

- Aquecimento adicional dos enrolamentos do estator;
- Perdas adicionais de potência ativa no estator;
- Torque adicional operando na direção oposta do movimento do rotor;
- Indução de correntes adicionais nos enrolamentos, e no ferro do rotor, e consequentemente gera perdas adicionais no rotor.

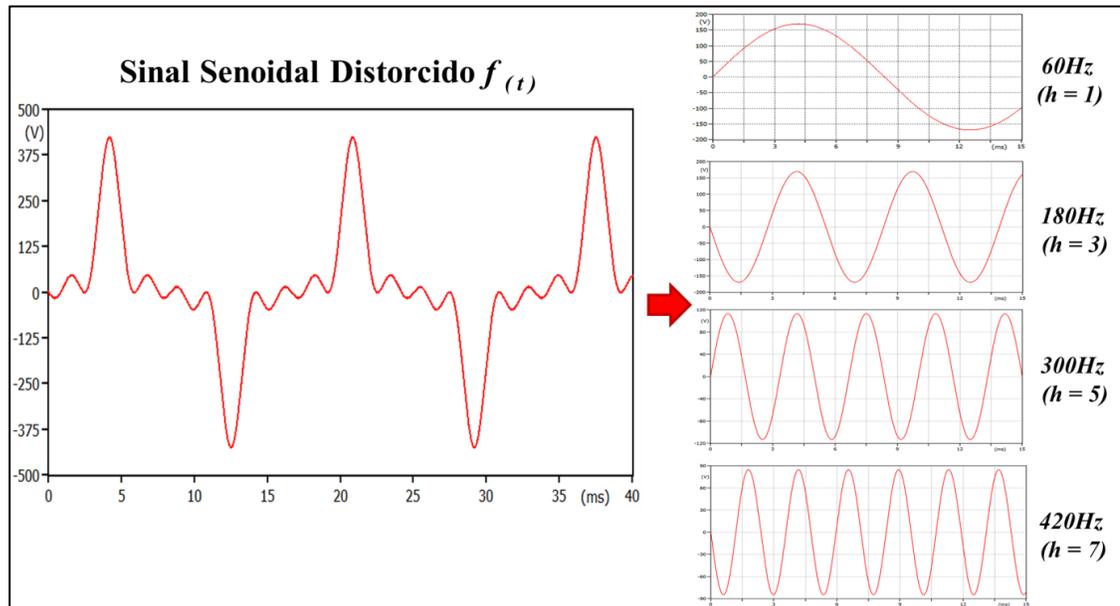
2.3 CARACTERIZAÇÃO DE HARMÔNICOS

Harmônicos é um conjunto de sinais com forma de onda senoidal e com frequências múltiplas inteiras de uma frequência de referência, chamada de frequência fundamental. Esses componentes se combinam com o componente fundamental e produzem as distorções harmônicas. As distorções harmônicas são fenômenos associados com a deformação na forma de onda da tensão e corrente em relação à onda senoidal da frequência fundamental (LEÃO; SAMPAIO; ANTUNES, 2014).

As distorções harmônicas em sistemas elétricos podem ser extremamente prejudiciais para o funcionamento adequado dos equipamentos elétricos (ALMEIDA, 2012). Alguns efeitos são ocasionados pela distorção harmônica, como: aquecimento dos condutores, alteração do torque e ruídos em motores e danos a bancos de capacitores devido a fenômenos de ressonância (ARRILLAGA, 1997; ARRILLAGA; WATSON, 2004).

No Brasil esta frequência fundamental é definida como 60 Hz. A Figura (2) apresenta a decomposição do sinal distorcido nas componentes harmônicas de 3^a, 5^a e 7^a ordem, que correspondem às frequências 180, 300 e 420 Hz, respectivamente.

Figura 2. Composição harmônica de um sinal distorcido.



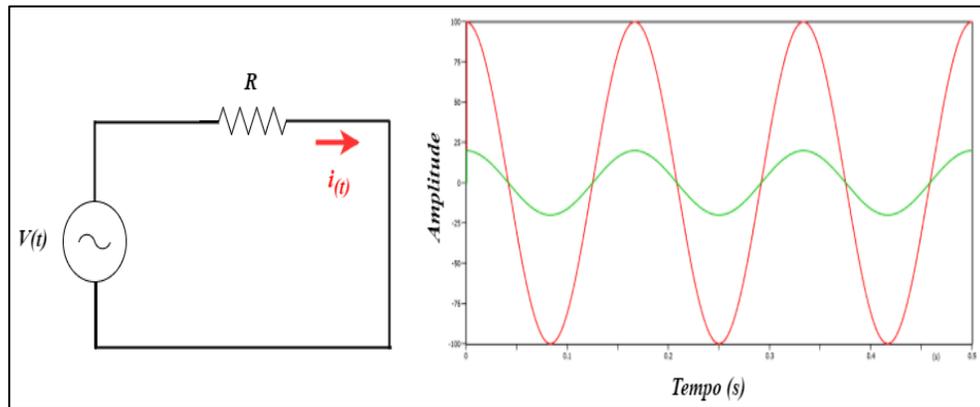
Fonte: LIMA, 2016.

As distorções harmônicas podem ser causadas tanto por cargas produzidas por equipamentos de consumidores, quanto por elementos do sistema elétrico como:

- Compensadores estáticos;
- Geradores e compensadores síncronos em menor escala;
- Conversores de Corrente Contínua (CC) / Corrente Alternada (CA);
- Transformadores.

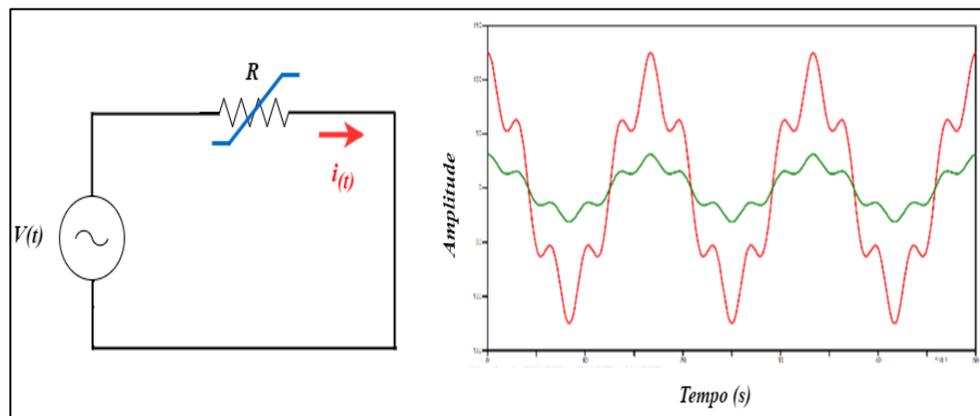
Os harmônicos podem ser gerados por cargas lineares e não lineares. Uma carga linear se caracteriza quando a corrente que circula na carga é diretamente proporcional à tensão aplicada. Logo, em geral são constituídas por resistores, indutores não saturáveis e capacitores de valores fixos. Assim, os mesmos drenam correntes de mesma forma de onda da fonte de alimentação (LIMA, 2016). Uma carga não linear é aquela que, quando alimentada com tensão senoidal, absorve uma corrente não senoidal, ou seja, a carga não-linear tem a característica de deformar a forma de onda da corrente absorvida. As Figuras (3) e (4) apresentam as características de uma carga linear e não linear respectivamente.

Figura 3. Características de uma carga linear.



Fonte: LIMA, 2016.

Figura 4. Características de uma carga não linear.



Fonte: LIMA, 2016.

A presença de harmônicos no sinal elétrico pode causar ressonância harmônica na rede do sistema elétrico. Ressonância harmônica é a condição na qual a frequência natural do sistema de energia está em ressonância com o harmônico de maior intensidade produzido por cargas não lineares. Dentre as consequências da ressonância harmônica estão:

- Sobretensões no Ponto de Acoplamento Comum (PAC), podendo causar danos a equipamentos ou reduzir sua vida útil.
- Aumento da distorção harmônica de tensão nos sistemas de energia elétrica, provocando redução da vida útil de equipamentos, falhas em Sistemas sensíveis, interferência em sistemas de comunicação.
- Sobrecorrentes nos equipamentos.
- Perdas elétricas pelo aquecimento de linha dos transformadores.

A forma mais comum de mitigação está relacionada à aplicação de filtros harmônicos. O objetivo é oferecer às componentes da onda de corrente e/ou tensão com frequências diferentes da fundamental um caminho alternativo de circulação, deixando fluir para o sistema alimentador somente a componente fundamental das mesmas, correspondente à frequência nominal da rede (PIRES, 2010).

As soluções mais comuns de filtragem harmônica utilizam filtros passivos sintonizados ou amortecidos. Sua principal função é reduzir a amplitude da corrente absorvida pela carga em uma frequência específica ou em uma determinada faixa de frequências a fim de evitar a sua propagação pelo sistema. Seu dimensionamento depende tanto do perfil da corrente produzida pelas cargas geradoras de harmônicos como da impedância do sistema.

Os níveis de distorção harmônica podem ser caracterizados pela análise do espectro harmônico considerando as amplitudes de cada componente harmônico individualmente. Também é comum utilizar uma medida global, a DHT (Distorção Harmônica Total), medida que considera a magnitude de todos os componentes harmônicos, conforme as Equações (6) e (7).

$$DHT_v (\%) = \frac{\sqrt{\sum_{h=2}^{h_{max}} V_h^2}}{V_1} \cdot 100\% \quad (6)$$

$$DHT_i (\%) = \frac{\sqrt{\sum_{h=2}^{h_{max}} I_h^2}}{I_1} \cdot 100\% \quad (7)$$

Onde:

V_h, I_h : são valores eficazes da tensão e da corrente harmônica de ordem h .

V_1, I_1 : são valores eficazes da componente fundamental da tensão e da corrente.

h_{max} : é a ordem do componente harmônico de maior frequência presente no sinal.

A Tabela 7 apresenta os limites permitidos para a DHT estipulados pelo PRODIST (ANEEL, 2010). No Brasil, estes valores servem para referência do planejamento elétrico em termos de QEE e que, serão estabelecidos em resolução específica, após período experimental de coleta de dados.

Tabela 7 - Valores de Referência globais das distorções harmônicas totais (em porcentagem da tensão fundamental).

Tensão Nominal do Barramento	Distorção Harmônica Total de Tensão DHT(%)
$VN \leq 1\text{kV}$	10
$1\text{kV} < VN \leq 13,8\text{kV}$	8
$13,8\text{kV} < VN \leq 69\text{kV}$	6
$69\text{kV} < VN \leq 230\text{kV}$	3

O indicador para avaliar o desempenho global quanto a harmônicos, em regime permanente, nos barramentos da Rede Básica corresponde à distorção harmônica de tensão.

Para efeito de comparação com o nível representativo do indicador padrão global, é necessário determinar o percentil de 95%, ou seja, considera-se o valor que foi superado em apenas 5% dos registros obtidos no período de 24 horas durante uma semana, e então é escolhido o maior valor dentre os sete obtidos anteriormente, para cada dia. Chama-se este novo valor de DTHTS95% (ANEEL, 2010)

O valor de cada indicador de harmônico (distorção total ou individual) é o valor máximo, dentre as sete amostras obtidas para um período de sete dias consecutivos, do percentil de 95% dos valores diários de cada indicador. Em outras palavras, o indicador que deve ser comparado aos padrões estabelecidos nos Procedimentos de Rede é o valor máximo semanal do percentil de 95% diário dos indicadores (ANEEL, 2010).

Segundo a Oracle, o percentil é uma função de distribuição inversa que assume um modelo de distribuição contínua. O percentil é uma medida da posição relativa de uma unidade de observação em relação a todas as outras de um conjunto de valores dispostos em ordem crescente, onde o p -ésimo percentil tem no máximo $p\%$ dos valores abaixo daquele ponto e no mínimo $(100 - p\%)$ dos valores acima. Para ser calculado é preciso um valor percentual e o resultado será o valor que ocupar a posição percentual fornecida no conjunto de dados através de (8), onde L_p é a posição da observação, p é o valor do percentil e n é o número de observações.

$$L_p = (n - 1) \frac{p}{100} + 1 \quad (8)$$

Caso a posição calculada não seja um número inteiro, é necessário realizar uma média dos valores que ocuparem as posições imediatamente acima e abaixo através de (9), onde V_{inf} é o valor abaixo, V_{sup} é o valor acima e β representa a parte decimal de V_{inf} .

$$Percentil(p) = V_{inf} + (V_{sup} - V_{inf}) \cdot \beta \quad (9)$$

2.4 INTELIGÊNCIA COMPUTACIONAL

2.4.1 Aprendizado de Máquina

Diversos domínios do conhecimento humano tendem a concentrar grandes volumes de dados, como por exemplo: astronomia, climatologia, geologia e outros. Todas essas informações precisam ser armazenadas e gerenciadas, para que seja possível retirar o máximo de conhecimento (MITRA; ACHARYA, 2003).

Com o passar dos anos, cada vez mais a tecnologia teve espaço nesses diversos domínios, disponibilizando um número maior de possibilidades no que se refere à aquisição de dados e, a partir da década de 1970, houve uma maior disseminação da utilização de técnicas de computação baseadas em inteligência artificial, voltadas para solucionar problemas reais, que tendem a crescer, seja em complexidade, seja em volume de dados. Segundo (LUGER; STUBBLEFIELD, 1998), define-se como inteligência artificial o ramo da ciência da computação que se preocupa com a automatização do comportamento inteligente. Com isso, necessita-se que as técnicas sejam capazes de criar por si próprias, a partir do conhecimento adquirido por experiências passadas, uma hipótese que consiga resolver o problema em questão. Esse processo é conhecido como Aprendizado de Máquina (AM) (FACELI *et al.*, 2011).

Baseando-se em Mitchell (1997), o AM se constitui como a capacidade de aprendizagem a partir de experiências passadas, para realização de uma determinada tarefa, focando-se em melhorar seu desempenho (TOM M. MITCHELL, 2009). Para o processo de aprendizagem, é aplicado o princípio de inferência denominado indução, no qual se obtêm conclusões genéricas a partir de um conjunto particular de exemplos (FACELI *et al.*, 2011). Para Rezende (2003), a indução é a forma de inferência lógica que permite conclusões genéricas a partir de um conjunto particular de exemplos (REZENDE, 2003).

Os modelos gerados por algoritmos de AM são treinados a partir de um conjunto de dados, com o objetivo de induzir hipóteses válidas capazes de lidar com situações não apresentadas durante seu desenvolvimento, para outros objetos do mesmo domínio, sem haver necessidade de refazer o modelo. A essa propriedade se dá o nome capacidade de generalização da hipótese.

De acordo com Faceli *et al.* (2011), quando uma hipótese apresenta uma baixa capacidade de generalização, significa que a mesma está *superajustada* aos dados existentes na base, o que caracteriza-se como *overfitting* (FACELI *et al.*, 2011). No caso inverso, o *underfitting*, o

algoritmo induz hipóteses com baixa taxa de acerto, configurando uma condição de *subajustamento*. Segundo Monard e Baranauskas (2003), um dos motivos para que essa situação ocorra, é quando os exemplos de treinamento disponíveis na base são pouco representativos ou também pelo modelo ser muito simples, e não capture os padrões existentes (MONARD; BARANAUSKAS, 2003).

Tanto o *overfitting* como o *underfitting* deve ser evitado quando o modelo criado pelo algoritmo de AM for treinado, pois não representam bem o modelo real. Em outras palavras, o *overfitting* significa que houve um sobre-treinamento, ou seja, o modelo aprendeu os padrões da base de treino e consegue representar muito bem apenas a base de dados treinada, ou seja, não consegue se adaptar a novos dados, já o *underfitting* significa que o modelo criado não aprendeu direito o padrão dos dados treinados e, portanto, também não consegue se adaptar a novos dados.

A forma como é representada e pesquisada, a hipótese induzida é particular de cada algoritmo de AM, agindo como uma restrição que implicará na hipótese que melhor se ajusta ao conjunto de treinamento. Sem esses procedimentos, os modelos serão especializados para cada situação presente na base, com isso, não conseguiriam aplicar o conhecimento adquirido para novos dados.

Outro aspecto importante para a construção de um bom modelo é o processo de tratamento dos dados. O modelo precisa ser capaz de lidar com a presença de diversos tipos de imperfeições, como, dados ausentes, inconsistentes, redundantes e ruidosos. Para solucionar esses problemas, é necessário, muitas vezes, que a base de dados original passe por uma etapa de pré-processamento. Em tal etapa, são utilizadas diversas técnicas para que o modelo não tenha dificuldade em induzir hipóteses corretamente. Segundo Faceli et al. (2011), essas técnicas podem ser agrupadas da seguinte forma (FACELI et al., 2011):

- Eliminação manual de atributos;
- Integração de dados;
- Amostragem de dados;
- Balanceamento de dados;
- Limpeza de dados;
- Redução de dimensionalidade;
- Transformação de dados.

Os algoritmos de AM podem ser organizados de acordo com o paradigma de aprendizado a ser adotado para a tarefa, sendo que podem ser divididas em: preditivas e descritivas, conforme mostrado na Figura (5).

Figura 5. Hierarquia de aprendizagem.



Fonte: FACELI et al., 2011.

2.4.1.1 Modelos Preditivos

São modelos que correspondem um conjunto de amostras rotuladas a uma função que irá estimar seus valores, sejam eles de natureza discreta ou contínua. Quando o rótulo é formado por um domínio com valores nominais, um estimador criado é um **classificador**. Quando o rótulo é constituído por valores numéricos, ordenado e infinito, o estimador é conhecido como **regressor**. Tanto modelos de classificação quanto de regressão possuem aprendizado supervisionado. Segundo Bigus (1996), o aprendizado supervisionado é utilizado quando, em um banco de dados, se tem tanto as perguntas como as respostas (BIGUS, 1996). Com isso, faz-se necessário o uso de atributos de entrada e de saída (também conhecido como atributo alvo), para que uma hipótese seja capaz de relacionar os valores dos atributos de entrada de uma determinada amostra presente no conjunto de treinamento, ao seu valor correspondente ao seu atributo de saída. De acordo com Faceli et al. (2011), os principais métodos preditivos são organizados em (FACELI et al., 2011):

- **Métodos baseados em distância:** consideram a proximidade entre os dados na realização de predições. São afetados pela medida ou função de distância utilizada, e necessitam que os atributos sejam normalizados, para que todos os valores estejam na mesma escala. Um método baseado em distância utilizado com frequência é o algoritmo dos vizinhos mais próximos.
- **Métodos probabilísticos:** são algoritmos baseados no teorema de Bayes, utilizados para

tarefas de previsão e diagnóstico. Representam a distribuição de probabilidade conjunta de um determinado grupo de variáveis aleatórias, ou seja, as informações disponíveis são de natureza imprecisa. O algoritmo mais conhecido desse método é o Naive-Bayes.

- **Métodos baseados em procura:** são construídos com base na caracterização do problema, de acordo com a busca em um espaço de possíveis soluções. Representam estratégias capazes de encontrar uma solução em uma base conhecida, de forma rápida, eficaz e generalizada. Os modelos que são frequentemente apontados na literatura são a árvore de decisão e os modelos baseados em regras.
- **Métodos baseados em otimização:** são problemas de maximização ou minimização, onde buscam uma hipótese que descreva as informações de acordo com a função que se pretende otimizar, levando em consideração um conjunto de restrições. Um dos modelos mais conhecidos são os chamados de algoritmos genéticos.

2.4.1.2 Modelos Descritivos

As tarefas do aprendizado descritivo, ou não supervisionado, se referem à identificação de informações relevantes nos dados sem a presença de um elemento externo para guiar o aprendizado. Essencialmente, o aprendizado reside na identificação de propriedades intrínsecas aos dados de entrada, de maneira a construir representações desses dados que possam servir a diversos propósitos como auxílio à tomada de decisões ou descoberta de conhecimento. Essas técnicas são utilizadas principalmente quando o objetivo do aprendizado é encontrar padrões ou tendências que auxiliem no entendimento dos dados (DE SOUTO et al., 2003). Mais precisamente no aprendizado não supervisionado não existem atributos meta. A partir do conjunto de dados x , um algoritmo de AM não supervisionado aprende a representar as entradas submetidas segundo algum critério de qualidade. As tarefas descritivas podem ser divididas em: sumarização, associação e agrupamento (FACELI et al., 2011).

- **Métodos de sumarização:** tem o objetivo de encontrar uma descrição simples e compacta dos dados. Para isso, podem ser utilizadas desde medidas estatísticas simples, como *mínimo*, *média* e *desvio padrão*, até técnicas sofisticadas de visualização e de determinação de relações funcionais entre atributos (HAN; PEI; KAMBER, 2011; MIRKIN, 2011).
- **Métodos de associações:** refere-se à busca de padrões frequentes de associações entre os

atributos de um conjunto de dados. O agrupamento, por sua vez, lida com a identificação de grupos nos dados de acordo com a similaridade entre os objetos.

- **Métodos de agrupamento:** é encontrar uma estrutura de *clusters* (grupos) nos dados em que os objetos pertencentes a cada *cluster* compartilham alguma característica ou propriedade relevante para o domínio do problema em estudo, ou seja, são de alguma maneira similar (JAIN; DUBES, 1988).

2.4.2 Regressão Linear

Nesta seção, são abordados alguns conceitos da estatística que são fundamentais para que se tenha um bom entendimento dos métodos de regressão, especificamente, a regressão linear simples, dentre os quais se podem citar a probabilidade, variável aleatória, função discreta de probabilidade, função de distribuição de probabilidade, função de densidade de probabilidade, esperança, variância, covariância e correlação.

2.4.2.1 Probabilidade

Todos os dias, as pessoas lidam com uma variedade de fenômenos ou experimentos aleatórios como: o simples lançamento de uma moeda no início das partidas de futebol, a cotação das bolsas de valores, as mudanças no clima, o carregamento dos sistemas elétricos e outros, cuja característica principal é a imprevisibilidade de ocorrência dos eventos desses fenômenos, o que torna impossível afirmar com toda certeza se determinado evento de um fenômeno aleatório qualquer irá ocorrer.

Uma forma de lidar com esse tipo de fenômeno se baseia na determinação de um valor numérico que indica a chance de ocorrência de seus eventos. Esse valor numérico é denominado de probabilidade.

A probabilidade pode ser definida como sendo uma função P que atribui valores numéricos, definidos no intervalo $[0,1]$, aos eventos de um espaço amostral Ω , sendo que esses valores numéricos estão relacionados à frequência de ocorrência deste evento. Considere-se que A seja um evento de um espaço amostral qualquer, a probabilidade de ocorrência de A , indicado por $P(A)$, deve satisfazer as seguintes condições:

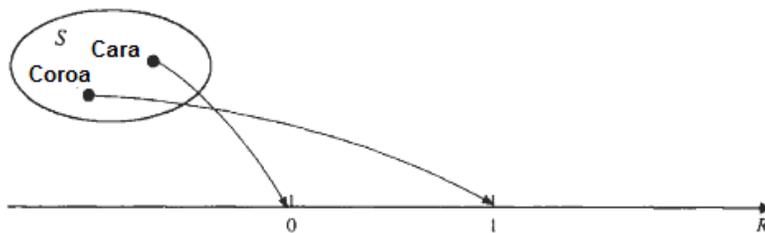
- 1) $P(\Omega)=1$;
- 2) $P(\cup_{j=1}^n A_j) = \sum_{j=1}^n P(A_j)$, sendo que A_j são eventos disjuntos.

Deve-se entender como espaço amostral o conjunto de todos os resultados possíveis de um fenômeno ou experimento aleatório, que por sua vez pode ser definido como uma situação ou um acontecimento cujos resultados não podem ser previstos com certeza.

2.4.2.2 Variável Aleatória

A variável aleatória é definida como uma função que associa cada resultado possível de um espaço amostral a um valor numérico. Por exemplo, considere o experimento de lançar uma moeda, cujo espaço amostral é $S=\{\text{cara, coroa}\}$. Pode-se definir uma variável aleatória X , de tal forma que $X(\text{cara})=0$ e $X(\text{coroa})=1$. Este exemplo é ilustrado melhor na Figura (6).

Figura 6. Variável aleatória associada com o lançamento de uma moeda.



Fonte: HSU, 2013.

As variáveis aleatórias podem ser classificadas de duas formas: discretas e contínuas. As variáveis aleatórias discretas assumem valores num conjunto enumerável, ou seja, mensurável, enquanto que as variáveis aleatórias contínuas assumem valores num conjunto não enumerável.

Normalmente, as informações sobre comportamento das variáveis aleatórias podem ser caracterizadas com a utilização de algumas medidas como: o valor esperado, variância, covariância e os coeficientes de correlação.

O valor esperado, também conhecido como esperança ou média, de uma variável aleatória é uma medida de tendência central que representa o ponto de equilíbrio da distribuição de seus valores. Considerando que X é uma variável aleatória, o valor esperado dessa variável pode ser calculado utilizando a equação (10).

$$E(X) = \mu_X = \sum_{i=1}^k x_i p_i \quad (10)$$

Onde:

x_i : são os possíveis valores da variável aleatória X ;

p_i : são as correspondentes probabilidades dos valores da variável aleatória.

A variância de uma variável aleatória é uma medida de dispersão que está relacionada com a variabilidade dos valores da variável. Considerando que X é uma variável aleatória, a variância dessa variável pode ser calculada utilizando a equação (11).

$$Var(X) = \sigma_X^2 = E[(X - \mu_X)^2] = \sum_{i=1}^k (x_i - \mu_X)^2 p_i \quad (11)$$

Onde:

x_i : são os possíveis valores da variável aleatória X ;

μ_X : é a esperança da variável aleatória X ;

p_i : são as correspondentes probabilidades dos valores da variável aleatória.

A covariância é uma medida de dependência linear entre duas variáveis aleatórias. Quando duas variáveis aleatórias são independentes entre si, o valor da covariância é zero. Considerando que X e Y são duas variáveis aleatórias quaisquer, a covariância entre essas variáveis pode ser calculada utilizando a equação (12).

$$Cov(X, Y) = \sigma_{XY} = E[(X - \mu_X)(Y - \mu_Y)] = E(XY) - E(X)E(Y) \quad (12)$$

O coeficiente de correlação, assim como a covariância, também mede o grau de dependência linear entre duas variáveis aleatórias, porém, diferentemente da covariância, assume valores no intervalo $[-1, 1]$ ou $[-100\%, 100\%]$, o que torna mais perceptível a avaliação do grau de dependência existente entre as variáveis aleatórias. Os graus de correlação podem ser classificados da seguinte forma:

- Correlação Forte: Ocorre quando o coeficiente de correlação assume valores nos seguintes intervalos: $[80\%, 100\%]$ e $[-80\%, -100\%]$;
- Correlação Intermediária: Ocorre quando o coeficiente de correlação assume valores nos seguintes intervalos: $[40\%, 80\%[$ e $[-40\%, -80\%[$;
- Correlação Fraca: Ocorre quando o coeficiente de correlação assume valores nos seguintes intervalos: $[10\%, 40\%[$ e $[-10\%, -40\%[$;
- Inexistência de Correlação: Ocorre quando o coeficiente de correlação assume valores no seguinte intervalo: $] -10\%, 10\%[$.

Considerando que X e Y são duas variáveis aleatórias quaisquer, a correlação entre essas variáveis pode ser calculada utilizando a equação (13).

$$\rho_{XY} = \frac{Cov(X, Y)}{\sqrt{\sigma_X^2} \sqrt{\sigma_Y^2}} \quad (13)$$

2.4.2.3 Função Discreta de Probabilidade

A função discreta de probabilidade é uma função que atribui a cada valor da variável aleatória a sua probabilidade. Considerando que X é uma variável aleatória qualquer, a função discreta de probabilidade pode ser calculada com equação (14).

$$P(X = x_i) = p(x_i) = p_i \quad (14)$$

Onde:

p_i é a probabilidade do valor x_i da variável aleatória X ;

x_i é o valor da variável aleatória X .

2.4.2.4 Função de Distribuição de Probabilidade

A função de distribuição de probabilidade (também conhecida como função acumulada de probabilidade) de uma variável aleatória discreta X é definida, para qualquer número real, de acordo com a equação (15).

$$F(x) = P(X \leq x) \quad (15)$$

Existem algumas distribuições de probabilidade padrões como a distribuição de Bernoulli, binomial, Poisson, geométrica e outros. Abaixo, são mostradas algumas dessas distribuições.

A distribuição de Bernoulli é utilizada para modelar um experimento com somente dois resultados possíveis, normalmente referidos como “sucesso” ou “fracasso”, e codificado como “1” ou “0”, respectivamente.

Uma variável aleatória discreta X tem uma distribuição de Bernoulli com parâmetro p , onde $0 \leq p \leq 1$, se a sua função de probabilidade discreta é dada pelas equações (16) e (17).

$$p_X(1) = P(X = 1) = p \quad (16)$$

$$p_X(0) = P(X = 0) = 1 - p \quad (17)$$

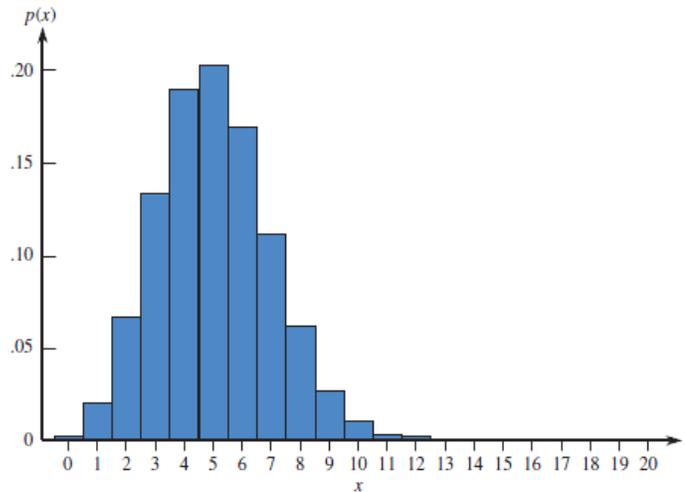
No caso da distribuição binomial, a mesma conta o número total de sucessos de n ensaios de Bernoulli independentes e todos com a mesma probabilidade de sucesso p . Logo a distribuição binomial de uma variável aleatória X com os parâmetros n e p , onde $n = 1, 2, 3, \dots$ e $0 \leq p \leq 1$, pode ser definida pela equação (18).

$$p_X(k) = P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k} \quad (18)$$

Onde: $n = 1, 2, 3, \dots, n$

A Figura (7) mostra um exemplo gráfico de uma função de distribuição de probabilidade binomial com $n = 20$ e $p = 0,25$.

Figura 7. Distribuição binomial com $n = 20$ e $p = 0,25$.



Fonte: PECK; OLSEN; DEVORE, 2011.

2.4.2.5 Função Densidade de Probabilidade

A função de densidade de probabilidade é uma função de distribuição para variáveis aleatórias contínuas que auxilia na atribuição de probabilidade. Considerando que $f(x)$ é uma função de densidade de probabilidade de uma variável contínua X , essa função deve satisfazer as seguintes condições:

- $f(x) \geq 0$, para todo $x \in (-\infty, \infty)$;
- A área definida por $f(x)$ é igual a 1, ou seja, $\int_{-\infty}^{\infty} f(x) dx = 1$;

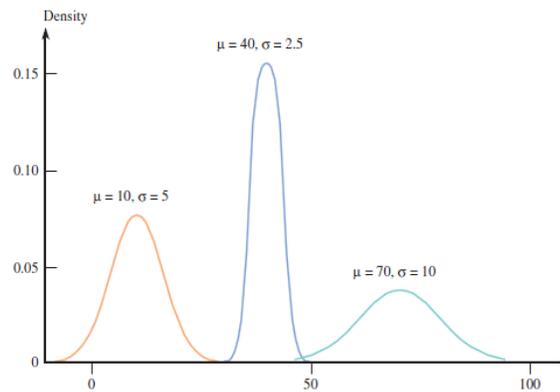
A distribuição normal, ou também conhecida como distribuição gaussiana, tem um importante papel nas análises estatísticas, uma vez que a mesma apresenta uma aproximação razoável para a distribuição de uma diversidade de variáveis.

A distribuição normal de uma variável aleatória contínua X pode ser definida conforme a equação (19).

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad (19)$$

A Figura (8) ilustra graficamente a distribuição normal de uma variável aleatória contínua qualquer.

Figura 8. Três distribuições normais.



Fonte: PECK; OLSEN; DEVORE, 2011.

2.4.3 Análise de Regressão

A análise de regressão, de um modo geral, é um método estatístico que consiste num conjunto de técnicas voltadas para investigar e descobrir a relação existente entre uma variável aleatória, denominada de variável resposta, e uma ou mais variáveis aleatórias, denominadas variáveis regressoras. O produto da análise de regressão é um modelo estatístico (também conhecido como modelo de regressão ou função de regressão), que contém uma simples descrição de um estado ou um processo (YAN, 2009). A Equação (20) ilustra o formato de um modelo de regressão.

$$Y = m(X) + \epsilon \quad (20)$$

Onde:

$m(x)$: é o modelo de regressão ou função de regressão;

ϵ : é a variável aleatória associada com o erro do modelo;

Os principais objetivos da análise de regressão são:

- Estabelecer uma relação causal entre duas ou mais variáveis;
- Prever o valor da variável resposta para um conjunto de valores das variáveis regressoras;
- Investigar as variáveis regressoras de tal forma a identificar quais variáveis são mais importantes para descrever a variável resposta.

Pode-se classificar a análise de regressão de duas formas: regressão paramétrica e regressão não paramétrica. A regressão paramétrica é um tipo de regressão cujo desenvolvimento do

modelo estatístico está condicionado a uma família de equações conhecida. No entanto, na regressão não paramétrica, desconhece-se o formato do modelo de regressão.

Nesta seção, é abordada um método de regressão conhecido como regressão linear simples, que é um tipo de regressão paramétrica cujo modelo estatístico é representado por uma reta.

2.4.3.1 Regressão linear simples

A regressão linear simples é um método estatístico que busca relacionar uma variável aleatória resposta a uma única variável aleatória regressora, condicionado à equação de uma reta, conforme mostrado na equação (21).

$$Y = \beta_0 + \beta_1 x + \epsilon \quad (21)$$

Onde:

Y : é uma variável aleatória resposta;

β_0, β_1 : são constantes;

ϵ : é o erro;

$E[\epsilon] = 0$;

$Var[\epsilon] = \sigma^2$.

No modelo de equação (19), supõe-se, normalmente, que a variável aleatória ϵ é possui uma distribuição normal com esperança igual a zero e variância igual à variância de um modelo gaussiano, ou seja, $\epsilon \sim N(0; \sigma^2)$.

O ajuste da curva de regressão é realizado com base no método dos mínimos quadrados, que, por sua vez, seleciona o modelo que apresenta a menor soma de quadrados dos erros. Então, com a aplicação do método dos mínimos quadrados, os parâmetros β_0 e β_1 do modelo de regressão mostrado na equação (21) são estimados de acordo com as equações (22) e (23), respectivamente.

$$\widehat{\beta}_0 = \bar{y} - \widehat{\beta}_1 \bar{x} \quad (22)$$

$$\widehat{\beta}_1 = \frac{\sum_{i=1}^n (y_i - \bar{y})(x_i - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad (23)$$

Onde:

$\widehat{\beta}_0$ e $\widehat{\beta}_1$: são os estimadores mínimos quadrados;

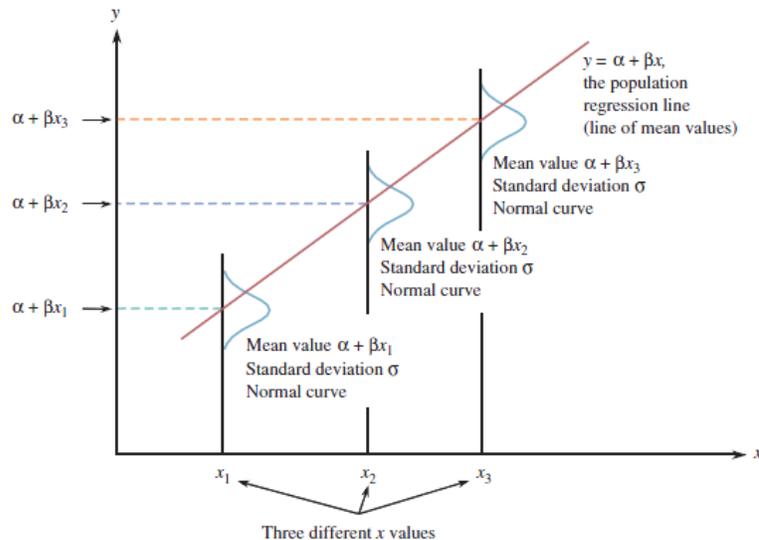
\bar{y} e \bar{x} : são as esperanças das variáveis aleatórias Y e X .

E a variância do modelo de regressão pode ser obtida pela equação (24).

$$\hat{\sigma} = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n - 2} \quad (24)$$

A Figura (9) ilustra graficamente o modelo estatístico mostrado na equação 3.10. De acordo com essa figura, pode-se observar que a parte determinística do modelo é numericamente igual à média da distribuição normal, e a variância do mesmo representa a variabilidade das variáveis aleatórias.

Figura 9. Modelo de regressão linear simples.



Fonte: PECK; OLSEN; DEVORE, 2011.

Uma vez obtido o modelo de regressão por meio da aplicação da regressão linear simples, avalia-se esse modelo por meio da análise da variância, de tal forma a verificar se o modelo de regressão obtido é adequado para explicar a relação entre as variáveis resposta e regressora.

A análise de variância é uma ferramenta estatística que, por meio de técnicas da inferência estatística, testa duas hipóteses principais: a hipótese nula, que representa a possibilidade do parâmetro β_1 (também conhecido como a declividade da reta) ser nulo; e a hipótese alternativa, que representa a possibilidade desse mesmo parâmetro não ser nulo.

Durante a análise de variância, são obtidos três parâmetros principais: a Soma de Quadrados Total (SQT), a Soma de Quadrados de Erro (SQE) e a Soma de Quadrados de Regressão (SQReg). Esses parâmetros são calculados por meio das equações (25), (26) e (27), respectivamente.

$$SQT = \sum_{i=1}^n (y_i - \bar{y})^2 \quad (25)$$

$$SQE = \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (26)$$

$$SQReg = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 \quad (27)$$

Onde:

\hat{y}_i é o valor estimado da variável y pelo modelo de regressão;

\bar{y} é a esperança da variável y ;

y_i é o valor da variável y utilizado para construir o modelo de regressão;

Com base nesses três parâmetros, calcula-se o valor da estatística do teste, conhecida com F_o , através da divisão do quadrado médio de regressão pelo quadrado médio do erro, conforme a equação (28). Com base no valor do parâmetro F_o , pode-se rejeitar ou não o modelo de regressão.

$$F_o = \frac{SQReg}{SQE / n - 2} \quad (28)$$

Além disso, há um outro indicador estatístico utilizado para avaliar o grau de confiabilidade do modelo proveniente de uma análise de regressão denominado de fator de determinação R^2 . A equação (29) mostra a expressão do fator de determinação.

$$R^2 = \frac{SQReg}{SQT} \quad (29)$$

2.4.4 Árvores de Decisão

A técnica de árvore de decisão possui aprendizado preditivo baseado em procura, criando uma hipótese baseada em instâncias particulares para a geração conclusões generalizadas, similar a regras de *if-then*. Ela abrange todo o espaço amostral, com isso, pode realizar predições para qualquer exemplo de entrada (FACELI et al., 2011).

Segundo Safavian e Landgrebe (1991), os principais objetivos da classificação por árvore de decisão são (SAFAVIAN; LANDGREBE, 1991):

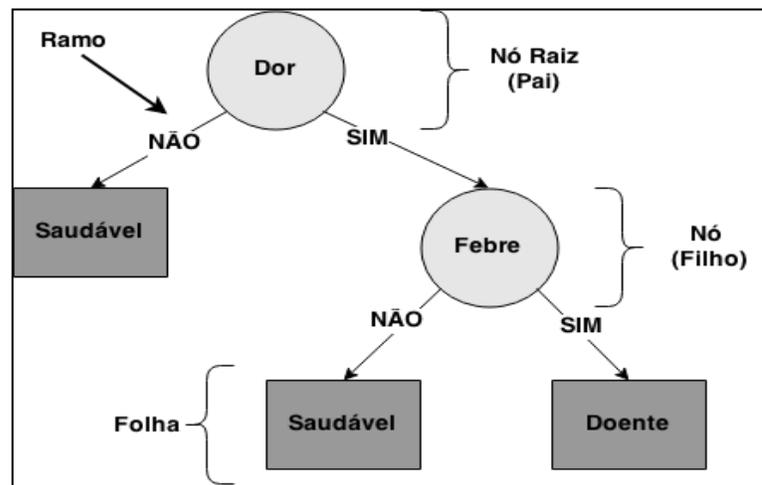
- Alta taxa de acurácia para classificação para amostras que não foram vistos na base de treinamento;
- Facilidade de atualização;
- Estrutura de simples compreensão;

Para Russell e Norvig (2003), uma árvore de decisão tem seus atributos de entrada divididos em contínuos ou discretos, no qual, para valores contínuos a aprendizagem é chamada de regressão, e para valores discretos são chamadas de classificação (RUSSELL; NORVIG, 2003).

Para Hosokawa (2011), uma árvore de decisão é uma estrutura composta por simples regras de decisão, onde são divididas sucessivamente uma coleção de registros em conjuntos menores (HOSOKAWA, 2011). Para cada divisão, os dados são separados de acordo com características em comum até chegar a pontos indivisíveis, que representam as classes. Os atributos do banco de dados são representados como nós internos (diferente do atributo classe), sendo que os valores do atributo classe são chamados de folhas. Para cada valor do atributo do nó pai a um nó filho, é rotulado formando ligações que são conhecidas como ramos. Seguindo o mesmo raciocínio, a estruturada da árvore, apresentada na Figura (10), é organizada de tal forma que:

- Cada nó interno (não-folha) seja rotulado com o nome de um dos atributos previsores;
- Cada ramo ou arestas saindo de um nó interno seja rotulado com valores do atributo naquele nó;
- Cada folha seja rotulada com uma classe.

Figura 10. Modelo de árvore de decisão para diagnóstico de pacientes.



Fonte: MONTEIRO, 2015.

Para a construção de uma Árvore de Decisão, é necessário que seja escolhido um critério para a seleção de atributo preditivo, que será utilizado em cada nó. Os critérios de seleção para a melhor divisão são baseados em diferentes medidas, tais como: impureza, distância e dependência.

A poda é considerada a parte mais importante do processo de construção de árvore, pelo razão da existência de ruídos na base de dados. Com isso, as árvores induzidas classificam novos

objetos em um modo não confiável e tendem a ter uma grande estrutura tornando difícil a compreensão. O método de poda pode evitar problemas com ruídos ou erros nos ramos da árvore de decisão que levam ao *overfitting* (FACELI et al., 2011) Existem diversas formas de realizar poda em uma árvore de decisão, e todas elas são classificadas como pré-poda ou pós-poda:

- Pré-poda realiza o procedimento durante a construção da árvore. Onde o processo para de dividir os atributos e transformará o nó corrente em um nó folha da árvore
- Pós-poda realiza o procedimento após a construção da árvore, eliminando ramos completos, fazendo a exclusão de todos os nós abaixo desse ramo é transformando o nó em folha quer representar a classe mais frequente no ramo.

2.4.4.1 CART – Classification and Regression Trees

O algoritmo CART (*Classification and Regression Trees*) consiste em uma técnica não paramétrica que estabelece o resultado através de partição binária recursiva, em que o nó pai se divide entre dois nós-filhos e o processo pode ser repetido tratando cada nó filho como um nó pai. Dependendo do tipo do atributo a árvore pode ser de classificação se o atributo for nominal, ou a árvore pode ser de regressão caso o atributo seja contínuo (BREIMAN et al., 1984).

Segundo Fonseca (1994) umas das virtudes do CART é a grande capacidade de pesquisa de relação entre os dados mesmo quando elas não são evidentes, bem como a produção de resultados sob a forma de árvores de decisão de grande simplicidade e legibilidade. As principais vantagens do Algoritmo de CART são:

- Poder utilizar variáveis independentes de diferentes tipos, desde contínuas, ordinais e nominais.
- Não obrigar a realização de transformações das variáveis iniciais independentes (como a logaritmização ou normalização), já que o método tem bom comportamento para qualquer tipo de dados.
- Poder usar a mesma variável em diferentes estágios do modelo, permitindo reconhecer efeitos que certas variáveis produzem sobre outras.
- Não necessitar satisfazer qualquer condição de aplicabilidade do modelo, o que não acontece nos modelos paramétricos.

Para executar a divisão de um nó em dois filhos o algoritmo utiliza questões cuja resposta só pode ser “sim” ou “não”. Quando encontrada a melhor divisão o processo é repetido para cada nó

filho, até que a divisão seja interrompida, ou seja, impossível de ser realizada. Após o fim do crescimento da árvore do algoritmo CART, é realizada a pós-poda para aumentar o poder de generalização (BREIMAN et al., 1984).

2.4.5 Redes Neurais Artificiais

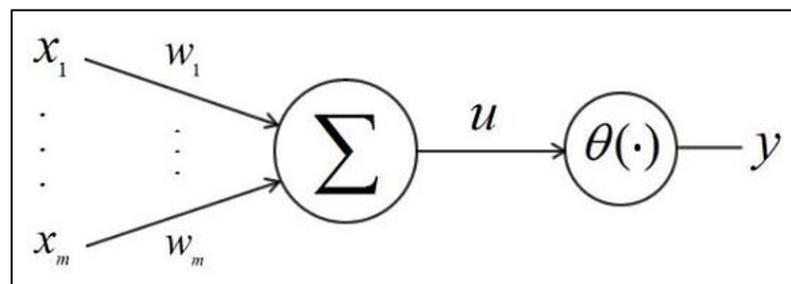
As redes neurais artificiais são modelos computacionais inspirados no sistema nervoso biológico, cujo funcionamento é semelhante a alguns procedimentos humanos, ou seja, aprendem pela experiência, generalizam exemplos através de outros, e abstraem características (WASSERMAN, 1989). A RNA pode exercer tarefas aplicadas tanto para a regressão quanto para a classificação, dependendo do algoritmo e da problemática escolhida.

O algoritmo de treinamento mais comum das RNAs envolve uma regra de correção de erros, na qual se recorre à otimização de uma função quadrática do erro entre as respostas da RNA e os rótulos dos exemplos (FACELI et al., 2011).

Vários algoritmos têm sido propostos na literatura para o ajuste dos parâmetros, para que o modelo obtenha um melhor desempenho, os principais modelos são: Perceptron, Perceptron Multicamadas, Redes de Kohonen e Redes de Hopfield, cada um destes com características direcionadas para diferentes aplicações.

As RNAs são sistemas computacionais distribuídos compostos de unidades de processamento simples (neurônios artificiais) que computam funções matemáticas, interconectadas e dispostas em uma ou mais camadas. Essas conexões, as sinapses, possuem um peso associado que ponderam a entrada recebida pelo neurônio. A estrutura básica de um neurônio artificial é exibida na Figura (11).

Figura 11. Estrutura básica de um neurônio artificial.



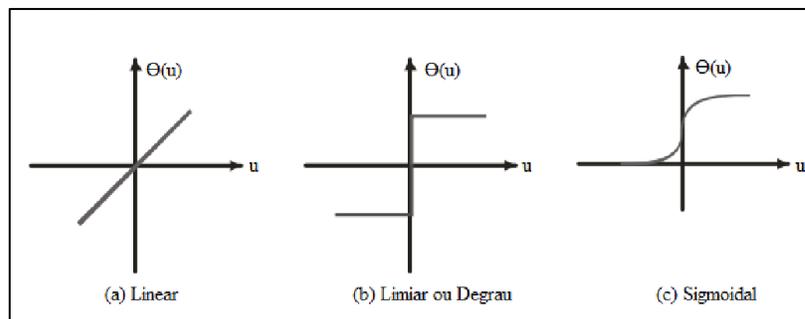
Fonte: SIQUEIRA, 2016.

O neurônio artificial computa uma função sobre a soma ponderada das m entradas gerando uma saída y dada pela equação (30):

$$y = \theta\left(\sum_{i=1}^m x_i w_i\right) \quad (30)$$

Sendo x_i a i -ésima entrada para o neurônio, w_i o peso sináptico que pondera a entrada x_i e θ a função de ativação ou função de transferência. As funções de ativação comumente utilizadas na literatura são: linear, limiar ou degrau e sigmoideal. O uso da função linear Figura (12.a) retorna como saída o valor de u . Na função limiar Figura (12.b) é definido o resultado da função como sendo 1 ou 0 de forma alternada (-1 também pode ser utilizado). A função sigmoideal Figura (12.c) é uma representação aproximada, contínua e diferenciável da função limiar.

Figura 12. Funções de ativação que podem ser usadas por um neurônio artificial.



Fonte: SIQUEIRA, 2016.

Outro ponto importante no que diz respeito às RNAs é sobre o processo de aprendizado da rede. Segundo Bernarda *et al* (2000) o aprendizado é um processo pelo qual os parâmetros de uma rede neural são ajustados com o intuito de fazer com que o modelo obtenha melhor desempenho. O tipo de aprendizagem é determinado pela maneira pela que o ajuste dos parâmetros ocorre (HAYKIN, 2009). Existem diversos algoritmos de treinamento que definem como o valor de cada peso deve ser alterado, geralmente sobre um conjunto específico de regras bem definidas.

Os principais paradigmas de aprendizado são supervisionado e não supervisionado. No aprendizado supervisionado as entradas e saídas são fornecidas por um supervisor cuja função é observar as respostas da rede para cada entrada e assim ajustar os parâmetros para que se possa encontrar uma ligação entre os dados de entrada e de saída. Já no aprendizado não supervisionado não há esse supervisor intervindo no processo, de tal forma que a rede deve aprender a partir de técnicas de agrupamento, por exemplo.

Um algoritmo supervisionado bastante relatado e utilizado na literatura para ajuste dos pesos é o algoritmo *backpropagation* que utiliza amostras rotuladas (entrada e saída esperada) por meio de um mecanismo de correção de erros em duas etapas.

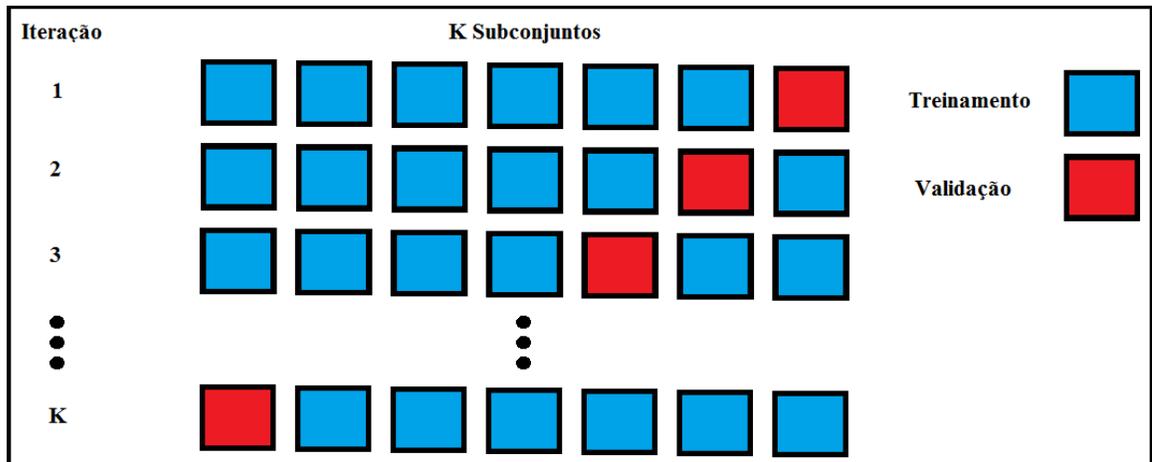
Na primeira etapa, chamada de *forward* ou ainda simplesmente de computação para frente, obtém-se a saída da rede através da amostra de entrada, processando os dados partindo-se da camada de entrada, passando pelas camadas intermediárias até alcançar a camada de saída. Na segunda etapa, chamada de *backward* ou computação para trás, é utilizado o erro (diferença entre a saída esperada e a saída obtida pela rede), para atualização dos pesos sinápticos da camada de saída até a primeira camada intermediária (HAYKIN, 2009; FACELI et al., 2011).

Um problema que pode ocorrer e deve ser evitado no processo de aprendizado é o de superespecialização, ou *overfitting*, no qual a rede fica especialista no conjunto de dados de treinamento e acaba perdendo o poder de generalização, gerando resultados ruins para novas amostras. Sendo assim, uma forma de se evitar o *overfitting* é dividir os dados em 3 subconjuntos diferentes entre si: treino, validação e teste.

A cada interação do processo de treinamento o subconjunto de validação é submetido à rede para que se avalie o erro quadrático médio gerado entre a saída obtida e a saída desejada de modo que se o erro de validação começar a aumentar muito sem diminuir em determinado número de gerações (perda do poder de generalização), pode-se interromper o processo, ficando a critério do supervisor as modificações dos parâmetros da rede.

Outra forma é utilizar validação cruzada, ou *cross validation*. O método de validação cruzada denominado *k-fold* consiste em dividir o conjunto total de dados em k subconjuntos mutuamente exclusivos do mesmo tamanho e, a partir disto, um subconjunto é utilizado para teste e os $k-1$ restantes são utilizados para estimação dos parâmetros e calcula-se a acurácia do modelo. Este processo é realizado k vezes alternando de forma circular o subconjunto de teste. A Figura (13) mostra o esquema realizado pelo *k-fold*.

Figura 13. Validação Cruzada.



Fonte: O Autor.

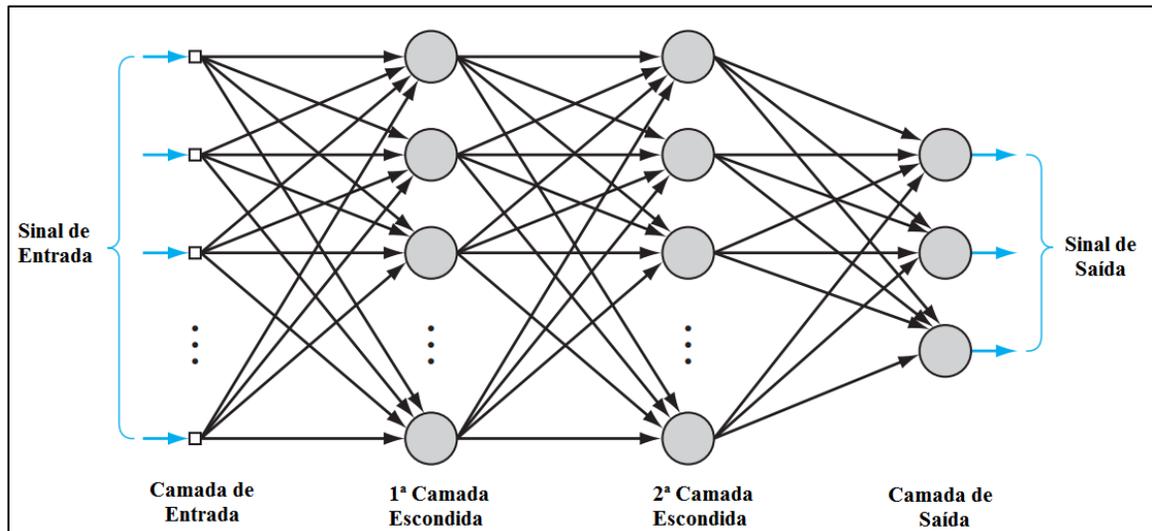
Ao final das k iterações, calcula-se a acurácia sobre os erros encontrados, através da equação descrita anteriormente, obtendo assim uma medida mais confiável sobre a capacidade do modelo de representar o processo gerador dos dados.

2.4.5.1 Perceptron Multicamadas

As redes Perceptron de Múltiplas Camadas (MLP do inglês Multilayer Perceptron) apresentam uma ou mais camadas intermediárias (ou ainda camada escondida) de neurônios e uma camada de saída. A função das camadas escondidas é intervir entre a camada de entrada e a de saída da rede de forma útil (HAYKIN, 2009). São consideradas redes aproximadoras universais, podendo realizar qualquer tipo de mapeamento linear ou não linear entre as entradas e saídas de um determinado sistema.

A Figura (14) ilustra uma rede MLP totalmente conectada, onde todos os neurônios de uma camada estão conectados com os neurônios da camada seguinte.

Figura 14. Arquitetura de uma rede MLP de duas camadas escondidas.



Fonte: Adaptado de HAYKIN, 2009.

A camada de entrada recebe o sinal de entrada e os envia para a camada seguinte. As camadas escondidas são responsáveis pelo processo de aprendizado que é feito através do processamento em cada neurônio. A função implementada por um neurônio de uma dada camada é uma combinação das funções processadas pelos neurônios da camada anterior que estão a ele ligados (FACELI et al., 2011). A camada de saída recebe o resultado do processamento das camadas anteriores e fornece uma resposta relacionada a uma das classes presentes no conjunto de dados.

2.5 CONSIDERAÇÕES FINAIS

A presença de harmônicos na rede elétrica pode causar ressonâncias que podem resultar no aumento de perdas elétricas e causar danos às cargas conectadas ao longo da rede. Inicialmente este capítulo apresentou medidas de mitigação de harmônicos com base em filtros harmônicos, suas fontes causadoras e efeitos gerados.

Também foram apresentadas as normas que regularizam o máximo de distorção harmônica permitida de tensão e corrente, e outros indicadores de qualidade de energia elétrica contidos no PRODIST.

Por fim, foi apresentado como funcionam as técnicas de IC que foram utilizadas no desenvolvimento da ferramenta computacional desenvolvida neste trabalho. No capítulo a seguir, será apresentada esta ferramenta, expondo sua arquitetura e interfaces gráficas dos módulos utilizados para análise de QEE e gerenciamento de dados de campanhas de medição.

3 DESENVOLVIMENTO DE SOFTWARE

3.1 CONSIDERAÇÕES INICIAIS

Neste capítulo é apresentada a ferramenta computacional SISQEE que foi desenvolvida, a estruturação do banco de dados e a arquitetura do sistema, bem como as principais telas da interface gráfica do sistema referentes aos módulos de banco de dados e de análise de medições.

3.2 BASE DE DADOS

Para formação das bases de dados, conforme a Figura (9), diversas campanhas de medição foram realizadas utilizando analisadores de QEE, de marca HIOKI modelo PW3198 e de marca RMS modelo MARH-32, devidamente instalados em pontos de interesse da rede elétrica de vários clientes comerciais e industriais, em grande parte essas campanhas envolveram medições em alimentadores que suprem o Polo Industrial de Manaus. Todas essas campanhas foram armazenadas no banco de dados de medições, gerenciado pelo *postgres*, o qual é acessado pelo software SISQEE, para realizar as análises de Regime Permanente, e de Correlação, utilizando as três técnicas disponíveis.

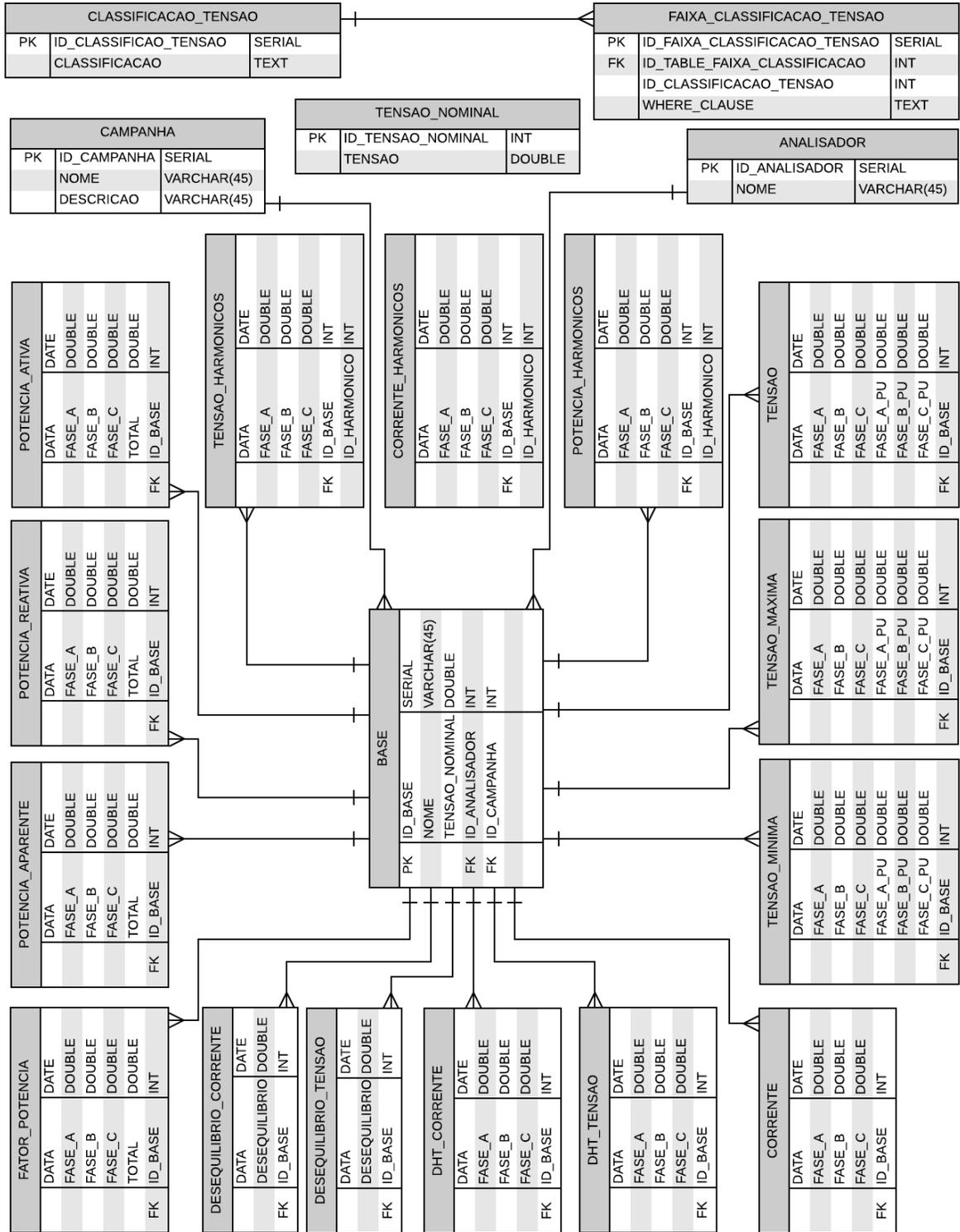
Cada analisador de QEE disponível no mercado possui seu próprio software com seus respectivos módulos de análise e geração de gráficos, logo cada um desses analisadores, embora possam analisar as mesmas coisas, apresentam arquivos de saídas com formatos e com padrões de exibição diferentes. Então, após realizar uma campanha de medição e antes de armazenar no banco de dados, é necessário fazer alguns procedimentos para normalizar as bases de dados.

Primeiramente é preciso definir quais dados compõe uma campanha de medição. Uma campanha é composta por várias bases de dados, onde cada base representa um ponto de interesse da rede elétrica onde foi instalado um analisador de QEE para realizar medições simultâneas de parâmetros elétricos, sendo que cada base de dados possui uma tensão nominal de referência no barramento.

É extraído um total de 15 (quinze) parâmetros elétricos dos analisadores de QEE, sendo: Tensão, Tensão Máxima, Tensão Mínima, Corrente, Potência Ativa, Potência Reativa, Potência Aparente, Fator de Potência, DHT_v , DHT_i , Desequilíbrio de Tensão e Desequilíbrio de Corrente, Tensão Harmônicos, Corrente Harmônicos e Potência Harmônicos. A Figura (15) apresenta o

Diagrama Entidade Relacionamento do banco de dados criado, contendo uma entidade para cada um destes parâmetros.

Figura 15. Diagrama Entidade Relacionamento.



Fonte: O Autor.

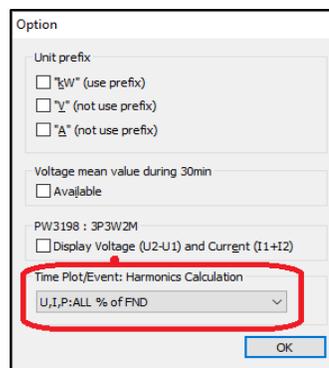
Além destes parâmetros, a Figura (15) também contém a entidade TENSÃO_NOMINAL utilizada para armazenar as tensões nominais de baixa e média tensão descritos a seguir na Tabela 8, a entidade CLASSIFICACAO_TENSAO que armazena as faixas de tensão *adequada*, *precária* e *crítica*, e a entidade FAIXA_CLASSIFICACAO_TENSAO, a qual armazena os limites de cada faixa estipulados pelo PRODIST descritos nas Tabelas 1-6.

Tabela 8 - TR armazenadas no banco de dados.

Tensão Nominal (V)
220
380
440
13.8 K
34.5 K
69 K
138 K
230 K

O analisador de QEE da marca HIOKI fornece todos os parâmetros para formar uma base de dados, com a possibilidade de gerar harmônicos até a 50ª ordem. Realizando a configuração no software do HIOKI para que os harmônicos exportados fiquem em porcentagem da fundamental, conforme Figura (16), todos os dados são exportados em um arquivo de saída no formato *csv*, conforme Figura (17).

Figura 16. Configuração HIOKI para porcentagem da fundamental.



Fonte: O Autor.

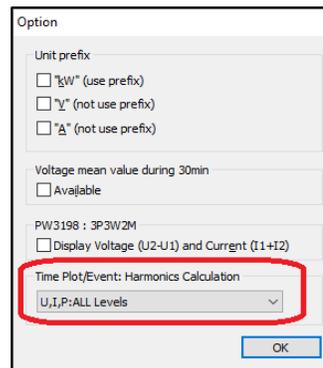
Figura 17. Arquivo de saída do HIOKI para harmônicos em porcentagem da fundamental.

	A	B	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
1	Date	Time	AvgU1(1)	AvgI1(1)	AvgP1(1)	AvgUI(2)	AvgI1(2)	AvgP1(2)	AvgU1(3)	AvgI1(3)	AvgP1(3)	AvgU1(4)	AvgI1(4)	AvgP1(4)	AvgU1(5)	AvgI1(5)	AvgP1(5)
2	22/05/2014	11:30:00	100.00E+0	100.00E+0	100.00E+0	0.01E+0	0.31E+0	0.00E+0	0.66E+0	0.48E+0	OVER	0.01E+0	0.10E+0	0.00E+0	1.16E+0	0.23E+0	0.00E+0
3	22/05/2014	11:31:00	100.00E+0	100.00E+0	100.00E+0	0.01E+0	0.30E+0	0.00E+0	0.68E+0	0.53E+0	OVER	0.01E+0	0.09E+0	0.00E+0	1.15E+0	0.25E+0	OVER
4	22/05/2014	11:32:00	100.00E+0	100.00E+0	100.00E+0	0.01E+0	0.31E+0	0.00E+0	0.68E+0	0.52E+0	OVER	0.01E+0	0.10E+0	0.00E+0	1.17E+0	0.19E+0	0.00E+0
5	22/05/2014	11:33:00	100.00E+0	100.00E+0	100.00E+0	0.01E+0	0.30E+0	0.00E+0	0.69E+0	0.53E+0	OVER	0.01E+0	0.10E+0	0.00E+0	1.18E+0	0.21E+0	OVER
6	22/05/2014	11:34:00	100.00E+0	100.00E+0	100.00E+0	0.01E+0	0.30E+0	0.00E+0	0.68E+0	0.55E+0	OVER	0.01E+0	0.10E+0	0.00E+0	1.14E+0	0.18E+0	OVER
7	22/05/2014	11:35:00	100.00E+0	100.00E+0	100.00E+0	0.01E+0	0.30E+0	0.00E+0	0.68E+0	0.56E+0	OVER	0.01E+0	0.11E+0	0.00E+0	1.11E+0	0.23E+0	OVER
8	22/05/2014	11:36:00	100.00E+0	100.00E+0	100.00E+0	0.01E+0	0.30E+0	0.00E+0	0.67E+0	0.59E+0	OVER	0.01E+0	0.10E+0	0.00E+0	1.09E+0	0.21E+0	OVER
9	22/05/2014	11:37:00	100.00E+0	100.00E+0	100.00E+0	0.01E+0	0.31E+0	0.00E+0	0.67E+0	0.59E+0	OVER	0.01E+0	0.11E+0	0.00E+0	1.06E+0	0.23E+0	OVER
10	22/05/2014	11:38:00	100.00E+0	100.00E+0	100.00E+0	0.01E+0	0.30E+0	0.00E+0	0.69E+0	0.58E+0	OVER	0.01E+0	0.10E+0	0.00E+0	1.03E+0	0.26E+0	OVER
11	22/05/2014	11:39:00	100.00E+0	100.00E+0	100.00E+0	0.01E+0	0.30E+0	0.00E+0	0.66E+0	0.54E+0	OVER	0.01E+0	0.11E+0	0.00E+0	1.06E+0	0.26E+0	OVER
12	22/05/2014	11:40:00	100.00E+0	100.00E+0	100.00E+0	0.01E+0	0.30E+0	0.00E+0	0.67E+0	0.53E+0	OVER	0.01E+0	0.12E+0	0.00E+0	1.05E+0	0.27E+0	OVER
13	22/05/2014	11:41:00	100.00E+0	100.00E+0	100.00E+0	0.01E+0	0.31E+0	0.00E+0	0.69E+0	0.55E+0	OVER	0.01E+0	0.12E+0	0.00E+0	1.03E+0	0.28E+0	OVER
14	22/05/2014	11:42:00	100.00E+0	100.00E+0	100.00E+0	0.01E+0	0.29E+0	0.00E+0	0.68E+0	0.57E+0	OVER	0.01E+0	0.11E+0	0.00E+0	1.03E+0	0.31E+0	OVER
15	22/05/2014	11:43:00	100.00E+0	100.00E+0	100.00E+0	0.01E+0	0.30E+0	0.00E+0	0.65E+0	0.59E+0	OVER	0.01E+0	0.12E+0	0.00E+0	1.06E+0	0.31E+0	OVER
16	22/05/2014	11:44:00	100.00E+0	100.00E+0	100.00E+0	0.01E+0	0.30E+0	0.00E+0	0.66E+0	0.61E+0	OVER	0.01E+0	0.11E+0	0.00E+0	1.06E+0	0.31E+0	OVER
17	22/05/2014	11:45:00	100.00E+0	100.00E+0	100.00E+0	0.01E+0	0.29E+0	0.00E+0	0.67E+0	0.62E+0	OVER	0.01E+0	0.11E+0	0.00E+0	1.07E+0	0.28E+0	OVER
18	22/05/2014	11:46:00	100.00E+0	100.00E+0	100.00E+0	0.01E+0	0.28E+0	0.00E+0	0.65E+0	0.61E+0	OVER	0.01E+0	0.11E+0	0.00E+0	1.07E+0	0.27E+0	OVER
19	22/05/2014	11:47:00	100.00E+0	100.00E+0	100.00E+0	0.01E+0	0.30E+0	0.00E+0	0.65E+0	0.58E+0	OVER	0.01E+0	0.12E+0	0.00E+0	1.06E+0	0.29E+0	OVER
20	22/05/2014	11:48:00	100.00E+0	100.00E+0	100.00E+0	0.01E+0	0.29E+0	0.00E+0	0.65E+0	0.55E+0	OVER	0.01E+0	0.12E+0	0.00E+0	1.02E+0	0.28E+0	OVER
21	22/05/2014	11:49:00	100.00E+0	100.00E+0	100.00E+0	0.01E+0	0.29E+0	0.00E+0	0.68E+0	0.51E+0	OVER	0.01E+0	0.12E+0	0.00E+0	1.03E+0	0.21E+0	OVER
22	22/05/2014	11:50:00	100.00E+0	100.00E+0	100.00E+0	0.01E+0	0.29E+0	0.00E+0	0.69E+0	0.52E+0	OVER	0.01E+0	0.12E+0	0.00E+0	1.00E+0	0.29E+0	OVER
23	22/05/2014	11:51:00	100.00E+0	100.00E+0	100.00E+0	0.01E+0	0.29E+0	0.00E+0	0.70E+0	0.50E+0	OVER	0.01E+0	0.12E+0	0.00E+0	1.02E+0	0.26E+0	OVER
24	22/05/2014	11:52:00	100.00E+0	100.00E+0	100.00E+0	0.01E+0	0.29E+0	0.00E+0	0.69E+0	0.51E+0	OVER	0.01E+0	0.12E+0	0.00E+0	1.00E+0	0.33E+0	OVER
25	22/05/2014	11:53:00	100.00E+0	100.00E+0	100.00E+0	0.01E+0	0.29E+0	0.00E+0	0.69E+0	0.50E+0	OVER	0.01E+0	0.11E+0	0.00E+0	1.06E+0	0.34E+0	OVER

Fonte: O Autor.

No entanto, pode-se observar que as colunas dos harmônicos que contém as fundamentais $AvgI(1)$, $AvgU(1)$ e $AvgP(1)$ de corrente, tensão e potência respectivamente, estão com todas os valores 100 (cem), pois o analisador exportou a componente fundamental em porcentagem dela mesma, quando na verdade deveria exportar seus valores reais. Para resolver isto é gerado um segundo arquivo contendo apenas os valores das fundamentais, utilizando a configuração mostrada na Figura (18).

Figura 18. Arquivo csv dos harmônicos em valores nominais.



Fonte: O Autor.

O analisador de QEE da marca RMS também pode gerar harmônicos até a 50ª ordem, porém não contém os parâmetros *tensão máxima*, *tensão mínima*, *desequilíbrio de tensão*, *desequilíbrio de corrente*, e *potência harmônicos*. Os dados coletados são exportados em dois arquivos no formato *csv*, um *contendo potência ativa*, *potência reativa*, *potência aparente* e *fator de potência*,

extraídos contidos nestes dois ou mais arquivos grandes, em vários arquivos menores no formato *csv*, padronizando o nome das colunas e a disposição delas, onde cada arquivo representa um parâmetro elétrico que foi criado uma entidade no banco de dados. Atualmente o SISQEE possui modelos apenas para dois analisadores de QEE disponível no mercado, mas conforme a necessidade é possível implementar mais classes modelo, afim de converter os padrões de estruturação e organização para o padrão uniforme adotado pelo SISQEE.

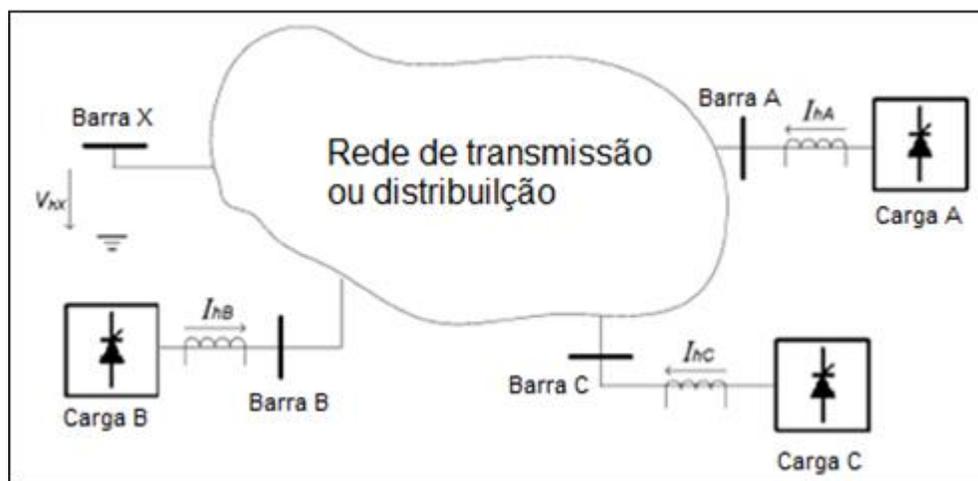
3.3 ARQUITETURA DE SOFTWARE

O software SISQEE é uma ferramenta capaz de oferecer suporte na realização de análises gráficas dos dados de grandezas elétricas obtidos de campanhas de medição por analisadores de QEE, calculando os respectivos indicadores de QEE estabelecidos pelo PRODIST.

Com o mesmo é possível estimar o possível grau do impacto harmônico individual de consumidores industriais nas redes de distribuição de energia elétrica, utilizando medições simultâneas de tensões e correntes harmônicas, bem como técnicas de estatística não paramétrica e de inteligência computacional.

A Figura (21) apresenta uma rede elétrica genérica demonstrando o problema de identificação das principais fontes harmônicas causadoras de distorções na tensão da rede elétrica, em um ponto de interesse. Deseja-se estimar os impactos que as cargas não lineares genéricas A, B, C têm na distorção harmônica da tensão no ponto de interesse, no caso a barra X.

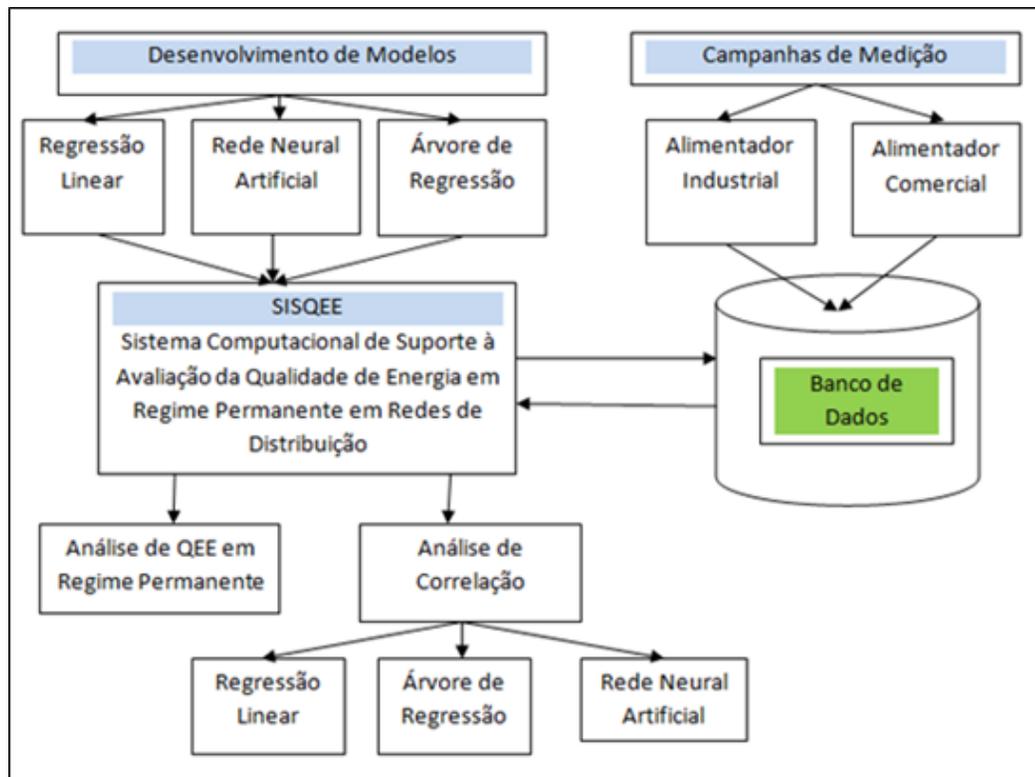
Figura 21. Caracterização do problema de identificação da contribuição de múltiplas fontes harmônicas na distorção harmônica da tensão em um ponto de interesse da rede elétrica.



Fonte: Adaptado de MAZIN; XU; HUANG, 2011.

Dessa forma, três modelos foram implementados, neste trabalho, para estabelecer esta correlação: Um modelo estatístico, baseado em Regressão Linear, um modelo causa-efeito, baseado em Redes Neurais Artificiais, e um modelo baseado em mineração de dados utilizando Árvore de Regressão. Definidas as técnicas, a metodologia que o SISQEE seguiu foi estruturada conforme apresentado na Figura (22).

Figura 22. Estruturação da Metodologia de Desenvolvimento do Projeto.



Fonte: Projeto P&D.

Em relação à análise de QEE em regime permanente, é comum que o processo de análise dos dados seja realizado manualmente por um especialista na área, exigindo atenção aos detalhes, demandando tempo e estando sujeito a falhas humanas. O software desenvolvido é capaz de reduzir o tempo de análise, abstraindo algumas etapas, focando na qualidade de experiência do usuário, oferecendo ao usuário uma plataforma onde o mesmo seja capaz de realizar análises em relação a tensões e correntes harmônicas, no intuito de avaliar os pontos de medições que apresentem níveis de harmônicos fora das faixas especificadas pelas normas, assim como, avaliar desequilíbrios de tensão e corrente entre as fases A, B e C, a fim de avaliar a distribuição das cargas entre as fases do sistema.

O SISQEE foi criado focado na tarefa de auxiliar os profissionais das áreas afins, fornecendo uma interface simples e agradável, permitindo que os usuários atinjam suas metas com eficiência e satisfação. É importante destacar, que para a utilização do software SISQEE são necessários apenas conhecimentos básicos de informática.

Na Tabela 9 abaixo, estão presente todos os recursos computacionais utilizados para o desenvolvimento do SISQEE, cada um com suas respectivas funções e versões.

Tabela 9 - Recursos Computacionais.

Recursos	Ferramentas
Linguagem	Java 8, JavaScript, HTML 5
Plataforma de Desenvolvimento	Netbeans 8.1
	pgAdmin 1.16.1
Banco de Dados	PostgreSQL 9.1
Sistema Operacional	Linux
	Windows

O SISQEE foi desenvolvido utilizando a linguagem de programação Java, com a metodologia denominada de Programação Orientada a Objetos. A linguagem Java é uma plataforma de computação gratuita lançada pela primeira vez pela Sun Microsystems em 1995. Segundo a Oracle Corporation, é a tecnologia que capacita muitos programas da mais alta qualidade, como utilitários, jogos e aplicativos corporativos, sendo executado em mais de 850 milhões de computadores pessoais e em bilhões de dispositivos em todo o mundo.

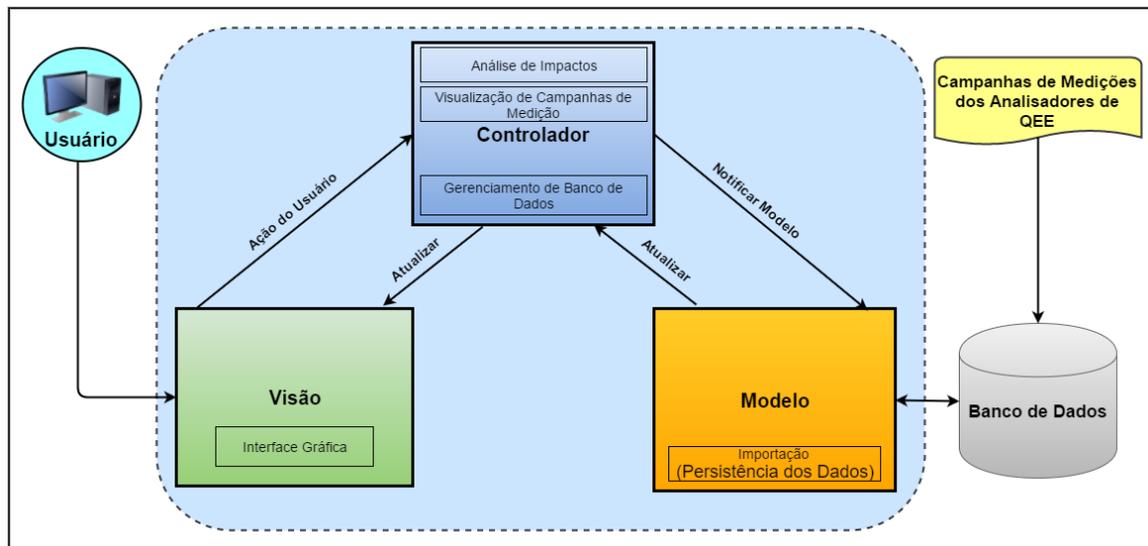
A metodologia Orientada a Objetos consiste na idealização de um sistema onde é possível abstrair detalhes que permitem uma alta taxa de reutilização de código e o fácil entendimento da arquitetura do sistema. É baseada na decomposição do sistema de acordo com os objetos que serão manipulados por este, oferecendo uma visão do sistema mais próxima do mundo real, juntamente com uma modelagem baseada nos dados e com uma maior transparência da análise do projeto.

Para o desenvolvimento do software foi adotado como padrão de design de arquitetura o *Model-View-Controller* (MVC) atualmente utilizado em muitos softwares e frameworks, dividindo o software em três camadas, de modo que é permitida a oportunidade de trabalhar com a área de portabilidade de software e reaproveitamento de código. Nesta arquitetura as funcionalidades do sistema são divididas nas seguintes camadas:

- **Modelo:** O modelo (Model) é utilizado para manipular informações de forma mais detalhada, sendo recomendado que, sempre que possível, se utilize dos modelos para realizar consultas, cálculos e todas as regras de negócio de um site ou sistema. É o modelo que tem acesso a toda e qualquer informação sendo essa vinda de um banco de dados, arquivo XML e afins.
- **Visão:** A visão (View) é responsável por tudo que o usuário final visualiza, toda a interface, informação, não importando sua fonte de origem, é exibida graças à camada de visão.
- **Controlador:** O controlador (Controller), como o nome já sugere, é responsável por controlar todo o fluxo de informação que passa pelo site ou sistema. É no controlador que se decide “se”, “o que”, “quando” e “onde” deve funcionar. Define quais informações devem ser geradas, quais regras devem ser acionadas e para onde as informações devem ir, sendo no controlador que essas operações devem ser executadas. Em resumo, é o controlador que executa uma regra de negócio (modelo) e repassa a informação para a camada de visualização (visão).

A Figura (23) apresenta a arquitetura MVC utilizada no SISQEE.

Figura 23. Arquitetura MVC.

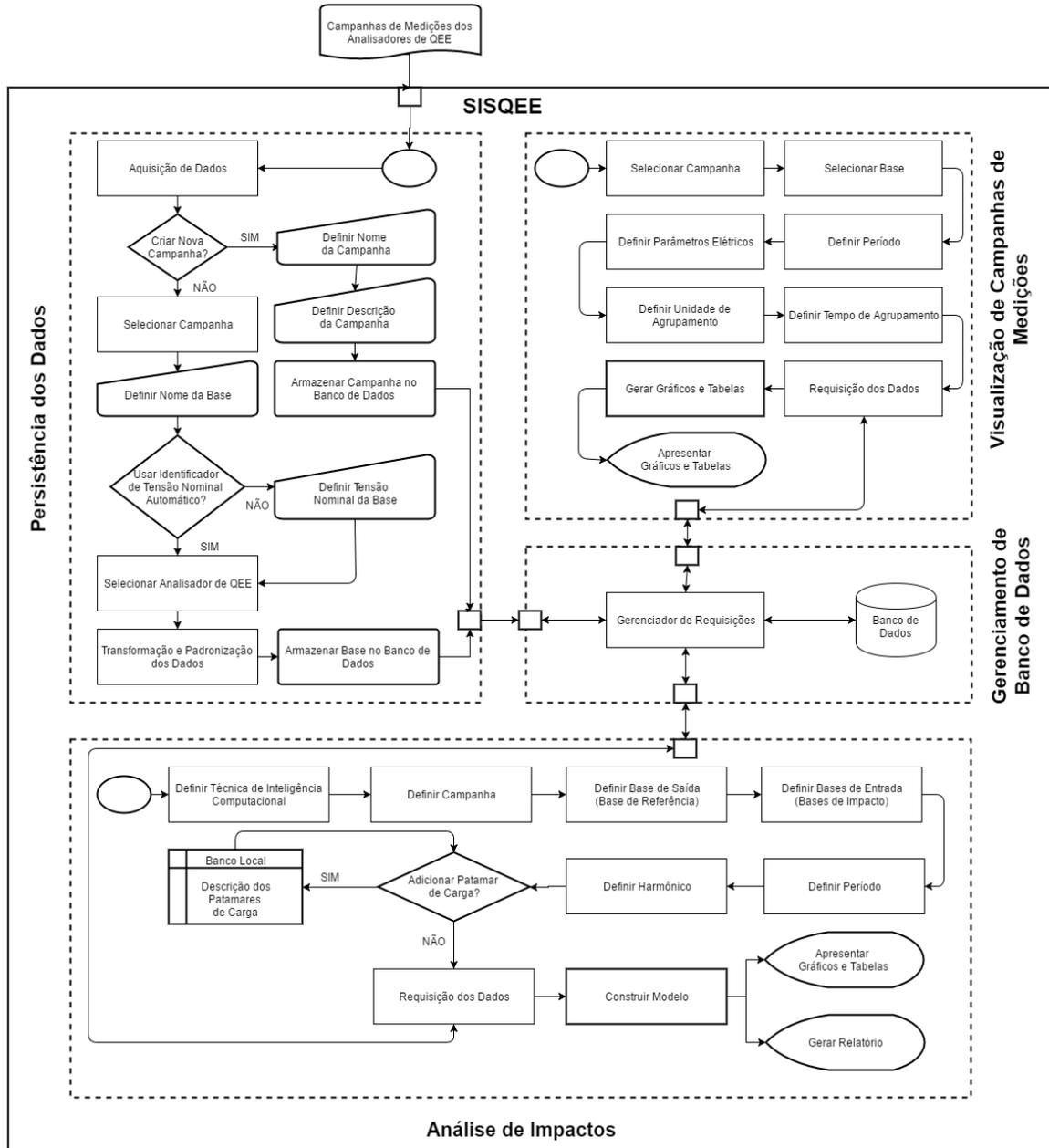


Fonte: O Autor.

Cada um destes blocos pode ser observado na estrutura do software apresentada na Figura (24), onde é possível observar a organização e o fluxo dos dados. Uma vez realizadas as campanhas de medições, é necessário primeiramente armazená-las no banco de dados, para isso é

utilizado do módulo de banco de dados que recebe os arquivos de saída contendo os dados de medições realizadas por diversos analisadores de QEE, sendo necessário convertê-los para um formato padrão, e transformar os dados para que fiquem homogêneos quando forem armazenados pelo módulo de gerenciamento de banco de dados, o qual está conectado com todos os outros módulos para que possa controlar o fluxo de informações que transitam pelo software, através do módulo de análise de impactos e através do módulo de visualização de campanhas de medições.

Figura 24. Estrutura do Software.



Fonte: O Autor.

A estrutura do SISQEE está dividida em quatro blocos:

- **Persistência dos Dados:** Responsável por receber os dados de campanhas de medições oriundas dos arquivos de saída de analisadores de QEE e armazená-los no banco de dados. Para isso é necessário primeiramente cadastrar uma nova campanha de medição, a qual poder conter qualquer quantidade de bases que representam um ponto da rede elétrica. Em seguida definir o analisador de QEE que extraiu estes dados, para que o SISQEE possa utilizar o modelo adequado para extrair os parâmetros elétricos, filtrando os campos necessários, e realizar a transformação dos dados em múltiplos arquivos contendo os dados que serão importados para o banco de dados.
- **Gerenciamento de Banco de Dados:** Responsável pelo fluxo de dados do sistema, modelando os dados dos parâmetros elétricos quando cada módulo requisitar acesso para visualização ou análise.
- **Visualização de Campanhas de Medição:** Responsável por fornecer os gráficos das series temporais dos parâmetros elétricos e suas respectivas tabelas que auxiliam na análise de QEE, contendo valores calculados dos indicadores recomendados pelo PRODIST.
- **Análise de Impactos:** Responsável por fornecer relatórios que contém os gráficos e tabelas dos módulos de Inteligência Computacional, que contém as parcelas de contribuição da distorção harmônica de cada ponto de rede elétrica considerando o background. Para isso, deve-se definir a ordem do harmônico que será utilizado para montar as bases para construção dos modelos que representem o sistema elétrico analisado, o período desejado, a quantidade de pontos que terão uma parcela de contribuição estimada em relação a outro ponto de interesse, e decidir se o período escolhido será reduzido através de patamares de carga.

3.4 INTERFACE GRÁFICA

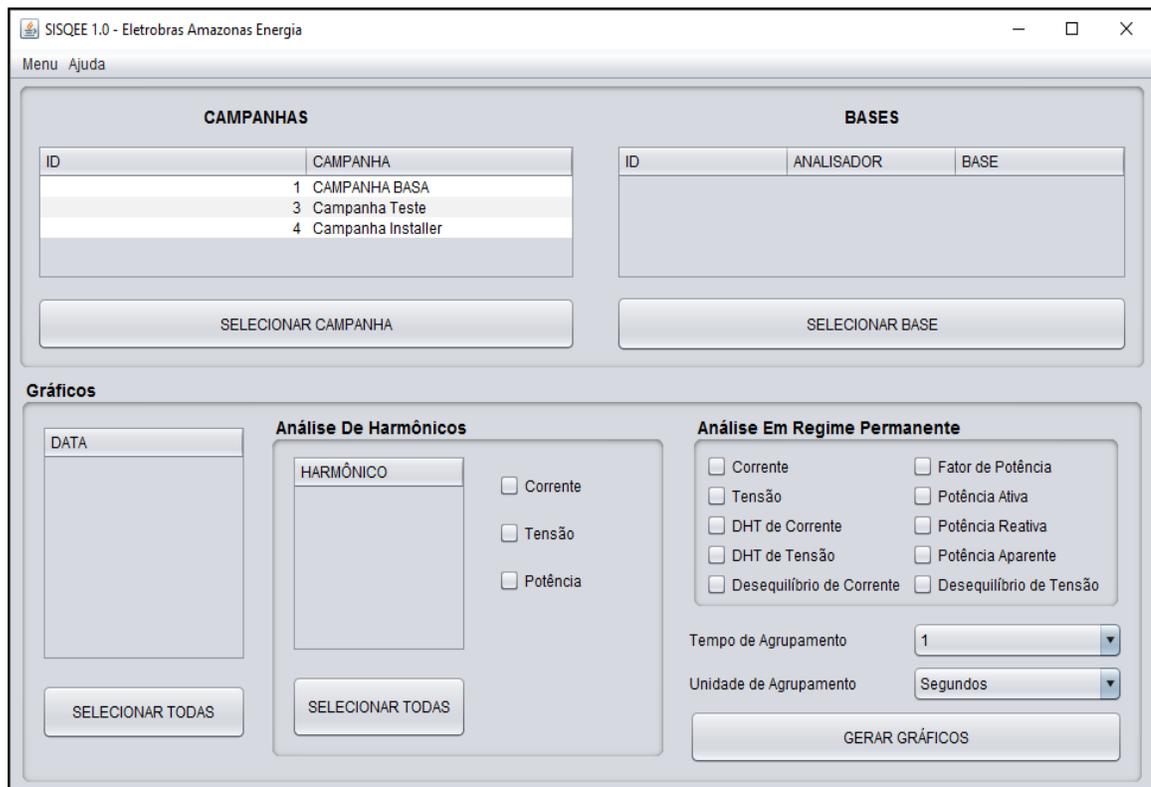
Com o objetivo de facilitar o manuseio do software, foi provida uma interface gráfica amigável para o usuário utilizando os recursos de JFrame do Java. O software está dividido em cinco módulos:

- Módulo de Banco de Dados
- Módulo de Análise de Medições
- Módulo de Inteligência Computacional – Árvores de Decisão

- Módulo de Inteligência Computacional – Redes Neurais
- Módulo de Inteligência Computacional – Regressão Linear

A tela inicial do software esta apresentada na Figura (25) corresponde ao Módulo de Análise de Medições. É possível acessar todos os demais módulos através das opções da barra de menu, assim como todas as rotinas necessárias para geração de gráficos para análise de Harmônicos. A interface fornece um ambiente simples e organizado, permitindo que usuário visualize todas as campanhas de medições armazenadas no banco de dados.

Figura 25. Tela Inicial do SISQEE.



Fonte: O Autor.

Com o objetivo de seguir um fluxo lógico de informação, os tópicos a seguir estão organizados considerando que este seja o primeiro contato do usuário com o sistema. Assim, primeiramente é apresentado o módulo de Banco de Dados, para o entendimento do processo de armazenamento e gerenciamento de campanhas de medições. Em seguida será descrito o funcionamento do módulo de Análise de Medições, e por último os módulos de Inteligência Computacional.

3.4.1 Módulo de Banco de Dados

Este módulo é responsável pelo gerenciamento dos dados que serão armazenados no banco de dados, e é o primeiro módulo a ser utilizado pelo usuário. Este módulo possui três etapas, onde cada etapa abre uma tela própria. A primeira etapa, conforme Figura (26), é criar uma campanha de medição que corresponde à inclusão dos analisadores utilizados e seus respectivos dados adquiridos. A segunda etapa consiste na importação da base de dados para ser incluído em uma campanha, onde cada base é na verdade um conjunto de arquivos de dados de medição de um analisador de QEE. A tela de importação de dados é mostrada na Figura (27).

Figura 26. Tela para criação de uma nova campanha de medição.

A tela 'Criar Nova Campanha' apresenta um formulário com dois campos de texto: 'Nome da Campanha' e 'Descrição'. Abaixo dos campos, há um botão 'CRIAR CAMPANHA'.

Fonte: O Autor.

Figura 27. Tela de importação de uma base de dados.

A tela 'CEAMAZON - Análise de Medições' é dividida em seções para gerenciamento de campanhas e importação de dados.

Campanhas

ID	CAMPANHA
1	CAMPANHA BASA
3	Campanha Teste
4	Campanha Inst...

Botões: SELECIONAR, ATUALIZAR

Bases de Dados

ID	ANALISADOR	BASE	TENSÃO BASE
----	------------	------	-------------

Importar Nova Base de Dados

Selecionar Arquivo

Nome da Base:

Analisador:

ABRIR

ABRIR

Tensao Nominal

Manual

Automático

- * 220 V
- * 380 V
- * 440 V
- * 13800 V
- * 34500 V

IMPORTAR

Fonte: O Autor.

Este módulo permite ao usuário armazenar a tensão de referência da base de dados referente à tensão utilizada para mudança de base, para que os dados de tensão sejam convertidos para *pu*. Para isso pode-se preencher o campo manualmente ou deixar que o SISQEE identifique de forma automática, desde que a tensão de referência esteja entre as tensões que estão armazenadas no banco de dados e que são exibidas no painel “*Tensão Nominal*”. Na Figura (28) a seguir é demonstrando o preenchimento dos dados necessários para o processo de importação de uma base de dados de medição.

Figura 28. Tela de Importação de Dados.

ID	CAMPANHA
16	Campanha 4
17	Campanha 5
18	Campanha 6
19	Campanha 7
20	Campanha 8
21	Campanha 9
22	Campanha 10
23	Campanha 11
24	Campanha 12

ID	ANALISADOR	BASE	TENSÃO BASE
97	HIOKI	Hotel manaos	220.0
98	HIOKI	Receita federal	34500.0
99	HIOKI	SE aparecida	13800.0
105	MARH-21 RMS	REFRIGERACAO	220.0

Fonte: O Autor.

Os analisadores de QEE utilizados não exportam os valores de tensão em *pu*, por isso é necessário realizar a conversão salvando ambas as versões no banco de dados. A opção de identificação automática da tensão nominal funciona seguindo alguns procedimentos. Realizando a leitura de uma amostra qualquer de tensão é possível identificar a tensão nominal base, no entanto é possível que ao escolher uma amostra ao acaso exista uma falha de leitura do medidor. A fim de evitar isto, primeiramente é realizada uma leitura prévia de todas as amostras de tensão da fase A, então é calculada uma média aritmética simples de acordo com a equação (31).

$$V_{média} = \frac{\sum_1^N V_i}{N} \quad (31)$$

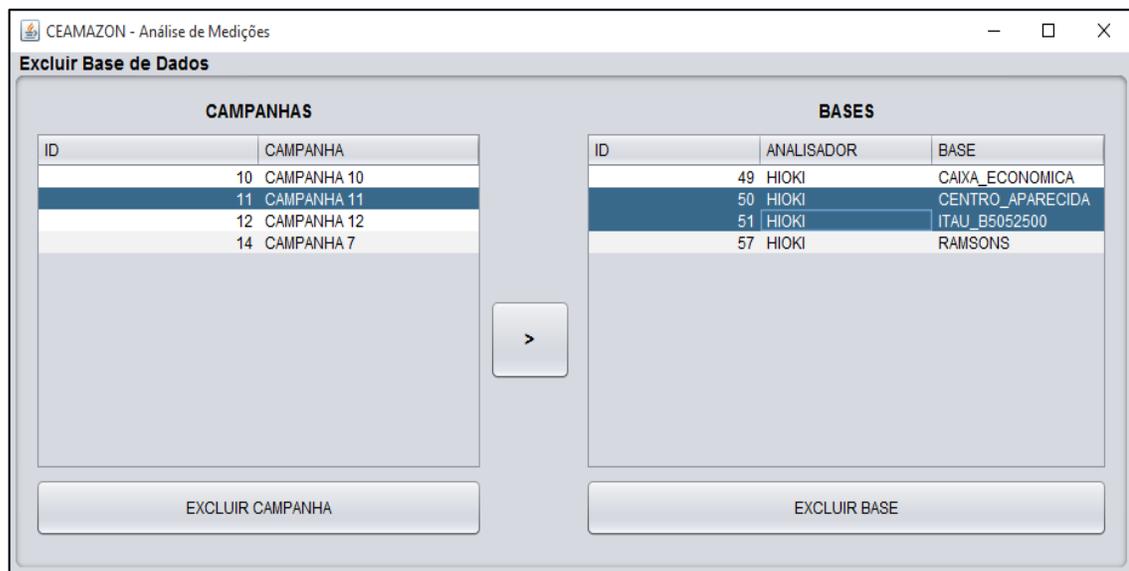
Uma vez obtido a tensão média é realizada uma consulta no banco de dados para verificar qual das tensões nominais contidas na Tabela 8 mais se aproxima da tensão calculada, para isso é utilizada a equação (32), onde é calculada a distância absoluta entre as tensões contidas no banco e a tensão média, por fim a consulta retorna a tensão que obtiver o menor valor.

$$\Delta(V_{m\acute{e}dia}, V_i) = |V_{m\acute{e}dia} - V_i| \quad (32)$$

Ao finalizar a importação será exibida uma mensagem de confirmação e os campos da tela voltarão ao estado inicial, e é criada dentro da pasta **DATABASES**, que pode ser acessada a partir da pasta raiz do software SISQEE, uma pasta com o nome da campanha e uma subpasta com o nome da base importada contendo outras três subpastas, conforme Figura (A-1). A partir deste ponto, tudo que o software gerar de relatórios e gráficos serão armazenados nestas pastas, caso alguma das técnicas de computação seja aplicada, outras pastas serão criadas dentro da pasta da base.

A última opção fornecida pelo módulo de Banco de Dados é a função de excluir uma ou mais bases ou ainda uma campanha inteira. A Figura (29) demonstra a tela para exclusão das bases de dados.

Figura 29. Tela para excluir uma base de dados.



Fonte: O Autor.

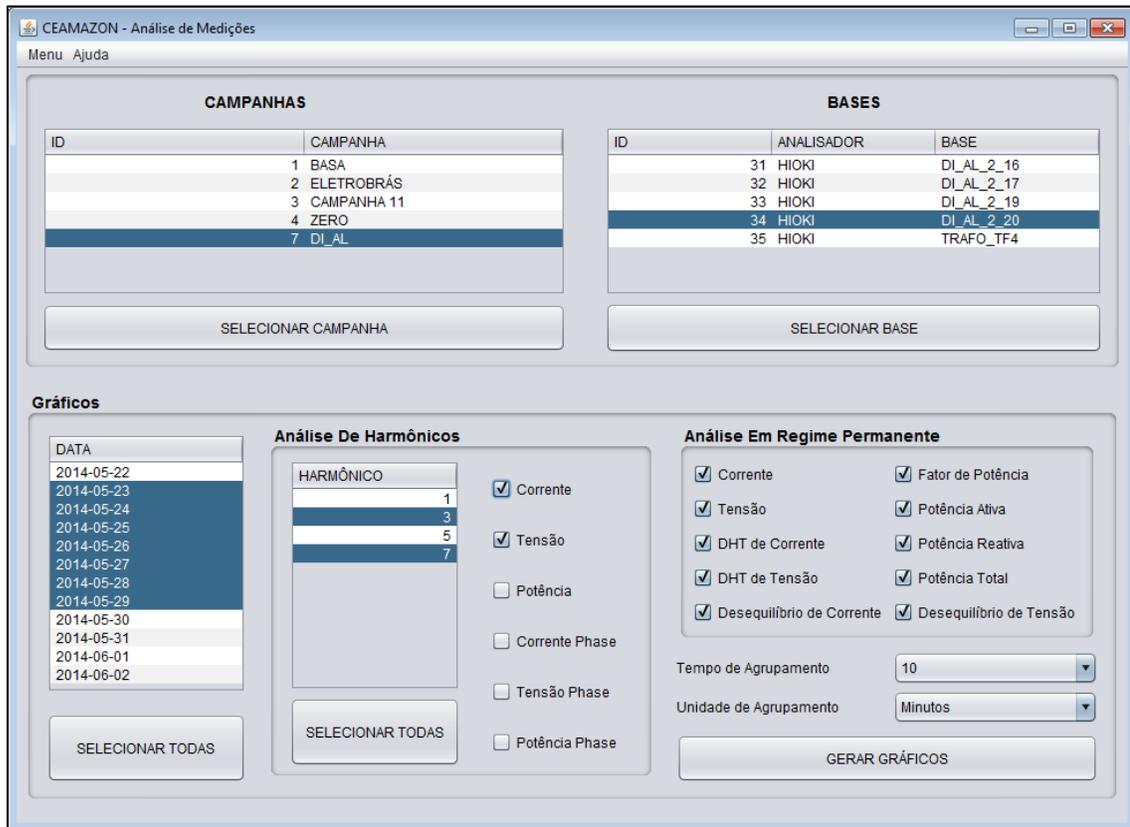
3.4.2 Módulo de Análise de Medições

É neste módulo que o usuário pode fazer as análises nos pontos de medição com relação à QEE, obedecendo aos critérios e limites estabelecidos pelo módulo 8 do PRODIST. Como mencionado anteriormente este módulo está contido na tela inicial do SISQEE, uma vez que uma

campanha foi criada e bases de dados foram importadas com sucesso, é possível gerar gráficos de análise. Para gerar os gráficos de uma base de dados é preciso escolher os parâmetros elétricos desejados, o intervalo de agrupamento dos dados e o período da análise, recomenda-se não utilizar a primeira e última data da base de dados, pois provavelmente a medição não possui dados de um dia completo, pois para isso, a medição deveria começar e terminar exatamente a meia noite.

O manual do PRODIST recomenda que o intervalo dos dados para análise seja de 10 minutos, no entanto é possível que na medição das campanhas, os analisadores de QEE tenham sido configurados para coletar dados em intervalos diferentes, até mesmo em segundos, por isso o SISQEE fornece esta opção de agrupar os dados, fornecendo opção de usuário escolher entre os múltiplos de 60 (sessenta) para agregar os dados (1, 2, 3, 4, 5, 6, 10, 12, 15, 20, 30, 60). Ressalta-se que se o intervalo escolhido pelo usuário for menor do que o intervalo que a base foi configurada, os gráficos serão gerados com o intervalo da base, caso contrário, o SISQEE realiza uma média aritmética simples utilizando todas as amostras contidas no intervalo selecionado pelo usuário, por exemplo, caso a base foi configurada para realizar a coleta de dados a cada 2 (dois) minutos, significa que a cada 5 (cinco) valores será tirado uma média obtendo-se um novo valor o qual será apresentado no gráfico gerado. A Figura (30) demonstra esta etapa para geração dos gráficos desejados.

Figura 30. Demonstração do processo de geração de gráficos pelo módulo de Análise de Medições.



Fonte: O Autor.

Os gráficos gerados podem ser acessados dentro da pasta *graph*, onde dentro dela são criadas outras subpastas, uma para cada parâmetro elétrico selecionado, conforme Figura (A-2).

Para a maioria dos parâmetros elétricos é criado um gráfico para cada dia dentro do período selecionado, um para o fim de semana e um para o período completo, conforme mostrado na Figura (A-3). Os gráficos gerados estão no formato *html* e serão abertos com um navegador de internet.

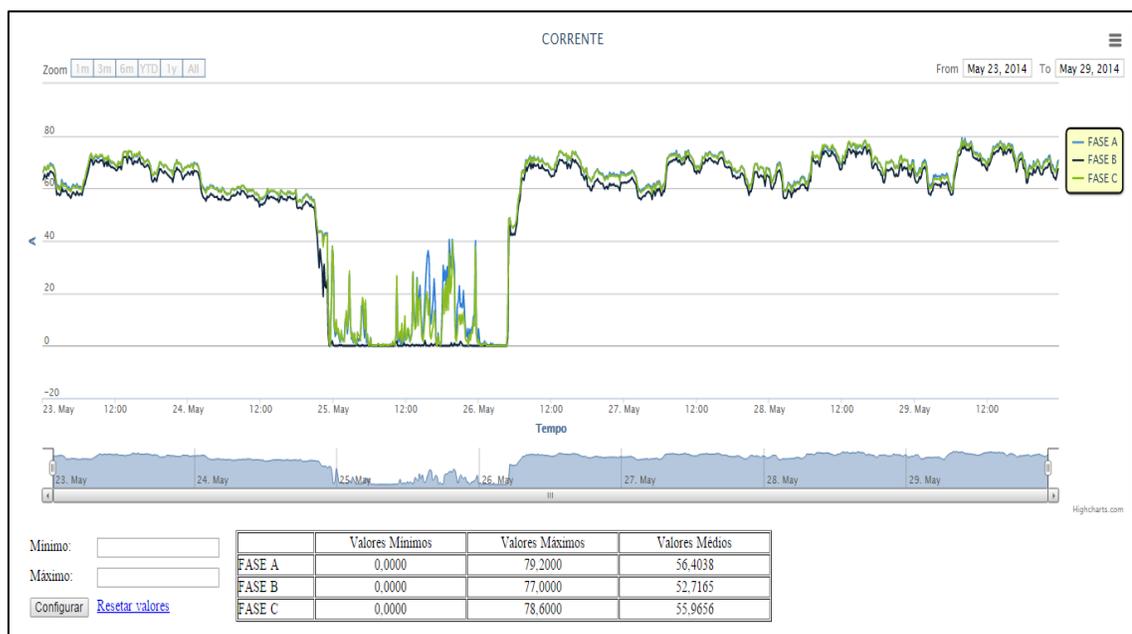
Para geração de gráficos foram utilizadas duas bibliotecas gratuitas desenvolvidas em JavaScript, a *jquery-2.0.3* responsável pelo carregamentos dos dados e a *highstock* responsável pela geração dos gráficos de forma responsiva, ou seja, os gráficos gerados podem se adaptar a tela para se adequar da melhor forma.

A vantagem de utilizar esta biblioteca é que ela possui diversos recursos adicionais para facilitar visualização e análise dos dados, como escolher quais das curvas plotadas estarão visíveis no gráfico; possui barra de rolagem inferior capaz de reduzir a amostra de dados no eixo x, já para o eixo y, a biblioteca não tem essa função, porém por se tratar de uma biblioteca para

web, foi possível implementar uma função também em JavaScript para realizar esse corte; a visualização dos dados de forma é precisa ao passar o mouse sobre uma das curvas plotadas; é possível realizar a exportação do gráfico para diversos formato de imagem como *jpg*, *png* e *svg* ou para *pdf*.

As desvantagens consistem na necessidade de possuir um navegador atualizado, e que a qualidade dos gráficos depende da qualidade da resolução de tela do computador em que o SISQEE é executado, além da necessidade de internet para exportar os gráficos. A Figura (31) mostra um gráfico de corrente gerado para um período de sete dias.

Figura 31. Gráfico de Corrente gerado pelo módulo de Análise de Medições.

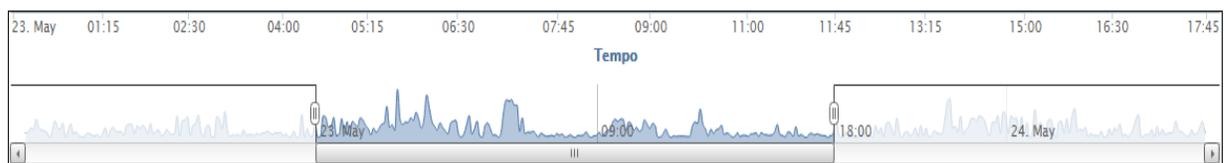


Fonte O Autor.

A maioria dos gráficos gerados apresenta uma tabela contendo dados para auxiliar análise, como valores mínimos, máximos e médios, alguns parâmetros contêm outros dados na tabela como percentil de 95%, já a tensão apresenta outras tabelas contendo dados de DRP e DRC.

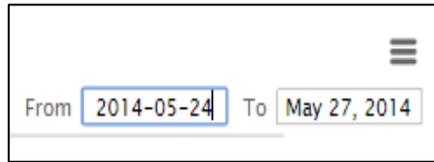
O SISQEE permite o ajuste dos eixos x e y, e para ajustar o eixo x basta ajustar a barra de rolagem abaixo do gráfico, conforme mostrado na Figura (32) ou ainda preenchendo os campos de datas no canto superior direito conforme Figura (33).

Figura 32. Demonstração do ajuste do eixo x no gráfico gerado.



Fonte: O Autor.

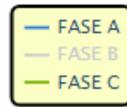
Figura 33. Demonstração alternativa do ajuste do eixo x no gráfico gerado



Fonte: O Autor.

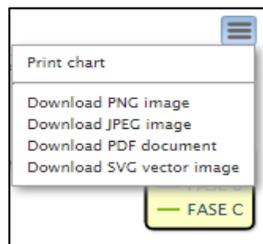
É possível ainda selecionar quais fases o gráfico exibirá, para isso basta desmarcar as fases desejadas na legenda a direita, mostrada na Figura (34). Para realizar a exportação ou salvar os gráficos basta acessar o menu no canto superior direito, conforme Figura (35).

Figura 34. Legenda do Gráfico.



Fonte: O Autor.

Figura 35. Exportar Gráfico.



Fonte: O Autor.

3.5 CONSIDERAÇÕES FINAIS

Neste capítulo foi mostrado como o SISQEE está estruturado em camadas e módulos, os recursos que foram utilizados para o desenvolvimento desta ferramenta, as funcionalidades que provêm ao usuário uma análise de QEE de melhor qualidade. Também foi apresentada a organização do banco de dados construído, sendo este um dos pontos forte da ferramenta, uma vez que permite a modelagem de dos dados de maneira uniforme, permitindo que outros analisadores de QEE possam ser integrados. Além disso, foram apresentadas as telas das interfaces gráficas que foram construídas para estes módulos. No capítulo a seguir, serão apresentados os módulos designados para realizar a análise de contribuição de impacto harmônico, mostrando suas respectivas telas gráficas e a metodologia aplicada em cada um.

4 MÓDULOS DE INTELIGÊNCIA COMPUTACIONAL

4.1 CONSIDERAÇÕES INICIAIS

Neste capítulo são apresentados os módulos de IC desenvolvidos no SISQEE, os quais realizam as análises de impacto de contribuição de injeção de harmônicos na rede elétrica, atribuindo uma porcentagem de responsabilidade para cada ponto de interesse em que foi realizada uma campanha de medição.

Estes módulos foram implementados utilizando as bibliotecas Weka e JFreeChart. Para que sejam utilizados é necessário que as campanhas de medições tenham sido importadas previamente no *Módulo de Banco de Dados*, e ainda que contenham dados de tensões e correntes harmônicas.

4.2 PATAMAR DE CARGA

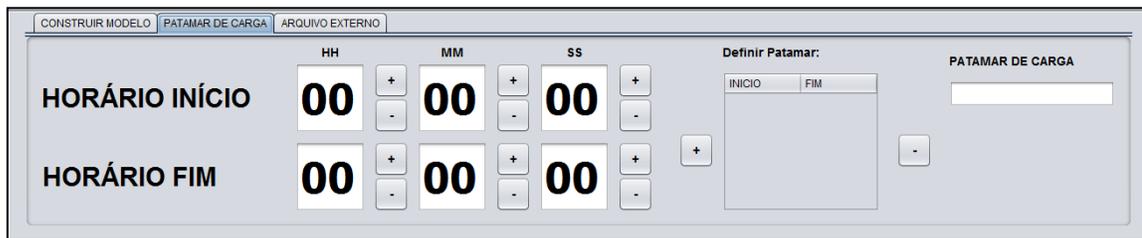
Cada uma destas técnicas é recomendada para uma situação diferente, a RL é recomendada para obter a porcentagem da contribuição de impacto harmônico causado pela tensão de *background* levando em consideração apenas um ponto da rede elétrica, a RNA e a AD são recomendadas para obter o impacto harmônico levando em consideração múltiplos pontos da rede elétrica, no entanto, ressalta-se que a RNA obtém melhores resultados considerando um período de análise grande, por exemplo, uma semana completa, e a AD obtém resultados melhores considerando intervalos pequenos, como por exemplo, patamares de carga.

Patamar de carga é a classificação das horas do mês, de acordo com o perfil de carga definido pelo ONS (Operador Nacional do Sistema Elétrico), o patamar de carga pode ser de três tipos:

- Leve (horários do dia em que se observam os menores consumos de energia elétrica);
- Médio (horários do dia em que o consumo de energia elétrica é mediano);
- Pesado (horários do dia nos quais o consumo de energia elétrica é maior e podem ocorrer picos).

Todos os módulos de inteligência computacional possuem uma aba para patamar de carga, conforme Figura (36), na qual o usuário pode acrescentar múltiplos horários para especificar um patamar de carga sobre o qual deseje realizar a análise de impactos.

Figura 36. Aba Patamar de Carga.



Fonte: O Autor.

Estes módulos também possuem uma aba para criação de modelos destas técnicas de IC utilizando como base um arquivo externo, conforme Figura (37) ao invés de montar a base com os dados que estão no banco de dados. O arquivo tem que estar no formato *csv* e no mesmo padrão dos arquivos que o SISQEE cria utilizando os dados armazenados das campanhas de medição. O propósito desta funcionalidade é realizar estudos de bases que não estejam vinculadas a uma campanha de medição armazenada no banco de dados, agindo semelhante à própria ferramenta do weka, com a ressalva de poder gerar um relatório contendo os resultados obtidos.

Figura 37. Aba Arquivo Externo.



Fonte: O Autor.

A principal aba, mostrada na Figura (38), destes módulos é utilizada para construir os modelos das técnicas de IC. Para a construção de um modelo é necessário selecionar uma campanha de medição, o ponto de medição que será usado como referência utilizando a tensão harmônica, os ponto de medição que serão avaliados os impactos utilizando as correntes harmônicas, o intervalo de tempo para realizar a análise e a ordem harmônica.

Figura 38. Aba Construir Modelo.

Definir Campanha		Definir Base de Saída (V):		Definir Bases de Entrada (I):		Definir Período:	Definir Harmônico:
ID	CAMPANHA	ID	BASE	ID	BASE	Data	Harmônico:
1	CAMPANHA BASA	1	C1_SALA_SERVIDOR	1	C1_SALA_SERVIDOR	2015-06-16	3
3	Campanha Teste	2	C2_SALA_SERVIDOR	2	C2_SALA_SERVIDOR	2015-06-17	
4	Campanha Installer	3	H1_SE_GERADOR_...	3	H1_SE_GERADOR E...	2015-06-18	
		4	H2_SALA_NOBREAKS	4	H2_SALA_NOBREAKS	2015-06-19	
		5	REFRIGERACAO	5	REFRIGERACAO	2015-06-20	
		6	SUBESTACAO_GERAL	6	SUBESTACAO_GERAL	2015-06-21	
						2015-06-22	
						2015-06-23	

Fonte: O Autor.

Então para cada fase é criada uma *thread*, um priorizador de tarefas. Cada *thread* é responsável por criar um conjunto de arquivos *csv's* para construção dos modelos de forma paralela a fim de reduzir o tempo de execução, estes arquivos são salvos nas respectivas pastas de cada técnica que são criadas dentro da pasta da base que está sendo usado como ponto de referência, conforme mostrado na Figura (A-4).

Ressalta-se que embora as medições sejam simultâneas, é possível que o usuário escolha as datas em que as bases não contenham a mesma quantidade de dados, por isso é recomendado que se possível não fossem utilizadas a primeira e a última datas contidas na campanha de medição, pois se considera altamente improvável que todos os analisadores de QEE tenham iniciado e terminado a medição exatamente no mesmo instante. Também é possível que por falta de atenção os analisadores tenham sido configurados para coletar os dados em intervalos diferentes.

Mesmo que uma dessas situações ocorra, o SISQEE se encarrega de resolver, pois no caso de haver analisadores com configurações diferentes, são realizadas consultas SQL para verificar o intervalo em que cada amostra foi realizada, então todas as bases são padronizadas com base no maior intervalo, realizando uma média aritmética simples nas bases que possuem intervalo menor, e no caso dos analisadores não terem iniciado e terminado a medição no mesmo instante, também são realizadas consultas no banco de dados, verificando o maior horário de início e o menor horário de término comum em todas em todos os pontos de medição.

Essas consultas constituem o ajuste das bases, ou seja, o tratamento automático para a montagem das bases de dados dos modelos, de modo que todas as entradas e a saída estejam com o mesmo intervalo de tempo entre as amostras.

O ajuste se inicia após o usuário definir a campanha de medição e as bases que serão utilizadas para construção dos modelos. Então é realizada uma série de consultas no banco de dados, a primeira consulta retorna a menor e a maior data de cada base, como mostrado a seguir.

ID_BASE	MIN_DATA	MAX_DATA
17	22/05/2014	02/06/2014
18	22/05/2014	02/06/2014
19	22/05/2014	24/05/2014
20	22/05/2014	02/06/2014
21	22/05/2014	02/06/2014

Em seguida, é consultado o maior período possível em comum entre as bases, retornando a maior data de início, e a menor data de término, resultando em:

DATA_INICIO	DATA_TERMINO
22/05/2014	24/05/2014

A partir desta faixa, o usuário escolhe o período para a construção dos modelos, então é consultado no banco de dados e apresentado para o usuário os possíveis harmônicos.

ID_HARMONICO
0
1
2
...
48
49
50

Então, para cada base é consultado o intervalo entre cada amostra realizando a diferença entre a primeira e a segunda amostras.

ID_BASE	DATA[0]	DATA[1]	INTERVALO
17	26/05/2014 18:51:00	26/05/2014 18:52:00	00:01:00
18	22/05/2014 11:30:00	22/05/2014 11:31:00	00:01:00
19	22/05/2014 13:22:00	22/05/2014 13:22:00	00:00:03
20	24/05/2014 03:57:00	24/05/2014 03:58:00	00:01:00
21	22/05/2014 18:20:00	22/05/2014 18:21:00	00:01:00

A partir desta consulta é verificado qual o maior intervalo dentre as bases para que possa ser obtidos seu valor em segundos, neste caso resultando em 60. Por fim, para cada fase é realizada uma consulta para cada base contendo os dados agregados pelo maior intervalo dentre as bases.

ID_BASE	PERIODO	COUNT	FASE_TEMP
17	22/05/2014 14:18:00	20	0.569
17	22/05/2014 14:19:00	20	0.5725
17	22/05/2014 14:20:00	20	0.5775
17	22/05/2014 14:21:00	20	0.5745
17	22/05/2014 14:22:00	20	0.57
17	22/05/2014 14:23:00	20	0.578

É nesta última consulta que os patamares de carga podem ser utilizados. Também é importante ressaltar que embora os valores de tensão e corrente harmônicos sejam importados para o banco de dados em porcentagem da componente fundamental, para criação dos modelos de IC eles são convertidos para valores nominais conforme as equações (33) e (34).

$$V_{h_nominal} = \frac{V_{h\%} \cdot V_1}{100} \quad (33)$$

$$I_{h_nominal} = \frac{I_{h\%} \cdot I_1}{100} \quad (34)$$

Onde V_1 e I_1 são as componentes fundamentais de corrente e tensão em valores nominais, $V_{h\%}$ e $I_{h\%}$ são as componentes harmônicas de mesma ordem de corrente e tensão em porcentagem da fundamental.

4.3 MÓDULO DE REGRESSÃO LINEAR

Este módulo permite a análise de regressão linear simples, a qual consiste em um método estatístico que relaciona uma variável aleatória, denominada de variável resposta (tensão harmônica para o caso em questão), a outra variável aleatória denominada de variável regressora (corrente harmônica para o caso em questão), condicionado ao modelo mostrado em (35) análogo ao mostrado anteriormente em (21) (SOARES, 2012).

$$V_h = \beta_0 + \beta_1 \cdot I_h + \xi \quad (35)$$

Onde V_h corresponde a tensão harmônica de um barramento qualquer a ser analisado, I_h é a corrente harmônica da fonte não linear considerada, na mesma frequência da tensão harmônica a ser analisada, β_0 e β_1 são constantes e ξ corresponde ao erro.

É importante ressaltar que para o caso em questão, β_0 é definido como uma tensão harmônica de *background*, ou seja, corresponde à distorção harmônica de tensão no barramento de interesse proveniente do sistema elétrico como um todo, desconsiderando a fonte não linear I_h selecionada para o estudo.

Como dito anteriormente a regressão linear é utilizada para verificar o impacto harmônico em um ponto de interesse considerando apenas um ponto de análise e o *background*, logo a base de dados para construção deste modelo possui apenas dois campos, corrente e tensão harmônicas de mesma ordem. Para construção dos modelos de regressão foi utilizado o método *getOLSRegression* da classe *org.jfree.data.statistics.Regression* da biblioteca *JFreeChart* do java,

o qual retorna os parâmetros a e b para uma equação $y = a + bx$, ajustada aos dados usando regressão por mínimos quadrados ordinários.

Uma vez obtidos os parâmetros da equação do modelo de regressão linear, calcula-se os impactos percentuais de *background* ponto a ponto, ou seja, para cada valor de tensão estimada, e da *fonte*. Primeiramente calcula-se o coeficiente de correlação de Pearson (36) entre a corrente e a tensão harmônicas medidas.

$$\rho = \frac{\sum_{i=1}^N (x_i - \hat{x})(y_i - \hat{y})}{\sqrt{\sum_{i=1}^N (x_i - \hat{x})^2} \cdot \sqrt{\sum_{i=1}^N (y_i - \hat{y})^2}} = \frac{cov(X, Y)}{\sqrt{var(X) \cdot var(Y)}} \quad (36)$$

Onde x_i e y_i são os valores medidos de ambas as variáveis e \hat{x} e \hat{y} são as médias de X e Y . O SISQEE calcula este coeficiente utilizando o método *correlation* da classe *PearsonsCorrelation*. Em seguida, define-se o coeficiente linear como sendo a tensão de *background* (V_{BG}), então a cada valor de tensão estimada da *fonte* (V_T) é calculado o impacto da tensão de *background*, caso o coeficiente de correlação de Pearson (ρ) seja maior ou igual a 0 (zero) o Fator de Impacto (IF - Impact Factor) pontual é calculado por (37), caso contrário assume-se que a correlação é inversa e o IF é calculado por (38). Por fim, o IF da fonte é obtido pelo completo do impacto da tensão de *background* conforme (39).

$$IF_{V_{BGi}}^h (\%) = \frac{V_{BG} \cdot 100\%}{V_{Ti}} \quad (37)$$

$$IF_{V_{BGi}}^h (\%) = \frac{V_{Ti} \cdot 100\%}{V_{BG}} \quad (38)$$

$$IF_{V_{Ti}}^h (\%) = 100,0 - IF_{V_{BGi}}^h (\%) \quad (39)$$

Os fatores de impacto gerais são obtidos pelas médias aritméticas simples destes impactos ponto a ponto, conforme (40) e (41).

$$IF_{V_{BG}}^h (\%) = \frac{\sum_{i=1}^N IF_{V_{BGi}}^h (\%)}{N} \quad (40)$$

$$IF_{V_T}^h (\%) = \frac{\sum_{i=1}^N IF_{V_{Ti}}^h (\%)}{N} \quad (41)$$

Para avaliar a confiabilidade dos resultados é utilizado o fator de determinação R^2 mostrado na equação (42) análoga a mostrada anteriormente em (29). Para isso, utilizam-se dois vetores, um para as tensões medidas V_M e um para as tensões estimadas V_E .

$$R^2 = \frac{\sum_{i=1}^N (V_{Ei} - \widehat{V}_E)^2}{\sum_{i=1}^N (V_{Mi} - \widehat{V}_M)^2} \quad (42)$$

Onde \widehat{V}_E corresponde à tensão média estimada pelo modelo de regressão, e \widehat{V}_M corresponde à tensão média medida pelos analisadores de QEE. Os resultados tem confiança se o valor obtido do r quadrado for pelo menos alta conforme a Tabela 10.

Tabela 10. Faixas de Classificação do R² Quadrado.

Valor de R ²	Intensidade de Correlação
0,00	Nula
0,00 - 0,09	Baixa
0,09 - 0,36	Moderada
0,36 - 0,81	Alta
0,81 - 0,98	Muito Alta
1,00	Perfeita

Todos estes passos são descritos de forma sucinta no Algoritmo 1 mostrado a seguir.

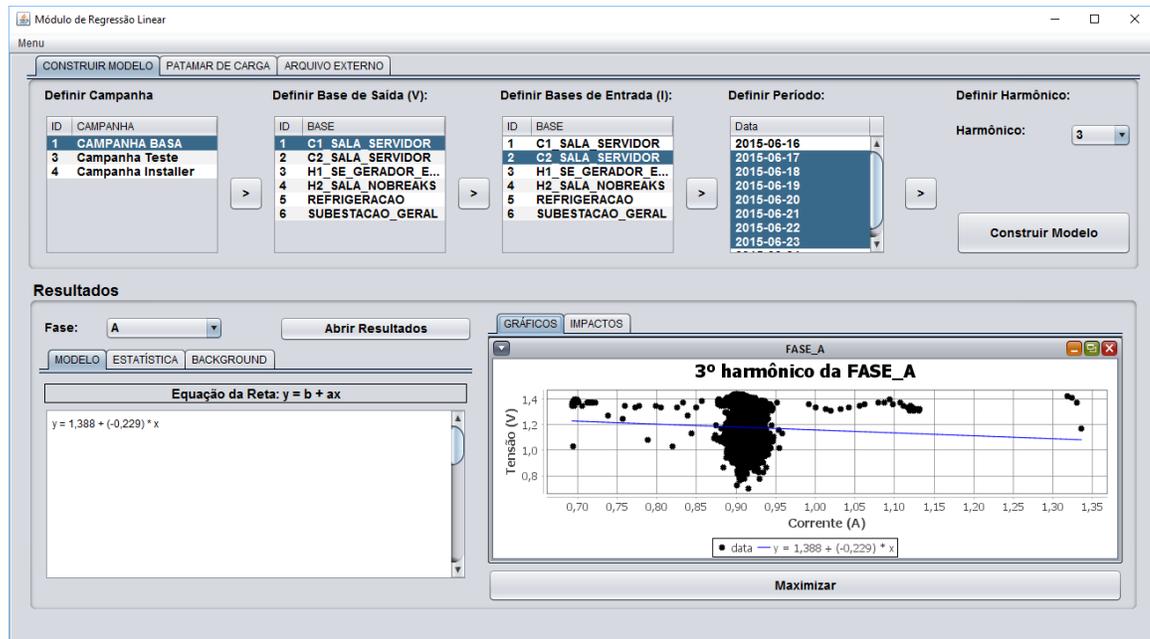
Algorithm 1 Análise de Impacto Harmônico de uma Carga Utilizando RL

Início

- 1: Obter parâmetros β_0 e β_1 do modelo de regressão linear
 - 2: $V_{BG} \leftarrow \beta_0$
 - 3: $\rho \leftarrow \frac{cov(V_T, I_T)}{\sqrt{var(V_T) \cdot var(I_T)}}$
 - 4: **para** $i \leftarrow 1$ **até** $size(V_E)$ **faça**
 - 5: **se** $\rho > 0$ **então**
 - 6: $IF_{V_{BG}i}^h (\%) \leftarrow \frac{V_{BG} \cdot 100\%}{V_{Ti}}$
 - 7: **senão**
 - 8: $IF_{V_{BG}i}^h (\%) \leftarrow \frac{V_{Ti} \cdot 100\%}{V_{BG}}$
 - 9: **fim se**
 - 10: $IF_{V_{Ti}i}^h (\%) \leftarrow 100,0 - IF_{V_{BG}i}^h (\%)$
 - 11: **fim para**
 - 12: $IF_{V_{BG}}^h (\%) \leftarrow média(IF_{V_{BG}i}^h (\%))$
 - 13: $IF_{V_T}^h (\%) \leftarrow média(IF_{V_{Ti}i}^h (\%))$
 - 14: $R^2 \leftarrow \frac{\sum_{i=1}^N (V_{Ei} - \widehat{V}_E)^2}{\sum_{i=1}^N (V_{Ti} - \widehat{V}_T)^2}$
 - 15: **se** $R^2 \geq$ Alta **então**
 - 16: Resultados apresentam boa confiabilidade
 - 17: **fim se**
- Fim**
-

Os resultados são apresentados na parte inferior da tela do módulo de RL, conforme mostrado na Figura (39).

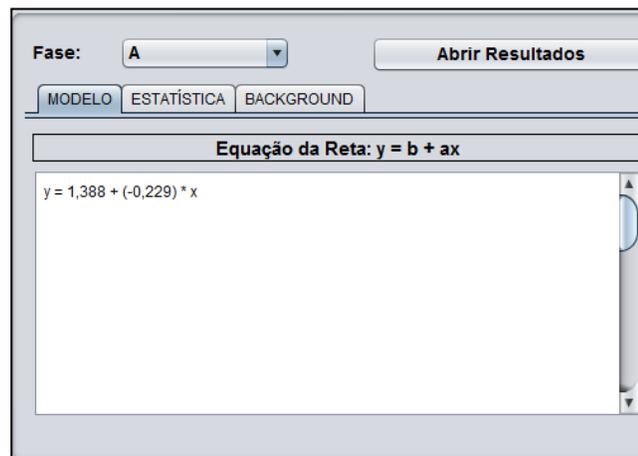
Figura 39. Tela Principal do Módulo de Regressão Linear.



Fonte: O Autor.

Os resultados são divididos em três abas à esquerda (*modelo*, *estatística* e *background*) e duas à direita (*gráficos* e *impactos*). Na aba *modelo* é exibida a equação da reta que descreve o modelo de regressão obtido, apenas da fase selecionada conforme Figura (40).

Figura 40. Resultados Regressão Linear: Aba Modelo.



Fonte: O Autor.

Na aba *estatística* é exibido o coeficiente de determinação (R^2) calculado para as três fases, o qual informa a qualidade do modelo gerado conforme Figura (41).

Figura 41. Resultados Regressão Linear: Aba Estatística.

Métrica	Fase A	Fase B	Fase C
R2	00,001	00,042	00,002

Fonte: O Autor.

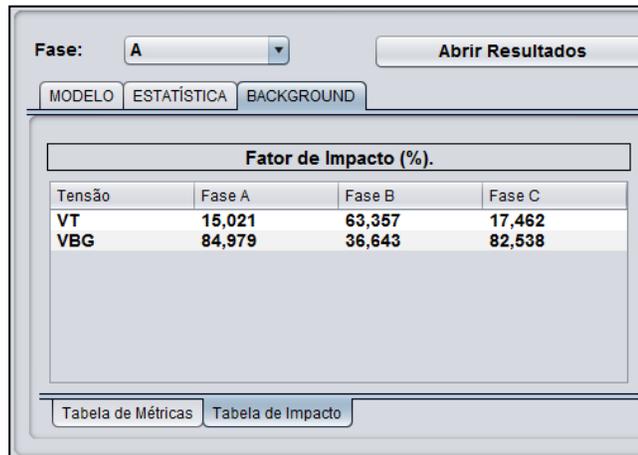
A aba *background* contém duas abas, *Tabela de Métricas* que contém as equações da reta geradas para as três fases, conforme Figura (42), e a *Tabela de Impacto* que contém os impactos de injeção de harmônicos calculados para a *fonte* (VT) e para o *background* (VBG) para cada uma das três fases conforme Figura (43).

Figura 42. Resultados Regressão Linear: Aba Background - Tabela de Métricas.

Métrica	Fase A	Fase B	Fase C
EQUAÇÃO	$y = 1,388 + (-0,22...$	$y = 2,143 + (-...$	$y = 1,300 + (-...$

Fonte: O Autor.

Figura 43. Resultados Regressão Linear: Aba Background - Tabela de Impacto.

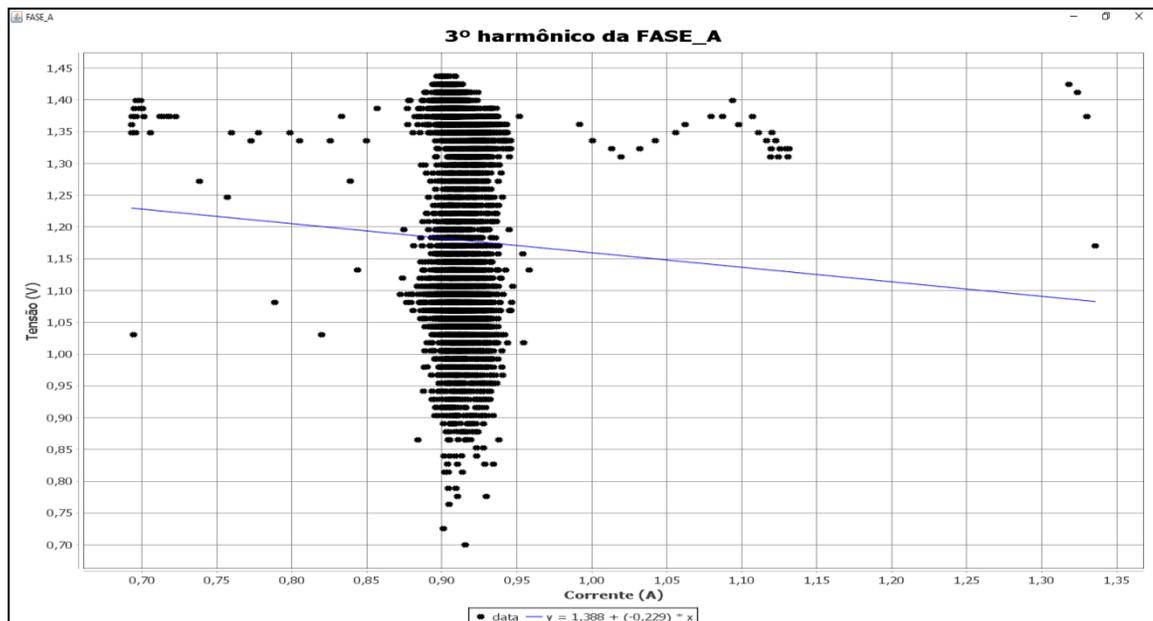


Fator de Impacto (%)			
Tensão	Fase A	Fase B	Fase C
VT	15,021	63,357	17,462
VBG	84,979	36,643	82,538

Fonte: O Autor.

Na aba *gráficos* estão contidos os gráficos de representação dos modelos de regressão criados para as três fases. A Figura (44) apresenta um destes modelos gerados.

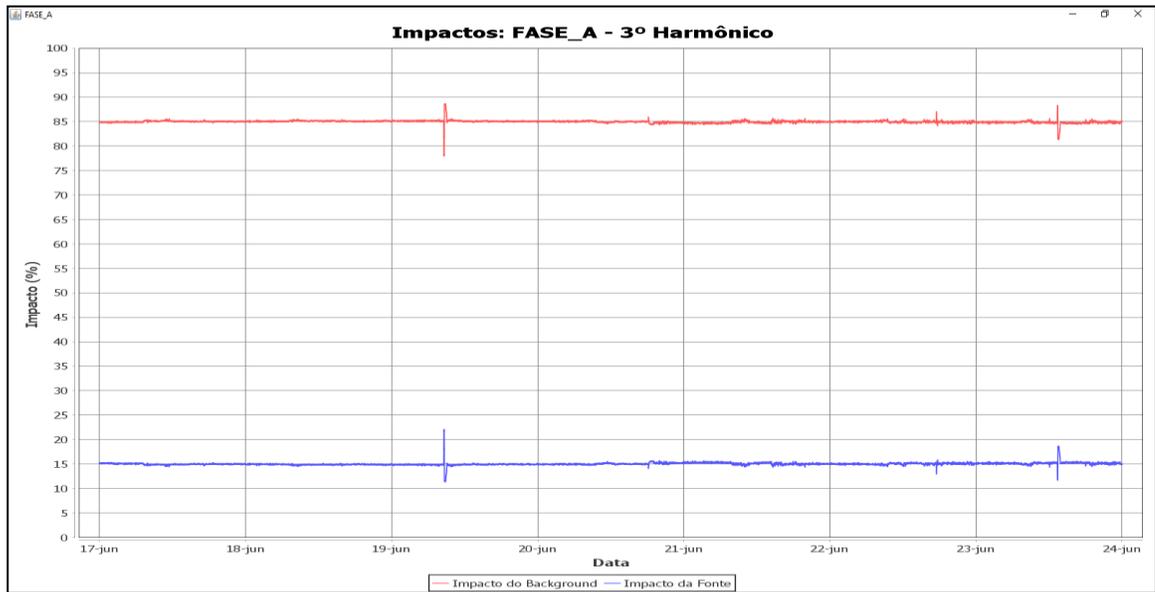
Figura 44. Resultados Regressão Linear: Gráfico do Modelo de Regressão Gerado.



Fonte: O Autor.

Por fim, a aba *impactos* contém os gráficos de impactos contendo os impactos pontuais da *fonte* e de *background*. A Figura (45) apresenta um destes modelos gerados.

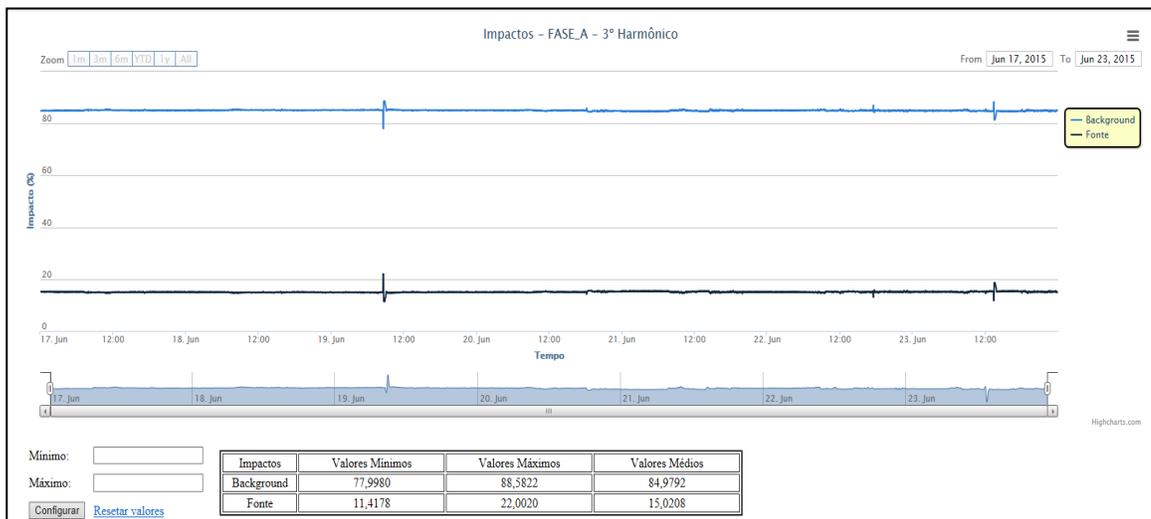
Figura 45. Resultados Regressão Linear: Gráfico de Impactos Pontuais.



Fonte: O Autor.

Ressalta-se que estes gráficos foram criados com a biblioteca *JFreeChart* e podem ser configurados com o botão direito do mouse caso o usuário deseje analisar apenas uma parte dos dados. Os gráficos de impacto também são gerados no formato *html* com as mesmas funcionalidades dos gráficos gerados pelo *Módulo de Análise de Medições*, a Figura (46) apresenta o mesmo gráfico da Figura (45) no formato *html*.

Figura 46. Resultados Regressão Linear: Gráfico de Impactos Pontuais HTML.



Fonte: O Autor.

Estes gráficos são salvos na pasta *REGRESSAO_LINEAR* que é criada dentro da pasta da base que foi utilizada como referência para o estudo, conforme mostrado na Figura (A-4). Ainda

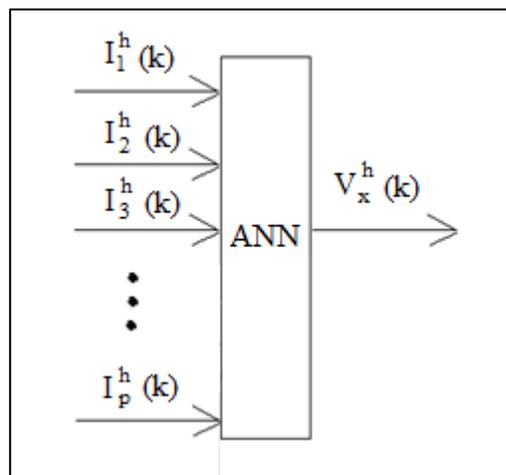
é possível gerar um relatório contendo todos os resultados, tabelas e gráficos, salvo dentro da pasta da base que foi utilizada como referência, conforme Figura (A-5), os gráficos também são salvos no formato *png*. Dentro dessa pasta é possível ver uma pasta para cada fase, onde são salvos arquivos que contém os dados de tensões e correntes harmônicas utilizados para construir os modelos de regressão, permitindo que o usuário utilize outras ferramentas para criar modelos de regressão usando os mesmo dados.

4.4 MÓDULO DE REDES NEURAS ARTIFICIAIS

Este módulo avalia os impactos causados por vários pontos de medição em relação a um ponto de referência utilizando longos períodos de medição. Para o processo de estimação pretende-se, com a utilização da rede neural, estimar a tensão harmônica na barra de interesse a partir do conhecimento adquirido no processo de aprendizagem, sendo utilizado no processo de estimação dados de entrada desconhecidos para o modelo neural desenvolvido.

A Figura (47) mostra o esquema geral do modelo proposto para a estimação da tensão harmônica em uma determinada barra de interesse, onde a entrada da rede é composta pelas correntes harmônicas oriundas das cargas não lineares que compõe o sistema em estudo.

Figura 47. Esquema geral do modelo proposto para a estimação da tensão harmônica em uma determinada barra de interesse.



Fonte: MANITO et al., 2014.

Para determinar o quanto cada carga não linear está impactando na distorção harmônica de tensão de um determinado barramento de interesse, segue-se o procedimento (MANITO et al., 2014):

Considere A como sendo a matriz que contém os padrões a serem aprendidos pela RNA, que correspondem à série temporal dos valores RMS das correntes harmônicas das barras j , na

frequência de interesse h , ou seja, $I_j^h(k)$, para $j = 1, 2, \dots, p$ e $h = 1, 2, 3, \dots, m$. Considere também C como o vetor de saída da RNA contendo a série temporal dos valores RMS da tensão harmônica no barramento de interesse V_x^h . Ambos os valores para I_j^h e V_x^h são series temporais de medidas para o período da campanha de medição. Para uma frequência harmônica específica h , A e C podem ser escritas como:

$$A = \begin{bmatrix} I_1^h(1) & I_1^h(2) & K & I_1^h(T) \\ I_2^h(1) & I_2^h(2) & K & I_2^h(T) \\ M & M & K & M \\ I_p^h(1) & I_p^h(2) & \Lambda & I_p^h(T) \end{bmatrix} \quad (43)$$

$$C = [V_x^h(1) \quad V_x^h(2) \quad \Lambda \quad V_x^h(T)] \quad (44)$$

Uma vez que a RNA esteja devidamente treinada, esta pode ser utilizada para estimar as tensões de saída V_x^h com certa precisão para cada vetor de entrada pertencente ao conjunto de dados de teste, ou seja, dados de medições de correntes que não foram usados durante a fase de treinamento da RNA.

Para determinar qual carga não linear está ocasionando maior impacto na distorção harmônica de tensão de um determinado barramento de interesse, a RNA treinada é submetida a novos vetores de entrada, os quais são modificados para incluir pequenas variações nas correntes harmônicas com relação aos valores medidos. O objetivo é determinar o quanto a tensão de saída da RNA é sensível às pequenas variações de corrente de cada carga não linear individualmente. Sendo assim, são criados n novos vetores de entrada, onde n é o número de entradas e cada entrada corresponde a uma serie temporal de corrente harmônica de uma respectiva carga.

Para cada novo vetor de entrada, mantendo as demais séries inalteradas em relação ao vetor de entrada original, uma das séries temporais de corrente harmônica é alterada pelo mesmo incremento ΔI ($0 < \Delta I \leq 1$), e a tensão harmônica correspondente é obtida pela RNA. Em outras palavras, este incremento corresponde a um fator percentual, em cada um dos novos vetores de entrada, uma das entradas terá todos os seus valores alterados multiplicados por este fator que varia de [1% a 10%] e que pode ser para mais ou para menos.

Com isso, um fator de sensibilidade pode ser definido para expressar a variabilidade da tensão harmônica com relação às correntes de carga individualmente. Matematicamente este procedimento pode ser expresso como:

$$I_{j_{new}}^h(k) = (\Delta I + 1)I_j^h(k) \quad (45)$$

$$A_{j_{new}} = \begin{bmatrix} I_1^h(k) \\ I_2^h(k) \\ \text{M} \\ I_{j_{new}}^h(k) \\ \text{M} \\ I_p^h(k) \end{bmatrix} \quad (46)$$

$$C_{j_{new}} = [V_{xj_{new}}^h(k)] \quad (47)$$

Onde,

$A_{j_{new}}$ – Nova matriz de entrada contendo todas as séries temporais de corrente inalteradas exceto a série temporal correspondente à $I_{j_{new}}^h(k)$;

$V_{xj_{new}}^h$ – Nova série temporal de tensão de saída estimada pela RNA treinada com $A_{j_{new}}$.

Comparando $V_{xj_{new}}^h(k)$ com $V_x^h(k)$, pode-se determinar a resposta $V_{xj_{new}}^h(k)$ que apresenta a maior variação com relação à série temporal original $V_x^h(k)$, e, conseqüentemente, a corrente da carga não linear associada com esta resposta pode ser classificada como a mais impactante com relação à distorção harmônica de tensão no ponto de interesse.

A fim de mensurar a contribuição relativa de cada carga não linear na distorção harmônica de tensão, um fator de impacto percentual $IF_{V_{ij}}^h(\%)$ é sugerido. Primeiramente, é necessário determinar o quanto a série temporal original V_{ih} varia quando todas as séries temporais de correntes de cada carga são variadas simultaneamente pelo mesmo incremento percentual $\Delta x\%$. Essa condição é expressa nas seguintes equações:

$$A_{new} = \begin{bmatrix} I_{1new}^h(k) \\ I_{2new}^h(k) \\ I_{jnew}^h(k) \\ \text{M} \\ I_{pnew}^h(k) \end{bmatrix} \quad (48)$$

$$C_{new} = [V_{xnew}^h] \quad (49)$$

Onde,

I_{jnew}^h - Nova série temporal de corrente para a variação percentual considerada ΔT ;

A_{new} - Nova matriz de entrada da RNA considerando a mesma variação percentual em todas as séries temporais de correntes harmônicas das cargas não lineares;

C_{new} - Nova série temporal de tensão de saída considerando a mesma variação percentual em todas as séries temporais de correntes harmônicas das cargas não lineares.

Para medir a contribuição relativa de cada carga não linear na distorção harmônica de tensão, um fator de impacto percentual IF_{VIj}^h (%) é sugerido. Estes fatores de impacto podem ser interpretados como valores percentuais relativos que são calculados para cada corrente individualmente, e é assumido que esses impactos somam 100%, ou seja:

$$IF_{VI1}^h(\%) + IF_{VI2}^h(\%) + \dots + IF_{VIp}^h(\%) = 100\% \quad (50)$$

Para obter o fator de impacto percentual, é necessário utilizar uma métrica para calcular o erro entre a série temporal de tensão estimada inicial e a série temporal de tensão estimada quando cada corrente é variada no tempo. A partir da formulação do Erro Absoluto Médio (MAE – Mean Absolut Error), pode se escrever que:

$$MAE_{jnew}^h = \frac{\sum_{k=1}^T |V_{xjnew}^h(k) - V_x^h(k)|}{T} \quad (51)$$

$$IF_{VIj}^h(\%) = \frac{MAE_{jnew}^h}{\sum_{j=1}^p MAE_{jnew}^h} \times 100\% \quad (52)$$

Ressalta-se que se podem utilizar outras métricas para calcular o erro, como o Erro o Erro Médio Quadrático (MSE - Mean Square Error) (53) ou a Raiz do Erro Médio Quadrático (RMSE - Root Mean Saquare Error) (54), ou outra métrica de interesse.

$$MSE_{jnew}^h = \frac{\sum_{k=1}^T (V_{xjnew}^h(k) - V_x^h(k))^2}{T} \quad (53)$$

$$RMSE_{jnew}^h = \sqrt{\frac{\sum_{k=1}^T (V_{xjnew}^h(k) - V_x^h(k))^2}{T}} \quad (54)$$

No Sistema elétrico real é extremamente difícil considerar simultaneamente todas as possíveis fontes harmônicas que impactam em um determinado barramento da rede elétrica. Sendo assim, é importante desenvolver metodologias de análises que possam determinar os impactos relativos na distorção de tensão da rede devido a essas fontes harmônicas que não são

explicitamente consideradas na análise. É uma prática nesses tipos de análises, considerar essas fontes não consideradas explicitamente como uma fonte de *background*.

A metodologia usando RNA, adotada neste trabalho, considera como entrada as correntes harmônicas e como saída as tensões harmônicas no barramento de interesse. Sendo assim, a estrutura da RNA está representando impedâncias equivalentes, relacionando as correntes harmônicas de entrada com as correspondentes tensões harmônicas de saída. Essa relação caracteriza um modelo causa-efeito, do qual é possível estabelecer uma estimativa da tensão de saída de background. Supondo que todas as correntes, incluindo as correntes de background, são inseridas como entrada da RNA, então a série temporal de tensão medida $V_x^h(k)$ seria estimada com precisão, com um pequeno erro de modelo da RNA. Portanto, uma estimativa da série temporal de tensão de background pode ser obtida pela diferença entre a série temporal de tensão medida original e estimativa desta obtida a partir da RNA usando apenas as correntes de interesse. Esta relação pode ser escrita como:

$$V_{BGx}^h(k) = V_x^h(k) - V_{Ex}^h(k) \quad (55)$$

Onde,

$V_{BGx}^h(k)$ - Tensão de background estimada no ponto de interesse;

$V_x^h(k)$ - Série temporal de tensão medida original;

$V_{Ex}^h(k)$ - Série temporal de tensão estimada usando apenas as correntes de interesse.

Para calcular o fator de impacto percentual da tensão de background é utilizado a norma do vetor (56) para a série temporal da tensão de background e da tensão do modelo gerado, mantendo a relação em (57).

$$\|x\| = \sqrt{\sum_{i=1}^N x_i^2} \quad (56)$$

$$\|V_{Ex}^h + V_{BGx}^h\| \geq \|V_{Ex}^h\| + \|V_{BGx}^h\| \quad (57)$$

Realizando a soma destes dois vetores norma, pode-se obter o fator de impacto do modelo criado e do background através de (58) e (59), tal que esta soma destes fatores representa cem por cento de impacto (60) e o fator de impacto do modelo representa os impactos calculados em (52), sendo necessário realizar ajustes fazendo a proporção dos impactos considerando agora o background conforme (61).

$$IF_{V_{Ex}}^h(\%) = \frac{\|V_{Ex}^h\| \cdot 100\%}{\|V_{Ex}^h\| + \|V_{BGx}^h\|} \quad (58)$$

$$IF_{V_{BGx}}^h(\%) = \frac{\|V_{BGx}^h\| \cdot 100\%}{\|V_{Ex}^h\| + \|V_{BGx}^h\|} \quad (59)$$

$$IF_{V_{Ex}}^h(\%) + IF_{V_{BGx}}^h(\%) = 100\% \quad (60)$$

$$IF_{VI}^h(\%) = IF_{VI}^h(\%) \cdot IF_{V_{Ex}}^h(\%) \quad (61)$$

Todos estes passos são descritos de forma sucinta no Algoritmo 2 mostrado a seguir.

Algorithm 2 Análise de Impacto Harmônico de Múltiplas Cargas

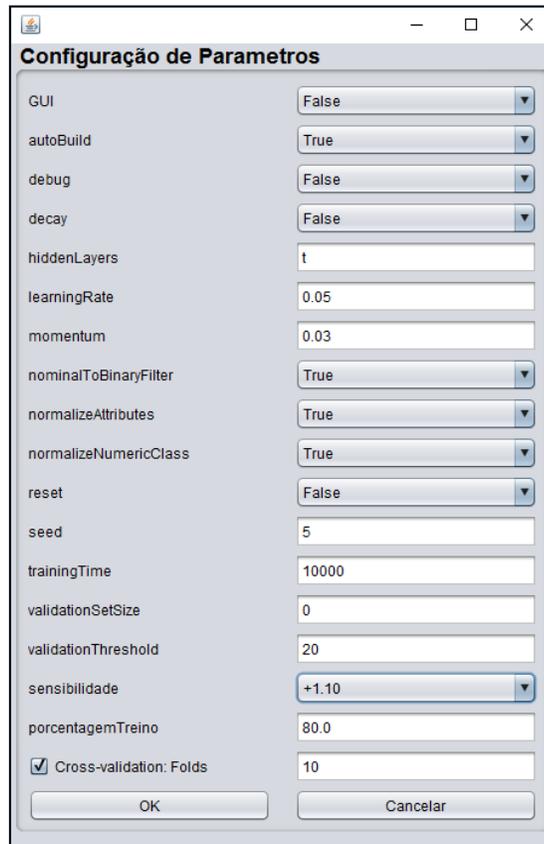
Início

- 1: $modelo \leftarrow (RNA, AR)$
- 2: $erro \leftarrow (MAE, MSE, RMSE)$
- 3: $\sigma \leftarrow [1 \text{ a } 10\%]$
- 4: $I_hModelo \leftarrow modelo.getEntradas()$
- 5: $V_EModelo \leftarrow modelo.estimar(I_hModelo)$
- 6: **para** $i \leftarrow 1$ **até** $size(I_hModelo)$ **faça**
- 7: $I_h \leftarrow I_hModelo.copia()$
- 8: $I_h[i] \leftarrow I_h[i] \cdot \sigma$
- 9: $V_E \leftarrow modelo.estimar(I_h)$
- 10: $erros[i] \leftarrow erro.calcular(V_EModelo, V_E)$
- 11: **fim para**
- 12: **para** $i \leftarrow 1$ **até** $size(I_hModelo)$ **faça**
- 13: $IF_{VI[i]}^h(\%) \leftarrow \frac{erros[i]}{\sum_{j=1}^p erros[j]} \cdot 100\%$
- 14: **fim para**
- 15: $V_{BGx}^h \leftarrow V_x^h(k) - V_{Ex}^h(k)$
- 16: $\|V_{Ex}^h\| \leftarrow \sqrt{\sum_{i=1}^N (V_{Ex[i]}^h)^2}$
- 17: $\|V_{BGx}^h\| \leftarrow \sqrt{\sum_{i=1}^N (V_{BGx[i]}^h)^2}$
- 18: $IF_{V_{Ex}}^h(\%) \leftarrow \frac{\|V_{Ex}^h\| \cdot 100\%}{\|V_{Ex}^h\| + \|V_{BGx}^h\|}$
- 19: $IF_{V_{BGx}}^h(\%) \leftarrow \frac{\|V_{BGx}^h\| \cdot 100\%}{\|V_{Ex}^h\| + \|V_{BGx}^h\|}$
- 20: **para** $i \leftarrow 1$ **até** $size(I_hModelo)$ **faça**
- 21: $IF_{VI[i]}^h(\%) \leftarrow IF_{VI[i]}^h(\%) \cdot IF_{V_{Ex}}^h(\%)$
- 22: **fim para**

Fim

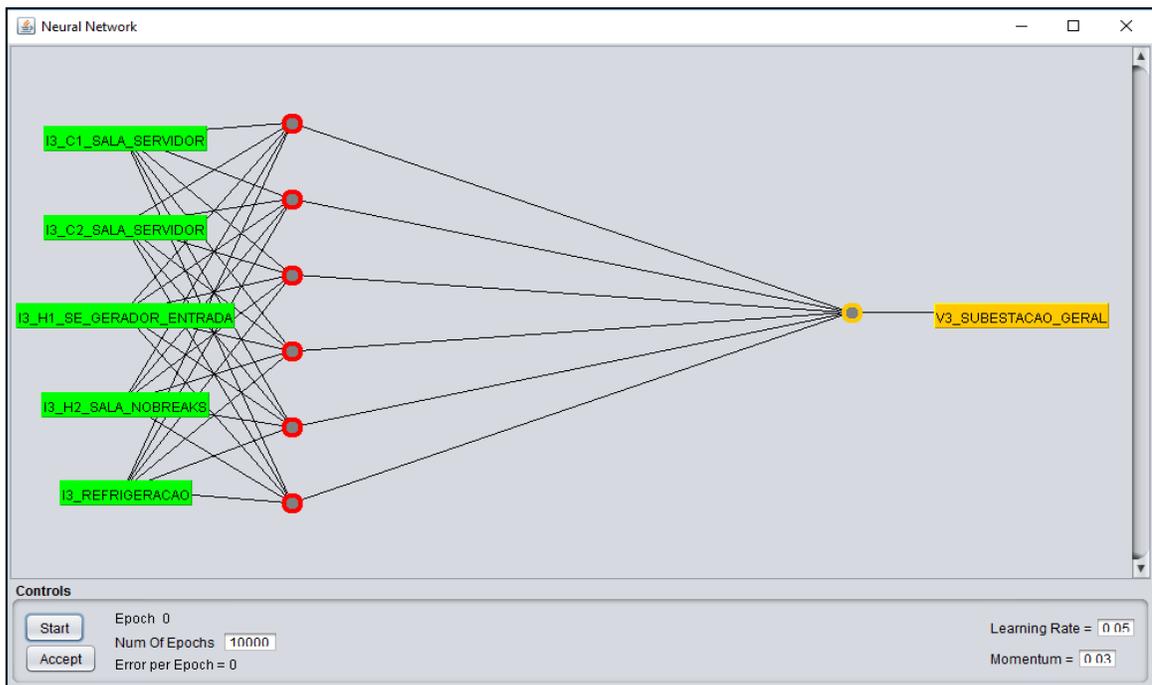
A RNA já vem previamente com uma configuração definida, a Figura (48) mostra a tela contendo estes parâmetros. Ressalta-se que o parâmetro sensibilidade corresponde ao incremento $(\Delta + 1)$ definido para alterar as séries temporais de corrente harmônica em 10 (dez) por cento. A Figura (49) mostra uma das RNAs que são criadas para cada fase conforme os parâmetros configurados.

Figura 48. Tela de Configuração de Parâmetros de RNA.



Fonte: O Autor.

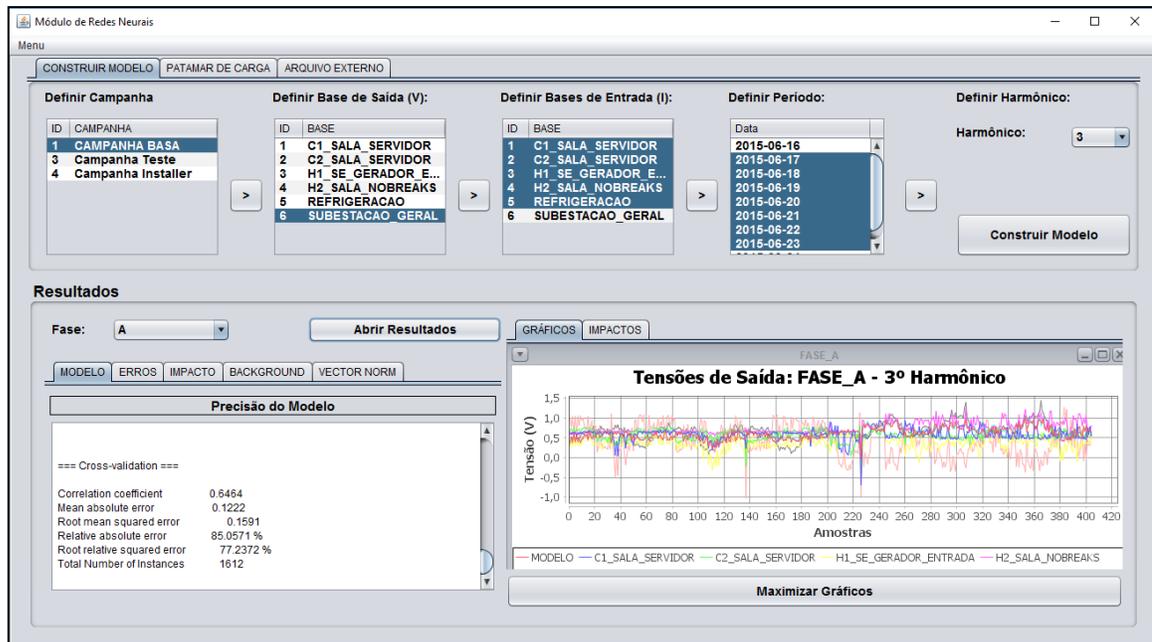
Figura 49. Topologia da RNA.



Fonte: O Autor.

Os resultados são apresentados na parte inferior da tela do módulo de RNA, conforme na Figura (50).

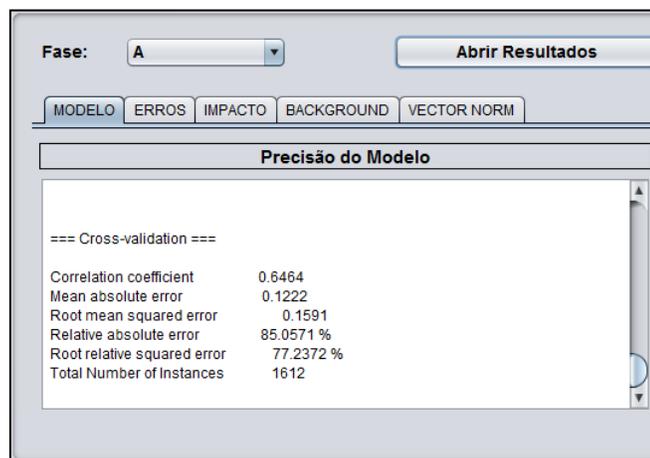
Figura 50. Tela Principal do Módulo de Redes Neurais Artificiais.



Fonte: O Autor.

Os resultados são apresentados em cinco abas a esquerda (*modelo*, *erros*, *impacto*, *background* e *vector norm*) e duas à direita (*gráficos* e *impactos*). Na aba *modelo* são exibidos os dados do modelo de RNA criado pelo *weka*, apenas da fase selecionada conforme Figura (51).

Figura 51. Resultados RNA: Aba Modelo.



Fonte: O Autor.

Na aba *erros* estão contidos o MAE, MSE e RMSE obtidos para cada entrada do modelo. A aba *impacto* contem as parcelas percentuais de impacto de cada entrada sem levar em conta o

background. Na aba *background* estão contidas as parcelas de impacto das entradas e do background. A Figura (52) apresenta a aba de background utilizando o MAE.

Figura 52. Resultados RNA: Aba Background - Tabela de MAE.

Base de Entrada	Fase A	Fase B	Fase C
C1_SALA_SER...	12,566	21,902	06,255
C2_SALA_SER...	16,940	33,232	08,234
H1_SE_GERA...	26,084	24,221	73,910
H2_SALA_NO...	22,562	14,584	05,568
REFRIGERACAO	02,742	01,570	00,999
BACKGROUND	19,106	04,491	05,034

Fonte: O Autor.

A aba *vector norm* contém as normas calculadas em função das séries temporais da tensão fonte e da tensão de background, e as parcelas de impacto referentes ao modelo e ao background, que podem ser vistas conforme Figura (53).

Figura 53. Resultados RNA: Aba Vector Norm - Impacto.

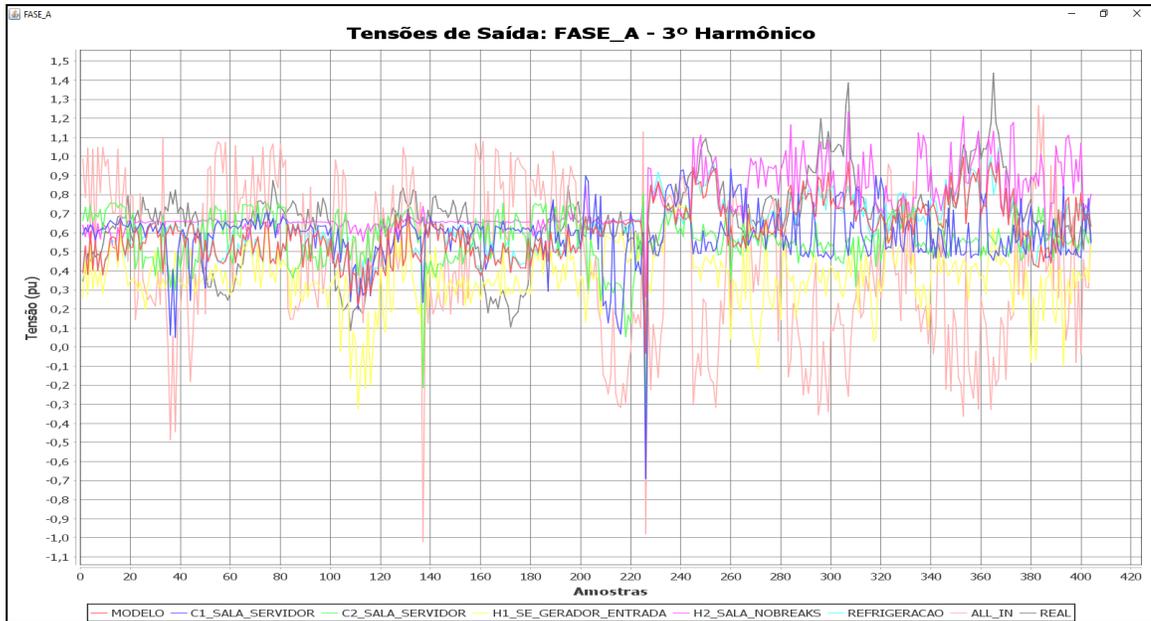
Base de Entrada	Fase A	Fase B	Fase C
VE	80,894	95,509	94,966
VBG	19,106	04,491	05,034

Fonte: O Autor.

A aba *gráficos* contém os gráficos de representação dos modelos de RNA criados para as três fases, os quais, além das curvas de tensão estimada (*modelo*) e medida (*real*), contém as curvas de tensão estimada dos vetores de entrada cuja uma das correntes foi alterada pelo fator de sensibilidade, e também a curva *all_in*, que representa o somatório dos erros de todas estas curvas alteradas. Por fim, a aba *impactos* contém os gráficos contendo os impactos pontuais das *fontes* e

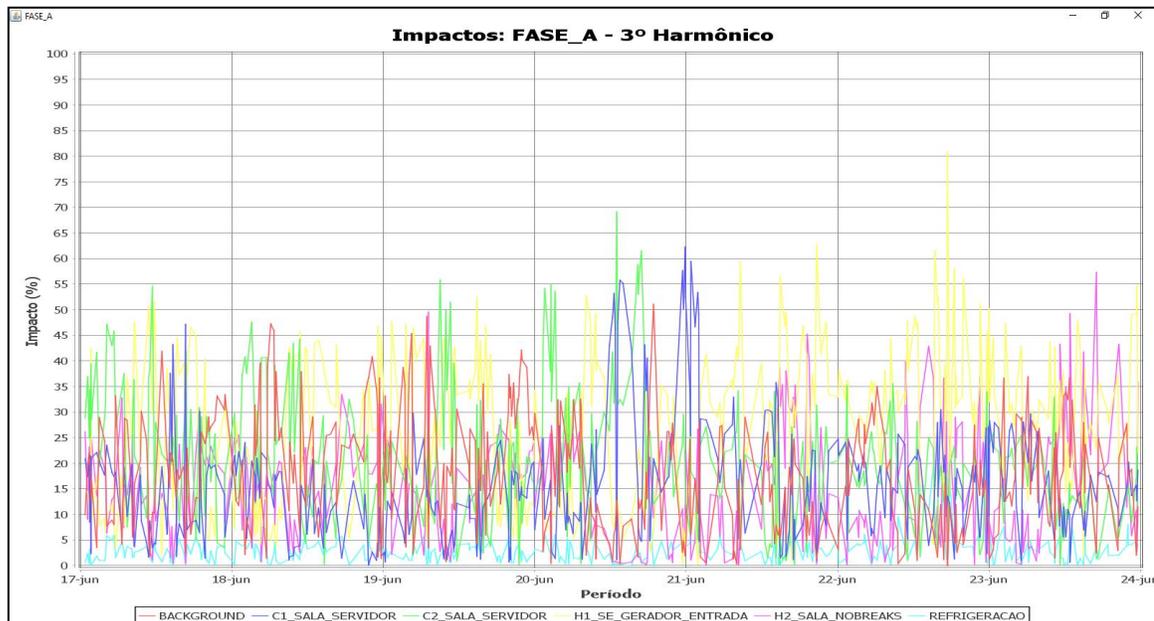
de *background*. As Figuras (54) e (55) apresentam um modelo destes gráficos de modelo e de impacto respectivamente.

Figura 54. Gráfico das séries temporais das tensões estimadas pelo modelo de RNA.



Fonte: O Autor.

Figura 55. Gráfico dos fatores de impacto do modelo de RNA.

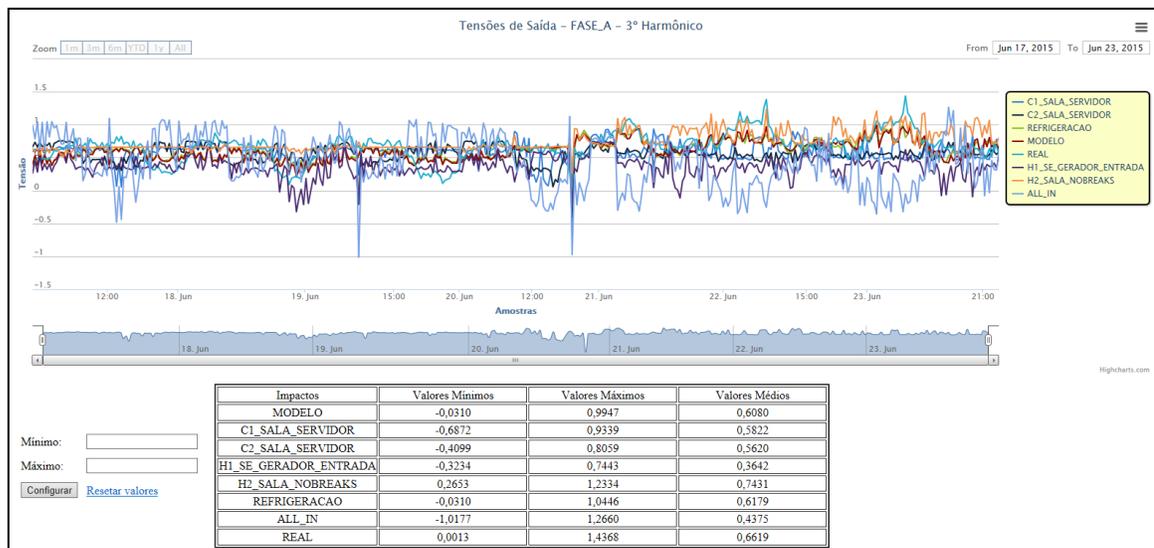


Fonte: O Autor.

Ressalta-se que estes gráficos foram podem ser configurados da mesma forma que os de RL, com a ressalva que ambos os gráficos são gerados no formato *html* e salvos na pasta

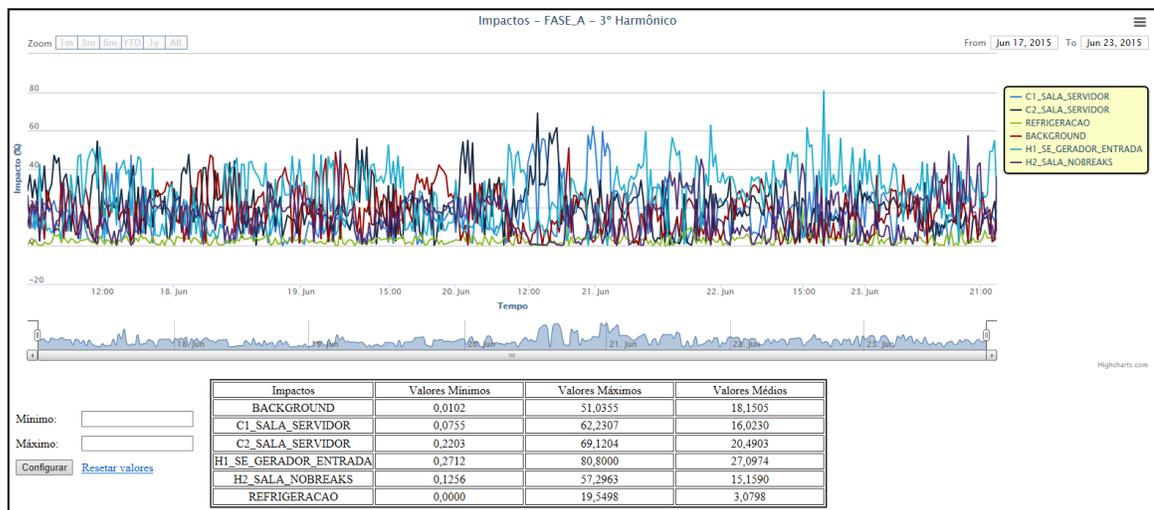
REDES_NEURAIIS que é criada dentro da pasta da base que foi utilizada como referência para o estudo, onde também é salvo os relatórios dos resultados que foram gerados e os arquivos que contendo os dados que foram usados para gerar os modelos. As Figuras (56) e (57) apresentam estes mesmo gráficos gerados em *html*, permitindo ao usuário a vantagem de poder remover as curvas que desejar e dispor de uma tabela contendo valores mínimos, máximos e médios, para que possa realizar uma melhor análise.

Figura 56. Gráfico das séries temporais das tensões estimadas pelo modelo de RNA em html.



Fonte: O Autor.

Figura 57. Gráfico das séries temporais das tensões estimadas pelo modelo de RNA em html.



Fonte: O Autor.

4.5 MÓDULO DE ÁRVORES DE REGRESSÃO

Assim como o módulo de RNA, este módulo avalia os impactos causados por vários pontos de medição em relação a um ponto de referência, com a ressalva que os resultados se mostram melhores utilizando períodos menores para o treinamento. Para o processo de estimação pretende-se, com a utilização de uma árvore de regressão, estimar a tensão harmônica na barra de interesse a partir do conhecimento adquirido no processo de aprendizagem, sendo utilizados, no momento de estimação, dados de entrada desconhecidos para o modelo desenvolvido, da mesma forma que os modelos de RNA.

Uma árvore de regressão (um tipo de técnica de aprendizado supervisionado) é um modelo de predição, o qual é estruturado em árvore, para um dado objetivo numérico caracterizado por um vetor de atributos também numéricos. A árvore de regressão é uma categoria da árvore de decisão, portanto, ela é treinada e testada usando o algoritmo chamado de Classification and Regression Tree (CART). Uma árvore de regressão possui estrutura binária com dois tipos de nós, o nó interno com dois nós filhos e o nó terminal sem qualquer nó filho. Cada nó terminal fornece um resultado de predição. Para propósitos de treinamento de uma boa árvore de regressão, os conjuntos de dados para treino e teste devem possuir o mesmo formato.

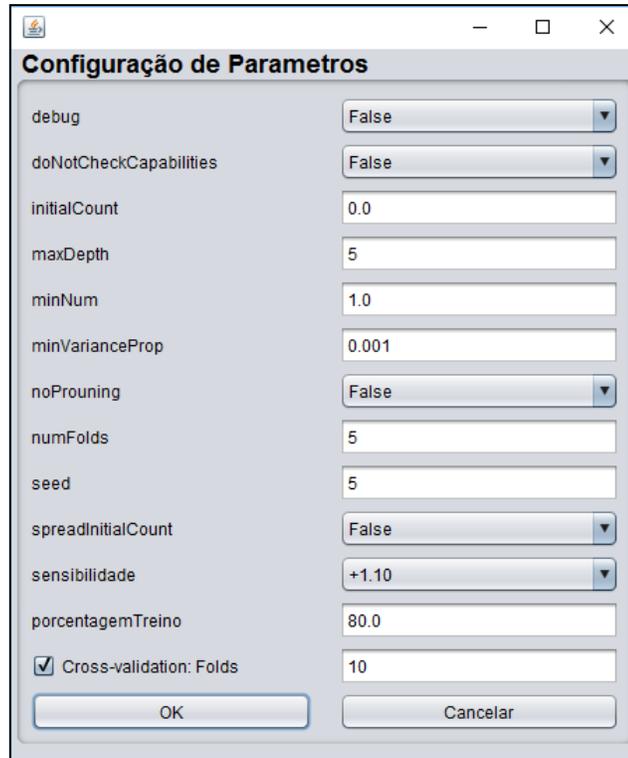
O conjunto de dados para treino é primeiramente usado para treinar uma árvore máxima pela partição recursiva de um nó pai em dois nós filhos até que a partição não melhore a precisão. Essa árvore máxima é, então, podada, passo a passo, até gerar uma série de árvores menores, com o objetivo de atenuar o problema do overfitting (diminuição da capacidade de generalização). Maiores detalhes sobre partição, critério de parada e algoritmos de poda de árvores de decisão são fornecidas em (BREIMAN et al., 1984).

O desempenho dessas árvores menores na etapa de teste é medido pelo erro quadrático médio para árvores de regressão. A árvore ótima é, portanto, definida como o modelo de árvore que apresenta menor erro de predição na etapa de teste. A vantagem do método de árvore de regressão consiste na sua estrutura simplificada do modelo, a qual facilita a análise devido à redução do número de atributos de entrada. O método é particularmente atraente porque a árvore de regressão usa um modelo que torna os resultados fáceis de interpretar e replicar.

Este módulo segue exatamente o mesmo raciocínio no módulo de RNA para calcular os fatores de impacto e, portanto, também utiliza o Algoritmo 2. A árvore de regressão já vem previamente com uma configuração definida. A Figura (58) mostra esta tela contendo os

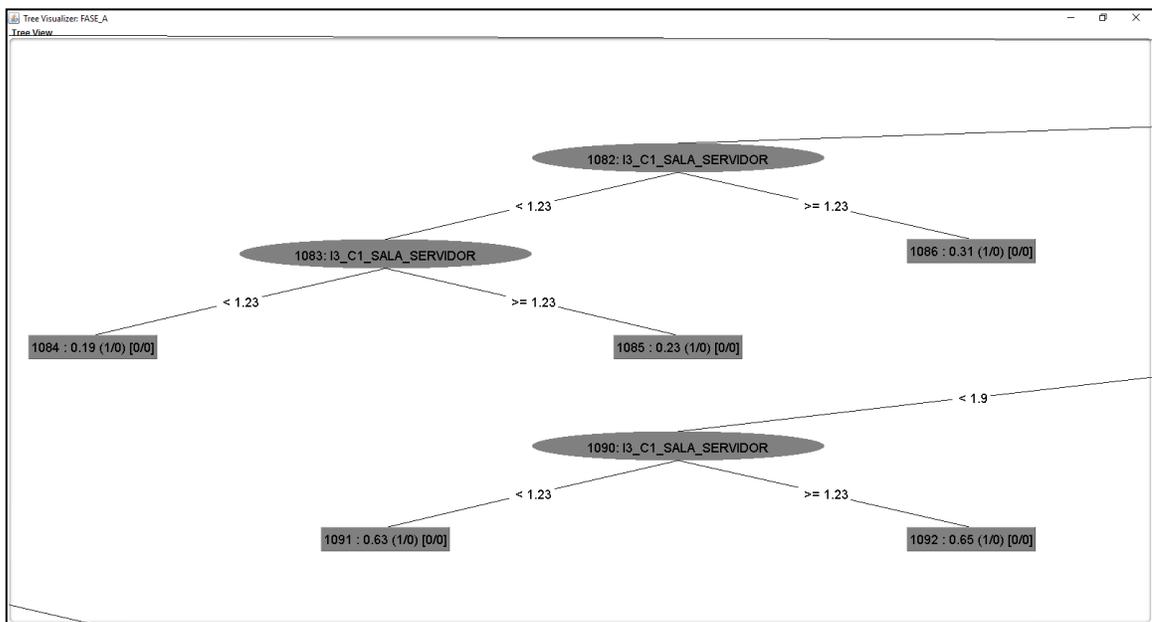
parâmetros previamente definido, e a Figura (59) mostra parte de uma das árvores que são criadas para cada fase conforme os parâmetros configurados.

Figura 58. Tela de Configuração de Parâmetros de AR.



Fonte: O Autor.

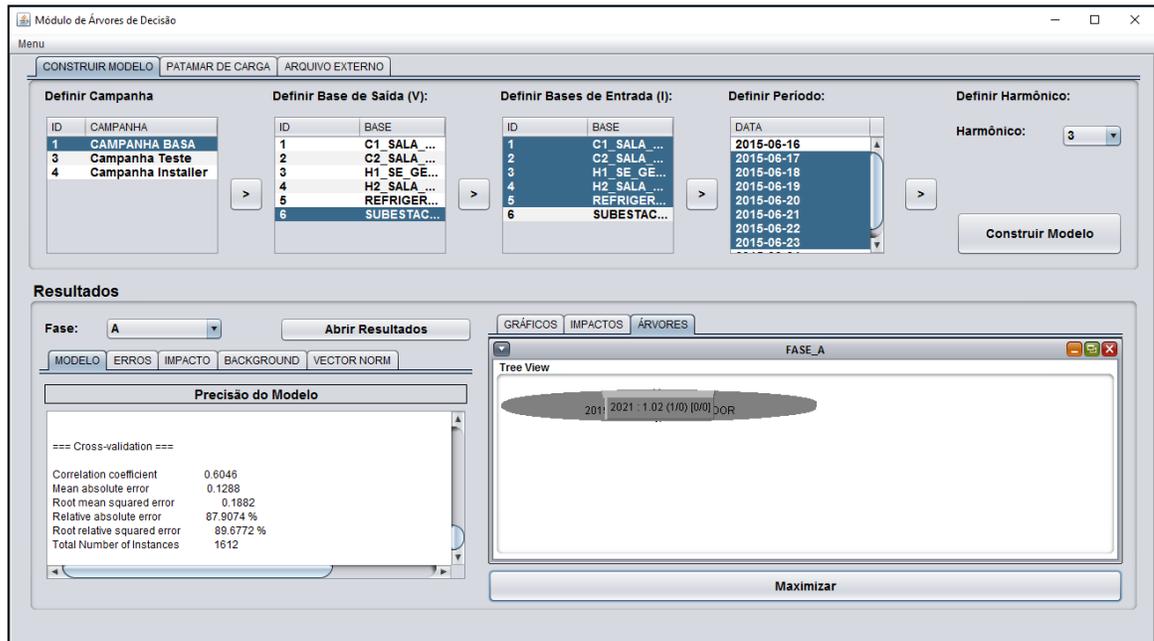
Figura 59. Exemplo contendo parte do ramo direito da AR.



Fonte: O Autor.

Os resultados são apresentados na parte inferior da tela do módulo de regressão, conforme na Figura (60). Este módulo fornece os mesmos tipos de resultados da RNA, com a ressalva de que ao invés de uma rede, fornece uma árvore contida na aba *árvores*, e suas respectivas regras contidas na aba *modelo*.

Figura 60. Tela Principal do Módulo de Árvores de Regressão.



Fonte: O Autor.

4.6 CONSIDERAÇÕES FINAIS

Neste capítulo foram apresentados os módulos de IC presentes no SISQEE, que têm como finalidade realizar a análise de contribuição de impacto harmônico, mostrando suas respectivas telas gráficas e a metodologia aplicada em cada um destes módulos, ressaltando as características de cada técnica e os tipos de resultados que cada uma apresenta.

O módulo de Regressão Linear tem a limitação de representar a influência das cargas de background, analisando o impacto harmônico devido a uma carga específica de cada vez, sempre mostrando o valor de R^2 para avaliar a confiabilidade dos resultados. O módulo de Rede Neural Artificial pode realizar análises de contribuição harmônica de múltiplas cargas considerando o background, apresentando gráficos de impacto e de tensões gerados pelos modelos criados, bem como as tabelas contendo as parcelas percentuais de cada ponto analisado. O módulo de Árvore de Regressão, além dos mesmos tipos de resultados gerados pelo módulo de RNA, fornece uma própria estrutura de classificação dividindo o conjunto de dados em intervalos menor e igual ou

maior e igual a valores limites obtidos no processo de classificação, exibidos em uma árvore na qual podem ser observadas hierarquicamente as cargas que tem maior influência no impacto harmônico.

Dessa forma, as três técnicas apresentam aspectos que podem se complementar, ampliando o espectro da aplicação das mesmas. No capítulo a seguir, serão apresentados os resultados de um estudo de caso realizado a partir de uma campanha de medição, utilizando os módulos desenvolvidos no SISQEE.

5 ESTUDOS DE CASO

5.1 CONSIDERAÇÕES INICIAIS

Para a realização dos estudos de casos foi utilizada uma das campanhas de medições realizadas no polo industrial de Manaus durante o andamento do projeto P&D. A Tabela 11 apresenta a descrição desta campanha, informando o período das medições de cada ponto analisado.

Tabela 11. Descrição da Campanha de Medição.

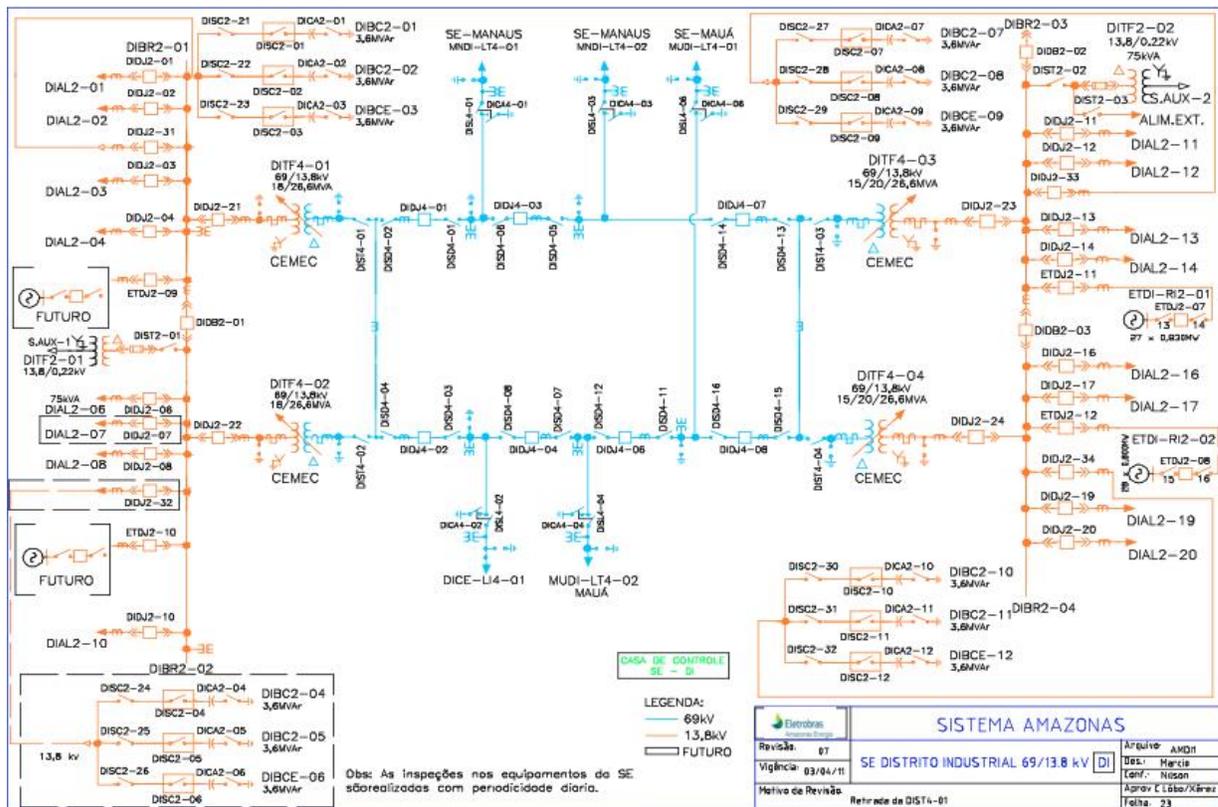
Período	Pontos Medidos	Tensão Base (V)
23/05/2014 a 29/05/2014	Secundário do transformador DITF4-04 da SE DI	13,8 k
23/05/2014 a 29/05/2014	Alimentador DIAL2-16 da SE DI	13,8 k
23/05/2014 a 29/05/2014	Alimentador DIAL2-17 da SE DI	13,8 k
22/05/2014 a 24/05/2014	Alimentador DIAL2-19 da SE DI	13,8 k
23/05/2014 a 29/05/2014	Alimentador DIAL2-20 da SE DI	13,8 k

A subestação (SE) distrito industrial é uma subestação que fornece energia elétrica a uma parte das indústrias localizadas no distrito industrial de Manaus. Essa subestação apresenta um arranjo em anel e é alimentada pelas subestações Manaus I (MN) por meio das linhas denominadas MNDI-LT4-01 e MNDI-LT4-02, e Mauá (MU) através das linhas MUDI-LT4-01e MUDI-LT4-02 no nível 69 kV.

Por meio do alimentador DICE-LI4-01, a SE DI (Sub Estação Distrito Industrial) alimenta 9 indústrias em 69 kV: Moto Honda, LG electronics, Videolar, Pastore, Bic Amazônia (Tecnocerio S/A), Showa, Reman (refinaria da Petrobras), Yamaha e Compaz. Além disso, a SE DI é composta de quatro transformadores de 69 Δ /13,8Y kV de 26,6 MVA e nove bancos de capacitores de 3,6 MVAr.

A partir dos quatro barramentos de 13,8 kV, partem 13 alimentadores que são responsáveis por fornecer energia elétrica a aproximadamente 81 consumidores industriais. A Figura (61) ilustra o diagrama unifilar a SE DI.

Figura 61. Diagrama Unifilar da Subestação Distrito Industrial.

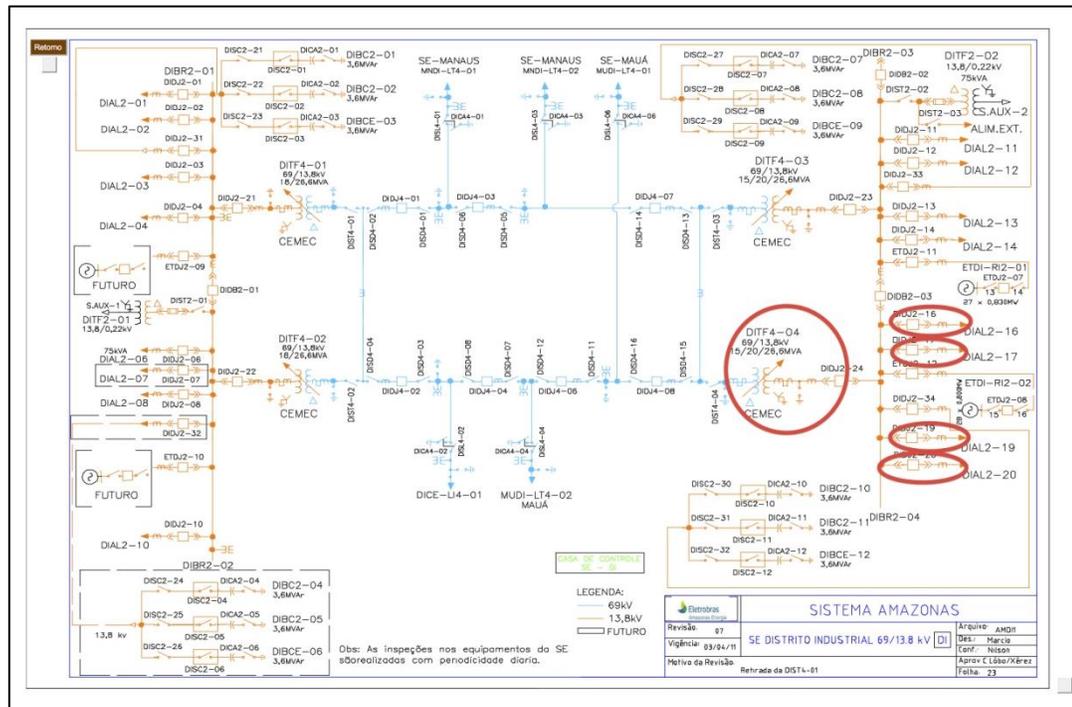


Fonte: Projeto P&D.

5.2 ANÁLISE DE QEE

Os resultados desta subseção bem como as das técnicas de IC foram obtidos utilizando o SISQEE e são oriundos de uma campanha de medição, cuja rede elétrica é de 13,8 kV conforme pode ser observado no diagrama unifilar mostrado na Figura (62). A obtenção dos dados se deu mediante a instalação do analisador de qualidade da energia PW3198 da HIOKI. A fim de não estender muito este tópico é apresentada apenas a análise das medições coletadas no lado de baixa tensão do transformador DITF4-04 da SE DI, o qual teve seus dados registrados cada um minuto. O período de análise é de 23 a 29 de maio de 2014.

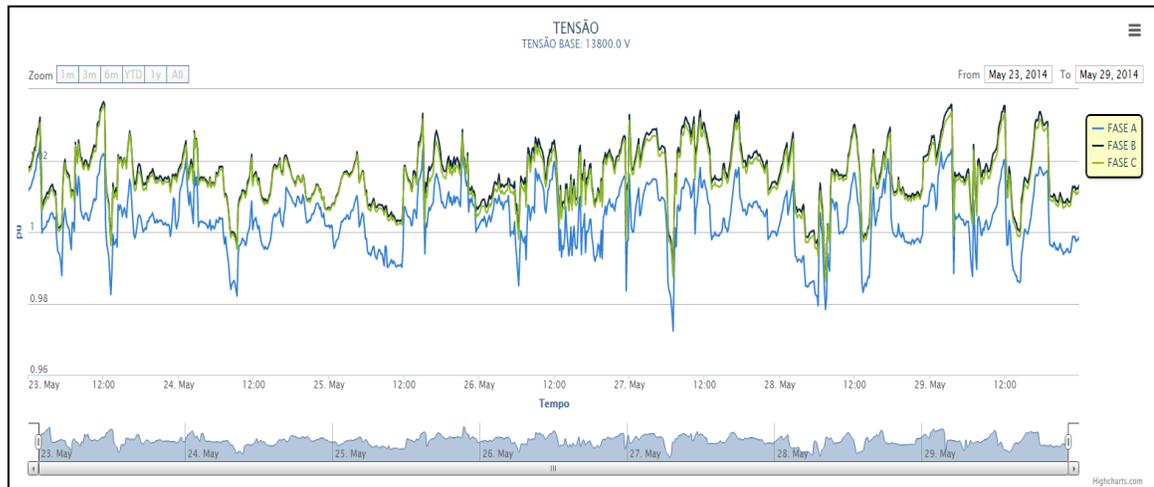
Figura 62. Diagrama Unifilar da Campanha de Medição 6.



Fonte: Projeto P&D.

Analisando os registros da Figura (63), que ilustram os perfis das tensões em pu nas fases A, B e C do transformador DITF4-04 num período de 7 (sete) dias com dados agrupados com intervalo de 10 (dez) minutos totalizando 1008 amostras, verificou-se que o valor da tensão rms não violou o nível adequado ($0,93 \text{ TR} \leq \text{TL} \leq 1,05 \text{ TR}$), conforme estabelecidos no módulo 8 do PRODIST mostrados na Tabela 4, durante todo período de medição. A Tabela 12 contém os valores máximos, mínimos e médios de tensões registrados neste período e as Tabelas 13-14 contém respectivamente os valores *nlp* e *nlc* das equações (1) e (2) usados para calcular o DRP e DRC, os quais foram nulos estando, portanto, dentro dos limites de 3% e 0,5% estipulados pelo PRODIST.

Figura 63. Tensão RMS medida no lado de baixa tensão do transformador DITF4-04 da SE DI.



Fonte: O Autor.

Tabela 12. Valores máximos, mínimos e médios de tensão do transformador DITF4-04 da SE DI.

	Valores Mínimos (pu)	Valores Máximos (pu)	Valores Médios (pu)
FASE A	0,9408	1,0309	1,0036
FASE B	0,9561	1,0478	1,0160
FASE C	0,9550	1,0460	1,0147

Tabela 13. DRP transformador DITF4-04 da SE DI.

	NLP	Total de Amostras
FASE A	0	1008
FASE B	0	1008
FASE C	0	1008
DRP	0	

Tabela 14. DRC transformador DITF4-04 da SE DI.

	NLC	Total de Amostras
FASE A	0	1008
FASE B	0	1008
FASE C	0	1008
DRC	0	

Analisando a Figura (64), que mostra a taxa de distorção harmônica total de tensão (DHTv) das três fases do transformador DITF4-04 da SE DI, observou-se que os valores de DHT de tensão atingiram um valor máximo igual 3,28% na fase A, 3,34% na fase B e 3,38% na fase C, estando, portanto, em conformidade com o limite estabelecido pelo módulo 8 do PRODIST, que

é de 8% para esse nível de tensão. A Tabela 15 contém os valores máximos, mínimos e médios de distorções harmônicas totais de tensão registrados neste período.

Figura 64. Taxa de distorção harmônica total de tensão medida do transformador DITF4-04 da SE DI.



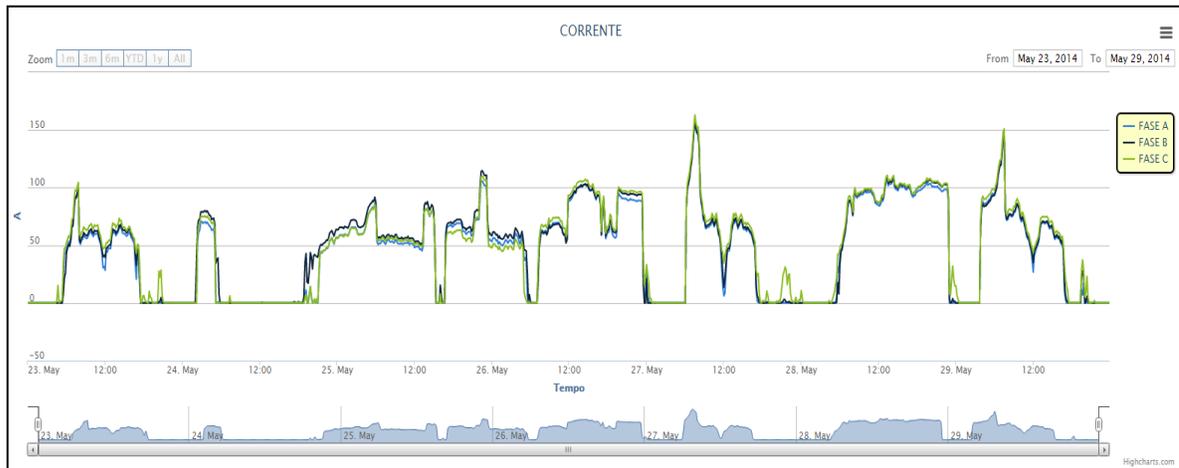
Fonte: O Autor.

Tabela 15. Valores máximos, mínimos, médios e percentis de DHTv medidos no lado de baixa tensão do transformador DITF4-04 da SE DI

	Valores Mínimos (%)	Valores Máximos (%)	Valores Médios (%)	Percentil 95%
FASE A	1,2000	3,1660	1,8936	2,7440
FASE B	1,0150	3,1980	1,8217	2,8345
FASE C	0,9360	3,2480	1,8601	2,8965

A Figura (65) apresenta um ciclo semanal do consumo do conjunto de cargas que são alimentadas pela linha DIAL2-16. De acordo com essa figura, verificou-se que os dias de semana não apresentaram grandes diferenças em relação ao final de semana. O máximo valor de corrente registrado no lado de baixa tensão do transformador DITF4-04 durante o período considerado foi igual a 158,1A na fase A, 158A na fase B e 163A na fase C. A Tabela 16 contém os valores registrados dos máximos, mínimos e médios da corrente neste período.

Figura 65. Corrente RMS medida no lado de baixa tensão do transformador DITF4-04 da SE DI.



Fonte: O Autor.

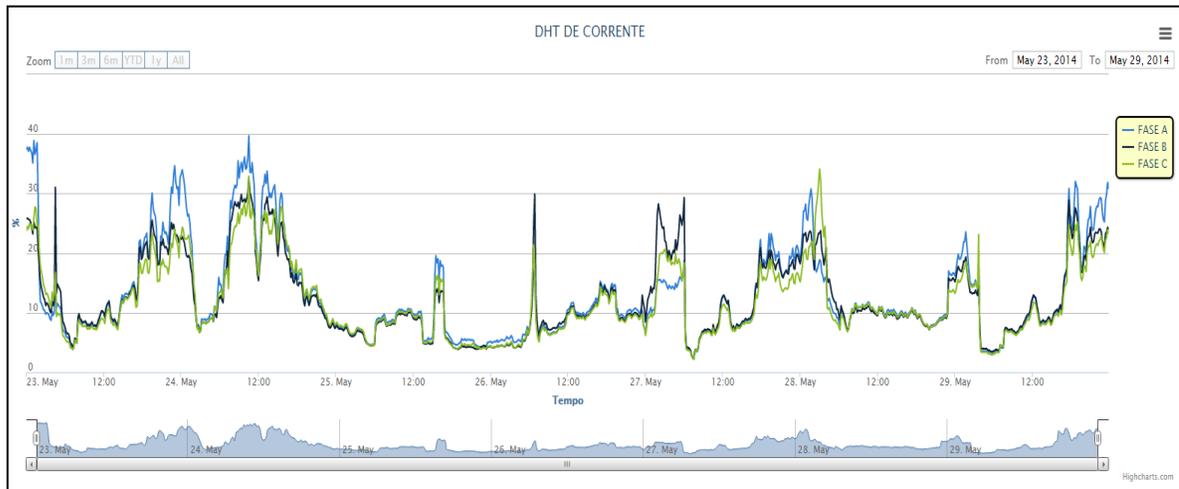
Tabela 16. Valores máximos, mínimos e médios de corrente medidos no lado de baixa tensão do transformador DITF4-04 da SE DI.

	Valores Mínimos (A)	Valores Máximos (A)	Valores Médios (A)
FASE A	0,0000	158,1000	44,0987
FASE B	0,0000	158,0000	46,5391
FASE C	0,0000	163,0000	47,0850

Pode-se identificar a corrente assumindo valores nulos muitas vezes, durante o ciclo semanal de consumo. Na verdade o valor nulo se deve a sensibilidade da garra de corrente que não apresenta precisão adequada para valores muito baixos de corrente. Pode-se concluir que estas correntes não são nulas observando os gráficos da Figura (66), o qual representa a DHT_i nesses mesmos intervalos, como sendo diferente de zero.

Analisando a Figura (66), que mostra a taxa de distorção harmônica total de corrente (DHT_i) das três fases do transformador DITF4-04 de 13,8 kV da SE DI, observa-se que os valores de DHT de corrente se encontram em níveis bem elevados, principalmente durante a madrugada, atingindo por muitas vezes valores acima de 20%. A Tabela 17 contém os valores máximos, mínimos, médios e percentis das distorções harmônicas totais de corrente, registrados neste período.

Figura 66. Taxa de distorção harmônica total de corrente medida no transformador DITF4-04 da SE DI.



Fonte: O Autor.

Tabela 17. Valores máximos, mínimos, médios e percentis de DHTi medidos no lado de baixa tensão do transformador DITF4-04 da SE DI.

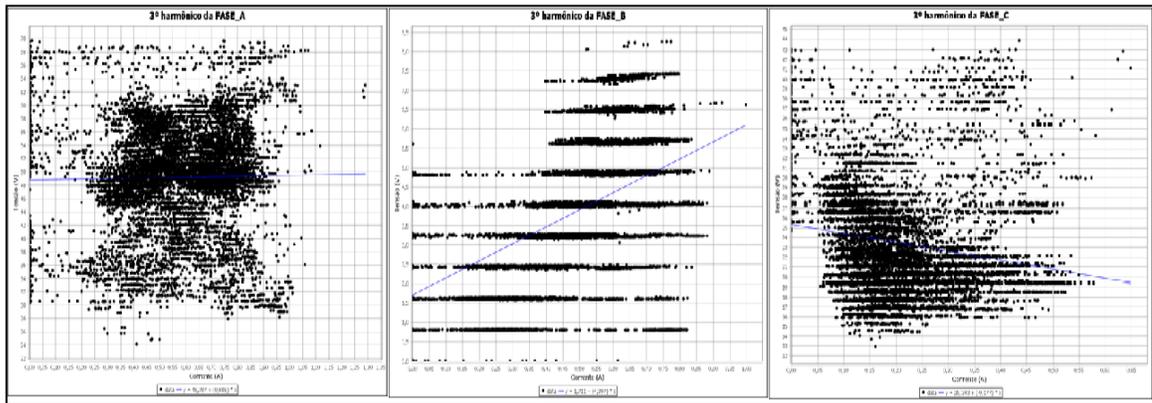
	Valores Mínimos (%)	Valores Máximos (%)	Valores Médios (%)	Percentil 95%
FASE A	2,1850	39,7070	13,1186	30,9524
FASE B	2,3130	32,6960	12,4332	25,5567
FASE C	2,2080	34,1120	11,8681	24,4777

5.3 REGRESSÃO LINEAR

A seguir serão apresentados os resultados obtidos utilizando a metodologia baseada em regressão linear, verificando a contribuição harmônica dos alimentadores DIAL2-16, DIAL2-17, DIAL2-19 e DIAL2-20 em relação ao background tendo como referência o transformador DITF4-04 da SE DI, utilizando dados do 3º harmônico.

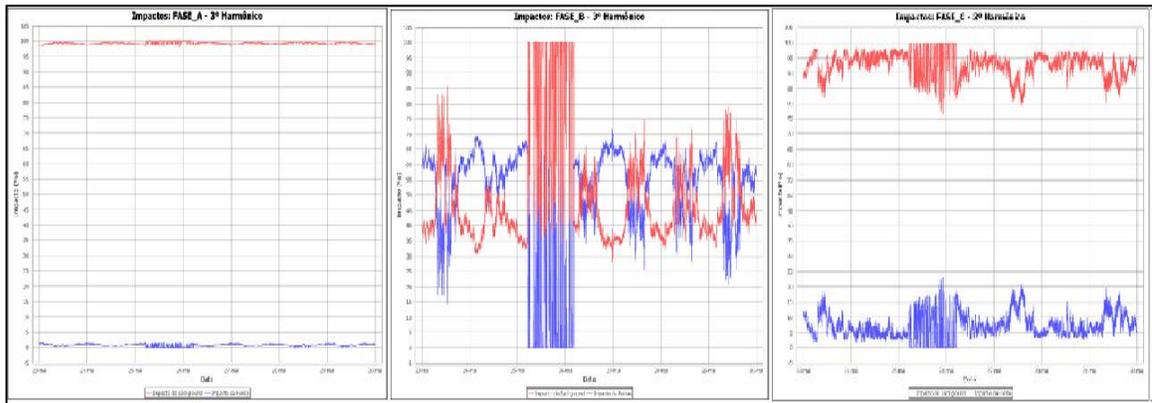
A Figura (67) mostra os modelos de regressão linear criados para as três fases contendo as equações da reta, e a Figuras (68) mostra os impactos pontuais considerando o alimentador DIAL2-16 (curva azul) em relação ao background (curva vermelha). O período de análise é de 23-05-2014 00h00min a 29-05-2014 23h59min.

Figura 67. Modelo de RL do 3º harmônico para as três fases da fonte DIAL2-16 em relação a transformador DITF4-04 da SE DI.



Fonte: O Autor.

Figura 68. Fator de Impacto Pontual do 3º harmônico para as três fases da fonte DIAL2-16 em relação a transformador DITF4-04 da SE DI.



Fonte: O Autor.

A Tabela 18 apresenta os fatores de impacto percentuais sobre o 3º harmônico de tensão na subestação do transformador devido a quatro alimentadores individualmente e suas respectivas fontes de background. É observado que em quase todos os casos a parcela percentual de impacto do background é sistematicamente maior do que as parcelas de impacto devido a si mesmas. O único caso em que a parcela de impacto da própria fonte foi maior do que a de background foi na fase B do alimentador DIAL2-19, ocorrendo da parcela de background ser negativa, levando a hipótese de que nesta fase background atuou como atenuador do impacto harmônico na barra do transformador.

Tabela 18. Comparação dos impactos harmônicos na distorção de tensão devido às fontes de corrente harmônica e de background usando modelo de RL construído para o estudo de caso.

FONTE		FASE A	FASE B	FASE C
DIAL2-16	IF _{VT} (%)	0,759	49,771	7,138
	IF _{VBG} (%)	99,241	50,229	92,862
	R²	0 / Nula	0,28 / Moderada	0,032 / Baixa
DIAL2-17	IF _{VT} (%)	9,018	35,028	23,012
	IF _{VBG} (%)	90,982	64,972	76,988
	R²	0,034 / Baixa	0,025 / Baixa	0,414 / Alta
DIAL2-19	IF _{VT} (%)	18,901	103,301	17,82
	IF _{VBG} (%)	81,099	-3,301	82,18
	R²	0,154 / Moderada	0,388 / Alta	0,215 / Moderada
DIAL2-20	IF _{VT} (%)	07,312	32,572	24,188
	IF _{VBG} (%)	92,688	67,428	75,812
	R²	00,041 / Baixa	00,126 / Moderada	00,427 / Alta

De acordo com os resultados da Tabela 18, as fontes mais impactantes de 3º harmônico comparados com os outros alimentadores analisados, foram DIAL2-17 e DIAL2-19 para a fase A, DIAL2-16 e DIAL2-20 para fase B, e DIAL2-17 e DIAL2-20 para a fase C. Isto, levando em consideração que são os resultados mais confiáveis, uma vez que valores de R^2 foram os maiores obtidos.

5.4 REDES NEURAS ARTIFICIAIS

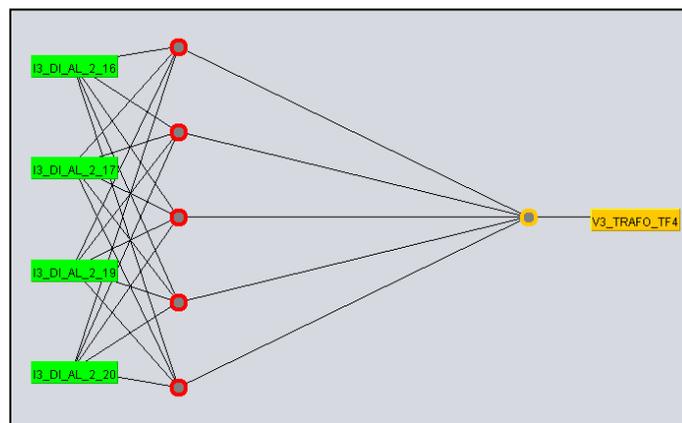
A seguir são apresentados os resultados referentes à análise da contribuição das correntes harmônicas de 3ª ordem dos alimentadores DIAL2-16, DIAL2-17, DIAL2-19 e DIAL2-20, utilizando dados de tensão harmônica também de 3ª ordem da barra localizada no lado de baixa tensão do transformador DITF4-04. A RNA utilizada é a perceptron de múltiplas camadas (MLP - Multi Layer Perceptron).

Devido a medição do ponto DIAL2-19 possuir período menor em relação aos demais, o período de análise é de 22-05-2014 14:18:00 a 24-05-2014 07:47:00. Os parâmetros utilizados para criação do modelo de RNA são os definidos como os padrões pelo SISQEE, os quais foram descritos anteriormente na Figura (48), a porcentagem de dados utilizados para treino e teste são

de 80% e 20% respectivamente, sendo 1993 amostras para treino e 499 amostras para teste, totalizando 2492 amostras.

A fim de evitar o *overfitting* foi utilizada validação cruzada dividindo a base de treino em 10 *folds*. O fator de sensibilidade ΔT utilizado foi 1.10, o qual corresponde a um aumento das entradas de correntes harmônicas em 10%. É adotada a mesma topologia para cada uma das RNAs criadas, a qual consiste em uma camada escondida contendo $n + 1$ neurônios (configuração padrão do software), onde n é o número de entradas, resultando em cinco neurônios, conforme Figura (69).

Figura 69. Topologia das RNAs do Estudo de Caso.



Fonte: O Autor.

Os resultados a seguir foram obtidos utilizando o MAE como métrica de avaliação de erro das series temporais de corrente harmônicas geradas com o fator de sensibilidade. A Tabela 19 contém os fatores de impactos dos barramentos dos alimentadores considerando o background.

Tabela 19. Fator de Impacto (%) com Background do Estudo de Caso.

FONTE	FASE A	FASE B	FASE C
DIAL2-16	21,377	11,310	07,313
DIAL2-17	21,512	15,116	10,549
DIAL2-19	13,415	26,514	21,130
DIAL2-20	33,493	32,490	49,756
BACKGROUND	10,204	14,570	11,252

Como pode ser visto na Tabela 19, considerando o período de medição, o alimentador que apresentou maior influência na distorção de tensão do barramento estudado em questão foi DIAL2-20 em todas as fases, alcançando um valor igual a 33,493% na fase A, 32,490% na fase B e 49,756% na fase C. Adicionalmente, pode-se constatar que a parcela de contribuição (fator de impacto) do background é significativa nas três fases, atingindo valores acima de 10%, ou seja, a

parcela de carga não linear não representada neste caso (background) impactou significativamente na distorção do barramento sob análise, principalmente sobre a fase B que atingiu o valor 14,570%. A Figura (70) apresenta as tensões de saída da RNA bem como os valores de tensão medidos no barramento de interesse.

Figura 70. Tensões de saída da RNA de 3º harmônico para as três fases na barra de 13,8 kV do transformador DITF4-04.

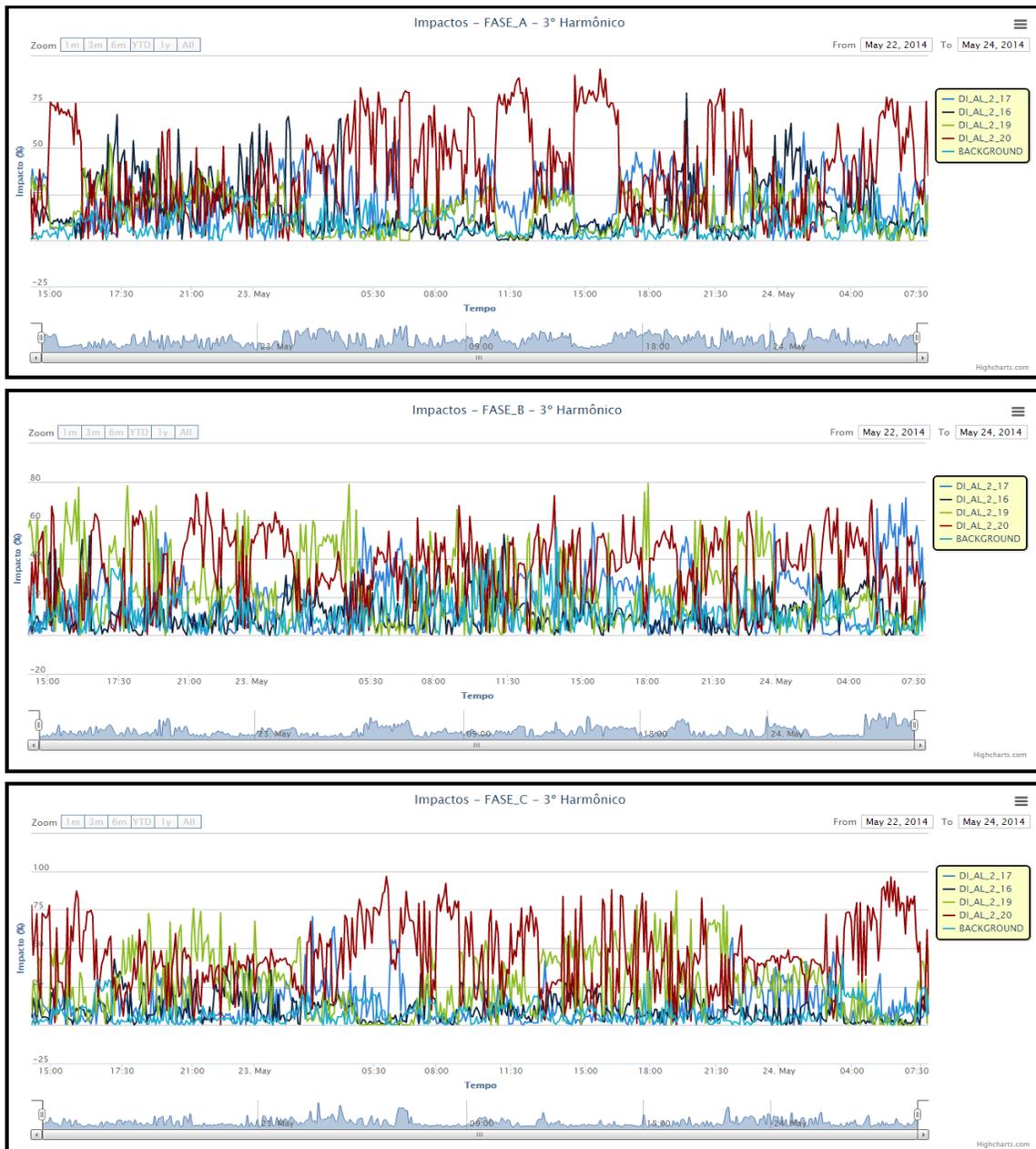


Fonte: O Autor.

Como pode ser visto nestes gráficos, não há uma dominância em nenhuma das fases, ao longo de todo o período, por parte de um alimentador somente, ou seja, em determinados

momentos um alimentador específico apresenta maior variação na saída da rede neural (maior sensibilidade), enquanto que em outro momento outro alimentador é responsável por apresentar maior sensibilidade. A Figura (71) apresenta os impactos na distorção de tensão do barramento estudado decorrentes de cada alimentador bem como do background ao longo do período.

Figura 71. Impactos de 3º harmônico do modelo de RNA para as três fases na barra de 13,8 kV do transformador DITF4-04.



Fonte: O Autor.

Como pode ser visto, existem variações dos impactos ao longo do tempo, alternando os alimentadores. Contudo, quando o impacto é analisado de maneira geral, fica evidente, em alguns casos, a dominância de certa carga harmônica, DIAL2-20.

Para avaliar a confiabilidade dos resultados é preciso observar a qualidade do modelo criado. A Tabela 20 contém as métricas e erros fornecidos pela biblioteca do weka, nela é possível observar que o coeficiente de correlação das entradas em relação à saída é acima de 0.7 para as fases A e C, e acima de 0.9 para a fase B, sendo este considerado um valor muito alto.

Tabela 20. Métricas de avaliação do modelo de RNA do Estudo de Caso.

MÉTRICAS	FASE A	FASE B	FASE C
Correlation coefficient	0.7424	0.9222	0.776
Mean absolute error	4.0171	0.9222	1.7993
Root mean squared error	5.4567	0.7878	2.6261
Relative absolute error	65.4322 %	33.2853 %	61.1654 %
Root relative squared error	67.3283 %	39.0389 %	66.0971 %
Total Number of Instances	1993	1993	1993

Também se observa que o Root Relative Squared Error foi próximo de 30% para a fase B e superior a 60% para as fases A e C, é importante ressaltar que embora este número pareça significativo, é considerado aceitável, pois a metodologia adotada realiza um estudo de sensibilidade para verificar qual das entradas é mais impactante na saída, além disso, boa parte deste erro está associada ao erro de background, ou seja, todos os pontos da rede elétrica estudada que não possuem campanha de medição, pois neste caso, quanto maior o número de entradas da RNA e o período da campanha de medição, menor será a parcela de erro associada ao background.

5.5 ÁRVORES DE REGRESSÃO

A seguir são apresentados os resultados referentes à análise da contribuição das correntes harmônicas de 3ª ordem dos alimentadores DIAL2-16, DIAL2-17, DIAL2-19 e DIAL2-20, em relação à barra do transformador DITF4-04. A árvore de regressão foi treinada e testada usando o algoritmo CART.

O período de que possuem medições simultâneas é o mesmo utilizado na técnica de RNA, 22-05-2014 14h18min a 05-05-2014 07h47min, porém fazendo uso do patamar de carga leve estipulado pela ONS, que é de 00h00min a 06h00min, faz com que o período de análise seja reduzido para os dias 23 e 24 de maio de 2014.

A porcentagem de dados utilizados para treino e teste, 80% e 20% respectivamente, e o fator de sensibilidade que incrementa as entradas de correntes harmônicas utilizado também foi de 10%. Os parâmetros utilizados para criação do modelo de AR são os definidos como os padrões pelo SISQEE, descritos na Figura (58).

Os resultados a seguir foram obtidos utilizando o MAE como métrica de avaliação de erro das series temporais de corrente harmônicas geradas com o fator de sensibilidade. A Tabela 21 contém os fatores de impactos dos barramentos dos alimentadores considerando o background.

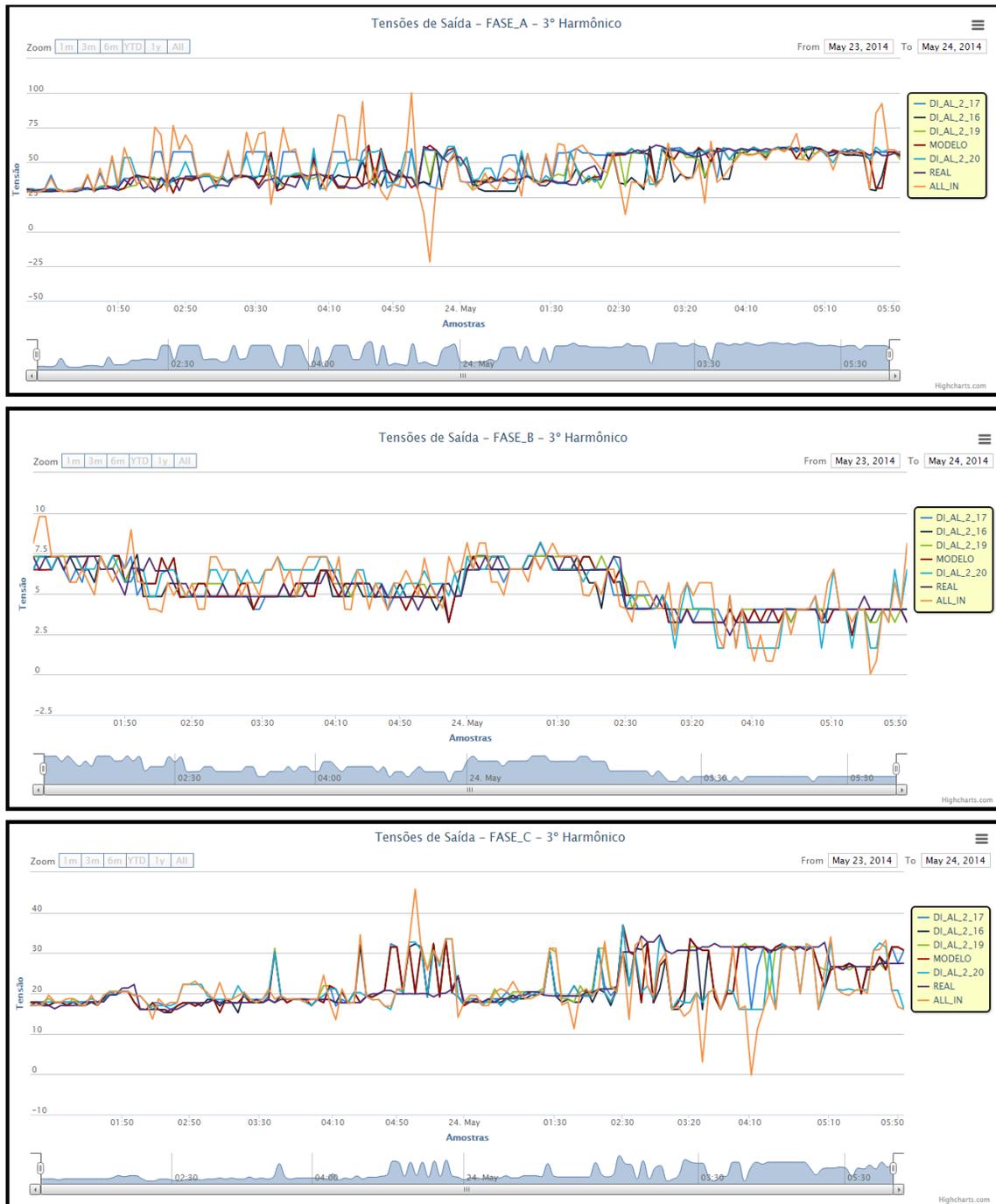
Tabela 21. Fator de Impacto (%) com Background do Estudo de Caso.

FONTE	FASE A	FASE B	FASE C
DIAL2-16	22,217	14,632	03,967
DIAL2-17	39,073	03,154	02,891
DIAL2-19	08,533	16,716	01,623
DIAL2-20	17,387	57,433	75,921
BACKGROUND	12,790	08,064	15,599

Como pode ser visto na Tabela 21, considerando o período de medição, o alimentador que apresentou maior influência na distorção de tensão do barramento estudado em questão foi DIAL2-17 para a fase A com impacto de 39,073%, e o alimentador DIAL2-20 para as fases B e C, alcançando um valor igual a 57,433% e 75,921% respectivamente. Adicionalmente, pode-se constatar que a parcela de contribuição (fator de impacto) do background é significativa nas três fases, principalmente sobre a fase C que atingiu o valor de 15,599%.

A Figura (72) apresenta as tensões de saída da AR bem como os valores de tensão medidos no barramento de interesse. Como pode ser visto nessas figuras, não há uma dominância em nenhuma das fases, ao longo de todo o período, por parte de um alimentador somente.

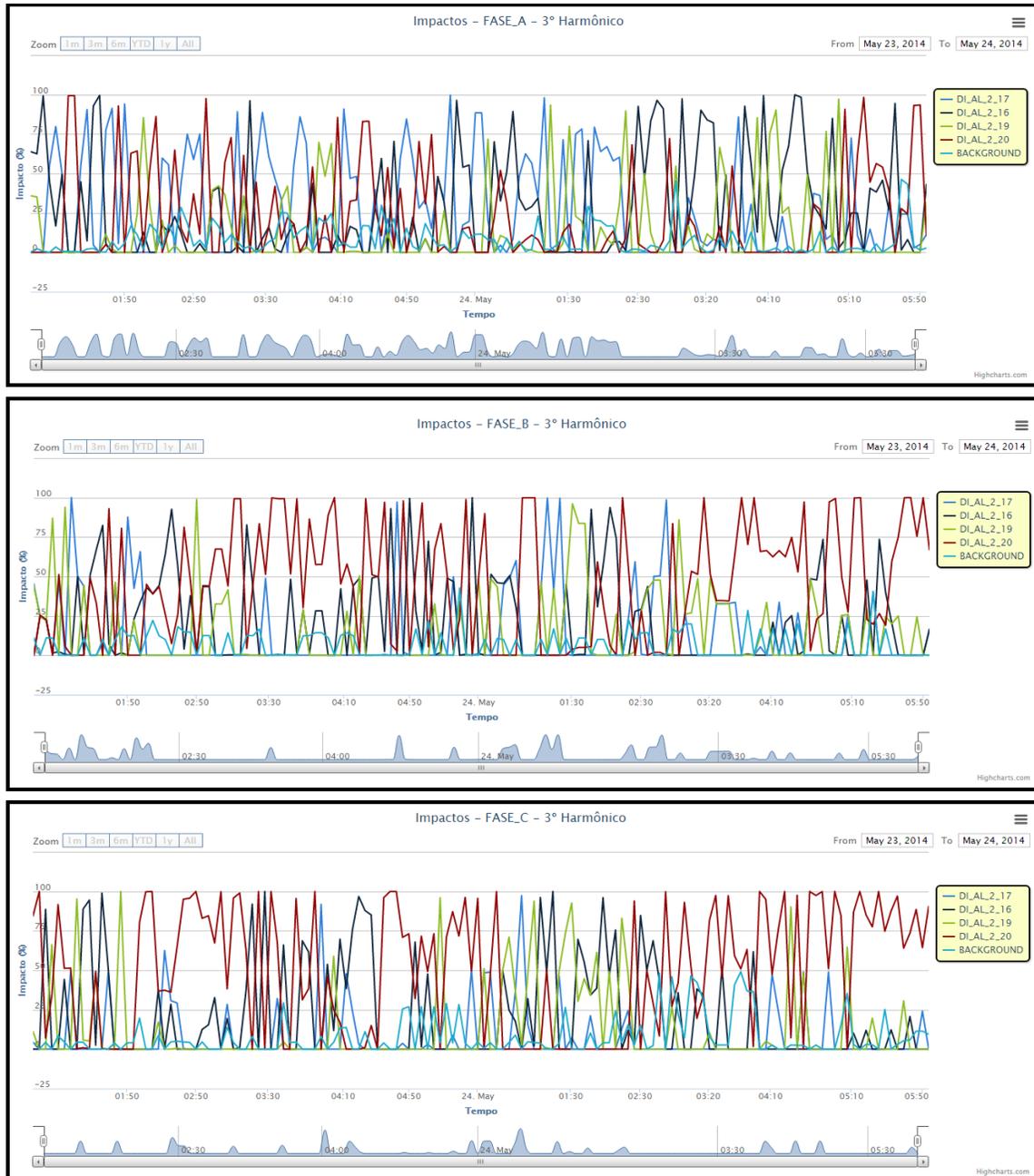
Figura 72. Tensões de saída do modelo de AR de 3º harmônico para as três fases na barra de 13,8 kV do transformador DITF4-04.



Fonte: O Autor.

A Figura (73) apresenta os impactos na distorção de tensão do barramento estudado em decorrentes de cada alimentador bem como do background ao longo do período. Observa-se que existem variações dos impactos ao longo do tempo, alternando o alimentador que possui maior influência. Contudo, quando o impacto é analisado de maneira geral, fica evidente, em alguns casos, a dominância de certa carga harmônica, DIAL2-20.

Figura 73. Impactos de 3º harmônico do modelo de AR para as três fases na barra de 13,8 kV do transformador DITF4-04.



Fonte: O Autor.

Para avaliar a confiabilidade dos resultados é preciso observar a qualidade do modelo criado, a Tabela 22 contém as métricas e erros fornecidos pela biblioteca do weka, nela é possível observar que o coeficiente de correlação das entradas em relação à saída é 0.8176 para a fase A, 0.8821 para a fase B e 0.6825 para a fase C, sendo estes valores considerados altos, também se observa que o Root Relative Squared Error se mostrou menor conforme quanto maior o coeficiente de correlação, sendo o melhor resultado o da fase B.

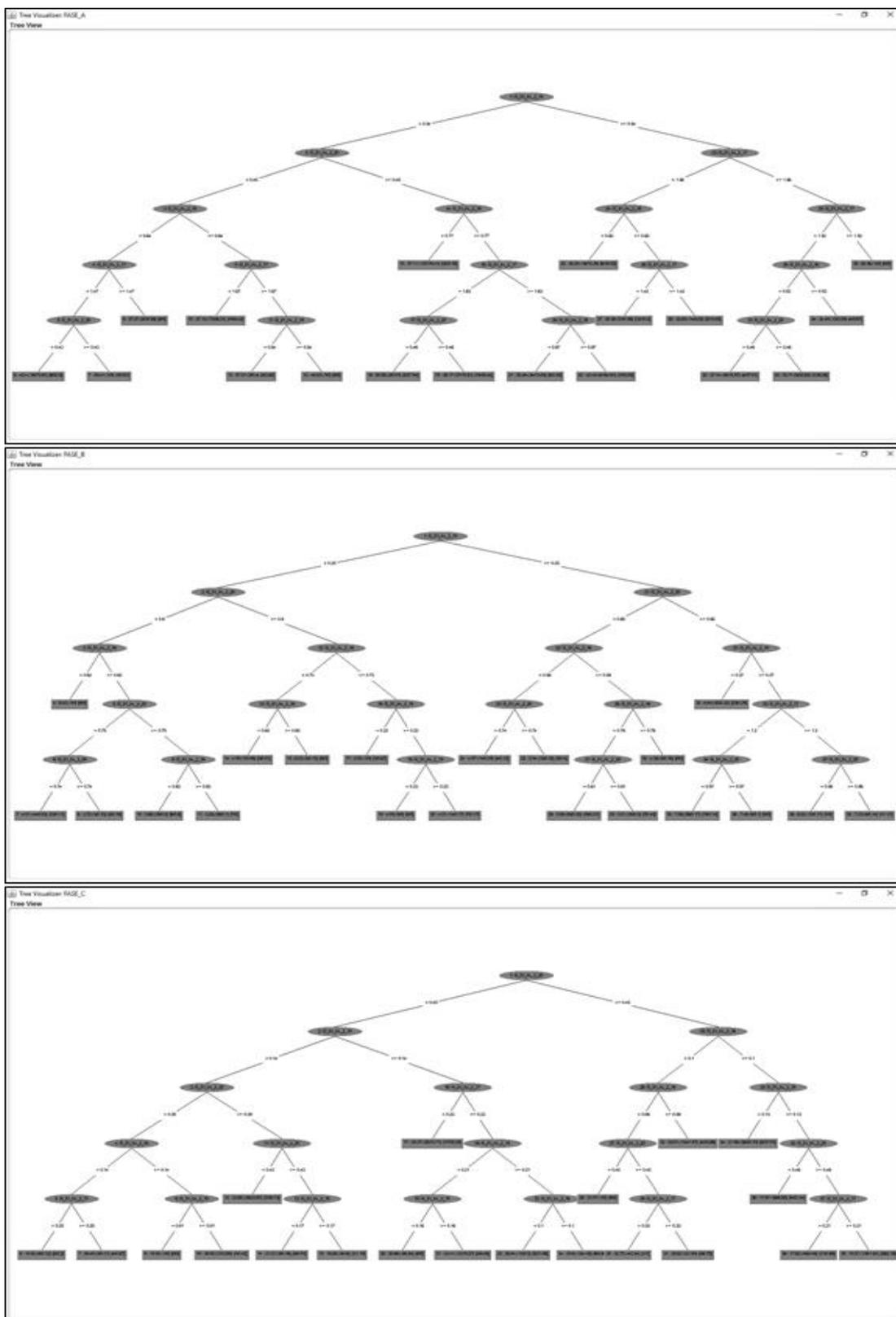
Tabela 22. Métricas de avaliação do modelo de AR do Estudo de Caso.

MÉTRICAS	FASE A	FASE B	FASE C
Correlation coefficient	0.8157	0.8885	0.6766
Mean absolute error	4.3549	0.4363	2.793
Root mean squared error	6.6607	0.6129	4.2136
Relative absolute error	40.9729 %	37.5746 %	57.1488 %
Root relative squared error	58.3168 %	46.0396 %	75.2364 %
Total Number of Instances	576	576	576

O SISQEE fornece as regras que compõe o modelo de AR criado, bem como a própria árvore em sim, a qual pode ser explorada dentro do ambiente do SISQEE. As Figura (74) apresentam as árvores das três fases, as quais foram criadas seguindo os parâmetros mostrados anteriormente na Figura (58), limitando a profundidade da árvore para cinco e realizando a poda, a fim de evitar ou pelo menos atenuar o *overfitting*.

Ressalta-se que ao se criar um modelo de AR, talvez seja necessário verificar se é possível eliminar regras redundantes fora do ambiente do SISQEE, e até mesmo realizar novamente poda na árvore. Em cada árvore a corrente que possui maior correlação com a tensão V3_DITF4-04 é dado como o nó raiz, ou seja, I3_DIAL2-19 para as fases A e B, e I3_DIAL2-20 para a fase C.

Figura 74. Árvores de regressão das três fases.



Fonte: O Autor.

Nestas árvores fica explícito o nível de corrente harmônica de cada nó que define o grau de impacto harmônico em cada folha. O SISQEE permite a manipulação do tamanho dos nós para uma melhor visualização, além disso, fornece as regras que compõe cada árvore. Foi criado um total de 35 nós para a fase A e 39 para as fases B e C. As regras destes nós estão dispostas nas Tabelas B-1 a B-3.

5.6 CONSIDERAÇÕES FINAIS

Este capítulo apresentou os resultados obtidos das análises utilizando os módulos da ferramenta criada SISQEE, utilizando dados reais obtidos de campanha de medição. O módulo de Análise de QEE apresentou gráficos de alguns dos parâmetros elétricos que o SISQEE fornece, assim como suas respectivas tabelas contendo os indicadores que regem a qualidade segundo o PRODIST, permitindo ao usuário verificar se os valores estão ou não de acordo com as normas.

O módulo de RL permitiu avaliar a contribuição harmônica de um alimentador e de em relação a um ponto de interesse, o módulo de RNA permitiu uma análise de contribuição de múltiplos alimentadores assim como de background, podendo utilizar períodos grandes, pois quanto mais robusta a base de treino melhor o aprendizado, e o módulo de AR também permitiu uma análise com vários alimentadores, utilizando curtos períodos de análise, mostrando o uso de patamar de carga, o qual é possível ser utilizado em qualquer um dos módulos de IC.

Ressalta-se que embora os estudos de AR e RNA tenham sido realizados utilizando as configurações padrões, o SISQEE permite que o usuário possa alterá-los, de modo a procurar pelo modelo que melhor represente o sistema elétrico estudo, obtendo erros de modelo menores aumentando a qualidade dos resultados.

Também é importante ressaltar que embora os erros dos modelos obtidos de RNA e AR sejam altos, a metodologia aplicada nessas técnicas não exige muita precisão, e o peso dos resultados é relativo, não implicando em uma determinação exata dos impactos das fontes de alimentadores. A metodologia utilizada realiza estudos de sensibilidade para saber qual das entradas do modelo tem mais influência na saída, fornecendo um norte não conclusivo para que a concessionária de energia possa realizar estudos mais específicos, a fim de tomar as devidas medidas mitigadoras.

Além disso, parte dos erros de modelo se deve ao background, ou seja, todos os pontos da rede elétrica que não foram levados em consideração, quanto maior o número de entradas, maior quantidade de dados e quanto mais representativos eles forem, melhor será o modelo construído.

6 CONSIDERAÇÕES FINAIS

6.1 CONCLUSÕES

O SISQEE surge como uma ferramenta inovadora bastante robusta, capaz de unificar dados de campanhas de medições de diversos analisadores de QEE disponíveis no mercado, padronizando e armazenando os dados em um único servidor de banco de dados.

Seu uso permitirá que a concessionária de energia elétrica faça uma gestão mais efetiva das condições de operação da sua rede elétrica, com relação aos níveis presentes de distorção harmônica de tensão, identificando as cargas suspeitas de estarem contribuindo de forma mais significativa para esta distorção. Isso permite a adoção de medidas remediais para a correção das transgressões verificadas, o que certamente melhorará a qualidade da energia, reduzindo perdas técnicas, e evitando multas da ANEEL.

Além do mais, o SISQEE viabiliza de forma simples, via interface gráfica amigável, que sejam avaliadas as condições de fornecimento em regime permanente às UCs. Isso permite o levantamento dos indicadores de QEE estabelecidos pelo PRODIST, módulo 8. Dessa forma as empresas poderão responder tecnicamente, via laudos gerados a partir dos relatórios gerados pelo SISQEE, a eventuais reclamações e processos judiciais dos consumidores a respeito de prejuízos provocados pela perda de qualidade no fornecimento de energia. Esse procedimento certamente evitará prejuízos financeiros advindos desses processos. Percebe-se que evitar multas e reduzir perdas técnicas são os dois aspectos principais advindos da utilização do SISQEE.

O SISQEE proveu a criação de um banco de dados robusto que pode armazenar os imensos volumes de dados que as concessionárias lidam. O banco de dados foi estruturado de forma homogênea capaz de receber dados de medições de mais de um analisador de QEE disponível no mercado, tornando-o flexível.

Os gráficos gerados pelo SISQEE possuem recursos que facilitam análise do usuário, pois além de serem responsivos, ou seja, se adaptarem de acordo com a tela, também é possível configurá-los para exibir apenas parte do período analisado ou para que sejam exibidas apenas as curvas desejadas, além da possibilidade de exportá-los para diferentes formatos de imagem. Além disso, os gráficos possuem tabelas contendo valores que auxiliam na realização dessas análises, apresentando valores referentes às curvas exibidas e aos indicadores contidos no PRODIST.

Os módulos de análise de contribuição de impactos harmônicos baseados em técnicas de IC permitem a investigação dos alimentadores que mais contribuem para a distorção da tensão, para posterior identificação dos consumidores poluidores pertencentes a esses alimentadores. Um aspecto marcante que esta abordagem exige como ferramenta de auxílio à gestão da qualidade da energia elétrica, é que ela fornece as variáveis críticas e os seus limites que afetam a distorção de tensão.

Além disso, estes módulos possibilitam a criação de relatórios contendo os resultados obtidos devidamente formatados com espaços reservados para as considerações do usuário. Embora o SISQEE utilize dados reais obtidos através de campanhas de medições, é possível realizar análises de contribuição de impactos utilizando dados de simulação, sem a necessidade de armazená-los no banco de dados, desde que estejam no formato e padrões adotados pelo SISQEE.

Essas informações são estratégicas para o planejamento da Empresa, que a partir dessas análises poderá implantar medidas mitigadoras para a correção e melhoria da qualidade, aumentando a satisfação de seus consumidores e evitando as penalizações por parte da Agência Reguladora, por problemas de perda de qualidade no fornecimento da energia.

Todos esses aspectos relatados mostram a funcionalidade e aplicabilidade do SISQEE como ferramenta de suporte ao usuário para realizar as análises de QEE e de contribuição de impacto harmônico, cumprindo o propósito com que foi criado se mostrando funcional e efetivo, podendo no futuro ficar disponível para o mercado do setor elétrico tornando-se cabeça de série.

6.2 PROPOSTAS DE TRABALHOS FUTUROS

Como trabalhos futuros, pretende-se focar em aspectos de modo a deixar o sistema mais robusto, através da:

- Criação de modelos para outros analisadores de QEE disponíveis no mercado.
- Expansão da base de dados inserindo outros parâmetros elétricos.
- Implementação de outros indicadores de qualidade.
- Realização de um diagnóstico automático nas análises de QEE a partir das condições de operações contidas no banco de dados.

6.3 RESUMO DAS PUBLICAÇÕES

Conferências Nacionais:

Rocha, G. V. S.; Tostes, M. E. de L.; Manito, A. R. A, Monteiro, F. P.; Bezerra U. H.; Soares, T. M.; Araújo, F. G. M.. Sistema Inteligente de Suporte a Análise da Qualidade da Energia Elétrica em regime Permanente. Em: Conferência Brasileira sobre Qualidade da Energia Elétrica – CBQEE, 2015, Campina Grande – PB.

Araújo, F. G. M ; Tostes, M. E. de L.; Vieira, J. P. A.; Bezerra U. H.; Manito, A. R. A; Soares, T. M.; Araújo, F. G. M.; **Rocha, G. V. S.**. Utilização da Regressão Linear para Analise da Contribuição de uma Carga não linear em um Sistema Elétrico. Em: Conferência Brasileira sobre Qualidade da Energia Elétrica – CBQEE, 2015, Campina Grande – PB.

Simpósios Nacionais:

A. R. A, Monteiro; Bezerra U. H.; M. E. de L.; Soares, T. M.; **Rocha, G. V. S.**. Estimação da Contribuição de Cargas Não Lineares na Distorção Harmônica de Tensão de um Barramento de Interesse do Sistema Elétrico Utilizando Rede Neural Artificial. Em: Simpósio Brasileiro de Sistemas Elétricos – SBSE, 2014, Foz do Iguaçu – PR.

REFERÊNCIAS BIBLIOGRÁFICAS

ALMEIDA, C. F. M. Fontes Distribuídas de Harmônicos em Sistemas Elétricos de Potência. p. 258, 2012.

ANEEL, A. N. de E. E. Procedimentos de Distribuição de Energia Elétrica no Sistema Elétrico Nacional (PRODIST) Módulo 8: Qualidade da Energia Elétrica. **Procedimentos de Distribuição de Energia Elétrica no Sistema Elétrico Nacional – PRODIST**, p. 1–76, 2010. Disponível em: <http://www2.aneel.gov.br/arquivos/PDF/Modulo8_Revisao_7.pdf>.

ARRILLAGA, J. **Power system harmonic analysis**. [s.l.] Wiley, 1997.

ARRILLAGA, J.; WATSON, N. R. **Power system harmonics**. [s.l.] Wiley, 2004.

BIGUS, J. P. **Data mining with neural networks: solving business problems from application development to decision support**. [s.l.] McGraw-Hill, 1996.

BREIMAN, L.; FRIEDMAN, J.; STONE, C. J.; OLSHEN, R. A. **Classification and regression trees**. [s.l.] Taylor & Francis, 1984.

DE SOUTO, M. C. P.; LORENA, A. C.; DELBEM, A. C. B.; DE CARVALHO, A. Técnicas de Aprendizado de Máquina para problemas de Biologia Molecular. **Sociedade Brasileira de Computação**, 2003.

DUGAN, R. C.; MCGRANAGHAN, M. F.; SANTOSO, S.; BEATY, H. W. Electrical Power Systems Quality. **Power**, p. 373–435, 2002. Disponível em: <<http://spectrum-analysis.info/wp-content/uploads/pdfs/Electrical Power Systems Quality by Mark F McGranaghan - Great Book.pdf>>.

ESTEVES, E. A. C. IMPACTOS NA QUALIDADE DA ENERGIA ELÉTRICA PROVENIENTES DO RÁPIDO CRESCIMENTO DA DEMANDA NO SUDESTE DO ESTADO DO PARÁ. p. 39, 2012. Disponível em: <http://www.repositorio.ufpa.br/jspui/bitstream/2011/2981/1/Dissertacao_ImpactosQualidadeEnergia.pdf>.

FACELI, K.; LORENA, A. C.; GAMA, J.; CARVALHO, A. Inteligência Artificial: Uma abordagem de aprendizado de máquina. **Rio de Janeiro: LTC**, 2011.

HAN, J.; PEI, J.; KAMBER, M. **Data mining: concepts and techniques**. [s.l.] Elsevier Science, 2011.

HAYKIN, S. S. **Neural networks and learning machines, 3rd edition**. [s.l.: s.n.]

- HOSOKAWA, E. O. Técnica de Árvore de Decisão em Mineração de Dados. p. 40, 2011. Disponível em: <<http://www.fatecsp.br/dti/tcc/tcc0003.pdf>>.
- HSU, H. **Probability, random variables and random processes**. [s.l: s.n.]v. 53
- JAIN, A. K.; DUBES, R. C. **Algorithms for Clustering Data**Prentice Hall, 1988. . Disponível em: <<http://portal.acm.org/citation.cfm?id=SERIES10022.42779>>.
- LEÃO, R. P. S.; SAMPAIO, R. F.; ANTUNES, F. L. M. **Harmônicos em Sistemas Elétricos**Rio de Janeiro: Elsevier, , 2014. .
- LIMA, D. D. S. DESENVOLVIMENTO DE UM SIMULADOR DE REDES ELÉTRICAS PARA ESTUDOS DA PROPAGAÇÃO DE HARMÔNICOS. 2016.
- LUGER, G. F.; STUBBLEFIELD, W. A. **Artificial Intelligence : Structures and Strategies for Complex Problem Solving**, 1998. .
- MANITO, A. R. A.; BEZERRA, U. H.; EMÍLIA, M.; TOSTES, D. L.; SOARES, T. M. Estimação da Contribuição de Cargas Não Lineares na Distorção Harmônica de Tensão de um Barramento de Interesse do Sistema Elétrico Utilizando Rede Neural Artificial. **Simpósio Brasileiro de Sistemas Elétricos – SBSE, Foz do Iguaçu – PR**, 2014.
- MATOS, E. O. de; SOARES, T. M.; BEZERRA, U. H.; TOSTES, M. E. de L.; MANITO, A. R. A.; JR, B. C. C. Using linear and non-parametric regression models to describe the contribution of non-linear loads on the voltage harmonic distortions in the electrical grid. *IET Generation, Transmission & Distribution*, v. 10, n. 8, p. 1825–1832, 2016. Disponível em: <<http://digitalibrary.theiet.org/content/journals/10.1049/iet-gtd.2015.0948>>.
- MAZIN, H. E.; XU, W.; HUANG, B. Determining the harmonic impacts of multiple harmonic-producing loads. **IEEE Transactions on Power Delivery**, v. 26, n. 2, p. 1187–1195, 2011.
- MIRKIN, B. **Core concepts in data analysis: summarization, correlation and visualization**. [s.l.] Springer London, 2011.
- MITRA, S.; ACHARYA, T. **Data mining multimedia, soft computing, and bioinformatics**. [s.l: s.n.]
- MONARD, M. C.; BARANAUSKAS, J. A. Conceitos sobre Aprendizado de Máquina. **Sistemas inteligentes: fundamentos e aplicações**, p. 89–114, 2003.

- MONTEIRO, F. P. *ClasSIS : UMA METODOLOGIA PARA CLASSIFICAÇÃO* ClasSIS : UMA METODOLOGIA PARA CLASSIFICAÇÃO. 2015.
- PECK, R.; OLSEN, C.; DEVORE, J. L. **Introduction to statistics and data analysis**. [s.l.] Cengage Learning, 2011.
- PIRES, I. A. O Setor Elétrico - Capítulo III Efeitos de harmônicos no sistema de distribuição e limites segundo as principais normas nacionais e internacionais – Parte II. p. 40–43, mar. 2010. Disponível em: <http://www.osetoreletrico.com.br/web/documentos/fasciculos/Ed50_fasc_harmonicos_cap3.pdf>.
- REZENDE, S. O. **Sistemas inteligentes: fundamentos e aplicações**. [s.l.] Manole, 2003.
- RUSSELL, S. J.; NORVIG, P. **Artificial intelligence: a modern approach**. [s.l.] Prentice Hall, 2003.
- SAFAVIAN, S. R.; LANDGREBE, D. A Survey of Decision Tree Classifier Methodology. **IEEE Transactions on Systems, Man and Cybernetics**, v. 21, n. 3, p. 660–674, 1991.
- SHOJAIE, M.; MOKHTARI, H. A method for determination of harmonics responsibilities at the point of common coupling using data correlation analysis. **IET Generation, Transmission & Distribution**, v. 8, n. 1, p. 142–150, 2014. Disponível em: <<http://digital-library.theiet.org/content/journals/10.1049/iet-gtd.2013.0111>>.
- SIQUEIRA, A. L. C. RECONHECIMENTO DE ATIVIDADES HUMANAS UTILIZANDO REDES NEURAS AUTO-ASSOCIATIVAS E DADOS DE SMARTPHONE. 2016.
- SOARES, T. M., 2012. “Estimação da Contribuição de Múltiplas Fontes Harmônicas Utilizando Regressão Não Paramétrica”. Dissertação de Mestrado, Faculdade de Engenharia Elétrica, UFPA, 2012.
- SOTO, A.; CANELOI, R. Implementação de um Sistema Real Time Baseado em Redes Neurais Artificiais para Monitoramento Remoto de Tensão. 2015.
- TOM M. MITCHELL. **Machine learning**. [s.l.: s.n.]
- UNSAR, O.; SALOR, O.; CADIRCI, I.; ERMIS, M. Identification of harmonic current contributions of iron and steel plants based on time-synchronized field measurements - Part I: At PCC. **IEEE Transactions on Industry Applications**, v. 50, n. 6, p. 4336–4347, 2014.
- WASSERMAN, P. D. **Neural computing: theory and practice**. [s.l.] Van Nostrand Reinhold, 1989.

XU, W.; BAHRY, R.; MAZIN, H. E.; TAYJASANANT, T. A method to determine the harmonic contributions of multiple loads. **2009 IEEE Power and Energy Society General Meeting, PES '09**, p. 1–6, 2009.

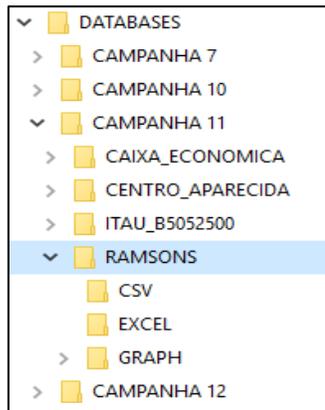
YAN, X. **Linear regression analysis: theory and computing**. [s.l.] World Scientific Publishing Company Pte Limited, 2009.

APÊNDICES

Este tópico é destinado para apresentar as imagens e tabelas adicionais para uma melhor visualização da estrutura dos diretórios criados pelo SISQEE, assim como as regras criadas pelos modelos de AR.

APÊNDICE A – Organização de Arquivos e Diretórios

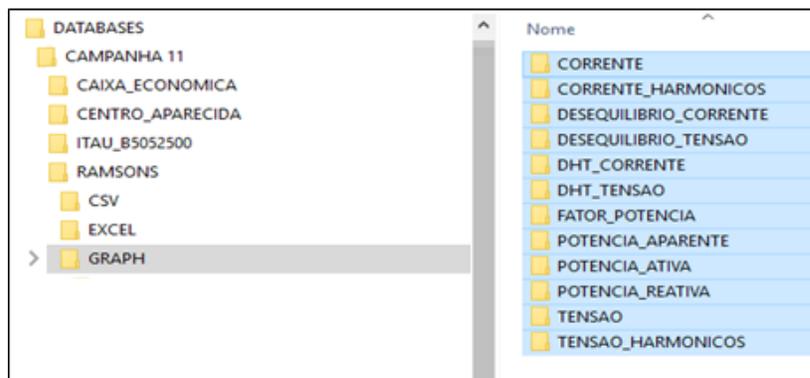
Figura 75. Estrutura dos diretórios quando uma base de dados é importada.



Fonte: O Autor.

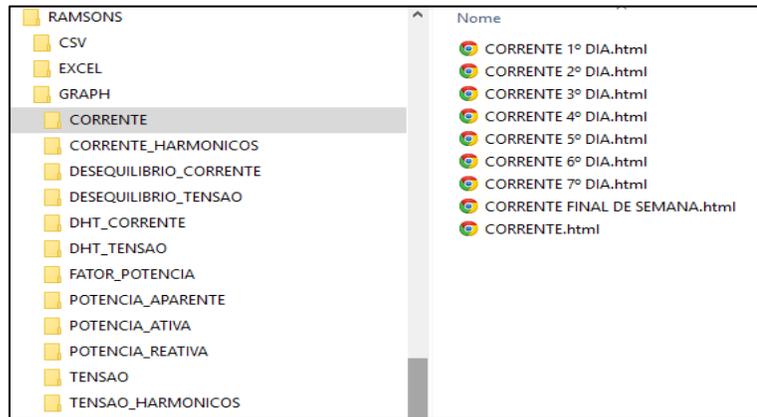
Dentro das subpastas CSV e EXCEL são salvos os arquivos padronizados e transformados que foram criados a partir dos arquivos de saída dos analisadores de QEE, para cada entidade no banco de dados que corresponde a um parâmetro elétrico é criado um arquivo contendo apenas os dados referentes a si mesmo. Por fim, a subpasta GRAPH é o diretório onde serão armazenados os gráficos gerados pelo *Módulo de Análise de Medições*.

Figura 76. Estrutura de diretórios dos gráficos.



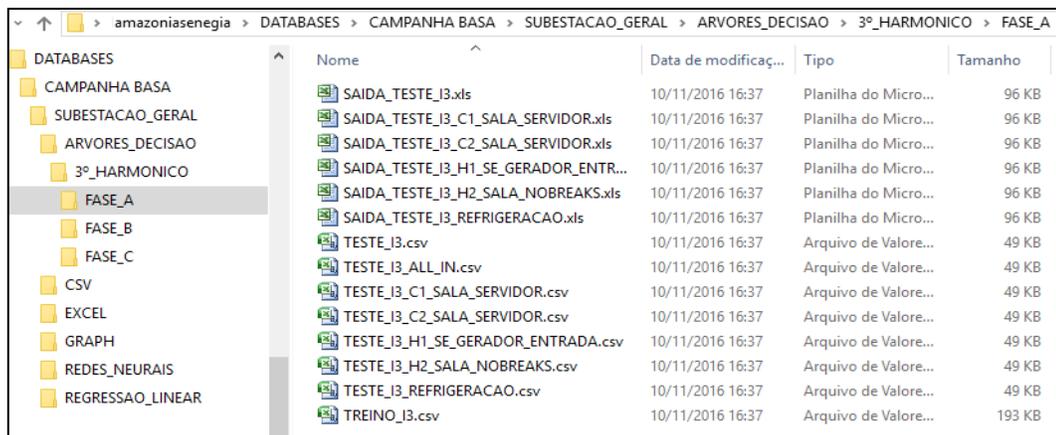
Fonte: O Autor.

Figura 77. Arquivos html criados no processo de geração de gráficos pelo Módulo de Análise de Medições.



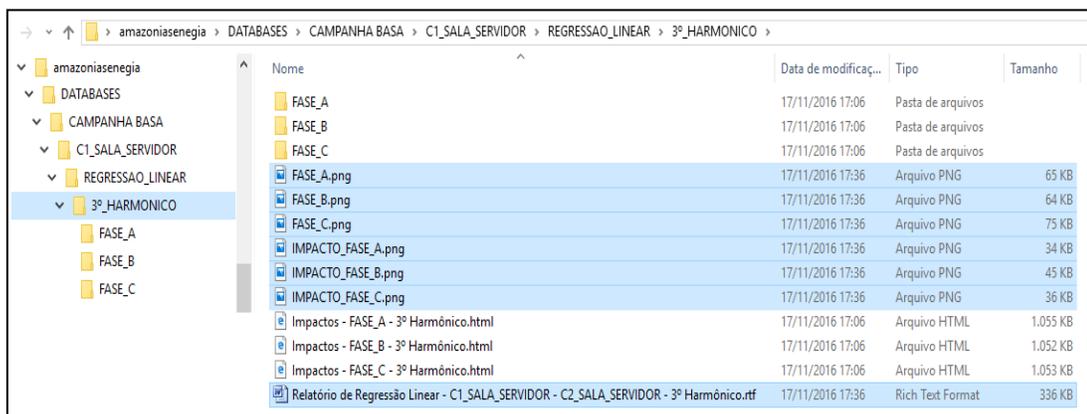
Fonte: O Autor.

Figura 78. Conjunto de Arquivos CSV's.



Fonte: Autor.

Figura 79. Resultados Regressão Linear: Diretório de Resultados.



Fonte: O Autor.

APÊNDICE B – Regras das Árvores de Regressão

Tabela 23. Relações causa-efeito da árvore de regressão que levam ao menor valor de distorção harmônica de tensão de 3ª ordem da fase A.

Nó da Árvore	Regras
1	Se $I3_DI_AL_2_19 < 0,54$ Então Nó 2 Senão Nó 23
2	Se $I3_DI_AL_2_20 < 0,45$ Então Nó 3 Senão Nó 14
3	Se $I3_DI_AL_2_16 < 0,64$ Então Nó 4 Senão Nó 9
4	Se $I3_DI_AL_2_17 < 1,47$ Então Nó 5 Senão Nó 8
5	Se $I3_DI_AL_2_20 < 0,43$ Então Nó 6 Senão Nó 7
6	$V3_DITF4-04 = 42,40$ V
7	$V3_DITF4-04 = 59,41$ V
8	$V3_DITF4-04 = 57,27$ V
9	Se $I3_DI_AL_2_17 < 1,67$ Então Nó 10 Senão Nó 11
10	$V3_DITF4-04 = 37,13$ V
11	Se $I3_DI_AL_2_19 < 0,54$ Então Nó 12 Senão Nó 13
12	$V3_DITF4-04 = 57,21$
13	$V3_DITF4-04 = 40,03$
14	Se $I3_DI_AL_2_16 < 0,77$ Então Nó 15 Senão Nó 16
15	$V3_DITF4-04 = 57,11$ V
16	Se $I3_DI_AL_2_17 < 1,63$ Então Nó 17 Senão Nó 20
17	Se $I3_DI_AL_2_20 < 0,46$ Então Nó 18 Senão Nó 19
18	$V3_DITF4-04 = 50,26$ V
19	$V3_DITF4-04 = 38,17$ V
20	Se $I3_DI_AL_2_16 < 0,87$ Então Nó 21 Senão Nó 22
21	$V3_DITF4-04 = 55,46$ V
22	$V3_DITF4-04 = 42,45$ V
23	Se $I3_DI_AL_2_17 < 1,56$ Então Nó 24 Senão Nó 29
24	Se $I3_DI_AL_2_16 < 0,83$ Então Nó 25 Senão Nó 26
25	$V3_DITF4-04 = 35,29$ V
26	Se $I3_DI_AL_2_17 < 1,42$ Então Nó 27 Senão Nó 28
27	$V3_DITF4-04 = 30,38$ V
28	$V3_DITF4-04 = 32,65$ V
29	Se $I3_DI_AL_2_17 < 1,92$ Então Nó 30 Senão Nó 35
30	Se $I3_DI_AL_2_16 < 0,93$ Então Nó 31 Senão Nó 34
31	Se $I3_DI_AL_2_20 < 0,46$ Então Nó 32 Senão Nó 33
32	$V3_DITF4-04 = 37,14$ V
33	$V3_DITF4-04 = 53,71$ V
34	$V3_DITF4-04 = 34,49$ V
35	$V3_DITF4-04 = 62,36$ V

Tabela 24. Relações causa-efeito da árvore de regressão que levam ao menor valor de distorção harmônica de tensão de 3ª ordem da fase B.

Nó da Árvore	Regras
1	Se $I3_DI_AL_2_19 < 0,25$ Então Nó 2 Senão Nó 21
2	Se $I3_DI_AL_2_20 < 0,80$ Então Nó 3 Senão Nó 12
3	Se $I3_DI_AL_2_16 < 0,62$ Então Nó 4 Senão Nó 5
4	$V3_DITF4-04 = 6,42$ V
5	Se $I3_DI_AL_2_20 < 0,75$ Então Nó 6 Senão Nó 9
6	Se $I3_DI_AL_2_20 < 0,74$ Então Nó 7 Senão Nó 8
7	$V3_DITF4-04 = 4,01$ V
8	$V3_DITF4-04 = 3,72$ V
9	Se $I3_DI_AL_2_16 < 0,83$ Então Nó 10 Senão Nó 11
10	$V3_DITF4-04 = 3,66$ V
11	$V3_DITF4-04 = 3,28$ V
12	Se $I3_DI_AL_2_16 < 0,73$ Então Nó 13 Senão Nó 16
13	Se $I3_DI_AL_2_16 < 0,65$ Então Nó 14 Senão Nó 15
14	$V3_DITF4-04 = 4,95$ V
15	$V3_DITF4-04 = 6,23$ V
16	Se $I3_DI_AL_2_19 < 0,22$ Então Nó 17 Senão Nó 18
17	$V3_DITF4-04 = 2,05$ V
18	Se $I3_DI_AL_2_19 < 0,23$ Então Nó 19 Senão Nó 20
19	$V3_DITF4-04 = 4,09$ V
20	$V3_DITF4-04 = 4,51$ V
21	Se $I3_DI_AL_2_20 < 0,85$ Então Nó 22 Senão Nó 31
22	Se $I3_DI_AL_2_16 < 0,58$ Então Nó 23 Senão Nó 26
23	Se $I3_DI_AL_2_20 < 0,74$ Então Nó 24 Senão Nó 25
24	$V3_DITF4-04 = 4,97$ V
25	$V3_DITF4-04 = 5,94$ V
26	Se $I3_DI_AL_2_16 < 0,78$ Então Nó 27 Senão Nó 30
27	Se $I3_DI_AL_2_20 < 0,81$ Então Nó 28 Senão Nó 29
28	$V3_DITF4-04 = 5,06$ V
29	$V3_DITF4-04 = 5,51$ V
30	$V3_DITF4-04 = 4,38$ V
31	Se $I3_DI_AL_2_19 < 0,37$ Então Nó 32 Senão Nó 33
32	$V3_DITF4-04 = 6,49$ V
33	Se $I3_DI_AL_2_17 < 1,20$ Então Nó 34 Senão Nó 37
34	Se $I3_DI_AL_2_20 < 0,97$ Então Nó 35 Senão Nó 36
35	$V3_DITF4-04 = 7,08$ V
36	$V3_DITF4-04 = 7,48$ V
37	Se $I3_DI_AL_2_20 < 0,96$ Então Nó 38 Senão Nó 39
38	$V3_DITF4-04 = 6,52$ V
39	$V3_DITF4-04 = 7,23$ V

Tabela 25. Relações causa-efeito da árvore de regressão que levam ao menor valor de distorção harmônica de tensão de 3ª ordem da fase C.

Nó da Árvore	Regras
1	Se $I3_DI_AL_2_20 < 0,45$ Então Nó 2 Senão Nó 25
2	Se $I3_DI_AL_2_19 < 0,14$ Então Nó 3 Senão Nó 16
3	Se $I3_DI_AL_2_20 < 0,39$ Então Nó 4 Senão Nó 11
4	Se $I3_DI_AL_2_20 < 0,54$ Então Nó 5 Senão Nó 8
5	Se $I3_DI_AL_2_17 < 0,25$ Então Nó 6 Senão Nó 7
6	$V3_DITF4-04 = 19,55$ V
7	$V3_DITF4-04 = 26,49$ V
8	Se $I3_DI_AL_2_19 < 0,01$ Então Nó 9 Senão Nó 10
9	$V3_DITF4-04 = 19,92$ V
10	$V3_DITF4-04 = 30,05$ V
11	Se $I3_DI_AL_2_20 < 0,43$ Então Nó 12 Senão Nó 13
12	$V3_DITF4-04 = 23,65$ V
13	Se $I3_DI_AL_2_16 < 0,17$ Então Nó 14 Senão Nó 15
14	$V3_DITF4-04 = 21,21$ V
15	$V3_DITF4-04 = 16,85$ V
16	Se $I3_DI_AL_2_17 < 0,23$ Então Nó 17 Senão Nó 18
17	$V3_DITF4-04 = 29,37$ V
18	Se $I3_DI_AL_2_19 < 0,21$ Então Nó 19 Senão Nó 22
19	Se $I3_DI_AL_2_19 < 0,16$ Então Nó 20 Senão Nó 21
20	$V3_DITF4-04 = 30,88$ V
21	$V3_DITF4-04 = 23,41$ V
22	Se $I3_DI_AL_2_16 < 0,10$ Então Nó 23 Senão Nó 24
23	$V3_DITF4-04 = 26,04$ V
24	$V3_DITF4-04 = 29,80$ V
25	Se $I3_DI_AL_2_16 < 0,10$ Então Nó 26 Senão Nó 33
26	Se $I3_DI_AL_2_16 < 0,08$ Então Nó 27 Senão Nó 32
27	Se $I3_DI_AL_2_20 < 0,45$ Então Nó 28 Senão Nó 29
28	$V3_DITF4-04 = 21,07$ V
29	Se $I3_DI_AL_2_17 < 0,33$ Então Nó 30 Senão Nó 31
30	$V3_DITF4-04 = 32,72$ V
31	$V3_DITF4-04 = 35,62$ V
32	$V3_DITF4-04 = 25,01$ V
33	Se $I3_DI_AL_2_16 < 0,13$ Então Nó 34 Senão Nó 35
34	$V3_DITF4-04 = 21,98$ V
35	Se $I3_DI_AL_2_20 < 0,48$ Então Nó 36 Senão Nó 37
36	$V3_DITF4-04 = 17,67$ V
37	Se $I3_DI_AL_2_17 < 0,21$ Então Nó 38 Senão Nó 39
38	$V3_DITF4-04 = 17,62$ V
39	$V3_DITF4-04 = 19,37$ V