

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

FÁBIO CORRÊA DOS SANTOS

METODOLOGIA DE AUXÍLIO NA TOMADA DE DECISÃO NA
FORMAÇÃO DE CONJUNTOS DE UNIDADES CONSUMIDORAS DE
ENERGIA ELÉTRICA

DM 21/2010

UFPA / ITEC / PPGE
Campus Universitário do Guamá
Belém-Pará-Brasil
2010

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

FÁBIO CORRÊA DOS SANTOS

METODOLOGIA DE AUXÍLIO NA TOMADA DE DECISÃO NA
FORMAÇÃO DE CONJUNTOS DE UNIDADES CONSUMIDORAS DE
ENERGIA ELÉTRICA

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

UFPA / ITEC / PPGEE
Campus Universitário do Guamá
Belém-Pará-Brasil
2010

S237m

Santos, Fábio Corrêa dos

Metodologia de auxílio na tomada de decisão na formação de conjuntos de unidades consumidoras de energia elétrica / Fábio Corrêa dos Santos; orientadora, Maria Emília de Lima Tostes.-2010.

Dissertação (Mestrado) – Universidade Federal do Pará, Instituto de Tecnologia, Programa de Pós-graduação em Engenharia Elétrica, Belém, 2010.

1. Energia elétrica – distribuição. 2. Energia elétrica – consumo. Política energética. I. orientador. II. título.

CDD 22. ed. 621.3191

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

**METODOLOGIA DE AUXÍLIO NA TOMADA DE DECISÃO NA
FORMAÇÃO DE CONJUNTOS DE UNIDADES CONSUMIDORAS DE
ENERGIA ELÉTRICA**

AUTOR: FÁBIO CORRÊA DOS SANTOS

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 SISTEMAS DE ENERGIA.

APROVADA EM 08/10/2010

BANCA EXAMINADORA:

Prof.^a. Maria Emília de Lima Tostes, Dr.^a.
ORIENTADOR – UFPA

Prof. Ádamo Lima de Santana, Dr
CO-ORIENTADOR – UFPA

Prof. Carlos Renato Lisboa Francês, Dr
MEMBRO – UFPA

Cláudio Luciano da Rocha Conde, Dr
MEMBRO EXTERNO

VISTO:

Prof. Dr. Marcus Vinícius Alves Nunes
(COORDENADOR DO PPGEE/ITEC/UFPA)

UFPA / ITEC / PPGEE

Dedico esta dissertação aos membros da Família Santos: Mãe Iêda, Pai Iso e Tio Mauro.

AGRADECIMENTOS

À Deus, o Senhor da vida, criador de todas as coisas existentes nesse Universo e Único digno de receber todo louvor e toda glória.

À minha querida família, mãe Iêda, pai Iso e tio Mauro por terem proporcionado um lar repleto de amor e felicidade, onde encontrei a motivação necessária para vencer os desafios com os quais me deparei no mestrado.

À minha orientadora, Prof^a Dr^a Maria Emília de Lima Tostes, pela amizade, pelo incentivo, motivação e pela paciência em ensinar durante estes anos de mestrado.

Para mim, um exemplo de profissional a ser seguido.

Ao meu co-orientador Prof. Dr. Ádamo Lima de Santana, pelo apoio, incentivo e pelas contribuições durante a elaboração deste trabalho.

Ao Prof. Guilherme Conde, pela amizade, pelo incentivo, motivação e paciência em ensinar, principalmente na parte computacional deste trabalho.

Aos colegas Rogério Silva, Samile Cristina, Walenda Tostes e Prof^a Dr^a Teresinha participantes do projeto Novos Atributos da CELPA, pela amizade e aprendizado.

Aos Engenheiros da CELPA Edilberto Müller e Rodolfo Alves, por compartilharem sua experiência prática na área de distribuição de energia.

A todos aqueles que compartilharam das angústias e alegrias deste trabalho.

“Em Deus faremos proezas, porque Ele mesmo calca dos pés os nossos adversários”

Salmo 60:12

SUMÁRIO

LISTA DE FIGURAS	X
LISTA DE TABELAS.....	Xi
LISTA DE ABREVIACÕES E SIGLAS	Xiii
LISTA DE SÍMBOLOS	Xiv
RESUMO	Xv
ABSTRACT	Xvi
CAPÍTULO 1.....	1
INTRODUÇÃO	1
1.1 MOTIVAÇÕES	1
1.2 OBJETIVO	3
1.3 JUSTIFICATIVA	3
1.4 REVISÃO BIBLIOGRÁFICA	4
1.5 ESTRUTURA DA DISSERTAÇÃO.....	6
CAPÍTULO 2.....	8
A METODOLOGIA DA ANEEL	8
2.1 CONSIDERAÇÕES INICIAIS	8
2.2 UM BREVE HISTÓRICO DA LEGISLAÇÃO.....	8
2.2.1 Das Definições.....	11
2.2.2 Dos Indicadores de Continuidade.....	12
2.3 A METODOLOGIA DE DEFINIÇÃO DAS METAS.....	14
2.4 O MÓDULO 8 DOS PROCEDIMENTOS DE DISTRIBUIÇÃO (PRODIST).....	18
2.5 CONSIDERAÇÕES FINAIS	19
CAPÍTULO 3.....	20
ANÁLISE DE CLASSIFICADORES.....	20
3.1 CONSIDERAÇÕES INICIAIS	20
3.2 A CLASSIFICAÇÃO NO CONTEXTO DA APRENDIZAGEM DE MÁQUINA ..	20
3.3 REDES NEURAIS: PERCEPTRON DE MÚLTIPLAS CAMADAS	22
3.3.1 Definições e Características	23
3.3.2 Redes <i>Perceptron</i> de Múltiplas Camadas (MLP).....	26
3.3.3 O Algoritmo de Retropropagação de Erro (<i>Error back-propagation</i>).....	27
3.3.4 O Algoritmo de Retropropagação de Erro com Momentum.....	29
3.4 CLASSIFICADOR K-VIZINHOS MAIS PRÓXIMOS	30
3.5 CONSIDERAÇÕES FINAIS	32
CAPÍTULO 4.....	33
A METODOLOGIA PROPOSTA	33
4.1 CONSIDERAÇÕES INICIAIS	33
4.2 A METODOLOGIA PROPOSTA: LINHAS GERAIS	33
4.3 A CONSTRUÇÃO DO CLASSIFICADOR	36
4.3.1 O Conjunto de dados para o Aprendizado supervisionado.....	36
4.3.2 Avaliação do desempenho dos classificadores.....	37
4.3.3 O Arranjo do Classificador.....	40

4.4	METODOLOGIA PARA UNIÃO DE CONJUNTOS.....	41
4.4.1	Para os cinco atributos.....	42
4.4.2	Patamar inicial para os indicadores de Continuidade: Ano zero.....	42
4.5	CONSIDERAÇÕES FINAIS	44
CAPÍTULO 5.....		46
ESTUDO DE CASO.....		46
5.1	CONSIDERAÇÕES INICIAIS	46
5.2	O ESTADO DO PARÁ E A CELPA	46
5.3	DEFINIÇÃO DOS CENÁRIOS DE ESTUDO.....	50
5.4	ANÁLISE DOS CENÁRIOS	51
5.4.1	Cenário 1: Belém - Perimetral e Bolonha urbano.....	51
5.4.2	Cenário 2: Ulianópolis urbano e Ulianópolis não urbano	56
5.4.3	Cenário 3: Paragominas urbano e Paragominas não urbano	61
5.4.4	Cenário 4: Água Azul do Norte e Canaã dos Carajás.....	66
5.4.5	Cenário 5: Santa Maria não urbano e São Miguel do Guamá	71
5.4.6	Cenário 6: Curionópolis e Eldorado dos Carajás	76
5.4.7	Cenário 7: Abaetetuba e Igarapé Mirim	81
5.5	CONSIDERAÇÕES FINAIS	86
CAPÍTULO 6.....		87
CONCLUSÃO		87
6.1	REALIZAÇÕES E CONTRIBUIÇÕES.....	87
6.2	SUGESTÃO DE TRABALHOS FUTUROS	89
REFERÊNCIAS BIBLIOGRÁFICAS.....		91
ANEXO 1		95
Metas dos Indicadores DEC e FEC para conjuntos do tipo interligado e isolado.		95

LISTA DE FIGURAS

Figura 3.1: Modelo de um neurônio artificial.....	24
Figura 3.2: Arquitetura de uma rede <i>feed forward</i> MLP.....	26
Figura 3.3: Influência do parâmetro k na classificação de um novo elemento, utilizando a distância euclidiana .	31
Figura 4.1: Estrutura do Software ANABENCH	33
Figura 4.2: Construção das bases treino-teste para validação cruzada	37
Figura 4.3: Arranjo de redes neurais para a classificação.....	41
Figura 5.1: Mesorregiões do estado do Pará.....	47
Figura 5.2: Divisão do estado do Pará em 188 conjuntos.....	48
Figura 5.3: Conjuntos alocados no <i>cluster</i> 17 (a) e <i>cluster</i> 5 (b) segundo ANEEL em 2007	49
Figura 5.4: Metas do cenário 1 ao longo de 8 anos para o indicado DEC	52
Figura 5.5: Taxa de redução anual do cenário 1 para o indicador DEC	53
Figura 5.6: Metas do cenário 1 ao longo de 8 anos para o indicado FEC.....	54
Figura 5.7: Taxa de redução anual para o cenário 1 para o indicador FEC	55
Figura 5.8: Metas do cenário 2 ao longo de 8 anos para o indicado DEC	57
Figura 5.9: Taxa de redução anual para o cenário 2 para o indicador DEC	58
Figura 5.10: Metas do cenário 2 ao longo de 8 anos para o indicado FEC.....	59
Figura 5.11: Taxa de redução anual para o cenário 2 para o indicador FEC	60
Figura 5.12: Metas do cenário 3 ao longo de 8 anos para o indicado DEC.....	62
Figura 5.13: Taxa de redução anual para o cenário 3 para o indicador DEC	63
Figura 5.14: Metas do cenário 3 ao longo de 8 anos para o indicado FEC.....	64
Figura 5.15: Taxa de redução anual para o cenário 3 para o indicador FEC	65
Figura 5.16: Metas do cenário 4 ao longo de 8 anos para o indicado DEC	67
Figura 5.17: Taxa de redução anual para o cenário 4 para o indicador DEC	68
Figura 5.18: Metas do cenário 4 ao longo de 8 anos para o indicado FEC.....	69
Figura 5.19: Taxa de redução anual para o cenário 4 para o indicador FEC	70
Figura 5.20: Metas do cenário 5 ao longo de 8 anos para o indicado DEC.....	72
Figura 5.21: Taxa de redução anual para o cenário 5 para o indicador DEC	73
Figura 5.22: Metas do cenário 5 ao longo de 8 anos para o indicado FEC.....	74
Figura 5.23: Taxa de redução anual para o cenário 5 para o indicador FEC	75
Figura 5.24: Metas do cenário 6 ao longo de 8 anos para o indicado DEC.....	77
Figura 5.25: Taxa de redução anual para o cenário 6 para o indicador DEC	78
Figura 5.26: Metas do cenário 6 ao longo de 8 anos para o indicado FEC.....	79
Figura 5.27: Taxa de redução anual para o cenário 6 para o indicador FEC	80
Figura 5.28: Metas do cenário 7 ao longo de 8 anos para o indicado DEC.....	82
Figura 5.29: Taxa de redução anual para o cenário 7 para o indicador DEC	83
Figura 5.30: Metas do cenário 7 ao longo de 8 anos para o indicado FEC.....	84
Figura 5.31: Taxa de redução anual para o cenário 7 para o indicador FEC	85

LISTA DE TABELAS

Tabela 2.1: Valores dos Padrões Coletivos estabelecidos pela Portaria DNAEE nº 46/78	9
Tabela 2.2: Valores dos Padrões Individuais estabelecidos pela Portaria DNAEE nº 46/78.....	10
Tabela 3.1: Tabela Atributo-Valor	21
Tabela 4.1: Resultados das Simulações para Rede Neurais tipo MLP	39
Tabela 4.2: Resultados das Simulações para algoritmo kNN	40
Tabela 4.3: Atributos equivalentes na união de conjuntos	42
Tabela 4.4: Valores anuais dos indicadores DEC e FEC e média anual do número de consumidores.....	43
Tabela 5.1: Cenários com bons resultados de metas de DEC e FEC.....	50
Tabela 5.2: Respostas das redes neurais e <i>cluster</i> escolhido	51
Tabela 5.3: Atributos dos conjuntos Bolonha urbano e Belém - Perimetral.....	51
Tabela 5.4: Ano zero e valor da meta final dos indicadores DEC e FEC para o cenário 1	52
Tabela 5.5: Metas para o cenário 1 ao longo dos 8 anos para o indicador DEC.....	53
Tabela 5.6: Comparação do DEC agregado com o obtido pela Metodologia Proposta para o cenário 1	54
Tabela 5.7: Metas para o cenário 1 ao longo dos 8 anos para o indicador FEC	55
Tabela 5.8: Comparação do FEC agregado com o obtido pela Metodologia Proposta para o cenário 1	56
Tabela 5.9: Atributos dos conjuntos Ulianópolis urbano e Ulianópolis não urbano	56
Tabela 5.10: Ano zero e valor da meta final dos indicadores DEC e FEC para o cenário 2.....	57
Tabela 5.11: Metas para o cenário 2 ao longo dos 8 anos para o indicador DEC.....	57
Tabela 5.12: Comparação do DEC agregado com o obtido pela Metodologia Proposta para o cenário 2	58
Tabela 5.13: Metas para o cenário 2 ao longo dos 8 anos para o indicador FEC	59
Tabela 5.14: Comparação do FEC agregado com o obtido pela Metodologia Proposta para o cenário 2.....	60
Tabela 5.15: Atributos dos conjuntos Paragominas urbano e Paragominas não urbano.....	61
Tabela 5.16: Ano zero e valor da meta final dos indicadores DEC e FEC para o cenário 3.....	61
Tabela 5.17: Metas para o cenário 3 ao longo dos 8 anos para o indicador DEC.....	62
Tabela 5.18: Comparação do DEC agregado com o obtido pela Metodologia Proposta para o cenário 3.	63
Tabela 5.19: Metas para o cenário 3 ao longo dos 8 anos para o indicador FEC	64
Tabela 5.20: Comparação do FEC agregado com o obtido pela Metodologia Proposta para o cenário 3	65
Tabela 5.21: Atributos dos conjuntos Água Azul do Norte e Canaã dos Carajás.....	66
Tabela 5.22: Ano zero e valor da meta final dos indicadores DEC e FEC para o cenário 4.....	66
Tabela 5.23: Metas para o cenário 4 ao longo dos 8 anos para o indicador DEC.....	67
Tabela 5.24: Comparação do DEC agregado com o obtido pela Metodologia Proposta para o cenário 4	68
Tabela 5.25: Metas para o cenário 4 ao longo dos 8 anos para o indicador FEC	69
Tabela 5.26: Comparação do FEC agregado com o obtido pela Metodologia Proposta para o cenário 4.....	70
Tabela 5.27: Atributos dos conjuntos Santa Maria não urbano e São Miguel do Guamá	71
Tabela 5.28: Ano zero e valor da meta final dos indicadores DEC e FEC para o cenário 5.....	71
Tabela 5.29: Metas para o cenário 5 ao longo dos 8 anos para o indicador DEC.....	72
Tabela 5.30: Comparação do DEC agregado com o obtido pela Metodologia Proposta para o cenário 5	73
Tabela 5.31: Metas para o cenário 5 ao longo dos 8 anos para o indicador FEC	74

Tabela 5.32: Comparação do FEC agregado com o obtido pela Metodologia Proposta para o cenário 5.....	75
Tabela 5.33: Atributos dos conjuntos Curionópolis e Eldorado dos Carajás.....	76
Tabela 5.34: Ano zero e valor da meta final dos indicadores DEC e FEC para o cenário 6.....	76
Tabela 5.35: Metas para o cenário 6 ao longo dos 8 anos para o indicador DEC.....	77
Tabela 5.36: Comparação do DEC agregado com o obtido pela Metodologia Proposta para o cenário 6.....	78
Tabela 5.37: Metas para o cenário 6 ao longo dos 8 anos para o indicador FEC.....	79
Tabela 5.38: Comparação do FEC agregado com o obtido pela Metodologia Proposta para o cenário 6.....	80
Tabela 5.39: Atributos dos conjuntos Abaetetuba e Igarapé Mirim.....	81
Tabela 5.40: Ano zero e valor da meta final dos indicadores DEC e FEC para o cenário 7.....	81
Tabela 5.41: Metas para o cenário 7 ao longo dos 8 anos para o indicador DEC.....	82
Tabela 5.42: Comparação do DEC agregado com o obtido pela Metodologia Proposta para o cenário 7.....	83
Tabela 5.43: Metas para o cenário 7 ao longo dos 8 anos para o indicador FEC.....	84
Tabela 5.44: Comparação do FEC agregado com o obtido pela Metodologia Proposta para o cenário 7.....	85

LISTA DE ABREVIACOES E SIGLAS

ANEEL	Agencia Nacional de Energia Eltrica.
DNAEE	Departamento Nacional de gua e Energia Eltrica.
CELPA	Centrais Eltricas do Par S.A.
PRODIST	Procedimentos de Distribuio de Energia Eltrica.
CEPEL	Centro de Pesquisas de Energia Eltrica.
DEC	Durao Equivalente de Interrupo por Consumidor.
FEC	Frequncia Equivalente de Interrupo por Consumidor.
DIC	Durao de Interrupes por Unidade Consumidora.
FIC	Frequncia de Interrupes por Unidade Consumidora.
DMIC	Durao Mxima das Interrupes por Unidade Consumidora.
RNA	Redes Neurais Artificiais.
kNN	<i>k-Nearest Neighbors.</i>
ERAP	Extenso da Rede Primria.
AREA	rea do conjunto.
PNI	Potncia Instalada.
CMM	Consumo Mdio Mensal.
NUC	Nmero de Unidades Consumidoras.

LISTA DE SÍMBOLOS

$ca(i)$	Número de consumidores atingidos na interrupção (i) do conjunto considerado.
$t(i)$	Duração de cada evento (i), no período de apuração.
C_S	Número total de consumidores do conjunto considerado.
T	Duração em anos do período de transição.
t	Ano variando de 1 a 8 anos.
DEC_0	Valor do patamar inicial do indicador DEC.
FEC_0	Valor do patamar inicial do indicador FEC.
$Meta_{DEC}$	Valor do patamar final do indicador DEC.
$Meta_{FEC}$	Valor do patamar final do indicador FEC.
$Meta_{DEC}^t$	Valor da meta anual, no ano t , do indicador DEC.
$Meta_{FEC}^t$	Valor da meta anual, no ano t , do indicador FEC.
$x_j(n)$	Padrão de entrada apresentado ao neurônio j na iteração n .
$e_j(n)$	Erro do neurônio j na iteração n .
$d_j(n)$	Saída esperada para o neurônio j na iteração n .
$o_j(n)$	Saída do neurônio j pertencente à camada de saída na iteração n .
$w_{ji}(n)$	Peso sináptico da saída do neurônio j para o neurônio i , na iteração n .
$y_i(n)$	Saída do neurônio i na iteração n .
$v_j(n)$	Sinal de ativação para o neurônio j na iteração n .
$E(n)$	Erro quadrático médio.
$\delta(n)$	Gradiente local.
$\omega_{ji}(n)$	Peso sináptico que conecta o neurônio i ao neurônio j na iteração n .
$\Delta\omega_{ji}(n)$	Correção do peso sináptico $\omega_{ji}(n)$.
η	Constante de aprendizado.
β	Constante de <i>momentum</i> .

RESUMO

A edição dos Procedimentos de Distribuição de Energia Elétrica (PRODIST) confirmou a análise comparativa entre empresas de distribuição de energia elétrica estabelecida pela Resolução 024/2000 como base para o estabelecimento das metas dos indicadores de continuidade DEC (Índice de Duração Equivalente de Interrupção por Consumidor) e FEC (Índice de Frequência Equivalente de Interrupção por Consumidor). O estabelecimento das metas é influenciado diretamente pela definição dos conjuntos de unidades consumidoras das empresas de distribuição de energia elétrica, portanto, é de interesse das empresas distribuidoras uma boa definição desses conjuntos. Este trabalho apresenta o desenvolvimento de uma metodologia baseada em técnicas de aprendizado de máquina que auxilie as empresas de distribuição de energia elétrica na tomada de decisão da definição de conjuntos de unidades consumidoras objetivando melhor compatibilidade das metas para os indicadores de DEC e FEC com a realidade dos conjuntos em relação às características ambientais e de infraestrutura da área de concessão da empresa distribuidora.

PALAVRAS-CHAVES: Qualidade de Energia Elétrica, Continuidade de Fornecimento, Indicadores de Continuidade, Aprendizado de Máquina.

ABSTRACT

The issue of Proceedings of Electric Power Distribution (PRODIST) confirmed the comparative analysis of the distribution companies of electricity established by Resolution 024/2000 as the basis for establishing the goals of continuity indicators DEC (Customer Average Duration Interruption Index) and FEC (Customer Average Frequency Interruption Index). The setting of targets is directly influenced by the definition of the sets of consumer units of the distribution companies of electricity is therefore of interest to the distribution companies a good definition of these sets. This work presents the development of a methodology based on machine learning techniques to assist companies in electric power distribution in decision making of the definition of sets of consumer units aimed at improving compatibility of the goals for the indicators of DEC and FEC with reality assemblies in relation to environmental characteristics and infrastructure of the concession area of the distribution company.

KEYWORDS: Power Quality, Distribution Reliability, Continuity Index, Learning Machine.

CAPÍTULO 1

INTRODUÇÃO

1.1 MOTIVAÇÕES

O contexto mundial de comercialização de energia elétrica identifica a energia elétrica como um produto ou serviço, partindo desta consideração, deve atender a níveis mínimos de qualidade. Kotler (1994) define qualidade como a soma das características e aspectos de um produto ou serviço que proporcionam a satisfação de necessidades, exigências e expectativas declaradas e implícitas dos consumidores. Tanure (2000) e Kagan (2009) destacam três aspectos que estão relacionados à qualidade da energia elétrica: qualidade do atendimento, qualidade do produto e qualidade do serviço. O primeiro aspecto diz respeito ao relacionamento comercial concessionária-clientes. Neste aspecto são levados em consideração os tempos de resposta às solicitações dos clientes, elaboração de estudos e orçamentos de serviços na rede de distribuição, etc. O segundo aspecto diz respeito à conformidade do produto, ou seja, o grau de perfeição com que a forma de onda de tensão é disponibilizada aos consumidores. O último aspecto trata da continuidade do fornecimento de energia elétrica que é mensurada pela duração e pela frequência com que ocorrem as interrupções no fornecimento de energia elétrica. Este trabalho tratará deste último aspecto.

O setor elétrico brasileiro possui hoje aproximadamente 61,5 milhões de unidades consumidoras em 99% dos municípios brasileiros, atendidos por 63 concessionárias de distribuição. As 63 empresas encontram-se divididas em 4008 conjuntos de unidades consumidoras (ANEEL, 2008). No Brasil, a Agência Nacional de Energia Elétrica (ANEEL) exige que as concessionárias executem a distribuição de energia elétrica, em suas respectivas áreas de concessão, com qualidade. Os níveis mínimos de qualidade exigidos pela ANEEL, no caso a continuidade do fornecimento de energia elétrica, são representados principalmente por dois indicadores: DEC (Duração Equivalente de Interrupção por Unidade Consumidora) e FEC (Frequência Equivalente de Interrupção por Unidade Consumidora) que representam a duração e a frequência com que ocorrem as interrupções de energia elétrica, respectivamente. Caso estes níveis mínimos de qualidade não sejam obedecidos, a concessionária é penalizada pela ANEEL, obrigando a concessionária a investir na qualidade do serviço prestado a sociedade.

Com a edição da Resolução nº 024, de 27 de janeiro de 2000, a Agência Nacional de Energia Elétrica estabeleceu um novo referencial no tratamento dos aspectos relativos à continuidade do fornecimento de energia elétrica aos consumidores finais de energia elétrica. Ao adotar a comparação entre as empresas como principal elemento para definição das metas dos indicadores de desempenho a serem praticadas, criou-se o ambiente necessário para que as empresas passem a buscar continuamente a evolução destes indicadores.

Para a criação de uma metodologia possível para o estabelecimento das metas dos indicadores de continuidade a serem cumpridas pelas concessionárias de distribuição a ANEEL levou em consideração o estudo realizado por Tanure (2000), onde as metas a serem propostas para as empresas de distribuição de energia elétrica são fixadas para cada conjunto de unidades consumidoras e a meta global da empresa definida como a média do desempenho dos conjuntos, ponderados pelo número de consumidores de cada conjunto. A metodologia empregada baseia-se em técnicas de agrupamentos, ou de formação de *clusters*, que permite identificar áreas geográficas com características técnicas, físicas e econômicas homogêneas, que, no entanto, podem apresentar padrões de desempenho diferenciados. Com base nestas diferenças de desempenho, podem ser identificados melhores padrões praticados pelos conjuntos agrupados em um mesmo *cluster* e tomando-o como referência, podem-se definir as metas de desempenho a serem adotadas para os demais conjuntos, com desempenhos inferiores. A ANEEL utiliza na caracterização dos conjuntos de unidades consumidoras 5 atributos: área do conjunto, extensão de rede primária, potência instalada, consumo médio mensal e número de unidades consumidoras. Esses cinco atributos são utilizados pelo algoritmo *k-means* na definição de 30 agrupamentos (30 *clusters*) de conjuntos de unidades consumidoras. Uma vez definidos os 30 *clusters*, são estabelecidos os padrões de referência dos indicadores de continuidade DEC e FEC para cada um dos 30 *clusters* e a partir destes as metas anuais dos indicadores de continuidade para cada um dos conjuntos de unidades consumidoras de energia elétrica pertencentes a um mesmo *cluster*.

A definição dos conjuntos de unidades consumidoras de energia elétrica por parte da empresa de distribuição de energia elétrica é uma etapa que exerce forte impacto na definição das metas dos indicadores de continuidade do fornecimento de energia elétrica, sendo interessante o desenvolvimento de metodologias que auxiliem na tomada de decisão nesta importante etapa.

1.2 OBJETIVO

Este trabalho tem o objetivo de desenvolver uma metodologia que auxilie as empresas de distribuição de energia elétrica na tomada de decisão da definição dos conjuntos de unidades consumidoras de energia elétrica.

A metodologia utilizará técnicas de aprendizado de máquinas para capturar as características de cada um dos agrupamentos de conjuntos de unidades consumidoras para a construção de um modelo de classificador que auxiliará na melhor configuração de conjuntos que gere metas para os indicadores DEC e FEC mais condizentes com a realidade de cada empresa distribuidora.

A metodologia possibilitará às empresas de distribuição de energia a realização de estudos de configurações de conjuntos de unidades consumidoras sem ter que ficar consultando a ANEEL para a definição das metas para cada nova configuração de conjuntos.

1.3 JUSTIFICATIVA

Os conjuntos de unidades consumidoras de energia elétrica exercem forte impacto na definição das metas dos indicadores de continuidade e por esta razão merecem muita atenção em seu processo de definição. Em muitos casos esse processo de definição é guiado pelo conhecimento do especialista do domínio da empresa de distribuição. O software ANABENCH utilizado pela ANEEL na definição das metas dos indicadores de continuidade é um software de uso exclusivo do órgão regulador, por esta razão não é possível realizar análises de configurações de conjuntos de unidades consumidoras por parte das empresas distribuidoras de energia elétrica que visem encontrar as melhores configurações que proporcionem metas para os indicadores de continuidade mais condizentes com a realidade de cada empresa. Por este razão, fica evidente o desenvolvimento de uma metodologia que auxilie as empresas de distribuição de energia elétrica na tomada de decisão na formação de conjuntos de unidades consumidoras nestas áreas.

1.4 REVISÃO BIBLIOGRÁFICA

Os estudos na área de qualidade da energia elétrica abordando o tema da qualidade do fornecimento são muito recorrentes como mostram os seguintes trabalhos.

Tanure (2000) apresenta a metodologia que foi adotada pela ANEEL na definição das metas dos indicadores de continuidade DEC e FEC. Neste trabalho é apresentado um pequeno histórico da legislação, a regulação do fornecimento de energia elétrica em outros países, descreve a metodologia com o uso de técnicas de agrupamento e análise comparativa de conjuntos de unidades consumidoras e, por fim, mostra um estudo de caso aplicando a metodologia em 4135 conjuntos de 56 concessionárias existentes na época do estudo.

Queiroz (2002) mostra em um estudo realizado na área de concessão de uma empresa de distribuição de energia elétrica que os conjuntos de unidades consumidoras desta empresa são mais bem caracterizados quando as variáveis de mercado são separadas das variáveis do sistema elétrico. Esta separação entre as variáveis é utilizada pelo órgão regulador da Argentina que de modo similar à ANEEL também utiliza o algoritmo *k-means* para realizar a tarefa de agrupamento.

Neto *et al* (2003) discute os impactos dos fatores não gerenciáveis sobre os indicadores de continuidade DEC e FEC que não foram levados em consideração na definição dos atributos que caracterizam os conjuntos de unidades consumidoras. Também são apresentados casos de estabelecimento de metas distintas para conjuntos atendidos por circuitos de uma mesma família e que apresentam características geográficas, sócio-econômicas e de densidade populacional semelhantes.

Sperandio (2003) propõe o uso de mapas auto-organizáveis na identificação de agrupamentos de unidades consumidoras. Os mapas auto-organizáveis são um tipo de rede neural onde é possível extrair informações intrínsecas de uma grande quantidade de dados por meio de mapas topológicos de duas dimensões. Este estudo foi realizado na área de concessão de uma empresa de distribuição de energia elétrica.

Tanure (2004) desenvolveu uma metodologia para o estabelecimento de metas dos indicadores de continuidade baseada nas técnicas de *cluster* dinâmico e análise de desempenho DEA (*Data Envelopment Analysis* – Análise Envoltória de Dados). A técnica de *cluster* dinâmico é aplicada na etapa de agrupamento de conjuntos de unidades consumidoras onde foram utilizadas 3 variáveis para caracterização dos conjuntos: área, número de unidades consumidoras e consumo total. Uma vez obtidos os agrupamentos de conjuntos é realizada a

análise de desempenho via metodologia DEA para a definição das metas dos indicadores de continuidade. As variáveis utilizadas nesta etapa foram: comprimento de rede, capacidade instalada e recursos reconhecidos para operação e manutenção (O&M).

Sperandio (2004) mostra um estudo realizado na área de concessão de uma empresa de distribuição onde foram utilizados diversos atributos além dos utilizados pela ANEEL na caracterização dos conjuntos de unidades consumidoras e a utilização de outro algoritmo para a etapa de agrupamento de conjuntos: os mapas auto-organizáveis (SOM – *Self Organizing Maps*). As variáveis escolhidas para as análises refletiam o mercado e o sistema elétrico da empresa de distribuição em estudo. Estas variáveis foram submetidas à análises estatísticas (descritiva, correlações e análise fatorial) com o objetivo que encontrar as variáveis mais relevantes para o estudo. A técnica SOM foi utilizada em uma metodologia de validação cruzada com o algoritmo *k-means* durante a etapa de agrupamento.

Pessanha *et al* (2004) descreve o software ANABENCH que utiliza a metodologia proposta por Tanure (2000) para definição das metas dos indicadores de continuidade DEC e FEC. Neste estudo foram utilizadas as informações de 5750 conjuntos de unidades consumidoras de todo o Brasil.

Pessanha (2006) propõe uma metodologia de definição de metas dos indicadores de continuidade coletivos baseada nas técnicas de análise envoltória de dados (DEA) composta por duas fases: na primeira fase é realizada a comparação entre as empresas distribuidoras definindo-se por meio da técnica DEA as metas globais dos indicadores de continuidade para cada empresa distribuidora. A segunda fase consiste no uso de técnicas de otimização via programação linear para definição das metas dos conjuntos de unidades consumidoras a partir da meta global da empresa distribuidora estabelecida na fase anterior.

Hunzilier *et al* (2008) propõe uma metodologia que permite a definição de metas dos indicadores de continuidade a partir dos planos reais de investimentos das empresas distribuidoras e dos custos operacionais. A metodologia foi aplicada à área de concessão da empresa CELPA, concessionária atuante na região norte do Brasil.

Rosa *et al* (2009) apresenta duas sugestões para o aprimoramento da metodologia empregada pela ANEEL na definição das metas dos indicadores de continuidade: aumento do número de atributos que caracterizam os conjuntos de unidades consumidoras e a substituição da técnica de agrupamento dos conjuntos. Os autores adicionaram aos 5 atributos já utilizados 3 novos atributos à análise: densidade de raios, média de precipitação anual e envelhecimento de ativos (grau de deterioração das redes de distribuição). Quanto à técnica o método *k-means* foi substituído pelo método de *Ward* na análise de agrupamento por os autores argumentarem

que o método *k-means* ser muito instável e depender de um parâmetro chamado de “semente na inicialização” na formação dos agrupamentos de conjuntos.

Tostes *et al* (2009) analisam uma metodologia de inclusão de novos atributos voltados para as peculiaridades da Região Norte, mais especificamente o estado do Pará. O estudo foi realizado na área de concessão das Centrais Elétricas do Pará (CELPA) e consistia na adição de cinco novos atributos que abrangem características de infra-estrutura, ambiental e topologia da rede elétrica. O estudo mostrou que os 5 atributos utilizados atualmente explicam apenas 59,48% do modelo proposto pela ANEEL para a área de concessão da CELPA e a inclusão de 5 novos atributos no modelo proporcionaram uma melhor caracterização nos conjuntos de unidades consumidoras da CELPA onde 99,92% do modelo da ANEEL foram explicados.

Santos (2009) apresenta os instrumentos utilizados pela ANEEL para garantir o nível de qualidade desejada do serviço de distribuição de energia elétrica e faz um histórico e uma avaliação da construção de alguns desses instrumentos. Neste trabalho, o autor divide os instrumentos em indicadores e indutores de qualidade. Os indicadores são definidos como índices que permitem a mensuração da qualidade do serviço de distribuição de energia elétrica por meio da comparação com uma meta estabelecida ou como resultados de outras empresas. Os indutores são definidos como instrumentos que devem levar os serviços de distribuição de energia elétrica a uma qualidade considerada aceitável.

1.5 ESTRUTURA DA DISSERTAÇÃO

A presente dissertação está dividida em 6 capítulos. O primeiro capítulo, Introdução, discute questões acerca das motivações, objetivos, justificativa, revisão bibliográfica e a forma de organização deste trabalho.

No capítulo 2, um breve histórico da legislação que norteia o assunto é apresentado: os indicadores de continuidade de caráter coletivo e individual. Os principais pontos da metodologia utilizados atualmente pela ANEEL na definição das metas dos indicadores de continuidade são descritos. Por fim, são expostas as principais alterações realizadas na metodologia com a edição dos Procedimentos de Distribuição de Energia Elétrica (PRODIST).

No capítulo 3, serão apresentadas as principais características da tarefa de classificação mostrando sua importância dentro da área de aprendizado de máquina. A seguir, serão apresentadas duas técnicas que realizam a tarefa de classificação: As redes neurais do

tipo *Perceptron* de Múltiplas Camadas e o Algoritmo k-vizinhos mais próximos (kNN - *k-Nearest Neighbors*).

No capítulo 4, propõe-se o uso de uma metodologia que faça uso da tarefa de classificação para auxiliar as empresas de distribuição de energia elétrica na tomada de decisão na formação de conjuntos de unidades consumidoras. A metodologia consiste da construção de um classificador por meio de uma das técnicas apresentadas no capítulo 3.

No capítulo 5, será apresentado um estudo de caso onde a metodologia proposta no capítulo anterior será utilizada. O estudo de caso utilizará dados de conjuntos de unidades consumidoras pertencentes a área de concessão da CELPA, empresa distribuidora de energia elétrica cuja área de concessão engloba todo o estado do Pará.

No capítulo 6 são resumidas as principais contribuições do trabalho e feitas algumas sugestões para trabalhos futuros.

CAPÍTULO 2

A METODOLOGIA DA ANEEL

2.1 CONSIDERAÇÕES INICIAIS

Este capítulo apresenta a metodologia que atualmente a Agência Nacional de Energia Elétrica (ANEEL) adota para definição das metas dos indicadores de continuidade no fornecimento de energia elétrica. Em um primeiro momento, será apresentado um breve histórico da legislação que norteia o assunto. Logo após, os principais pontos da metodologia atualmente vigente serão descritos. Por fim, serão mostrados os principais pontos tratados pelos Procedimentos de Distribuição de Energia Elétrica (PRODIST) na definição das metas dos indicadores de continuidade.

2.2 UM BREVE HISTÓRICO DA LEGISLAÇÃO

A preocupação com a regulação da continuidade do fornecimento de energia elétrica não é um fato recente. A primeira legislação que abordou o assunto foi a Portaria nº 46, de 17 de abril de 1978, do antigo Departamento Nacional de Águas e Energia Elétrica (DNAEE). A portaria tratava das disposições referentes à continuidade de serviço do fornecimento de energia elétrica a serem observadas pelas empresas distribuidoras de energia elétrica. A Portaria definia dois índices de continuidade: o Índice de Duração Equivalente de Interrupção por Consumidor (DEC), medido em horas, e o Índice de Frequência Equivalente de Interrupção por Consumidor (FEC). Estes índices mensuram a duração e a frequência das ocorrências de interrupção de energia elétrica, respectivamente, e eram calculados pelas seguintes expressões:

$$DEC = \frac{\sum_{i=1}^n ca(i)t(i)}{Cs} \quad (2.1)$$

$$FEC = \frac{\sum_{i=1}^n ca(i)}{Cs} \quad (2.2)$$

Onde:

$ca(i)$: Número de consumidores atingidos na interrupção (i) do conjunto considerado;

$t(i)$: Duração de cada evento (i), no período de apuração;

i : Número de interrupções com duração superior a 3 minutos, variando de 1 a n ;

C_s : Número total de consumidores do conjunto considerado.

A apuração dos indicadores era realizada trimestralmente e anualmente e para efeito de apuração destes indicadores deveriam ser levadas em conta todas as interrupções com duração maior que 3 minutos e independente se eram oriundas de natureza programada, acidental, manobras, etc. Interrupções com duração menores que 3 minutos, ou provocadas devido à falha nos equipamentos das instalações do consumidor, ou decorrentes de racionamento de energia elétrica não eram apuradas.

Os indicadores DEC e FEC são de caráter coletivo e neste sentido a Portaria define o conceito de conjunto de consumidores que para Tanure (2000, p. 13) são agrupamentos de consumidores vinculados a uma dada área física, o que corresponde a um subconjunto da área de concessão. Consumidores localizados em áreas urbanas não contíguas não poderiam ser agrupados em um mesmo conjunto com exceção dos consumidores atendidos em tensão igual ou superior a 69 kV e por sistemas subterrâneos de distribuição com secundário reticulado ou radial.

A Portaria definia os valores máximos anuais para os indicadores coletivos de continuidade. Para consumidores atendidos por tensão de transmissão ou subtransmissão igual ou superior a 69 kV foram estabelecidos os valores de 15 horas, para o indicador DEC, e 25 interrupções para o indicador FEC. Para consumidores alimentados por tensões de transmissão ou subtransmissão inferiores a 69 kV ou rede primária ou secundária de distribuição foram estabelecidos os valores apresentados na Tabela 2.1.

Tabela 2.1: Valores dos Padrões Coletivos estabelecidos pela Portaria DNAEE nº 46/78

Características dos conjuntos de unidades consumidoras Atendidas em tensão inferior a 69 kV	DEC (horas)	FEC (ocorrências)
Atendidos por sistema subterrâneo com secundário reticulado	15	20
Atendidos por sistema subterrâneo com secundário radial	20	25
Atendido por sistema aéreo, com mais de 50.000 consumidores	30	45
Atendido por sistema aéreo, com número de consumidores entre 15.000 e 50.000	40	50
Atendido por sistema aéreo, com número de consumidores entre 5.000 e 15.000	50	60
Atendido por sistema aéreo, com número de consumidores entre 1.000 e 5.000	60	70
Atendido por sistema aéreo, com menos de 1.000 consumidores	120	90

Fonte: Pessanha (2006, p.40)

A Portaria estabelecia os valores máximos anuais por interrupções no fornecimento de energia elétrica a cada consumidor. Estes valores estão descritos na Tabela 2.2.

Tabela 2.2: Valores dos Padrões Individuais estabelecidos pela Portaria DNAEE nº 46/78

Características das unidades consumidoras	DEC (horas)	FEC (ocorrências)
Consumidor atendido por sistema subterrâneo	30	35
Consumidor atendido em tensão superior a 69 kV	30	40
Consumidor atendido em tensão de transmissão ou subtransmissão inferior a 69 kV ou em tensão primária de distribuição, cuja unidade de consumo não se situe na zona rural	80	70
Consumidor atendido em tensão de secundária de distribuição e pertencente a conjunto com mais de 1000 consumidores, cuja unidade de consumo não se situe em zona rural	100	80
Consumidor localizado em zonal rural atendido por sistema aéreo de distribuição, ou pertencente a qualquer conjunto com menos de 1.000 consumidores.	150	120

Fonte: Pessanha (2006, p.40)

Apesar da Portaria DNAEE nº46/78 ser um marco na regulamentação da continuidade do fornecimento de energia elétrica, ela apresentava diversas limitações dentre as quais podem ser citadas, segundo Tanure (2000):

- Não existência de punições para as empresas de distribuição que não cumprissem os padrões propostos;
- Dificuldade de acompanhar, fiscalizar e detalhar os indicadores de continuidade por parte do DNAEE devido à falta de recursos;
- Os procedimentos adotados pelas concessionárias para coleta e tratamento dos indicadores não eram isentos de erro resultando em baixa confiabilidade dos índices apurados.

Essas limitações aliada a não reivindicação por parte dos consumidores de seus direitos com relação ao padrão de qualidade no fornecimento de energia elétrica criaram um cenário onde os indicadores de continuidade pouco contribuíram para uma real melhoria do sistema (TANURE, 2000).

O DNAEE, em 22 de março de 1993, edita a Portaria nº 163 que constituiu um grupo de trabalho cujo objetivo era a ampliação da abrangência dos indicadores de qualidade. Na proposta do grupo de trabalho os indicadores DEC e FEC foram mantidos, mas a apuração foi refinada e passou-se a contabilizar interrupções a partir de 1 minuto. Além disso, propunha-se

um índice que medisse a satisfação do consumidor por meio de uma pesquisa de opinião e foram estabelecidos mais dois novos indicadores que avaliariam a potência interrompida por conjunto. Tais propostas deveriam ser implantadas em três etapas, mas sua implantação foi suspensa devido ao processo de reestruturação do setor elétrico brasileiro (TANURE, 2000).

Em meados dos anos 90 iniciou-se o processo de reestruturação do setor elétrico brasileiro. Neste novo cenário ocorreu a desverticalização do setor elétrico brasileiro onde geração, transmissão, distribuição e comercialização de energia elétrica passaram a ser negócios separados. Na distribuição de energia elétrica foram realizadas as privatizações das empresas distribuidoras de energia elétrica onde foi mantido o regime de monopólio regulado. Neste novo cenário, foi necessária a elaboração de novos contratos de concessão bem como a revisão da legislação de regulação da continuidade do fornecimento de energia elétrica de modo a corrigir as limitações constatadas na Portaria DNAEE nº 46/1978.

Em 27 de janeiro de 2000, a ANEEL edita a Resolução nº 024 que estabelece as disposições relativas à continuidade da distribuição de energia às unidades consumidoras. A Resolução incorporou todos os avanços dos regulamentos e contratos assinados anteriormente (TANURE, 2000). Nela são estabelecidos:

- Os critérios para apuração com maior precisão dos indicadores de continuidade, que passaram a ser apurados com maior confiabilidade por meio de procedimentos sujeitos a auditorias;
- As condições da análise comparativa entre as empresas com a definição de uma metodologia para determinação das metas dos indicadores e;
- O cálculo das penalidades aplicadas às empresas de distribuição de energia elétrica pelo não cumprimento das metas.

A Resolução nº24/2000 passou por duas alterações de seus dispositivos: a primeira em 13 de fevereiro de 2003 com a edição da Resolução nº 75 e a última em 28 de novembro de 2005 com a publicação da Resolução nº 177. Na sua forma final a Resolução nº 24/2000 trata dos seguintes aspectos:

2.2.1 Das Definições

A Resolução nº24/2000 define os conceitos de unidade consumidora e conjunto de unidades consumidoras. Uma unidade consumidora é um *“conjunto de instalações e equipamentos elétricos caracterizado pelo recebimento de energia elétrica em um só ponto de*

entrega, com medição individualizada e correspondente a um único consumidor". Um conjunto de unidades consumidoras é *“qualquer agrupamento de unidades consumidoras, global ou parcial, de uma mesma área de concessão de distribuição, definido pela concessionária ou permissionária e aprovado pela ANEEL”*.

A Resolução define os seguintes critérios para a formação dos conjuntos de unidades consumidoras:

- A definição de conjunto deverá permitir a identificação geográfica das unidades consumidoras;
- Unidades consumidoras localizadas em áreas não contíguas não poderão ser agrupadas num mesmo conjunto;
- Para o estabelecimento dos padrões dos indicadores de continuidade, para cada conjunto, deverão ser considerados os seguintes atributos: a área do conjunto (em km²), extensão da rede primária (em km). Média mensal da energia consumida nos últimos 12 meses (em kWh), total de unidades consumidoras atendidas, potência instalada (em kVA) e se o conjunto pertence ao sistema interligado ou isolado.

2.2.2 Dos Indicadores de Continuidade

Quanto aos indicadores a Resolução estabelece dois tipos: coletivo e individual.

Os indicadores de caráter coletivo são utilizados pelo órgão regulador para avaliar o desempenho das várias empresas de distribuição existentes no país e expressam a qualidade média da continuidade do serviço de distribuição de energia para todas as unidades consumidoras de um determinado conjunto.

Os indicadores de caráter individual apontam diretamente aos consumidores o padrão de desempenho da concessionária quanto à prestação do serviço adequado, uma vez que esses indicadores são aqueles efetivamente percebidos pelos mesmos.

Os indicadores coletivos que deverão ser apurados para cada conjunto são: Duração Equivalente de Interrupção por Unidade Consumidora (DEC) e Frequência Equivalente de Interrupção por Unidade Consumidora (FEC).

Para o cálculo desses indicadores são utilizadas as mesmas expressões definidas pela Portaria nº 46/1978, do DNAEE. A única diferença está no denominador das expressões que com a definição do termo unidade consumidora contempla a possibilidade de apuração individualizada e não mais por blocos de consumidores afetados pelo desligamento (Tanure, 2000, p. 20).

Para a apuração destes indicadores são consideradas interrupções com duração maior ou igual a 3 minutos.

O indicador DEC representa o intervalo de tempo que, em média, cada unidade consumidora do conjunto considerado ficou sem fornecimento de energia elétrica, durante o período de observação. A duração medida por meio deste indicador reflete as facilidades existentes para se recuperar um sistema após uma interrupção (veículos, comunicação, qualificação do pessoal, possibilidade de recomposição, grau de automação da rede, etc).

O indicador FEC representa o número de interrupções que, em média, cada unidade consumidora do conjunto em questão sofreu durante o período de observação. Este indicador reflete a fragilidade do sistema de distribuição em relação ao meio ambiente no qual está inserido e a degradação do sistema por envelhecimento e/ou falta de manutenção adequada.

Os indicadores individuais a serem apurados são: Duração de Interrupção por Unidade Consumidora ou por Ponto de Conexão (DIC), Frequência de Interrupção por Unidade Consumidora ou por Ponto de Conexão (FIC) e Duração Máxima de Interrupção Contínua por Unidade Consumidora ou por Ponto de Conexão (DMIC).

Para os cálculos destes indicadores são utilizadas as seguintes expressões:

$$DIC = \sum_{i=1}^n t(i) \quad (2.3)$$

$$FIC = n \quad (2.4)$$

$$DMIC = t(i)_{\max} \quad (2.5)$$

Onde:

i: índice de interrupções da unidade consumidora ou do ponto de conexão, ponto de apuração, variando de 1 a n;

n: número de interrupções da unidade consumidora ou do ponto de conexão considerada, por período de apuração;

t(i): tempo de duração da interrupção (i) da unidade consumidora ou do ponto de conexão considerada, no período de apuração;

t(i)_{max}: valor correspondente ao tempo da máxima duração de interrupção (i), no período de apuração, verificada na unidade consumidora ou no ponto de conexão considerado, expresso em horas e centésimos de horas.

2.3 A METODOLOGIA DE DEFINIÇÃO DAS METAS

A Resolução nº 024/2000 estabelece as linhas gerais para definição das metas dos indicadores de continuidade. Esta definição está baseada em uma análise comparativa dos desempenhos dos conjuntos das unidades consumidoras haja vista que o setor de distribuição de energia elétrica é monopolista e as concessionárias não enfrentam qualquer tipo de contestação quanto à qualidade do serviço que fornece aos consumidores sendo necessário que o órgão regulador crie mecanismos que emulem um ambiente competitivo entre os agentes que atuam neste segmento.

A metodologia atualmente utilizada pela ANEEL está baseada no trabalho realizado por Tanure (2000) e consiste em agrupar conjuntos semelhantes com base na técnica de análise de agrupamentos (*cluster analysis*) e dentro de cada agrupamento formado (*cluster*), identificar o conjunto de melhor desempenho, tomando-o como referência. A partir do desempenho deste conjunto serão estabelecidas as metas de desempenho para os demais conjuntos pertencentes ao mesmo agrupamento.

Para o desenvolvimento de seu estudo Tanure (2000) utilizou um banco de dados contendo as informações de 4135 conjuntos de 56 concessionárias que existiam na época do estudo. Para cada um desses conjuntos foram informados os dados de cinco atributos: a área do conjunto (em km²), extensão da rede primária (em km), Média mensal da energia consumida nos últimos 12 meses (em kWh), total de unidades consumidoras atendidas, potência instalada (em kVA) e mais a informação se os conjuntos pertenciam ao sistema interligado ou isolado. O uso de apenas cinco atributos para caracterizar os conjuntos foi justificado pelo fato que na época do estudo muitas concessionárias não tinham bancos de dados consolidados com informações detalhadas sobre seus sistemas em suas áreas de concessão. A informação se o conjunto pertence a um sistema interligado ou isolado não constitui um atributo sendo levado em consideração apenas na etapa de definição das metas dos conjuntos.

Antes de serem apresentados à etapa de agrupamento (ou clusterização) os dados são submetidos a análise de *outliers*, normalização e transformação logarítmica.

Segundo Pessanha *et al* (2004), a análise de *outliers* consiste na identificação de dados discrepantes que possam prejudicar o processo de formação de agrupamento. Na literatura técnica existem muitas técnicas de análise de *outliers* sendo empregado o uso do box-plot uma técnica explanatória de dados para identificação de dados discrepantes. No *box-plot* são definidos dois limites denominados cercas inferior e superior. Qualquer dado

localizado fora destes limites é considerado um dado discrepante. Na metodologia da ANEEL a análise de *outliers* é feita em cada um dos cinco atributos caracterizadores de um conjunto de unidade consumidora. Se algum dado pertencentes aos cinco atributos for considerado discrepante o conjunto é considerado atípico e é retirado da etapa de formação dos agrupamentos. Também são considerados conjuntos atípicos conjuntos que apresentarem indicadores de continuidade nulos.

Para Sperandio (2004), o processo de normalização é realizado para compatibilizar a variação e ordem de grandeza entre elas. Tal medida é importante, pois grandes variações podem ser interpretadas como um peso ou ponderador diferenciado entre os atributos. A normalização por variância foi utilizada nesta metodologia.

Por fim, os dados foram submetidos à transformação logarítmica, pois nos primeiros experimentos realizados por Tanure (2000) constatou-se que a utilização dos dados sem transformação resultava na formação de poucos agrupamentos com muitos conjuntos e diversos agrupamentos com poucos conjuntos.

O algoritmo escolhido por Tanure (2000) para a realização da tarefa de agrupamento foi o conhecido algoritmo estatístico *k-means* que é um método de agrupamento não hierárquico, pois não existe encadeamento entre os dados agrupados, por esta razão neste tipo de método não há a formação de dendogramas ao final da execução da técnica. De acordo com Mingoti (2007, p. 192), os algoritmos utilizados nas técnicas de agrupamento não hierárquicas são do tipo iterativos e possuem uma maior capacidade de análise de dados e por esta razão são recomendadas quando a quantidade de dados a serem agrupados é grande.

Para Pessanha *et al* (2004) e Mingoti (2007), o algoritmo *k-means* tem como objetivo encontrar diretamente uma partição de n elementos (neste caso os elementos são os conjuntos de unidades consumidoras), caracterizados por p atributos, em k subconjuntos, mutuamente exclusivos, denominados de *clusters*, de modo que a partição satisfaça dois requisitos: coesão interna e isolamento, ou seja, os elementos de um mesmo subconjunto sejam semelhantes entre si, mas ao mesmo tempo diferentes dos elementos pertencentes aos outros subconjuntos.

O algoritmo *k-means* é composto de quatro etapas:

- Etapa 1: Processo de partição é iniciado com a escolha dos os k centróides;
- Etapa 2: Cada elemento do conjunto de dados é comparado com cada centróide inicial. Tal comparação é feita por meio de uma medida de distância, que geralmente é a distância Euclidiana. O elemento é alocado ao agrupamento cuja distância é menor.

- Etapa 3: Depois de aplicar a etapa anterior para cada um dos elementos do conjunto de dados recalculam-se os valores dos centróides para cada novo grupo formado, e repete-se o passo anterior, considerando os centróides destes novos agrupamentos.
- Etapa 4: As etapas 2 e 3 devem ser repetidas até que todos os elementos estejam “bem alocados” em seus grupos, isto é, até que nenhuma realocação de elementos seja necessária.

Por ser uma técnica de agrupamento do tipo não hierárquica o algoritmo *k-means* necessita que o número de partições *k* seja conhecido a priori. Pessanha *et al* (2004) recomenda testar diferentes valores de *k* e observar os valores da dispersão intra-cluster e da dispersão inter cluster. A melhor solução é quando cada elemento corresponde a um único *cluster*, o que corresponde à máxima dispersão inter cluster, entretanto, tal proposta não atende ao objetivo do estudo que seria a redução do volume de dados (TANURE, 2000). No seu trabalho, Tanure (2000) avalia apenas a razão da dispersão intra cluster sobre a dispersão total dos dados com o aumento do número de *cluster* e estabelece 30 partições para a metodologia.

Uma vez realizada a etapa de agrupamento passa-se para a etapa de definição das metas dos conjuntos de unidades consumidoras. O modelo adotado para análise comparativa de desempenho foi o *Yardstick Competition* que de acordo com Tanure:

Modelo no qual as empresas são estimuladas a adotarem um desempenho baseado em um comportamento padrão. A fixação deste padrão decorre das análises de desempenho dos agentes que compõem o universo considerado e do padrão que se julga razoável atingir. Nesta abordagem, as empresas que se situam abaixo da referência têm uma margem de folga para seu desempenho. As empresas situadas acima desta marca devem buscar melhorar seus indicadores para aproximar-se do valor de referência. (TANURE, 2000, p.79).

A definição da marca de referência pode-se dá de diversas formas, pode-se, por exemplo, adotar a média do desempenho dos elementos do universo considerado. Entretanto, para o caso das distribuidoras de energia elétrica constatou-se uma assimetria na distribuição dos valores de DEC e FEC das famílias de *clusters*, daí justificou-se o uso de medidas de posição (mediana, percentis, decis, etc) para a escolha do padrão de desempenho. A ANEEL

adotou como marca de referência o segundo decil (percentil 20%) para conjuntos pertencentes ao sistema interligado e a mediana para conjuntos pertencentes ao sistema isolado.

Uma vez estabelecida a marca de referência deve-se determinar qual o horizonte de tempo que se pretende fazer a transição do patamar atual de desempenho para o patamar desejado. Tanure (2000) estabeleceu o período entre duas revisões tarifárias, ou seja, oito anos para o alcance do patamar desejado.

Para a definição das metas anuais durante o período de duas revisões tarifárias (8 anos) são utilizadas as seguintes expressões:

$$Meta_{DEC}^t = DEC_0 \left(\frac{Meta_{DEC}}{DEC_0} \right)^{\frac{t}{T}}, \text{ se } DEC_0 > Meta_{DEC} \quad (2.6)$$

$$Meta_{DEC}^t = Meta_{DEC} \quad , \text{ se } DEC_0 \leq Meta_{DEC} \quad (2.7)$$

$$Meta_{FEC}^t = FEC_0 \left(\frac{Meta_{FEC}}{FEC_0} \right)^{\frac{t}{T}}, \text{ se } FEC_0 > Meta_{FEC} \quad (2.8)$$

$$Meta_{FEC}^t = Meta_{FEC} \quad , \text{ se } FEC_0 \leq Meta_{FEC} \quad (2.9)$$

Onde:

T : é a duração em anos do período de transição, ou seja, 8 anos;

t : o ano em questão variando $t = 1, \dots, 8$ anos;

DEC_0 e FEC_0 : são os valores dos patamares iniciais dos indicadores de continuidade.

$Meta_{DEC}$ e $Meta_{FEC}$: são os valores do patamar final dos indicadores de continuidade.

$Meta_{DEC}^t$ e $Meta_{FEC}^t$: são valores das metas anuais dos indicadores de continuidade.

Os valores das metas anuais são obtidos pelas expressões (2.6) a (2.9) e arredondadas para o número inteiro imediatamente superior.

O Centro de Pesquisas de Energia Elétrica (CEPEL) desenvolveu o software ANABENCH baseado na metodologia proposta por Tanure (2000). O software é exclusivamente utilizado pela ANEEL na definição das metas dos indicadores de continuidade DEC e FEC, por esta razão não é possível realizar estudos de configurações de conjuntos de unidades consumidoras por parte das empresas distribuidoras de energia elétrica visando à diminuição das transgressões das metas dos indicadores de continuidade.

2.4 O MÓDULO 8 DOS PROCEDIMENTOS DE DISTRIBUIÇÃO (PRODIST)

Em 16 de dezembro de 2008, é aprovado por meio da Resolução nº 345 os Procedimentos de Distribuição de Energia Elétrica no Sistema Elétrico Nacional (PRODIST) que são documentos que tratam do relacionamento entre empresas distribuidoras de energia elétrica e a ANEEL no que diz respeito ao intercâmbio de informações e disciplinam o relacionamento entre as empresas distribuidoras e os demais agentes do setor elétrico (consumidores e centrais geradoras) conectados aos sistemas de distribuição, que incluem redes e linhas em tensão inferior a 230kV. O PRODIST é formado por 8 módulos e uma cartilha:

- Módulo 1 – Introdução;
- Módulo 2 – Planejamento da Expansão do Sistema de Distribuição;
- Módulo 3 – Acesso ao Sistema de Distribuição;
- Módulo 4 – Procedimentos Operativos do Sistema de Distribuição;
- Módulo 5 – Sistemas de Medição;
- Módulo 6 – Informações Requeridas e Obrigações;
- Módulo 7 – Cálculo de Perdas na Distribuição;
- Módulo 8 – Qualidade da Energia Elétrica; e
- Cartilha de Acesso ao Sistema de Distribuição.

O Módulo 8: Qualidade da Energia Elétrica estabelece procedimentos relativos à qualidade da energia elétrica e trata tanto da qualidade do produto como da qualidade do serviço prestado. Na qualidade do produto são definidos os termos, caracterizados os fenômenos, parâmetros e valores de referência que afetam a forma de onda senoidal e o valor da tensão em regime permanente senoidal. Na qualidade do serviço é estabelecida a metodologia para apuração dos indicadores de continuidade e dos tempos de atendimento às ocorrências emergenciais dos consumidores, definindo padrões, metas e responsabilidades. Muitos dispositivos da Resolução nº 024, de 27 de janeiro de 2000, foram confirmados no Módulo 8 do PRODIST com exceção dos artigos 3º, 7º, 8º, 10º, 13º, 17º e 21º que foram alterados.

Em 15 de dezembro de 2009, foi aprovada por meio da Resolução nº 395 a primeira revisão no PRODIST que revogou diversas Resoluções dentre as quais: a Resolução nº 345 de 16 de dezembro de 2008 e a Resolução nº 024, de 27 de janeiro de 2000, deste modo a regulamentação sobre qualidade da energia elétrica está disposta no Módulo 8 do PRODIST.

A primeira revisão do módulo 8 definiu um ponto da metodologia que era alvo de diversas críticas por parte das empresas distribuidoras: o critério de formação dos conjuntos de unidades consumidoras. Antes da revisão não existia um critério para formação de um conjunto, a única exigência dizia respeito a não se agrupar em um mesmo conjunto unidades consumidoras situadas em áreas não contíguas. Com a primeira revisão do PRODIST o critério de formação foi definido: as subestações que possuam primário em alta tensão e secundário em baixa tensão. Um conjunto abrange as redes de média tensão a jusante da subestação e de propriedade da distribuidora. As unidades consumidoras atendidas em baixa tensão e média tensão deverão pertencer ao mesmo conjunto da subestação que as atendam.

2.5 CONSIDERAÇÕES FINAIS

Este capítulo apresentou um breve histórico da legislação que rege a regulação da continuidade do fornecimento de energia elétrica no Brasil desde a criação do antigo Departamento Nacional de Águas e Energia Elétrica (DNAEE) até aos dias atuais com a Agência Nacional de Energia Elétrica (ANEEL). A metodologia utilizada pelo órgão regulador para a definição das metas dos indicadores de continuidade foi descrita e, por fim, foram apresentados os Procedimentos de Distribuição de Energia Elétrica no Sistema Elétrico Nacional (PRODIST) em especial o Módulo 8: Qualidade da Energia Elétrica que estabelece atualmente os procedimentos relativos à qualidade da energia elétrica no que tange tanto à conformidade (energia elétrica como produto) como para o fornecimento de energia (energia elétrica como serviço).

CAPÍTULO 3

ANÁLISE DE CLASSIFICADORES

3.1 CONSIDERAÇÕES INICIAIS

Neste capítulo, será abordada a tarefa da classificação. Classificar elementos sempre foi uma preocupação do ser humano desde o início da civilização, e trabalhos que propunham métodos matemáticos de classificação existiam desde o início do século, entretanto, o uso de tais métodos na prática demandava bastante tempo e o seu uso só se tornou viável a partir dos anos 60 com o uso crescente de computadores. O objetivo da tarefa de classificação é construção de um modelo, chamado classificador, que possua a capacidade de prever a classe de uma nova amostra, com resolução satisfatória. A tarefa de classificação é utilizada em diversos campos do conhecimento humano tais como: reconhecimento de padrões, identificação biométrica, visão de máquina, reconhecimento de caracteres, diagnóstico assistido por computador, mineração de dados e descoberta de conhecimento.

Neste capítulo serão apresentadas inicialmente as principais características da tarefa de classificação mostrando sua importância dentro da área de aprendizado de máquina. Em seguida, serão apresentadas duas técnicas que realizam a tarefa de classificação: As redes neurais do tipo *Perceptron* de Múltiplas Camadas e o Algoritmo k-vizinhos mais próximos (kNN - *k-Nearest Neighbors*).

3.2 A CLASSIFICAÇÃO NO CONTEXTO DA APRENDIZAGEM DE MÁQUINA

A tarefa de classificação é uma entre várias tarefas da área de aprendizado de máquina, segundo Mitchell (1997) *apud* Ferrero (2009) esta área tem por objetivo o desenvolvimento de técnicas computacionais que simulem o processo de aprendizagem e a construção de sistemas que possam adquirir conhecimento de forma automática. O classificador é um destes sistemas e adquire conhecimento (aprendizagem) por meio de um conjunto de exemplos chamado de conjunto de treinamento. Para Ferrero (2009), existem três tipos de aprendizagem: supervisionada, não-supervisionada e semi supervisionada. Na aprendizagem supervisionada têm-se um conjunto de exemplos com N elementos onde se sabe a priori a

qual classe cada uma dos exemplos pertence. No aprendizado não supervisionado não se conhece a priori a qual classe cada uma dos exemplos pertence. Já no aprendizado semi-supervisionado têm-se poucos exemplos em que se conhece a qual classe pertence e muitos sem classes definida.

Segundo Ferrero (2009), no aprendizado supervisionado o conjunto de treinamento é representado por uma tabela denominada tabela atributo-valor. Os atributos são um conjunto de informações que caracterizam um determinado exemplo, já o valor é a própria classes a qual pertence o exemplo. A Tabela 3.1 apresenta um exemplo de uma tabela atributo-valor.

Tabela 3.1: Tabela Atributo-Valor

Exemplo	A₁	A₂	...	A_M	Classe
1	X ₁₁	X ₁₂	...	X _{1M}	Y ₁
2	X ₂₁	X ₂₂	...	X _{2M}	Y ₂
...
N	X _{N1}	X _{N2}	...	X _{NM}	Y _N

A Tabela 3.1 apresenta um conjunto de treinamento com N exemplos, onde os termos A₁, A₂, ..., A_M são os M atributos que caracterizam um determinado exemplo. Os termos Y₁, Y₂, ..., Y_N são os valores das classes relativas aos N exemplos contidos no conjunto de treinamento e X_{NM} denotam o valor do atributo M referente ao exemplo N. Ferrero (2009) afirma que os valores de Y_N podem ser nominais ou numéricos. Quando os valores são nominais o processo é chamado de classificação, quando são numéricos, o processo é chamado de regressão.

Ferrero (2009) classifica os algoritmos de aprendizado supervisionados em *eager* ou *lazy*. Os algoritmos do tipo *eager* utilizam o conjunto de treinamento apenas para a construção do modelo do classificador e depois são descartados, pois apenas o modelo é necessário para realizar a tarefa de classificação. Os algoritmos do tipo *lazy* não constroem um modelo de classificador, sendo necessário que os exemplos do conjunto de treinamento sejam lembrados. As redes neurais do tipo *Perceptron* de Múltiplas Camadas são exemplos de classificadores do tipo *eager* enquanto que o algoritmo kNN é um exemplo de classificador tipo *lazy*.

Michie (1994) lista quatro características importantes de um classificador: acurácia, velocidade de resposta, interpretabilidade e tempo de aprendizado. A acurácia, ou taxa de erro, esta relacionada com a habilidade do classificador predizer corretamente a classe de amostras desconhecidas. Estas amostras desconhecidas estão reunidas em um conjunto disjunto dos elementos do conjunto de treinamento e é chamada de conjunto de teste. A

velocidade de resposta está relacionada com a velocidade com que um determinado tipo de classificador apresenta a classificação de um novo exemplo sendo em alguns casos, a característica mais relevante. (um classificador com 90% de acurácia pode ser preterido um classificador de 95% de acurácia se o primeiro for 100 vezes mais rápido em resposta que o segundo). A interpretabilidade diz respeito ao modelo que o classificador constrói e este deve ser compreensível ao ser humano. O tempo de aprendizado é importante em circunstâncias onde o ambiente no qual está inserido o classificador muda rapidamente, nestas situações é importante que o classificador aprenda em tempo real e com um número bem menor de exemplos.

3.3 REDES NEURAIAS: PERCEPTRON DE MÚLTIPLAS CAMADAS

Do ponto de vista biológico, o cérebro consiste de interconexões das células nervosas, os neurônios, que influenciam uns aos outros por meio de sinais elétricos (NAUCK, 2000). Sob outro ponto de vista, pode-se considerar o cérebro como um sistema de processamento de informação altamente complexo, não-linear, paralelo e com incomparável plasticidade e que processa as informações de uma forma totalmente diferente do computador digital (HAYKIN, 2001). Por causa destas características únicas o cérebro foi, é e será sempre alvo de estudos que possam trazer entendimento e explicações aos diversos mecanismos de seu funcionamento. Tal fato motiva não somente áreas do conhecimento particularmente interessadas, como a medicina, mas também áreas que possam utilizar os diversos mecanismos de funcionamento do cérebro para a solução de problemas da realidade. É neste sentido que surgiram os estudos relativos às Redes Neurais Artificiais (RNAs).

Os primeiros estudos sobre RNAs tiveram início na década de 40, com a publicação do artigo “*A Logical Calculus of the Ideas Immanent in Nervous Activity*” do neurofisiologista Warren McCulloch e do matemático Walter Pitts (CALDEIRAS, 2007). Este trabalho foi a primeira tentativa de entendimento da atividade neural a partir de unidades elementares de processamento e exerceu grande influência na comunidade científica da época.

Para Haykin (2001) e Azevedo (2000), outros marcos históricos importantes nos estudos sobre RNAs foram os estudos de Hebb, Rosembat, Bernard Widrow e Ted Hoff. O neuropsicólogo Hebb foi responsável pelo desenvolvimento da primeira regra de aprendizado, que ficou conhecida como aprendizado Hebbiano. Rosembat ficou conhecido por propor o *Perceptron* como o primeiro modelo de rede neural para aprendizagem por meio de treinamento supervisionado. Bernard Widrow e Ted Hoff propuseram a Regra Delta para

aprendizagem, também conhecida por Regra de aprendizado de Widrow-Hoff ou como algoritmo do mínimo quadrado médio (LMS - *Least Mean Square*). A Regra delta foi o ponto de partida para o desenvolvimento do algoritmo para aprendizado de RNAs mais conhecido e utilizado: A Regra Delta Generalizada ou Algoritmo de Retropropagação de Erro, o algoritmo *back-propagation*.

3.3.1 Definições e Características

Uma rede neural é definida como “*um processador maciçamente paralelamente distribuído de unidades de processamento simples, que têm a propensão natural para armazenar conhecimento experimental e torná-lo disponível para o uso*” (HAYKIN, 2001, p.28). Para Nauck (2000), as redes neurais são sistemas que tentam fazer uso de alguns dos conhecidos e esperados princípios de organização do cérebro humano. Dentre estes vários princípios pode-se citar a capacidade de aprendizagem, que pode ser definida como “*um processo pelo qual os parâmetros livres de uma rede neural são adaptados através de um processo de estimulação no qual a rede está inserida*”. (HAYKIN, 2001, p.75).

Durante a aprendizagem ocorre o ajuste dos parâmetros livres da rede neural em resposta a um estímulo externo. O algoritmo que realiza este ajuste de parâmetros é chamado de algoritmo de aprendizagem. Segundo Haykin (2001), existem cinco regras de aprendizado: aprendizado por correção de erro, aprendizado baseado em memória, aprendizado Hebbiana, aprendizagem competitiva e aprendizagem de Boltzmann. A maneira pela qual uma rede neural se relaciona com o seu ambiente é denominada de paradigma de aprendizado podendo ser do tipo supervisionado e não-supervisionado.

Outra característica importante citada por Haykin (2001) é a arquitetura de uma RNA e diz respeito ao arranjo com que os neurônios de uma RNA estão estruturados, estando intimamente ligada ao algoritmo de aprendizado que é utilizado durante o processo de aprendizagem da rede. Segundo Braga (2000), na definição de uma arquitetura de uma RNA são levados em consideração os seguintes parâmetros: número de camadas da rede, número de neurônios de cada camada, tipo de conexão entre os neurônios e a topologia da rede.

A unidade de processamento de informação de uma RNA é o neurônio artificial, que foi fortemente inspirado no neurônio biológico. O primeiro modelo de neurônio artificial foi desenvolvido na década de 40 e é conhecido pelo nome de neurônio de McCulloch-Pitts (CALDEIRAS, 2007). A Figura 3.1 ilustra o modelo de um neurônio artificial.

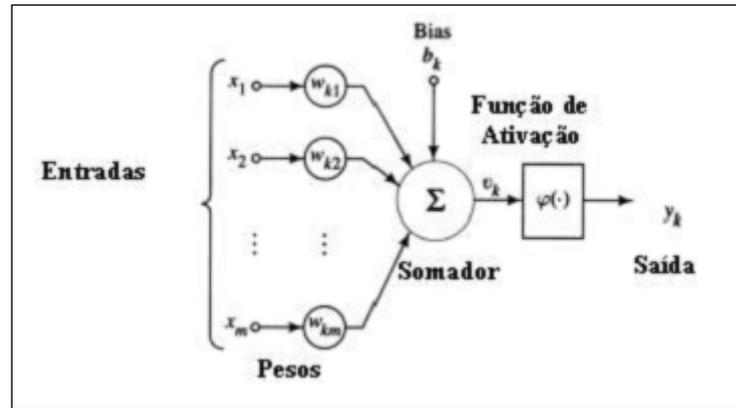


Figura 3.1: Modelo de um neurônio artificial

Fonte: Haykin (2001)

Haykin (2001) destaca três elementos no modelo de neurônio artificial genérico k que possui m entradas: as sinapses, o somador e função de ativação.

- **Sinapses** são caracterizadas por um peso ou força própria e possuem a função de multiplicar o valor do sinal de entrada (x_j) pelo peso sináptico (w_{kj}). Os pesos sinápticos são os parâmetros livres ajustados durante o processo de aprendizagem de uma RNA. Azevedo (2000) classifica os pesos sinápticos em excitatórios se $w_{kj} > 0$ e inibitórios se $w_{kj} < 0$.
- **Somador** bloco responsável por realizar uma combinação linear entre os m sinais de entrada já ponderados pelos respectivos pesos sinápticos. Em termos matemáticos tem-se:

$$u_k = \sum_{j=1}^m w_{kj} x_j \quad (3.1)$$

- **Função de ativação**, também chamada de função restritiva, tem a função de restringir a amplitude do sinal de saída do neurônio. A faixa de restrição da saída pode variar em um intervalo unitário fechado $[0,1]$ ou $[-1,1]$ dependendo da aplicação. Matematicamente tem-se:

$$y_k = \varphi(u_k) \quad (3.2)$$

Além destes três elementos básicos, o modelo neuronal inclui um *bias* (b_k) aplicado externamente que possui a função de variar a entrada da função de ativação, podendo aumentar quando o bias é positivo ou diminuir se negativo. Então a saída de um neurônio artificial tem-se a seguinte expressão:

$$y_k = \varphi(u_k + b_k) = \varphi(v_k) \quad (3.3)$$

O termo $u_k + b_k$ é conhecido como campo induzido e denotado por v_k .

Na modelagem matemática do neurônio k são utilizadas as expressões (3.1) e (3.3).

Segundo Haykin (2001) e Azevedo (2000), existem diversos tipos de função de ativação sendo as mais comuns: função limiar, função linear, função sigmóide e função tangente hiperbólico.

- **Função Limiar:** esta função foi utilizada no primeiro modelo de neurônio artificial, o modelo de McCulloh-Pitts. É definida pela seguinte expressão:

$$\varphi(v_k) = \begin{cases} 1, & \text{se } v \geq 0 \\ 0, & \text{se } v < 0 \end{cases} \quad (3.4)$$

Na engenharia esta função é conhecida pelo nome de função de Heaviside. Quando a faixa de restrição for o intervalo $[-1,1]$ a função limiar é definida pela função $sgn(x)$.

- **Função Linear:** Definida pela expressão:

$$\varphi(v_k) = ax \quad (3.5)$$

Onde a é um escalar

- **Função Sigmóide:** função mais utilizada como função de ativação em neurônios.

$$\varphi(v_k) = \frac{1}{1 + e^{-ax}} \quad (3.6)$$

Onde a é um escalar positivo. Quanto mais a cresce, mais a mudança de 0 para 1 se torna abrupta. Quando $a \rightarrow \infty$ a função se torna uma função limiar.

- **Função Tangente Hiperbólica**

$$\varphi(v_k) = \tanh(ax) = \frac{e^{ax} - e^{-ax}}{e^{ax} + e^{-ax}} \quad (3.7)$$

Onde a é um escalar positivo. Quanto mais a cresce, mais a mudança de -1 para 1 se torna abrupta. Quando $a \rightarrow \infty$ a função se torna uma função $sgn(x)$.

Ao longo do desenvolvimento das redes neurais diversos tipos foram criados, tais como: *Perceptron*, *Adaline*, *Madaline*, *Mapas de Kohonen*, *redes de base radial*, entre outras.

(CALDEIRAS, 2007). Este trabalho irá tratar das redes *Perceptron* de Múltiplas Camadas alimentadas adiante (*Feed-Foward Multilayer Perceptron*).

3.3.2 Redes *Perceptron* de Múltiplas Camadas (MLP)

Nas aplicações com redes neurais, as redes neurais tipo *Perceptron* de Múltiplas Camadas (MLP) são o tipo mais importante de RNA e são utilizadas com sucesso na solução de diversos problemas difíceis. (NAUCK, 2000, HAYKIN, 2001).

Uma rede neural MLP é uma rede que consiste de uma camada de entrada, uma ou mais camadas escondidas de neurônios e uma camada de saída de neurônios, onde o sinal de entrada se propaga para frente, camada por camada, através da rede. A Figura 3.2 apresenta uma rede MLP que é constituída de 4 unidades sensoriais de entrada, duas camadas escondidas sendo a primeira formada de 5 neurônios e a segunda por 3 neurônios e a camada de saída formada por 2 neurônios.

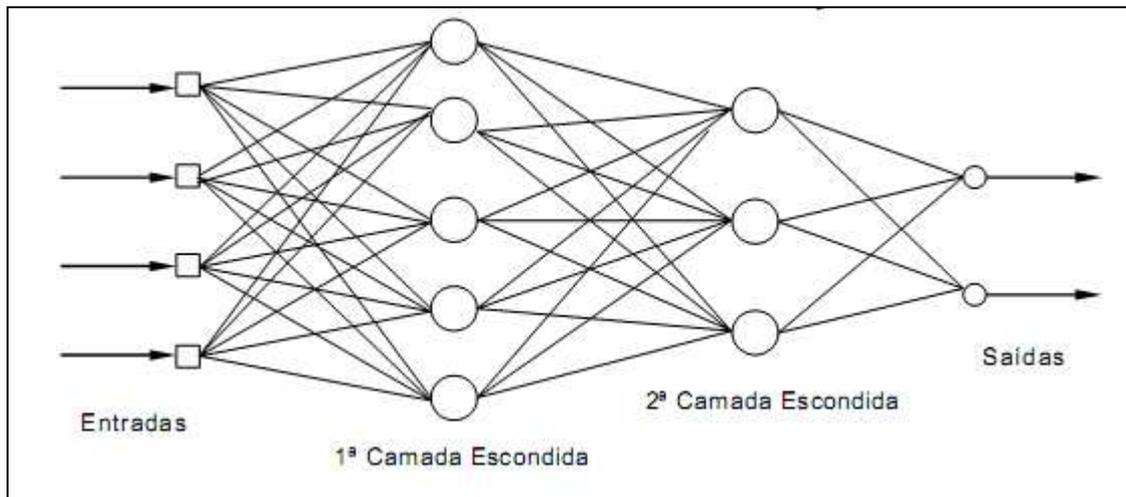


Figura 3.2: Arquitetura de uma rede *feed foward* MLP

Fonte: Dyminski (2000) apud Maciel (2005)

Haykin (2001) lista três importantes características que fazem com que uma rede neural tipo MLP tenha um grande poder computacional: Uso de uma função de ativação não-linear nos neurônios da rede, existência de camadas escondidas de neurônios e o alto grau de conectividade da rede. A primeira característica tem forte inspiração biológica. O uso de uma função não-linear que seja diferenciável em qualquer ponto é uma escolha muito melhor do que a observada com o uso de função tipo limiar, onde a mudança de um patamar ao outro é feita de forma abrupta. Uma das funções mais utilizada é a não linearidade tipo sigmóide. A

existência de camadas escondida de neurônios capacita a rede a extrair as características mais significativas dos padrões de entrada e o alto grau de conectividade da rede determinado pelas sinapses da rede exerce influência no conjunto de conexões sinápticas ou de seus pesos.

O aprendizado de uma rede neural tipo MLP é feito de modo supervisionado sendo o algoritmo de aprendizagem mais conhecido e utilizado: o algoritmo de retropropagação de erro, o algoritmo *back-propagation*.

3.3.3 O Algoritmo de Retropropagação de Erro (*Error back-propagation*)

Segundo Haykin (2001) e Michell (2000), o algoritmo de retropropagação de erro (*error back-propagation*) é baseado na regra de aprendizado por correção de erro e realiza a alteração no valor dos pesos sinápticos de uma RNA por meio do método do gradiente decrescente, de modo a minimizar o erro quadrático entre o valor da saída desejada e a saída fornecida pela rede neural. O algoritmo é também conhecido como Regra Delta Generalizada, por ser considerado uma generalização da Regra de Widrow-Hoff para redes neurais de múltiplas camadas.

Haykin (2001) divide o algoritmo de retropropagação em dois passos distintos: um passo para frente, ou propagação, e um passo para trás, ou retropropagação. Nesses dois passos, são distinguidos dois tipos de sinais: os sinais funcionais e os sinais de erro.

Segundo Parker (1987) apud Haykin (2001), os sinais funcionais são os sinais de entrada que se propagam para frente camada após camada até chegar à saída da RNA. Os sinais de erro são provenientes dos neurônios da camada de saída da RNA. Este sinal é propagado para trás, camada após camada.

Durante a fase de propagação, os pesos sinápticos da rede permanecem inalterados e os sinais funcionais são calculados individualmente, neurônios por neurônio. Durante a propagação para trás, os pesos sinápticos são alterados baseados no erro computado na saída.

Segundo Michell (1995) e Haykin (2001), dado um conjunto de treinamento onde cada exemplo é representado por um par na forma $(\mathbf{x}(n), \mathbf{d}(n))$ onde $\mathbf{x}(n)$ representa o valor do padrão de entrada da RNA e $\mathbf{d}(n)$ o valor da saída desejada e n é o número da iteração em curso. O algoritmo de retropropagação de erro segue os seguintes passos para uma rede com uma camada escondida:

Passo 1. Inicialização dos parâmetros livres da rede neural: cada peso sináptico e bias da rede neural são inicializados de forma aleatória.

Passo 2. Apresentação dos Exemplos de Treinamento: O padrão de entrada $\mathbf{x}(n)$ é apresentado à entrada da RNA. A saída desejada $\mathbf{d}(n)$ para o padrão $\mathbf{x}(n)$ também é apresentada à saída da RNA. Defini-se como uma época à apresentação completa do conjunto de treinamento.

Passo 3. Propagação para frente: Calculam-se os sinais funcionais de cada um dos neurônios pertencentes à rede por meio da expressão (3.8)

$$y_j(n) = \varphi(v_j(n)) \quad (3.8)$$

Onde o termo $v_j(n)$ é campo local induzido do neurônio e é calculado por:

$$v_j(n) = \sum_{i=0}^m w_{ji}(n) y_i(n) \quad (3.9)$$

Onde m é o total de entradas aplicadas ao neurônio j , $w_{ji}(n)$ é o peso sináptico que conecta o neurônio i ao neurônio j e $y_i(n)$ pode denotar tanto o sinal de entrada do neurônio j como o sinal funcional que surge na saída do neurônio i . Se o neurônio j pertencer à primeira camada oculta da RNA o índice i refere-se ao i -ésimo terminal de entrada da RNA, ou seja, $y_i(n) = x_i(n)$, onde $x_i(n)$ é o i -ésimo elemento do vetor de entrada. Caso o neurônio j pertencer à camada de saída da RNA o índice i refere-se ao j -ésimo terminal de saída da RNA, ou seja, $y_j(n) = o_j(n)$ onde $o_j(n)$ é o j -ésimo elemento do vetor de saída.

Passo 4. Calcula-se a diferença entre o valor da saída desejada $\mathbf{d}(n)$ e o valor da saída real da RNA denotada por $\mathbf{o}(n)$ conforme a expressão (3.10):

$$e_j(n) = d_j(n) - o_j(n) \quad (3.10)$$

Em seguida é calculado a função custo $E(n)$ que deverá ser minimizada. A função representa o valor instantâneo da energia total de erro dos neurônios pertencentes à camada de saída. A função custo é definida pela expressão (3.11):

$$E(n) = \frac{1}{2} \sum_{j \in C} e_j^2(n) \quad (3.11)$$

Onde C é o conjunto de todos os neurônios de saída da RNA.

Passo 5. Propagação para trás: Calculam-se os gradientes locais $\delta(n)$ da rede

1. Para cada neurônio pertencente à camada de saída da rede

$$\delta_j(n) = e_j(n)\varphi'(v_j(n)) \quad (3.12)$$

2. Para cada neurônio pertencente à camada oculta da rede

$$\delta_j(n) = \varphi'(v_j(n)) \sum_k \delta_k(n) w_{kj}(n) \quad (3.13)$$

Passo 6. Atualização de cada peso sináptico da rede por meio da expressão (3.14)

$$\omega_{ji}(n) = \omega_{ji}(n) + \Delta\omega_{ji}(n) \quad (3.14)$$

Onde o termo $\Delta\omega_{ji}(n)$ é a correção do peso sináptico $\omega_{ji}(n)$. Calculado pela expressão (3.15)

$$\Delta\omega_{ji}(n) = \eta\delta_j(n)y_i(n) \quad (3.15)$$

O termo η é conhecido pelo nome de constante de aprendizado e é utilizado como parâmetro para acelerar o processo de aprendizado.

Passo 7. Repita os passos 3 a 6 para próximas épocas até que o valor de $E(n)$ atinja certo critério de parada preestabelecido. A apresentação dos elementos do conjunto de treinamento de época para época é feita de forma aleatória.

3.3.4 O Algoritmo de Retropropagação de Erro com Momentum

Para Haykin (2001) e Azevedo (2000), a escolha do valor da taxa de aprendizado (η) pode acarretar alguns comportamentos não desejáveis durante o processo de aprendizado da rede neural. Se o valor de η for um valor pequeno ocorrerá uma pequena variação nos pesos sinápticos de uma iteração a outra e o processo de aprendizado se dará de forma muito lenta. Se o valor de η for grande o processo se dará de uma forma mais rápida, entretanto com um maior risco de instabilidade do processo de aprendizado. Uma forma de contornar esses inconvenientes é acrescentando um termo chamado de *momentum* na fórmula de atualização dos pesos sinápticos. Denotando como β o termo *momentum*, a fórmula de atualização dos pesos sinápticos é dada pela expressão (3.16).

$$\Delta\omega_{ji}(n) = \eta\delta_j(n)y_i(n) + \beta\Delta\omega_{ji}(n-1) \quad (3.16)$$

Nota-se pela expressão (3.16) que o termo *momentum* atua proporcionalmente no valor da variação dos pesos sinápticos da iteração anterior. O valor do *momentum* pode assumir a faixa de 0 a 1 podendo ser positivo ou negativo (na prática não se usam valores negativos). A adição do termo *momentum* garante uma melhor estabilidade durante o processo de aprendizado de uma rede neural tipo MLP.

3.4 CLASSIFICADOR K-VIZINHOS MAIS PRÓXIMOS

O algoritmo k-vizinhos mais próximos, também conhecido por kNN (*k nearest neighbors*), é um algoritmo de classificação cujo aprendizado é realizado de forma supervisionada. O algoritmo consiste em classificar um novo elemento levando em consideração os *k* elementos vizinhos mais próximos já classificados. Com base na classificação deste *k* vizinhos mais próximos é feita a decisão relativa a classificação do novo elemento.

O algoritmo kNN requer pouco esforço computacional durante a fase de treinamento. Entretanto, a classificação de um novo elemento requer um elevado custo computacional, pois no caso mais extremo esse novo elemento deverá ser comparado com todos os elementos já classificados.

Ferrero (2009) lista três parâmetros a serem observados na execução do algoritmo kNN: quais elementos devem ser lembrados, a medida de similaridade e o número de vizinhos mais próximos a serem considerados na comparação.

O primeiro parâmetro está diretamente relacionado com o esforço computacional demandado na classificação de um novo elemento, pois dependendo da quantidade de elementos do conjunto de treinamento o processo de classificação poderá ser lento. Uma possível metodologia para contornar este inconveniente é o não armazenamento de todos os elementos, mas dos elementos mais representativos de cada classe (FERRERO, 2009).

O segundo parâmetro está relacionado com a definição de uma medida que determine o grau de similaridade entre o exemplo a ser classificado e os elementos do conjunto de treinamento. Uma medida de similaridade bastante utilizada é a distância de *Minkowsky* que calcula a distância entre dois pontos E_i e E_j em um espaço *n*-dimensional de acordo com o parâmetro *d*. A distância é definida pela seguinte expressão:

$$\text{dist}(E_i, E_j) = \left(\sum_{l=1}^n |x_{il} - x_{jl}| \right)^{\frac{1}{d}} \quad (3.17)$$

A medida de *Minkowsky* é conhecida como distância de Manhattan/ *city-block* quando $d = 1$ e define a distância Euclidiana quando $d = 2$.

O terceiro parâmetro é a cardinalidade do conjunto de vizinhos mais próximos, ou seja, o parâmetro k . Se k for igual a 1 então o algoritmo kNN irá classificar o novo elemento com a classe do elemento mais próximo a ele. Se k for maior que 1 então o algoritmo irá classificar o novo elemento na classe predominante no conjunto dos k elementos vizinhos mais próximos. A Figura 3.3 ilustra a influência do parâmetro k no processo de classificação.

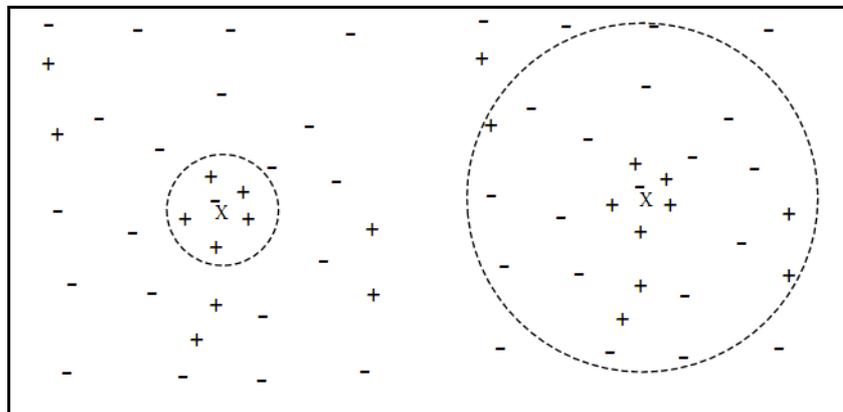


Figura 3.3: Influência do parâmetro k na classificação de um novo elemento, utilizando a distância euclidiana (Fonte: Wu, 2009)

A Figura 3.3 mostra um conjunto de treinamento que apresenta elementos classificados em duas classes: '+' e '-', e um novo elemento, denotado por 'x', que deverá ser classificado em uma dessas duas classes. Dependendo dos k vizinhos mais próximos considerados na classificação, o novo elemento pode ser classificado tanto em uma classe como outra. Para uma vizinhança com poucos elementos o novo elemento é classificado na classe '+', mas se a vizinhança for aumentada observa-se que o novo elemento já será classificado na classe '-'

Como se pode constatar o parâmetro k tem grande influência na classificação do novo elemento, entretanto, não existe um número k que seja apropriado para todos os problemas cuja solução seja pelo algoritmo kNN. Ferrero (2009) cita uma metodologia de atribuição de pesos à distância de similaridade dos elementos do conjunto do k vizinhos mais próximos

onde os elementos seriam ordenados em ordem crescente de distância de similaridade e os elementos mais próximos teriam um peso maior na atribuição da classe ao elemento a ser classificado.

3.5 CONSIDERAÇÕES FINAIS

Este capítulo tratou da análise de classificação de dados e sua importância dentro da área de aprendizado de máquina. Foram apresentadas as principais características que influenciam no desempenho da classificação de dados. Em seguida, foram abordadas duas técnicas computacionais que realizam a classificação de dados: as redes neurais tipo MLP e o algoritmo kNN. As redes neurais são modelos matemáticos que se inspiram na arquitetura e funcionamento do cérebro e são bastante utilizadas na tarefa de classificação de dados. O algoritmo kNN utiliza a ideia de classificar um novo elemento levando em consideração a classe dos k elementos mais próximos do novo elemento. No capítulo a seguir será definido qual técnica é mais apropriada na construção do modelo de classificador utilizado na metodologia de auxílio na tomada de decisão da formação de conjuntos de unidades consumidoras.

CAPÍTULO 4

A METODOLOGIA PROPOSTA

4.1 CONSIDERAÇÕES INICIAIS

Este capítulo tratará do desenvolvimento de uma metodologia a ser utilizada pelas empresas de distribuição de energia elétrica no auxílio à tomada de decisão na definição de conjuntos de unidades consumidoras de energia elétrica. Primeiramente, será apresentada uma visão geral da metodologia proposta. Em seguida, são mostrados os estudos de construção e análise de desempenho dos modelos de classificadores construído pelas técnicas apresentadas no capítulo anterior. Por último, será descrita a metodologia para a construção dos cenários de conjuntos de unidades consumidoras que serão avaliados no capítulo posterior.

4.2 A METODOLOGIA PROPOSTA: LINHAS GERAIS

A Metodologia utilizada pelo software ANABENCH para definição das metas dos indicadores DEC e FEC dos conjuntos de unidades consumidoras das empresas distribuidoras de energia elétrica pode ser resumida pela Figura 4.1

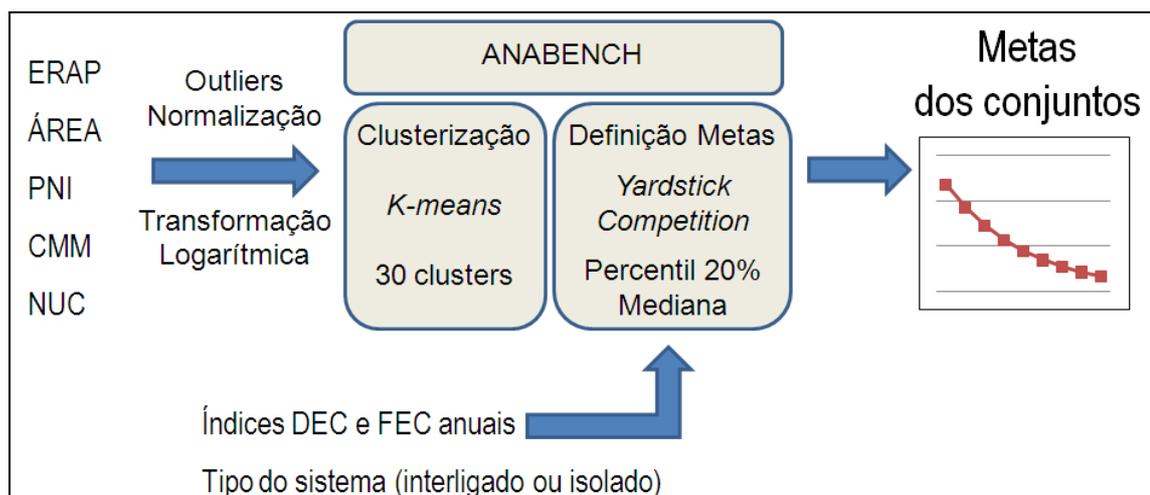


Figura 4.1: Estrutura do Software ANABENCH

Na Figura 4.1, observa-se os cinco atributos que caracterizam cada um dos conjuntos de unidades consumidoras. A seguir tem-se a etapa de tratamento dos dados com o expurgo de conjuntos que apresentam dados considerados *outliers*, passando pela normalização e posterior transformação logarítmica. Esses dados, já tratados, são submetidos ao processo de agrupamento por meio do algoritmo *k-means*. Ao final deste processo os conjuntos considerados como *outliers* são associados aos *clusters* com os centróides mais próximos. O resultado deste processo de agrupamento é a partição do conjunto de todos os conjuntos de unidades consumidoras do Brasil em 30 famílias, chamadas de *clusters*. Uma vez definidos os conjuntos que compõe de cada um dos 30 *clusters*, passa-se para a etapa de definição de metas dos indicadores de continuidade DEC e FEC. Para cada um dos *clusters* é aplicada a análise comparativa via técnica *Yardstick Competition*, onde são identificados os valores das marcas de referência para os indicadores DEC e FEC. Os valores de referência são definidos pelo percentil de 20% (para conjuntos pertencentes a sistemas interligados) ou pela mediana, (para conjuntos pertencentes ao sistema isolado) das distribuições dos índices DEC e FEC dos conjuntos de um mesmo *cluster*. A partir dos valores de referência e do patamar inicial são definidas as metas anuais para cada conjunto de unidade consumidor ao longo de duas revisões tarifárias.

A metodologia proposta neste trabalho é composta de duas fases: construção de um classificador por meio de técnicas de aprendizado de máquina e simulação de configurações de conjuntos de unidades consumidoras de energia elétrica.

A primeira fase consiste no uso de técnicas de aprendizado de máquina na construção de um classificador que capture as características intrínsecas de cada um dos 30 *clusters* de conjuntos de unidades consumidoras definidos pelo software ANABENCH utilizados na definição dos valores das metas dos indicadores de continuidade. Para construção deste classificador é necessário a utilização dos mesmos dados utilizados pela ANEEL na definição dos 30 *clusters*. O classificador tem como objetivo a simulação do comportamento da etapa de agrupamento do software ANABENCH. Essa primeira fase é de fundamental importância, pois o classificador construído servirá de base para a segunda fase.

A segunda fase consiste na realização de simulações de configurações de conjuntos de unidades consumidoras cujo objetivo é encontrar uma formação específica dos conjuntos em termos de quantidade e características (de acordo com as cinco variáveis) dos conjuntos, que levem as metas de DEC e FEC da empresa de distribuição de energia elétrica a um estado em que exista menos infrações ou que estas sejam menores. Cada formação de conjuntos é um cenário a ser explorado e analisado pelos especialistas da empresa de distribuição de energia

elétrica para se determinar se esta nova formação é melhor ou pior da atual empregada pela empresa. Para cada nova configuração, os novos conjuntos são apresentados ao classificador que irá informar para cada novo conjunto o seu referido *cluster*. Como cada *cluster* já possui uma meta final de continuidade de DEC e FEC, definido pelo software ANABENCH, é possível calcular as metas anuais para o novo cenário. Com base nessas novas metas para os novos conjuntos pode-se avaliar se a nova configuração de conjuntos, este novo cenário, deixa a empresa em uma situação com um nível de infração, das metas de continuidade, menor.

Neste trabalho, um cenário será a união de dois conjuntos já existentes na área de concessão da empresa distribuidora de energia elétrica em estudo. Entretanto, outros tipos de cenários poderão ser analisados pela Metodologia Proposta, tais que:

- **Redes em Malha:** nos centros urbanos há predominância de sistemas malhados e com elevado número de subestações, assim, fica difícil delimitar as áreas de abrangências das subestações para a definição de um conjunto de unidades consumidoras. Então, é de interesse da empresa de distribuição de energia elétrica a correta definição dos conjuntos nestas áreas.
- **Construção de novas subestações:** ao se construir uma nova subestação, a sua área de abrangência será um novo conjunto e um novo cenário para estudo. Portanto, podem-se realizar estudos de quais possíveis metas poderiam ser determinadas na aplicação da metodologia da ANEEL.
- **Influência da variação de algum dos cinco atributos:** com a aplicação da metodologia proposta pode-se realizar estudos sobre a influência da variação de algumas dos cinco atributos utilizados pela ANEEL na caracterização dos conjuntos de unidades consumidoras. Um desses estudos poderia ser o impacto de programas de universalização de energia elétrica nas metas dos indicadores de continuidade DEC e FEC.

O método proposto tem a vantagem de para se fazer esse estudo não é necessário executar o programa ANABENCH com os dados dos cinco atributos de todos dos conjuntos de consumidores de energia elétrica, pois o classificador já capturou as características de cada *cluster*. Assim, é possível realizar um estudo de qual configuração de conjunto é mais apropriado de modo a diminuir a violação das metas estabelecidas pelo órgão regulador. A desvantagem deste método reside no fato de ser necessário executar o ANABENCH pelo menos uma vez para se obter a classificação dos conjuntos em seus respectivos *clusters*.

4.3 A CONSTRUÇÃO DO CLASSIFICADOR

No capítulo anterior foram discutidas duas técnicas que fazem uso do aprendizado supervisionado para a construção de um modelo de classificador: as redes neurais tipo MLP e o algoritmo kNN. Nesta seção será apresentado um estudo de avaliação de desempenho dos modelos de classificadores construídos por meios das duas técnicas anteriormente citadas e será escolhida a técnica mais apropriada para solução do problema de classificação a ser utilizada na Metodologia Proposta. A escolha por essas técnicas de aprendizado de máquina se deve ao fato de serem duas técnicas já consagradas nos estudos envolvendo classificação de padrões. Além disso, o Laboratório de Redes de Alto Desempenho da Universidade Federal do Pará que ajudou na implementação das técnicas possui conhecimento consolidado no uso das técnicas utilizadas neste trabalho.

4.3.1 O Conjunto de dados para o Aprendizado supervisionado

O conjunto de dados utilizado neste trabalho foi construído a partir dos dados de 4008 conjuntos de unidades consumidoras existentes no ano de 2008, no Brasil, onde cada conjunto foi caracterizado com os cinco atributos (área do conjunto, extensão de rede primária, potência instalada, consumo médio mensal e número de consumidores), a informação se o conjunto é do tipo interligado ou isolado, os valores de DEC e FEC que foram utilizados como Ano Zero e o *cluster* onde foram alocados pelo software ANABENCH.

Desta base de 4008 conjuntos, 3547 conjuntos foram considerados como ativos por não apresentarem algum atributo ou indicador com valor nulo ou por estar fora da faixa determinada pela cerca inferior ou pela cerca superior na análise via *box-plot*.

A partir desta base de dados, foi construído o conjunto de dados para construção dos modelos de classificadores. O conjunto de dados é caracterizado pelos cinco atributos e as classes são o *cluster* onde os conjuntos foram alocados pelo *software* ANABENCH.

Para a construção do modelo do classificador através do uso de um conjunto de dados é necessário primeiramente o tratamento dos dados pertencentes a esta base. A normalização dos dados é um desses tratamentos e é requisito necessário para a aplicação das técnicas de redes neurais e kNN. A base de dados contendo 3547 conjuntos foi submetida ao processo de normalização por meio da divisão pelo valor máximo mais 20% deste valor para cada um dos cinco atributos. Esta folga de 20% no valor máximo justifica-se pelo fato de se trabalhar com

união de conjuntos de unidades consumidoras o que poderia ocasionar a extrapolação do limite superior obtido da normalização dos dados.

4.3.2 Avaliação do desempenho dos classificadores

As duas técnicas avaliadas neste trabalho foram submetidas ao processo de validação cruzada com k -folds com o objetivo de se avaliar o desempenho dos modelos de classificadores construídos por meio das técnicas.

Segundo Witten (2005), a validação cruzada com k -folds consiste na divisão do conjunto de dados em k subconjuntos disjuntos (*folds*) onde é mantida a mesma distribuição de classes da base original, ou seja, se a base original possuir duas classes com distribuição de 20% e 80%, então cada *fold* também terá esta mesma proporção de classes. Este processo é chamado de estratificação e é um procedimento importante, pois reduz a variância da estimativa. O conjunto de treinamento é construído tomando os $(k-1)$ *folds* enquanto que o *fold* remanescente é utilizado como conjunto de teste. Este processo é repetido k vezes, cada vez considerando um *fold* diferente para o conjunto de teste, como pode ser visto na Figura 4.2 para 10 *folds*.

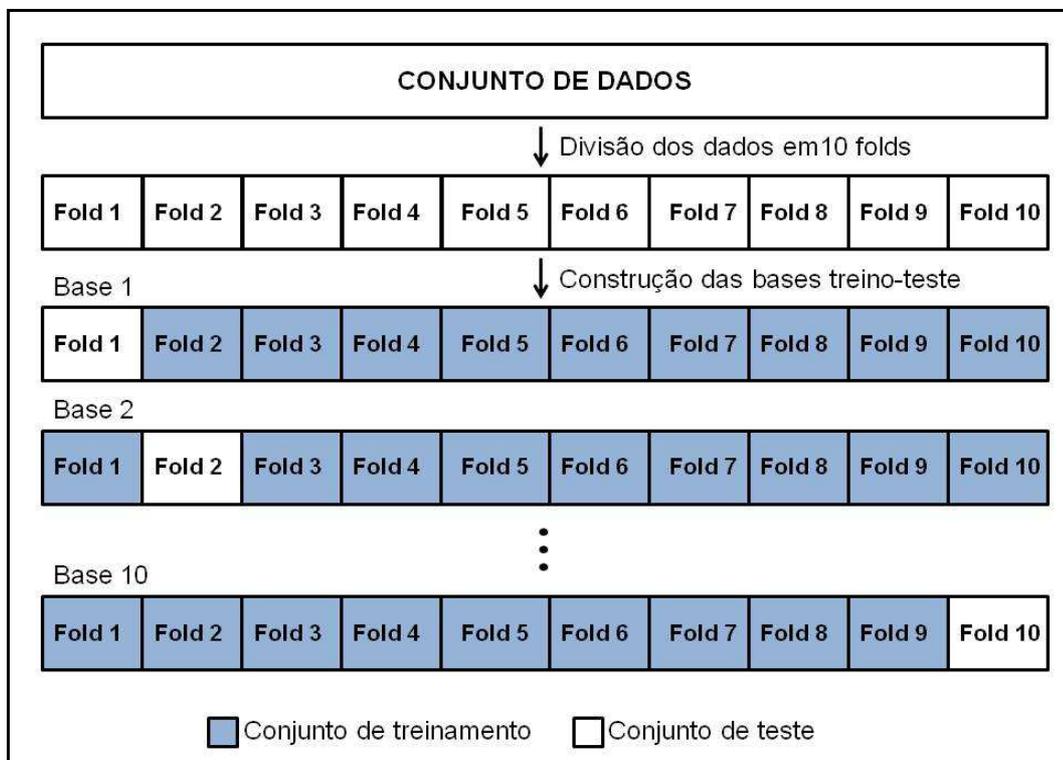


Figura 4.2: Construção das bases treino-teste para validação cruzada

A taxa de erro na validação cruzada é calculada pela média das taxas de erros em cada um dos pares de treino-teste.

Este trabalho utilizará a validação cruzada estratificada com *10-folds*. A escolha de 10 *folds* para divisão do conjunto de dados se deve ao fato de que extensivos testes em numerosos tipos de conjunto de dados aplicados a diferentes tipos de algoritmos de aprendizado de máquina mostraram que 10 *folds* são aproximadamente o número certo de *folds* para se obter a melhor estimativa da taxa de erro (WITTEN, 2005).

Para cada configuração de rede neural MLP e do algoritmo kNN foi apresentada uma base de treino para construção do classificador e teste para a mensuração da taxa de erro deste modelo de classificador. Ao final da apresentação das 10 bases de dados é calculada a média aritmética dos valores da taxa de erro.

A Ferramenta computacional utilizada para as simulações das técnicas de classificação foi o Weka (Waikato Environment for Knowledge Analysis) que consiste de um conjunto de programas em JavaTM, desenvolvida pela Waikato University na Nova Zelândia, com os principais algoritmos e técnicas de aprendizado de máquina, tais como: redes neurais, redes bayesianas, árvores de decisão, entre outras. Neste estudo foi utilizada versão 3.6 do Weka.

A rede neural tipo MLP consiste de cinco entradas (representando os cinco atributos) e 30 saídas (representando os 30 *clusters*) e optou-se por utilizar apenas uma camada escondida. A função de ativação utilizada nos neurônios foi a função sigmóide. Nas simulações das redes neurais, tipo MLP, foi realizado exaustivos testes de várias configurações de redes neurais, variando os parâmetros de número de neurônios na camada escondida, taxa de aprendizado (η) e *momentum* (α). O número de neurônios na camada escondida é obtido de forma empírica por meio de simulações, entretanto, Malhotra (2003) *apud* Maciel (2005) apresenta uma regra para definição do número de neurônios na camada escondida, definida pela expressão 4.1.

$$\text{número de neurônios} = \left(\frac{\sum \text{Entradas} + \sum \text{saídas}}{2} \right) \quad (4.1)$$

Optou-se por realizar simulações variando o número de neurônios na camada escondida em torno do valor de 18 neurônios, valor obtido pela aplicação da expressão 4.1 para o caso da rede em estudo.

Os melhores resultados foram obtidos com os parâmetros de 18 neurônios na camada escondida e as taxas de aprendizado e *momentum* na faixa de 0,2 a 0,3. A Tabela 4.1 apresenta os valores da taxa de erro usando a validação cruzada com 10 *folds* para as três melhores configurações encontradas.

Tabela 4.1: Resultados das Simulações para Rede Neurais tipo MLP

Parâmetros	$\eta=0.3$ $\alpha=0.2$	$\eta=0.2$ $\alpha=0.2$	$\eta=0.2$ $\alpha=0.3$
Base1	12,64%	13,82%	14,11%
Base2	9,11%	8,52%	8,52%
Base3	12,35%	13,23%	15%
Base4	14,41%	11,76%	15,29%
Base5	16,47%	15,58%	14,7%
Base6	14,7%	11,47%	11,17%
Base7	11,17%	12,05%	10,29%
Base8	15,29%	14,11%	12,64%
Base9	14,7%	13,52%	14,11%
Base10	13,52%	15,88%	15,29%
Média	13,44%	12,99%	13,11%

Na Tabela 4.1 estão destacados em negrito os valores das menores taxas de erros encontrados para as três melhores configurações de rede neural tipo MLP. Nota-se que as menores taxas de erro se encontram em torno de 10 % de erro.

Para as simulações com o algoritmo kNN foram testadas várias configurações variando o número de vizinhos mais próximos (k) e optou-se por utilizar a distância euclidiana como medida de similaridade sem o uso de pesos para ponderar elementos mais próximos do exemplo a ser classificado. Os melhores resultados foram obtidos variando o parâmetro k no intervalo de 1 a 4.

A Tabela 4.2 apresenta os valores da taxa de erro usando validação cruzada com 10 *folds* para as quatro melhores configurações encontradas.

Tabela 4.2: Resultados das Simulações para algoritmo kNN

Parâmetros	k=1	k=2	k=3	k=4
Base1	25,58%	28,52%	23,52%	24,41%
Base2	20,88%	24,11%	20,29%	21,76%
Base3	26,47%	28,23%	24,7%	26,17%
Base4	22,05%	22,05%	21,76%	21,76%
Base5	20,29%	23,82%	21,47%	22,94%
Base6	25,58%	29,11%	24,7%	27,35%
Base7	20,29%	23,82%	19,7%	19,7%
Base8	22,94%	27,94%	22,05%	27,05%
Base9	23,52%	27,05%	24,11%	24,7%
Base10	27,05%	29,7%	24,7%	23,52%
Média	23,47%	26,44%	22,70%	23,94%

Na Tabela 4.2 estão destacados em negrito os valores das menores taxas de erros encontrados para as quatro melhores configurações encontradas com o uso do algoritmo kNN. Nota-se que as menores taxas de erro se encontram em torno de 20 % de erro.

Comparando os valores médios da taxa de erro obtidos pela aplicação das duas técnicas nota-se que os modelos de classificadores construídos pela técnica de redes neurais tipo MLP obtiveram os melhores resultados de taxa de erro de classificação. Adiciona-se também o fato de que o algoritmo kNN ser do tipo *lazy*, ou seja, necessita que todos os elementos do conjunto de treinamento sejam armazenados, diferentemente das redes neurais tipo MLP, tal característica faz com que o esforço computacional para a classificação de um novo elemento seja reduzido. Portanto, ficam evidente as vantagens do uso das redes neurais tipo MLP para serem utilizadas na construção do classificador que será utilizado na Metodologia Proposta neste trabalho.

4.3.3 O Arranjo do Classificador

No item anterior foi decidido o uso das redes neurais tipo MLP na construção do modelo do classificador. Para melhorar ainda mais o processo de classificação de um cenário, optou-se em utilizar não apenas um classificador, mas um conjunto de classificadores, de tal modo que ao se apresentar um novo elemento cada um classificadores geraria uma resposta (neste estudo a resposta seria a que *cluster* o novo cenário pertenceria), e por meio de um

processo de eleição por maioria seria obtida a resposta final ao problema de classificação em algum dos 30 *clusters*. A Figura 4.3 mostra a idéia geral deste arranjo.

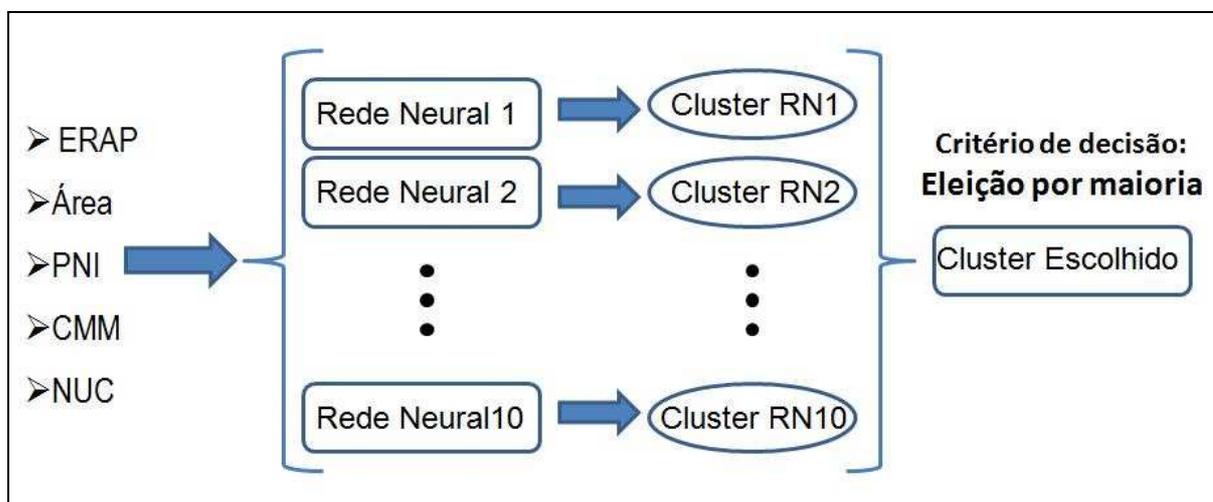


Figura 4.3: Arranjo de redes neurais para a classificação

O arranjo escolhido para este trabalho é constituído por 10 redes neurais tipo MLP. Essas 10 redes neurais foram obtidas dos resultados das três melhores configurações cujos resultados foram mostrados na Tabela 4.2. Para cada uma das 10 bases de dados utilizadas no processo de validação cruzada foi escolhida a rede neural com menor taxa de erro. Por exemplo, na base 1 foi escolhida a rede com 12,64% de taxa de erro, na base 2 foi escolhida a rede com 8,56% e assim por diante até a base 10 onde foi escolhida a rede com 13,52% de taxa de erro.

Para facilitar as análises que serão apresentadas no capítulo a seguir foi desenvolvido um programa em JavaTM que utilizasse os modelos de redes neurais construídos pelo Weka. Neste programa são apresentados os cinco atributos equivalentes da união e depois são apresentados os resultados das 10 redes neurais para que novo *cluster* deve pertencer este novo conjunto.

4.4 METODOLOGIA PARA UNIÃO DE CONJUNTOS

Neste trabalho um cenário é definido como a união de dois conjuntos já existentes na área de concessão de uma empresa de distribuição de energia elétrica. Uma vez definido, o cenário possuirá um novo conjunto de cinco atributos, que o caracterizam, bem como um novo ano zero a ser utilizado como patamar inicial no cálculo das metas anuais dos indicadores de continuidade ao longo do período de duas revisões tarifárias. A seguir é

apresentada a metodologia de cálculo para os cinco atributos e o ano zero dos indicadores de continuidade.

4.4.1 Para os cinco atributos

Os cinco atributos equivalentes da união de dois ou mais conjuntos corresponde a soma de cada um dos cinco atributos dos conjuntos envolvidos na união, com exceção do atributo consumo médio mensal que é obtido pela média ponderada pelo número de unidades consumidoras dos conjuntos envolvidos na união. A Tabela 4.3 ilustra o procedimento de união para dois conjuntos de unidades consumidoras.

Tabela 4.3: Atributos equivalentes na união de conjuntos

CONJUNTO	ERAP	AREA	PNI	CMM	NUC
Conjunto A	161,5	2114,98	22687,5	8564,91	16969
Conjunto B	432,98	68384,47	16402,5	901,44	3349
SOMA	594,48	70499,45	39090	7301,74	20318

Os valores obtidos pela soma são normalizados e são apresentados ao classificador que informará em que *cluster* este cenário será alocado. Conhecido o *cluster* em que foi alocado o cenário e sabendo se ele é pertence ao sistema interligado ou isolado é possível saber quais as metas finais dos indicadores de continuidade, pois estas metas já foram definidas pelo software ANABENCH. Os valores das metas finais dos indicadores DEC e FEC para conjuntos pertencentes à sistemas do tipo interligado e isolado são apresentados na Tabela do Anexo 1.

4.4.2 Patamar inicial para os indicadores de Continuidade: Ano zero

Para se efetuar o cálculo das metas dos indicadores de continuidade ao longo de duas revisões tarifárias são utilizadas as expressões 2.6 a 2.9, onde dois parâmetros devem ser conhecidos: o patamar final ou meta final e o valor do patamar inicial ou ano zero do indicador de continuidade em questão. A meta final é obtida segundo os procedimentos descritos no item anterior. O ano zero utilizado pelo ANABENCH é calculado por meio da média aritmética do desempenho dos indicadores anuais de DEC e FEC dos últimos três anos. Por esta razão ao se efetuar a união de dois ou mais conjuntos devem-se, primeiramente,

calcular os valores dos indicadores anuais DEC e FEC equivalentes da união desses dois ou mais conjuntos para os últimos 3 anos. Os valores dos indicadores DEC e FEC equivalentes da união são calculados por meio das expressões 4.2 e 4.3.

$$DEC_{eqv-anual} = \frac{Cm_1 \cdot DEC_1 + Cm_2 \cdot DEC_2 + \dots + Cm_n \cdot DEC_n}{Cm_1 + Cm_2 + \dots + Cm_n} \quad (4.2)$$

$$FEC_{eqv-anual} = \frac{Cm_1 \cdot FEC_1 + Cm_2 \cdot FEC_2 + \dots + Cm_n \cdot FEC_n}{Cm_1 + Cm_2 + \dots + Cm_n} \quad (4.3)$$

Nota-se que os indicadores DEC e FEC equivalentes da união são calculados por meio da média ponderada dos valores dos indicadores anuais dos conjuntos envolvidos na união e que os fatores de ponderação são os valores das médias do número de consumidores anuais dos conjuntos participantes da união. Os valores dos anuais dos indicadores DEC e FEC e da média anual do número de consumidores dos conjuntos de unidades consumidoras de uma determinada empresa distribuidora são encontrados no endereço eletrônico da ANEEL: www.aneel.gov.br.

Uma vez obtidos os valores dos indicadores DEC e FEC anuais equivalente da união para os últimos três anos calcula-se o valor do ano zero por meio da média aritmética dos valores dos últimos três anos.

Neste trabalho utiliza-se uma base de dados nacional do ano de 2008, portanto, serão necessárias as informações dos valores dos indicadores anuais DEC e FEC e média anual do número de consumidores de uma determinada empresa distribuidora dos anos de 2005, 2006 e 2007.

Por exemplo, para o caso da união dos conjuntos da Tabela 4.4 têm-se os seguintes passos para se obter o ano zero para os indicadores de DEC e FEC, toma-se como ano base o ano de 2008.

Tabela 4.4: Valores anuais dos indicadores DEC e FEC e média anual do número de consumidores

CONJUNTO	2005			2006			2007		
	NUC	DEC	FEC	NUC	DEC	FEC	NUC	DEC	FEC
Conjunto A	2058	49,57	47,81	2233	97,40	58,72	2851	225,83	131,27
Conjunto B	15979	16,76	33,58	15914	19,72	20,23	16451	32,38	28,29

A partir das informações da Tabela 4.4 têm-se:

- Cálculo dos Indicadores DEC e FEC equivalente da união para o ano de 2005.

$$DEC_{eqv-2005} = \frac{15979.16,76 + 2058.49,57}{15979 + 2058} = 20,50$$

$$FEC_{eqv-2005} = \frac{15979.33,58 + 2058.47,81}{15979 + 2058} = 32,20$$

- Cálculo dos Indicadores DEC e FEC equivalente da união para o ano de 2006.

$$DEC_{eqv-2006} = \frac{15914.19,72 + 2233.97,40}{15914 + 2233} = 29,28$$

$$FEC_{eqv-2006} = \frac{15914.20,23 + 2233.58,72}{15914 + 2233} = 24,97$$

- Cálculo dos Indicadores DEC e FEC equivalente da união para o ano de 2007.

$$DEC_{eqv-2007} = \frac{16451.32,38 + 2851.225,83}{16451 + 2851} = 60,95$$

$$FEC_{eqv-2007} = \frac{16451.28,29 + 2851.131,27}{16451 + 2851} = 43,50$$

- Cálculo do Ano zero para os indicadores DEC e FEC para o ano de 2008.

$$DEC_{Ano-zero} = \frac{20,50 + 29,28 + 60,95}{3} = 36,91$$

$$FEC_{Ano-zero} = \frac{35,20 + 24,97 + 43,50}{3} = 34,56$$

4.5 CONSIDERAÇÕES FINAIS

Este capítulo descreveu a Metodologia Proposta no auxílio a tomada de decisão na formação de conjuntos de unidades consumidoras. A metodologia possui duas fases: a construção de um modelo de classificador a partir dos mesmos dados utilizados na etapa de agrupamento do software ANABENCH e a análise de cenários visando o estudo de configurações de conjuntos de unidades consumidoras que levem a um estado em que exista menos infrações ou que estas sejam menores. Para a construção do modelo de classificador, foram avaliadas o desempenho de duas técnicas pelo método de validação cruzada com 10-*folds*: as redes neurais tipo MLP e o algoritmo kNN. A análise de desempenho dos modelos

mostrou que a técnica mais apropriada para o uso na Metodologia Proposta é o modelo construído via redes neurais tipo MLP. Para melhorar o processo de classificação foi feito um arranjo de classificadores onde cada um apresenta sua resposta de qual *cluster* o novo conjunto deve pertencer. O *cluster* é escolhido por meio de eleição simples das respostas dos classificadores. Por fim, apresentou-se a metodologia a ser empregada na construção de cenários formados pela união de dois ou mais conjuntos já existentes na área de concessão da empresa de distribuição em estudo.

CAPÍTULO 5

ESTUDO DE CASO

5.1 CONSIDERAÇÕES INICIAIS

Este Capítulo apresenta um estudo de caso da aplicação da Metodologia Proposta descrita no capítulo anterior. A metodologia foi aplicada aos conjuntos de unidades consumidoras pertencentes à área de concessão das Centrais Elétricas do Pará S.A.- CELPA. Em um primeiro momento serão apresentadas as características do estado do Pará, área de concessão da CELPA, mostrando as suas peculiaridades regionais que afetam no desempenho dos indicadores de continuidade. Logo depois, serão definidos os cenários a serem apresentados ao classificador. Cada cenário será analisado conforme a metodologia descrita no capítulo anterior.

5.2 O ESTADO DO PARÁ E A CELPA

O estado do Pará está localizado na região norte do Brasil com características climáticas tropical-úmida, possui área territorial de 1.247.689,515 km², sendo o segundo maior estado do país e o mais populoso da região norte com população estimada de 7.431.020 habitantes, resultando em uma densidade populacional de 5,66 Hab/km². O estado possui 144 municípios e está dividido em 6 mesorregiões e 22 microrregiões. As mesorregiões são: Baixo Amazonas, Sudoeste Paraense, Marajó, Sudeste Paraense, Nordeste do Pará e a Região Metropolitana de Belém. A mesorregião do Baixo Amazonas é formada por 14 municípios agrupados em 3 microrregiões: Almerim, Óbidos e Santarém. A mesorregião do Sudoeste Paraense é formada por 14 municípios agrupados em 2 microrregiões: Altamira e Itaituba. A mesorregião do Marajó é formada por 16 municípios agrupados em 3 microrregiões: Arari, Furo de Breves e Portel. A mesorregião do Sudeste Paraense é formada por 49 municípios agrupados em 7 microrregiões: Conceição do Araguaia, Marabá, Paragominas, Parauapebas, Redenção, São Felix do Xingu e Tucuruí. A mesorregião da Região Metropolitana de Belém é formada por 11 municípios agrupados em 2 microrregiões: Belém e Castanhal.

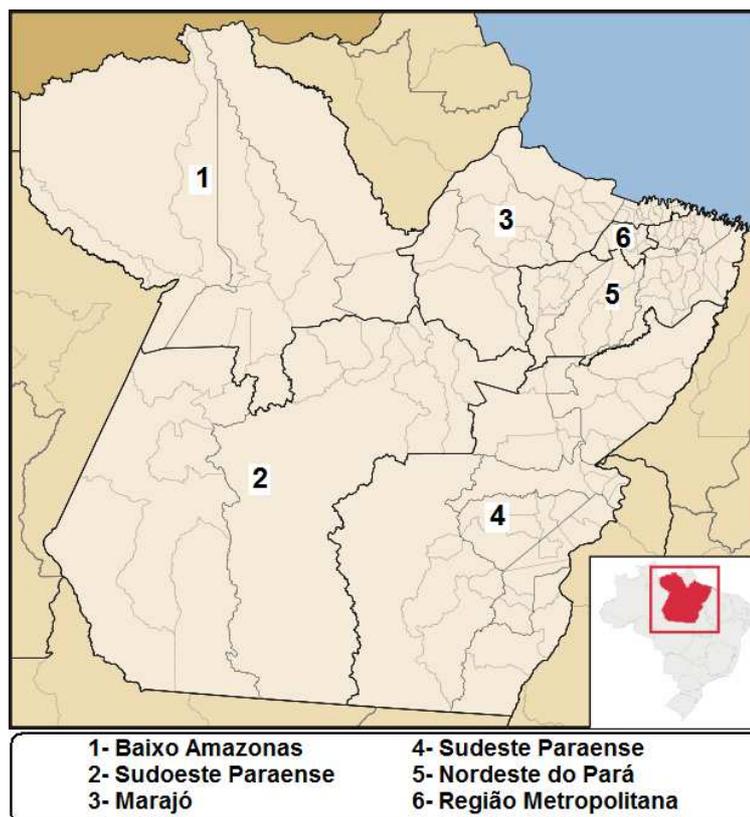


Figura 5.1: Mesorregiões do estado do Pará

A CELPA é uma concessionária pertencente ao Grupo Rede e que atua no estado do Pará desde a privatização das Centrais Elétricas do Pará em 1998. A área de concessão de Rede CELPA engloba todo o estado do Pará e atende cerca de 1.735.210 clientes que perfazem uma média de consumo anual de 3,38 MWh (dados de dezembro de 2006). O consumo residencial da CELPA responde por 37,63% do total da energia por ela vendida. A participação da classe industrial no total do consumo é de 20,95%. A classe comercial participa com 22,12%. O consumo rural, excluídos os consumidores rurais já classificados como residencial baixa renda, representa 1,93% do total. As demais classes, representadas pelos setores públicos, totalizam 17,37% do consumo total.

Atualmente, a área de concessão da CELPA encontra-se dividida em 188 conjuntos de unidades consumidoras. A Figura 5.2 apresenta a divisão da área de concessão da CELPA em 188 conjuntos de unidades consumidoras.

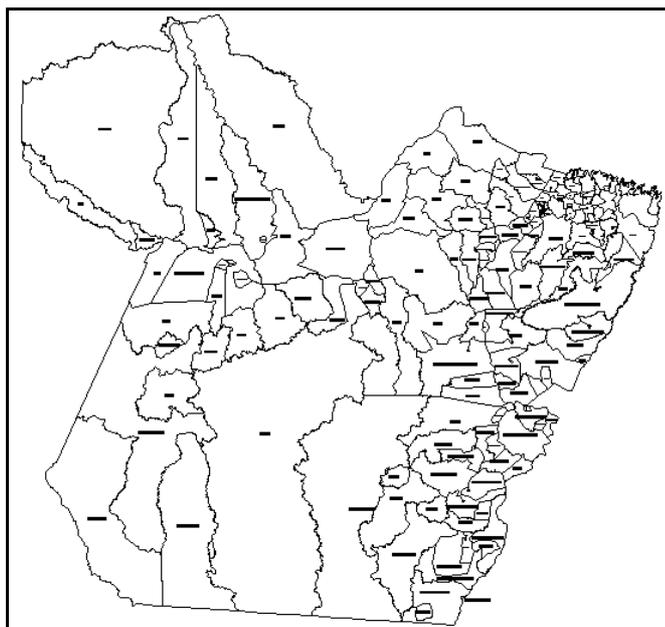


Figura 5.2: Divisão do estado do Pará em 188 conjuntos

A Rede CELPA possui as seguintes particularidades na sua área de concessão que repercutem nos índices de qualidade de fornecimento de energia elétrica:

- Dificuldades de comunicação de estações fixas para as móveis;
- Dificuldades de localização de endereços, por falta de identificação das ruas, de logradouros, de numeração de residências, etc;
- Distâncias maiores a serem percorridas em decorrência da dimensão do estado do Pará e a precariedade e sazonalidades dos acessos rodoviários que no período de inverno amazônico ficam totalmente inundadas sem a possibilidade de trafegabilidade;
- Áreas de preservação ambiental, propriedades particulares com densa arborização sem que seja permitida a execução de podagem e retirada de árvores que ao caírem provocam desligamentos prolongados devido ao nível de dano que provocam à rede de distribuição;
- Região com alto nível cerâmico, que provoca uma considerável queima de elos-fusíveis;
- Região com alto nível pluviométrico provocando interrupções freqüentes;
- Dificuldades na localização das falhas que ocorrem no sistema de distribuição elétrico;
- A implantação do Programa Luz para Todos (PLPT) num estado que possui uma das menores taxas de atendimento na área rural, com baixa densidade de carga, sem os requisitos mínimos necessários de infra-estrutura viária, que compromete o atendimento operacional. O PLPT implantou no período de 2004 a 2007 cerca de

19.700 km de redes de média tensão e este programa até a sua conclusão prevê a implementação de mais de 39.190 km;

- Condições sócio-econômicas desfavoráveis na região, inclusive na própria capital do estado, acarretando várias ligações clandestinas em áreas de invasão não regularizadas, que trazem como conseqüência desligamentos em circuitos projetados e operados adequadamente pela concessionária;
- Conjuntos periféricos com unidades consumidoras com baixo consumo per capita (kWh/UC). Estes conjuntos representam até 49% das unidades consumidoras com a classe de renda abaixo da linha da pobreza;
- Elevado índice de perdas no sistema elétrico da concessionária, onde é forte a cultura de desvio de energia pela população que gera cortes por sobrecarga nos condutores;
- Influência dos fatores não gerenciáveis tais como: pipa, terceiros/vandalismo, animal/ave, vegetação, fogo/queimadas, etc. No ano de 2007 a Rede CELPA totalizou uma porcentagem de 52,34% na apuração final dos indicadores de continuidade.

Devido essas características não serem levadas em conta na metodologia adotada pela ANEEL muitos conjuntos da CELPA acabam sendo agrupados com conjuntos que não compartilham as mesmas características de área de concessão da CELPA. A Figura 5.3 mostra conjuntos que apresentam características muito distintas serem agrupados em um mesmo *cluster*.

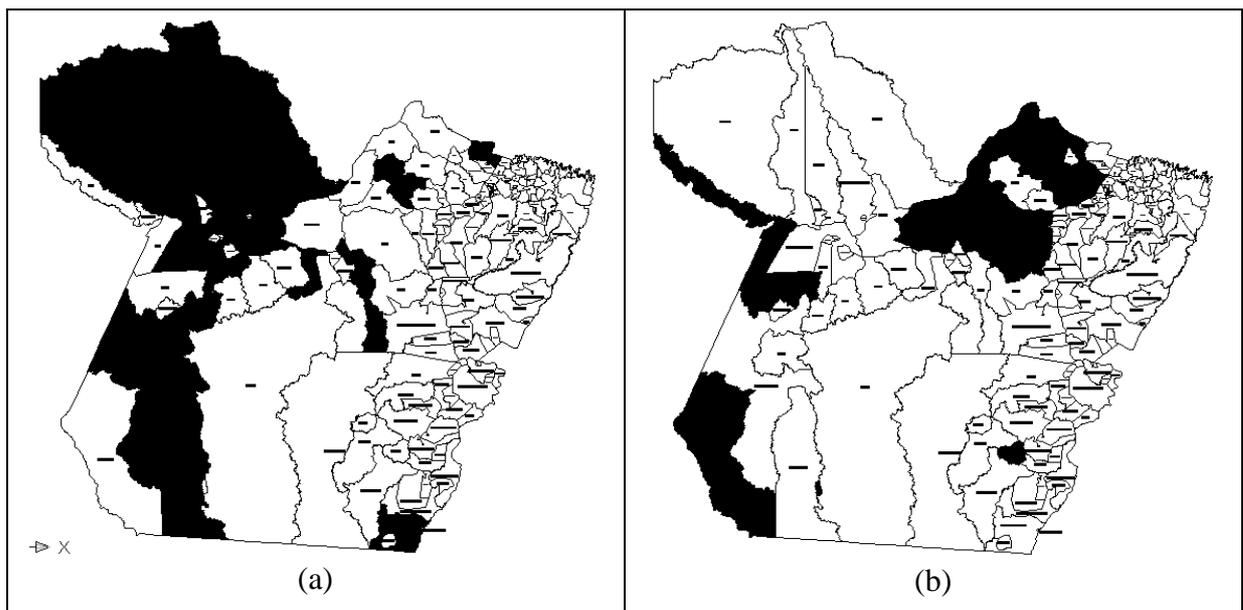


Figura 5.3: Conjuntos alocados no *cluster* 17 (a) e *cluster* 5 (b) segundo ANEEL em 2007

A Figura 5.3 mostra o efeito do uso de apenas cinco atributos na caracterização dos conjuntos de unidades consumidoras de energia elétrica pertencentes à área de concessão da empresa CELPA. A Figura 5.3(a) mostra que os conjuntos que foram alocados no *cluster* 17 estão bastante concentrados na mesorregião do Baixo Amazonas. No entanto, nota-se que neste mesmo *cluster* foram alocados conjuntos que não compartilham as mesmas características geográficas dos conjuntos da mesorregião do Baixo Amazonas. A Figura 5.3(b) mostra que a maioria dos conjuntos da CELPA alocados no *cluster* 5 estão concentrados na mesorregião do Marajó. No entanto, notam-se conjuntos localizados nas mesorregiões do Sudoeste Paraense, Baixo Amazonas e Sudeste Paraense alocados no *cluster* 5. Estes dois casos ilustram que a não inclusão de atributos que melhor caracterizem a área de concessão da CELPA ocasionam uma má classificação dos conjuntos e, com isso, uma má definição de metas a serem alcançadas por esses conjuntos anualmente.

5.3 DEFINIÇÃO DOS CENÁRIOS DE ESTUDO

Para realizar este estudo de caso foram definidos diversos cenários que consistiam da união de dois conjuntos de unidades consumidoras pertencentes à área de concessão da CELPA. A escolha dos conjuntos foi feita de forma aleatória, mas sempre respeitando o critério de contigüidade entre os conjuntos. Cada um desses cenários foi apresentado ao classificador e este informou em qual *cluster* este cenário pertence. Uma vez conhecido o *cluster* onde foi alocado e o tipo de sistema, se interligado ou isolado, é possível por meio da Tabela do Anexo 1 encontrar as metas finais dos indicadores de DEC e FEC para aquele cenário e por meio da metodologia descrita no item 4.4.2 calcular o ano zero do cenário a ser analisado. Uma vez conhecidas a meta final e o ano zero é possível calcular por meio das expressões (2.6) a (2.9) as metas anuais ao longo de duas revisões tarifárias. Este trabalho irá apresentar os resultados obtidos pela análise de 7 cenários apresentados na Tabela 5.1.

Tabela 5.1: Cenários com bons resultados de metas de DEC e FEC

Cenário	Conjuntos envolvidos
C1	Belém Perimetral e Bolonha urbano.
C2	Ulianópolis urbano e Ulianópolis não-urbano.
C3	Paragominas urbano e Paragominas não urbano.
C4	Santa Maria não-urbano e São Miguel do Guamá.
C5	Água Azul do Norte e Canaã dos Carajás.
C6	Curionópolis e Eldorado dos Carajás.
C7	Abaetetuba e Igarapé Mirim.

Cada cenário consiste de um novo conjunto que é caracterizado por novos 5 atributos que serão apresentados às 10 redes neurais. A Tabela 5.2 mostra os resultados obtidos ao se apresentar os novos atributos dos cenários. A escolha do *cluster* é obtida por meio de eleição simples das respostas das redes neurais.

Tabela 5.2: Respostas das redes neurais e *cluster* escolhido

Cenário	RN1	RN2	RN3	RN4	RN5	RN6	RN7	RN8	RN9	RN10	Escolhido
C1	7	23	7	23	23	23	23	23	23	23	23
C2	6	15	30	6	3	6	15	6	30	6	6
C3	26	9	26	18	17	9	9	9	18	26	9
C4	9	11	11	11	11	11	11	11	11	11	11
C5	29	17	17	17	17	17	17	17	29	29	17
C6	11	11	17	17	17	11	11	11	17	11	11
C7	17	17	17	17	17	17	17	17	17	17	17

A seguir serão mostradas as análises de cada um desses sete cenários.

5.4 ANÁLISE DOS CENÁRIOS

5.4.1 Cenário 1: Belém - Perimetral e Bolonha urbano

O primeiro cenário a ser analisado é a união dos conjuntos Bolonha urbano e Belém - Perimetral que são dois conjuntos do tipo interligado pertencentes à área metropolitana de Belém e que possuem as seguintes características resumidas na Tabela 5.3.

Tabela 5.3: Atributos dos conjuntos Bolonha urbano e Belém - Perimetral

CONJUNTO	ERAP (km)	AREA (km ²)	PNI (kVA)	CMM (MWh)	NUC
Bolonha urbano	170,62	197,12	51167,50	6297,67	27052
Belém - Perimetral	107,45	72,55	100947,50	15645,3	57712

O conjunto Belém - Perimetral apresenta maior consumo médio mensal, mais que o dobro de unidades consumidoras e quase o dobro de potência nominal instalada do que o conjunto Bolonha urbano. Na classificação feita pelo software ANABENCH o conjunto Belém - Perimetral foi alocado no *cluster* 23 cujas metas finais são, para conjuntos do tipo interligado, 6 horas/ano e 5 interrupções/ano para os indicadores DEC e FEC, respectivamente. O conjunto Bolonha urbano foi alocado no *cluster* 8 cujas metas finais são,

para conjuntos do tipo interligado, 8 horas/ano e 8 interrupções/ano para os indicadores DEC e FEC, respectivamente.

Ao se apresentar os novos cinco atributos resultantes da união destes dois conjuntos, o classificador alocou o cenário 1 no *cluster* 23 cujas metas finais são 6 horas/ano e 5 interrupções/ano para os indicadores DEC e FEC, respectivamente. A partir destas informações mais a informação do valor dos indicadores no ano zero utilizam-se as equações (2.6) a (2.9) para determinar as metas anuais no período compreendido entre duas revisões tarifárias (8 anos). A Tabela 5.4 resume as principais informações necessárias para definição das metas anuais.

Tabela 5.4: Ano zero e valor da meta final dos indicadores DEC e FEC para o cenário 1

Conjunto	DEC		FEC	
	Ano 0	Meta	Ano 0	Meta
Cenário 1	45,77	6	31,42	5
Bolonha urbano	72,62	8	55,11	8
Belém – Perimetral	34,70	6	21,67	5

A Figura 5.4 e a Tabela 5.5 mostram o gráfico das metas e os seus respectivos valores ao longo de duas revisões tarifárias para o cenário 1 e para os conjuntos formadores deste cenário para o indicador de continuidade DEC.

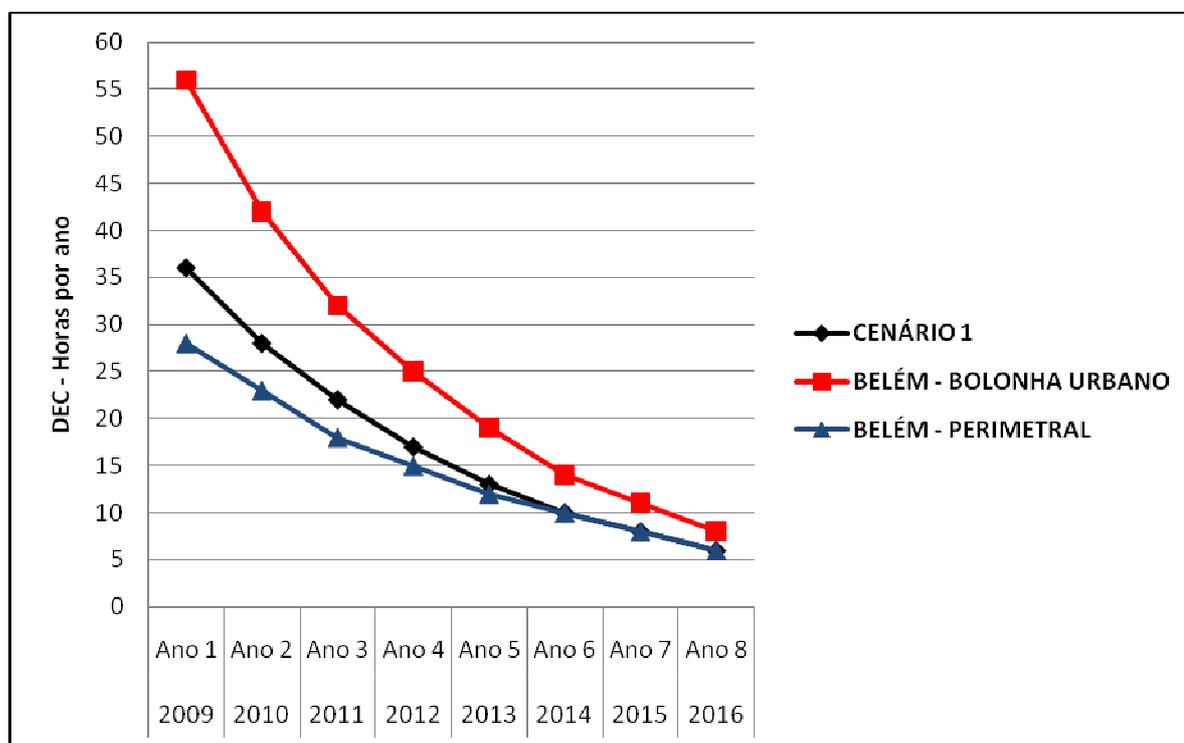


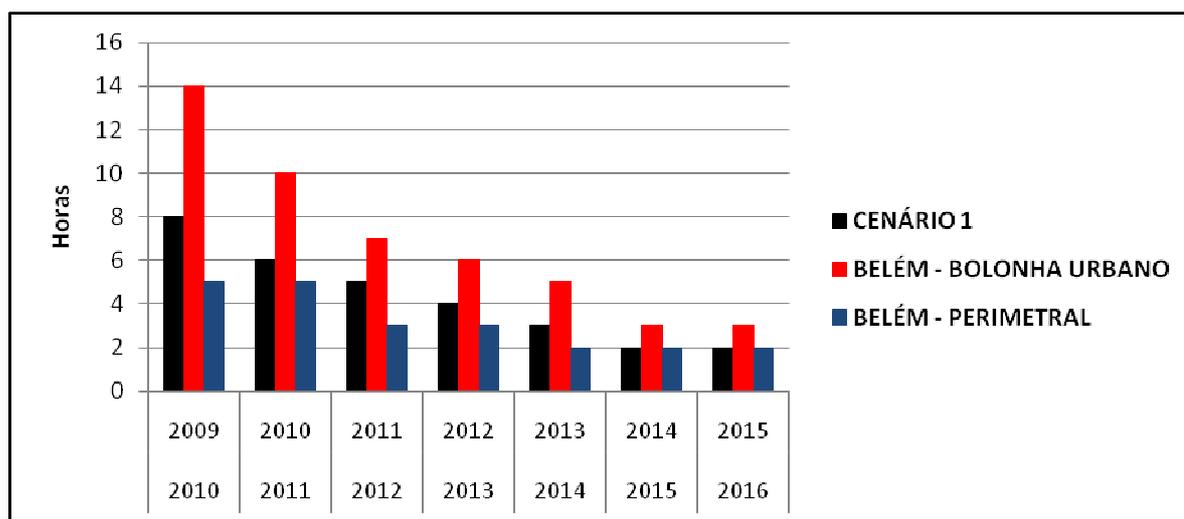
Figura 5.4: Metas do cenário 1 ao longo de 8 anos para o indicado DEC

Tabela 5.5: Metas para o cenário 1 ao longo dos 8 anos para o indicador DEC

Conjunto	Metas Anuais							
	2009	2010	2011	2012	2013	2014	2015	2016
Cenário 1	36	28	22	17	13	10	8	6
Bolonha Urbano	56	42	32	25	19	14	11	8
Belém – Perimetral	28	23	18	15	12	10	8	6

Comparando os gráficos de metas do indicador DEC do conjunto Bolonha urbano com o conjunto Belém – Perimetral nota-se que o conjunto Belém – Perimetral apresenta as metas mais rígidas se comparadas ao conjunto Bolonha urbano. O conjunto Bolonha urbano apresenta uma elevada taxa de redução anual entre as metas de dois anos consecutivos, observada principalmente nos 4 primeiros anos se comparada ao conjunto Belém - Perimetral. Essa característica não é muito desejada, pois exige que a concessionária alcance uma meta mais rígida em um período de tempo de um ano.

Ao se efetuar a união do conjunto Bolonha urbano com o conjunto Belém - Perimetral notam-se folgas de 8 horas no primeiro ano, 5 horas no segundo ano, 4 horas no terceiro ano, 2 horas no quarto ano e 1 hora no quinto ano nas metas do conjunto Belém - Perimetral que é o conjunto que possui o maior número de unidades consumidoras. Observou-se diminuição da taxa de redução anual entre metas de anos consecutivos para o conjunto Bolonha Urbano se comparadas ao cenário 1. A Figura 5.5 ilustra a taxa de redução anual do cenário 1 e dos conjuntos formadores deste cenário.

**Figura 5.5: Taxa de redução anual do cenário 1 para o indicador DEC**

Ainda na Figura 5.5 nota-se que a taxa de redução observada no conjunto Belém – Perimetral não sofreram grande variação se comparadas à taxa de redução do cenário 1.

A Tabela 5.6 apresenta uma comparação entre as metas obtidas por meio da aplicação da Metodologia Proposta para o indicador DEC e o valor do DEC agregado obtido por meio da média dos valores das metas do indicador DEC dos conjuntos envolvidos na união ponderados pelo número de unidades consumidoras para o primeiro e oitavo anos.

Tabela 5.6: Comparação do DEC agregado com o obtido pela Metodologia Proposta para o cenário 1

DEC	2009	2016
Agregado	36,94	6,64
Metodologia proposta	36	6

A Tabela 5.6 mostra que apesar da união ter proporcionado folgas nas metas para o conjunto Belém-Perimetral que é o conjunto com maior número de unidades consumidoras a união não é satisfatória para o indicador DEC, pois os valores de metas obtidos pela metodologia são similares aos obtidos pelo DEC agregado.

A Figura 5.6 e a Tabela 5.7 mostram o gráfico das metas e os valores das metas e ao longo de oito anos para o cenário 1 e para os conjuntos formadores deste cenário para o indicador de continuidade FEC.

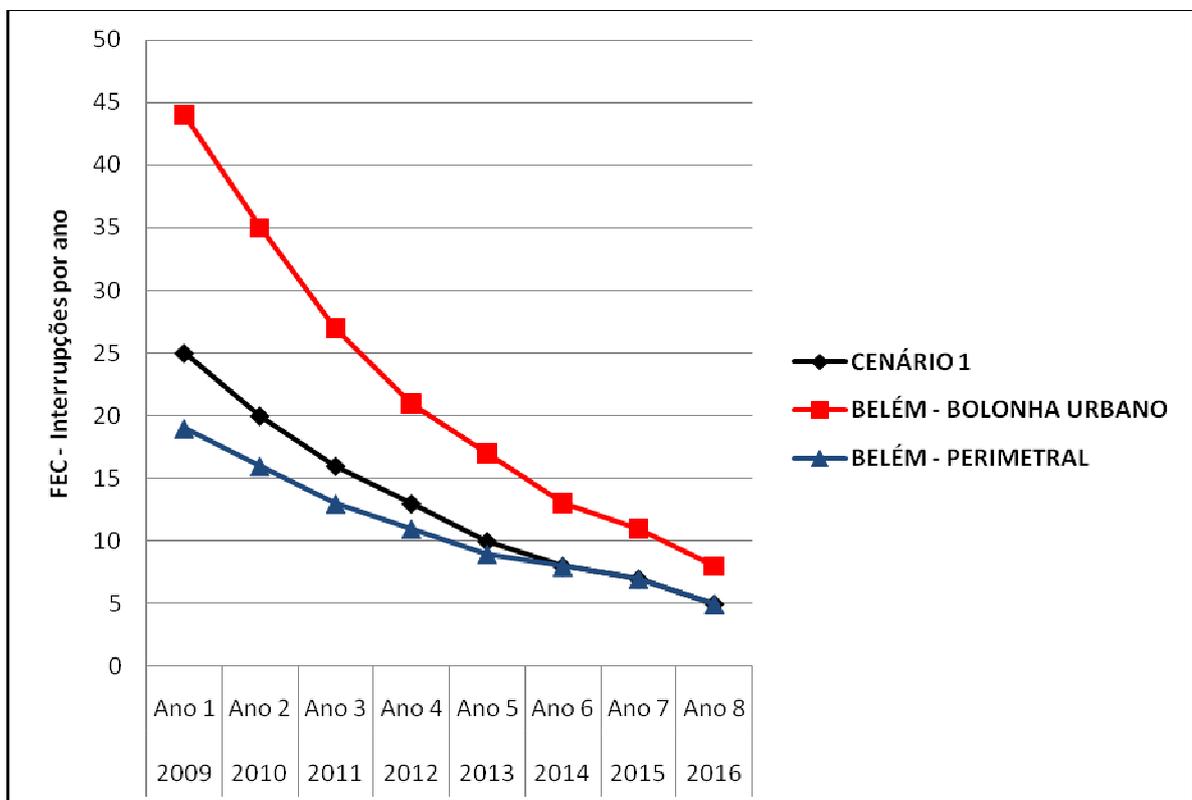


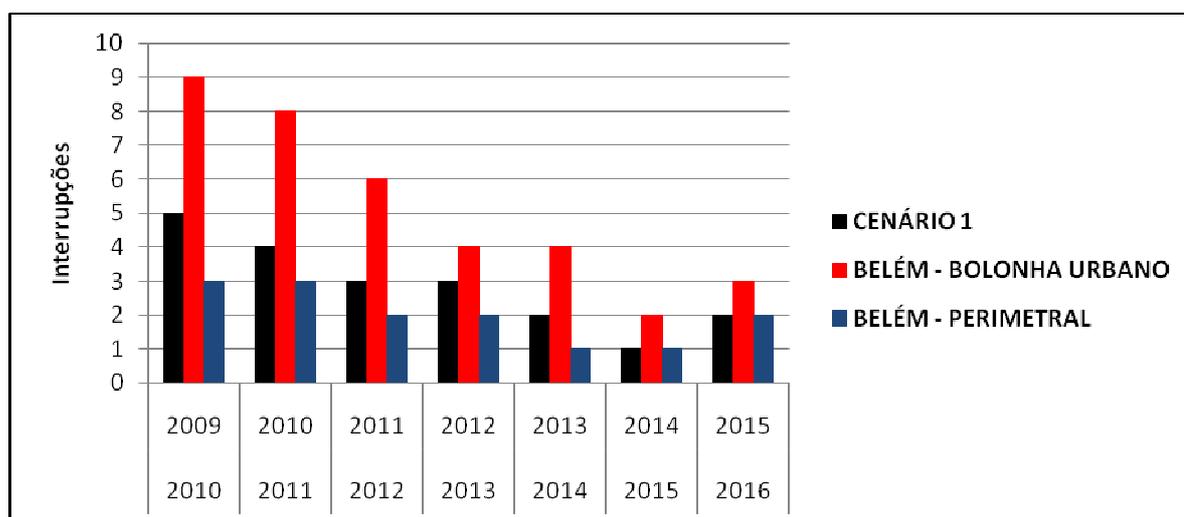
Figura 5.6: Metas do cenário 1 ao longo de 8 anos para o indicado FEC

Tabela 5.7: Metas para o cenário 1 ao longo dos 8 anos para o indicador FEC

Conjunto	Metas Anuais							
	2009	2010	2011	2012	2013	2014	2015	2016
Cenário 1	25	20	16	13	10	8	7	5
Bolonha Urbano	44	35	27	21	17	13	11	8
Belém – Perimetral	19	16	13	11	9	8	7	5

Analisando o gráfico 5.6 nota-se que o comportamento das metas do indicador DEC também foi observado no indicador FEC: metas mais rígidas para o conjunto Belém – Perimetral e elevada taxa de redução anual entre as metas de dois anos consecutivos, observada principalmente nos quatro primeiros anos no conjunto Bolonha urbano.

A união dos conjuntos proporcionou as mesmas características observadas para o indicador DEC: folgas nas metas do conjunto com maior número de unidades consumidoras, maior potência instalada e maior consumo médio mensal e diminuição da taxa de redução anual para o conjunto Bolonha urbano. As folgas nas metas de FEC no conjunto Belém – perimetral são de 6 interrupções para o primeiro ano, 4 interrupções para o segundo ano, 3 interrupções no terceiro ano, 2 interrupções para o quarto e 1 interrupção para o quinto ano. Não houve variação significativa na taxa de redução do cenário 1 se comparada com o conjunto Belém – Perimetral. A Figura 5.7 ilustra a diminuição da taxa de redução anual ao se efetuar a união dos conjuntos para o indicador FEC.

**Figura 5.7: Taxa de redução anual para o cenário 1 para o indicador FEC**

A comparação das metas obtidas pela Metodologia Proposta e o FEC agregado é apresentada na Tabela 5.8.

Tabela 5.8: Comparação do FEC agregado com o obtido pela Metodologia Proposta para o cenário 1

FEC	2009	2016
Agregado	26,98	5,96
Metodologia proposta	25	5

A Tabela 5.8 mostra que apesar da união ter proporcionado folgas nas metas do conjunto Belém–Perimetral nos cinco primeiros anos para o indicador FEC, a análise comparativa com o FEC agregado mostrou que a união não é satisfatória devido aos valores de FEC obtidos por meio da Metodologia serem similares aos obtidos por meio do FEC agregado.

5.4.2 Cenário 2: Ulianópolis urbano e Ulianópolis não urbano

O segundo cenário a ser analisado é a união dos conjuntos Ulianópolis urbano e Ulianópolis não urbano que são dois conjuntos do tipo interligado e possuem as seguintes características resumidas na Tabela 5.9.

Tabela 5.9: Atributos dos conjuntos Ulianópolis urbano e Ulianópolis não urbano

CONJUNTO	ERAP (km)	AREA (km²)	PNI (kVA)	CMM (MWh)	NUC
Ulianópolis não urbano	230,24	5095	2957,50	155,56	305
Ulianópolis urbano	25,91	18	10100	1135,49	3493

O conjunto Ulianópolis urbano apresenta cerca de 10 vezes mais unidades consumidoras, aproximadamente 7 vezes o consumo médio mensal e 3 vezes a potência instalada do que o conjunto Ulianópolis não urbano. Na classificação feita pelo software ANABENCH o conjunto Ulianópolis urbano foi alocado no *cluster* 13 cujas metas finais são, para conjuntos do tipo interligado, 10 horas/ano e 8 interrupções/ano para os indicadores DEC e FEC, respectivamente. O conjunto Ulianópolis não urbano foi alocado no *cluster* 28 cujas metas finais são, para conjuntos do tipo interligado, 32 horas/ano e 26 interrupções/ano para os indicadores DEC e FEC, respectivamente.

Ao se apresentar os novos cinco atributos resultantes da união destes dois conjuntos, o classificador alocou o cenário 2 no *cluster* 6 cujas metas finais são 23 horas/ano e 27 interrupções/ano para os indicadores DEC e FEC, respectivamente. A partir destas informações mais a informação do valor dos indicadores no ano zero utilizam-se as equações

(2.6) a (2.9) para determinar das metas no período compreendido entre duas revisões tarifárias. A Tabela 5.10 resume as principais informações necessárias à definição das metas.

Tabela 5.10: Ano zero e valor da meta final dos indicadores DEC e FEC para o cenário 2

Conjunto	DEC		FEC	
	Ano 0	Meta	Ano 0	Meta
Cenário 2	71,76	23	49,87	27
Ulianópolis não urbano	76,44	32	51,79	26
Ulianópolis urbano	71,34	10	49,69	8

A Tabela 5.11 e a Figura 5.8 mostram os valores das metas e o gráfico de decaimento das metas ao longo de oito anos para o cenário 2 e para os conjuntos formadores deste cenário para o indicador de continuidade DEC.

Tabela 5.11: Metas para o cenário 2 ao longo dos 8 anos para o indicador DEC

Conjunto	Metas Anuais							
	2009	2010	2011	2012	2013	2014	2015	2016
Cenário 2	63	54	47	41	36	31	27	23
Ulianópolis não urbano	69	62	56	50	45	40	36	32
Ulianópolis urbano	56	44	35	27	21	17	13	10

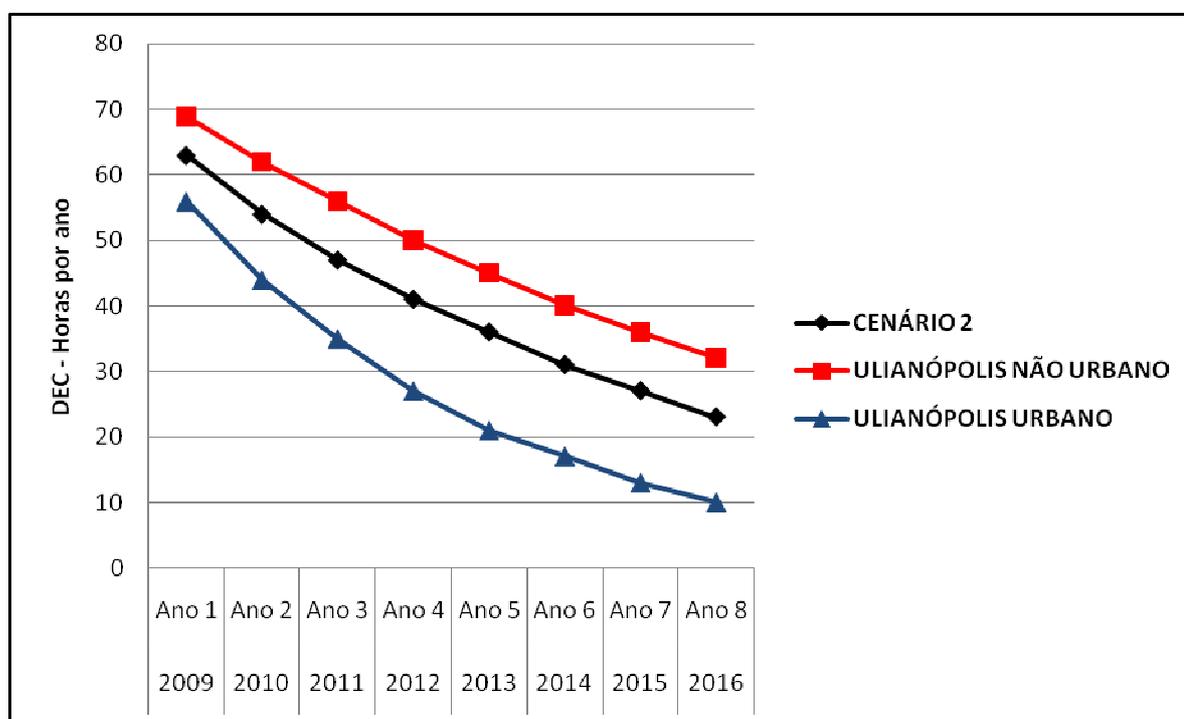


Figura 5.8: Metas do cenário 2 ao longo de 8 anos para o indicado DEC

Comparando os gráficos de metas do indicador DEC do conjunto Ulianópolis urbano com o conjunto Ulianópolis não urbano observa-se que o conjunto Ulianópolis urbano apresenta as metas mais rígidas e uma elevada taxa de redução anual entre as metas de dois anos consecutivos, observada principalmente nos 4 primeiros anos se comparadas ao conjunto Ulianópolis não urbano.

A união do conjunto Ulianópolis urbano com o conjunto Ulianópolis não urbano proporcionou folgas de 7 horas no primeiro ano, 10 horas no segundo ano, 12 horas no terceiro ano, 14 horas no quarto ano, 15 horas no quinto ano, 14 horas no sexto e sétimo anos e 13 horas no oitavo ano nas metas do conjunto Ulianópolis urbano que é o conjunto que possui o maior número de unidades consumidoras além da diminuição da taxa de redução anual entre metas de anos consecutivos para o mesmo conjunto. A taxa de redução do cenário 2 apresentou variação de 2 horas nos anos de 2009 à 2010 e 1 hora nos anos de 2010 à 2011 se comparada ao conjunto Ulianópolis não urbano. A Figura 5.9 ilustra a diminuição da taxa de redução anual ao se efetuar a união dos conjuntos em relação ao indicador DEC.

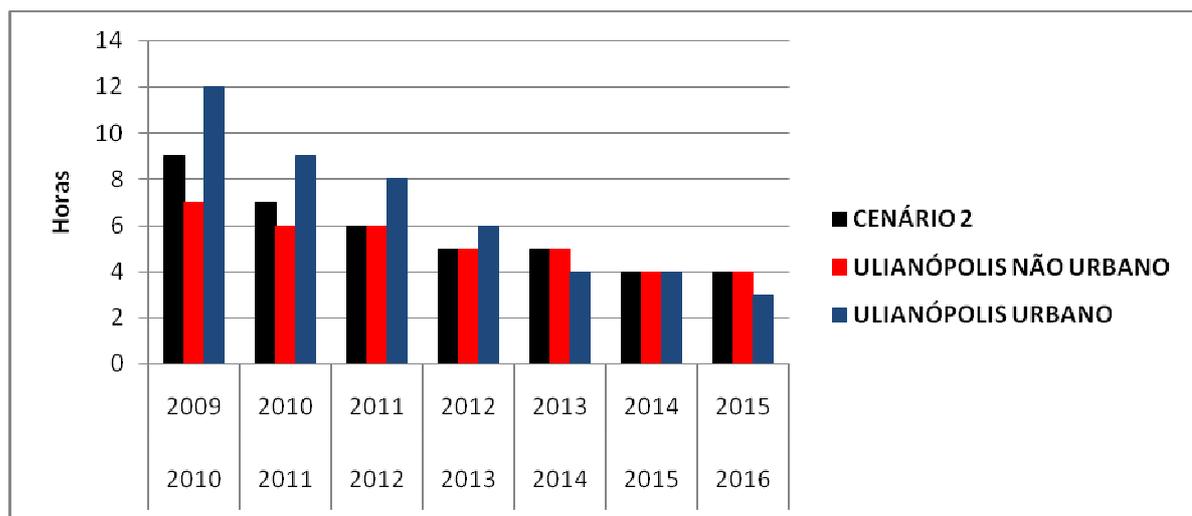


Figura 5.9: Taxa de redução anual para o cenário 2 para o indicador DEC

A comparação entre as metas obtidas pela Metodologia Proposta e o DEC agregado para o primeiro e oitavo anos esta apresentada na Tabela 5.12.

Tabela 5.12: Comparação do DEC agregado com o obtido pela Metodologia Proposta para o cenário 2

DEC	2009	2016
Agregado	57,04	11,77
Metodologia proposta	63	23

A comparação mostra que é vantajoso para o indicador DEC a união dos conjuntos Ulianópolis urbano e Ulianópolis não-urbano, pois as metas obtidas pela Metodologia são superiores às obtidas pelo DEC agregado.

A Tabela 5.13 e a Figura 5.10 mostram os valores das metas e o gráfico de decaimento das metas ao longo de oito anos para o cenário 2 e para os conjuntos formadores deste cenário para o indicador de continuidade FEC.

Tabela 5.13: Metas para o cenário 2 ao longo dos 8 anos para o indicador FEC

Conjunto	Metas Anuais							
	2009	2010	2011	2012	2013	2014	2015	2016
Cenário 2	47	43	40	37	34	32	30	27
Ulianópolis não urbano	48	44	40	37	34	31	29	26
Ulianópolis urbano	40	32	26	20	16	13	11	8

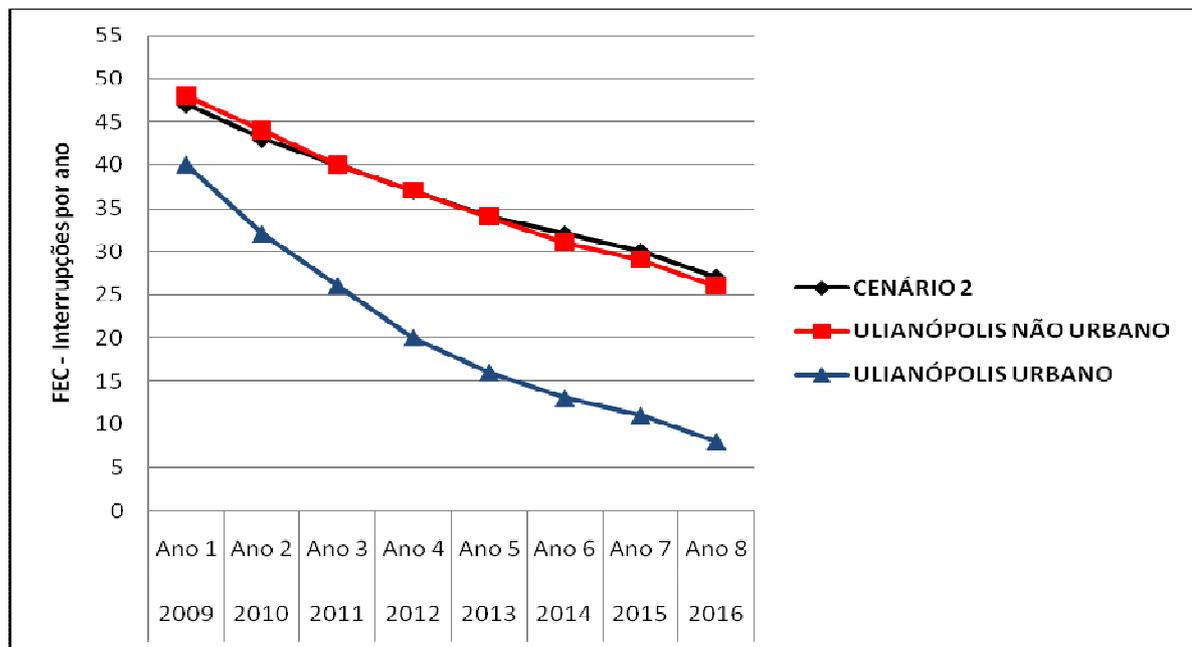


Figura 5.10: Metas do cenário 2 ao longo de 8 anos para o indicado FEC

A comparação dos gráficos de metas do indicador FEC do conjunto Ulianópolis urbano com o conjunto Ulianópolis não urbano mostra que o conjunto Ulianópolis urbano apresenta elevada taxa de redução anual e possui as metas mais rígidas se comparado ao conjunto Ulianópolis não urbano.

A união dos conjuntos fez com que a curva de metas do cenário 2 se assemelhasse à curva de metas do conjunto Ulianópolis não urbano o que proporcionou uma folga nas metas do conjunto Ulianópolis urbano que é o conjunto que apresenta maior número de unidades

consumidoras, maior potência instalada e maior consumo médio mensal. As folgas nas metas de FEC no conjunto Ulianópolis urbano são de 7 interrupções para o primeiro ano, 11 interrupção para o segundo ano, 14 interrupções no terceiro ano, 17 interrupções para o quarto ano, 18 interrupções para o quinto ano e 19 interrupções para os três últimos anos. Notou-se diminuição da taxa de redução anual para o conjunto Ulianópolis urbano. A Figura 5.11 ilustra a diminuição da taxa de redução anual ao se efetuar a união dos conjuntos para o indicador FEC.

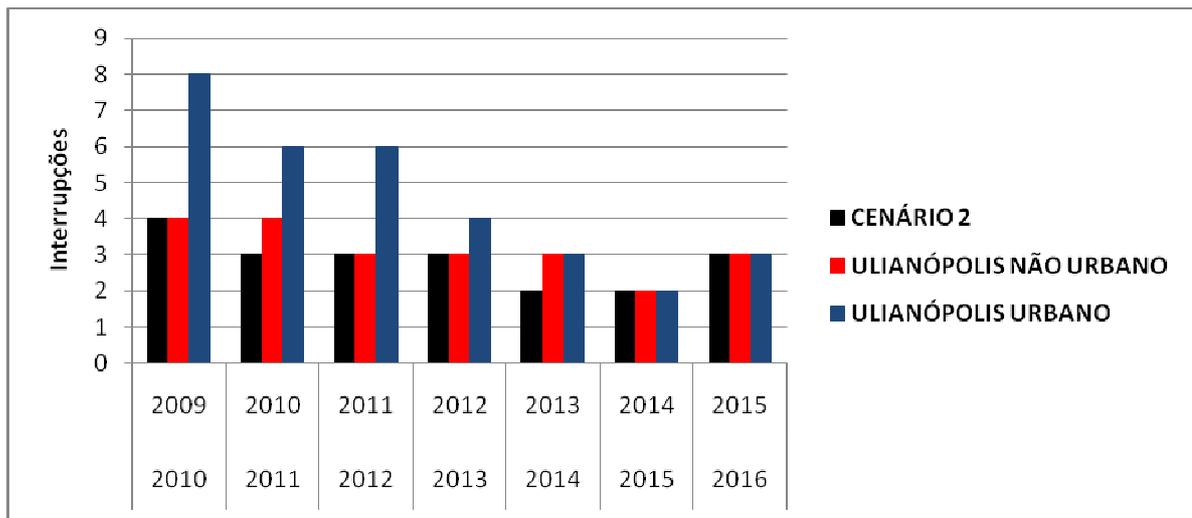


Figura 5.11: Taxa de redução anual para o cenário 2 para o indicador FEC

A Tabela 5.14 apresenta a comparação entre as metas obtidas pela Metodologia Proposta e o FEC agregado para o primeiro e último ano de metas.

Tabela 5.14: Comparação do FEC agregado com o obtido pela Metodologia Proposta para o cenário 2

FEC	2009	2016
Agregado	40,64	9,45
Metodologia proposta	47	27

A Tabela 5.14 mostra que é vantajoso para o indicador FEC a união dos conjuntos Ulianópolis urbano e Ulianópolis não urbano, pois os valores obtidos pela Metodologia Proposta para os anos de 2009 e 2016 são superiores aos obtidos por meio do FEC agregado para os mesmos anos considerados.

5.4.3 Cenário 3: Paragominas urbano e Paragominas não urbano

O terceiro cenário a ser analisado é a união dos conjuntos Paragominas urbano e Paragominas não urbano que são dois conjuntos do tipo interligado e possuem as seguintes características mostradas na Tabela 5.15.

Tabela 5.15: Atributos dos conjuntos Paragominas urbano e Paragominas não urbano

CONJUNTO	ERAP (km)	AREA (km ²)	PNI (kVA)	CMM (MWh)	NUC
Paragominas não urbano	1708,79	20272	33172,50	1637,48	4811
Paragominas urbano	76,11	45	36185	5981,79	19156

O conjunto Paragominas urbano apresenta aproximadamente 4 vezes mais unidades consumidoras e um consumo médio mensal 3,5 vezes maior do que o conjunto Paragominas não urbano. Na classificação feita pelo software ANABENCH o conjunto Paragominas urbano foi alocado no *cluster* 12 cujas metas finais são, para conjuntos do tipo interligado, 10 horas/ano e 8 interrupções/ano para os indicadores DEC e FEC, respectivamente. O conjunto Paragominas não urbano foi alocado no *cluster* 11 cujas metas finais são, para conjuntos do tipo interligado, 33 horas/ano e 19 interrupções/ano para os indicadores DEC e FEC, respectivamente.

Ao se apresentar os novos cinco atributos resultantes da união destes dois conjuntos, o classificador alocou o cenário 3 no *cluster* 9 cujas metas são 28 horas/ano e 18 interrupções/ano para os indicadores DEC e FEC, respectivamente. A partir destas informações mais a informação do valor dos indicadores no ano zero utilizam-se as equações (2.6) a (2.9) para determinar das metas no período compreendido entre duas revisões tarifárias. A Tabela 5.12 resume as principais informações necessárias à definição das metas.

Tabela 5.16: Ano zero e valor da meta final dos indicadores DEC e FEC para o cenário 3

Conjunto	DEC		FEC	
	Ano 0	Meta	Ano 0	Meta
Cenário 3	28,10	28	23,85	18
Paragominas não urbano	37,25	33	27,27	19
Paragominas urbano	22,82	10	21,67	8

A Tabela 5.17 e a Figura 5.12 apresentam os valores das metas ao longo de duas revisões tarifárias e gráfico de decaimento para o cenário 3 e para os conjuntos formadores deste cenário para o indicador de continuidade DEC.

Tabela 5.17: Metas para o cenário 3 ao longo dos 8 anos para o indicador DEC

Conjunto	Metas Anuais							
	2009	2010	2011	2012	2013	2014	2015	2016
Cenário 3	29	29	29	29	29	29	29	28
Paragominas não urbano	37	37	36	36	35	35	34	33
Paragominas urbano	21	19	17	16	14	13	12	10

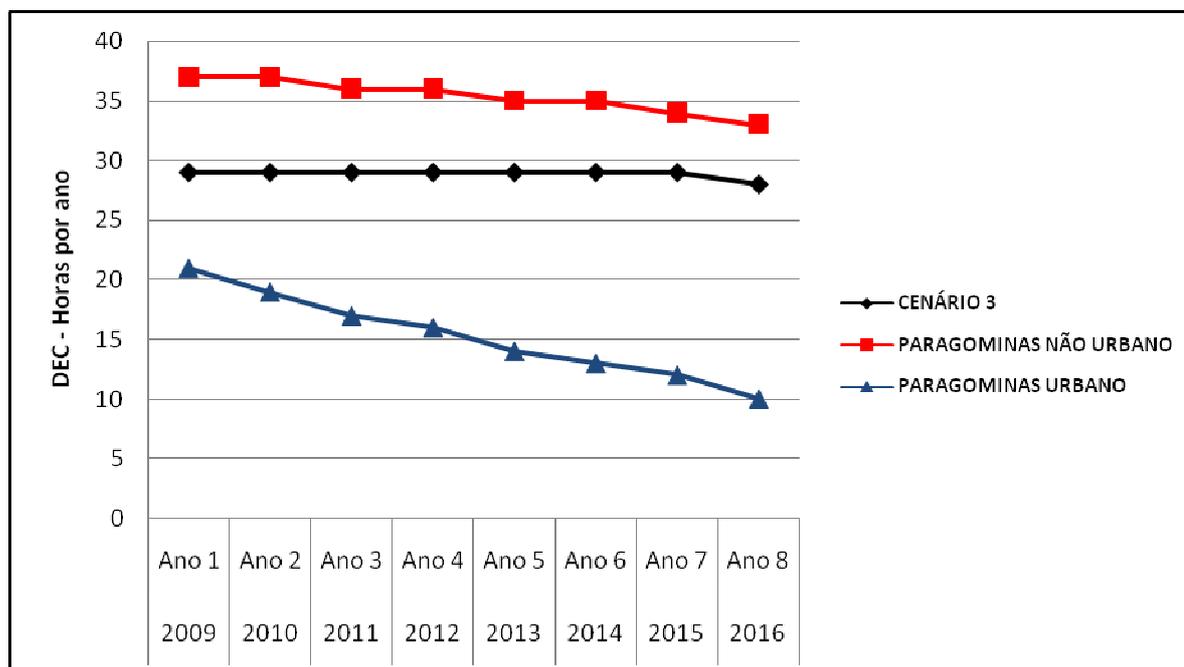


Figura 5.12: Metas do cenário 3 ao longo de 8 anos para o indicado DEC

Comparando os gráficos de metas do indicador DEC do conjunto Paragominas urbano com o conjunto Paragominas não urbano observa-se que o conjunto Paragominas urbano apresenta as metas mais rígidas e uma elevada taxa de redução anual.

Ao se efetuar a união do conjunto Paragominas urbano com o conjunto Paragominas não urbano constatam-se folgas de 8 horas no primeiro ano, 10 horas no segundo ano, 12 horas no terceiro ano, 13 horas no quarto ano, 15 horas no quinto ano, 16 horas no sexto, 17 horas no sétimo ano e 18 horas no oitavo ano nas metas do conjunto Paragominas urbano que é o conjunto que possui o maior número de unidades consumidoras além da diminuição da taxa de redução anual tanto para o conjunto Paragominas urbano como para o conjunto

Paragominas não urbano. A Figura 5.13 ilustra a diminuição da taxa de redução anual ao se efetuar a união dos conjuntos para o indicador DEC.

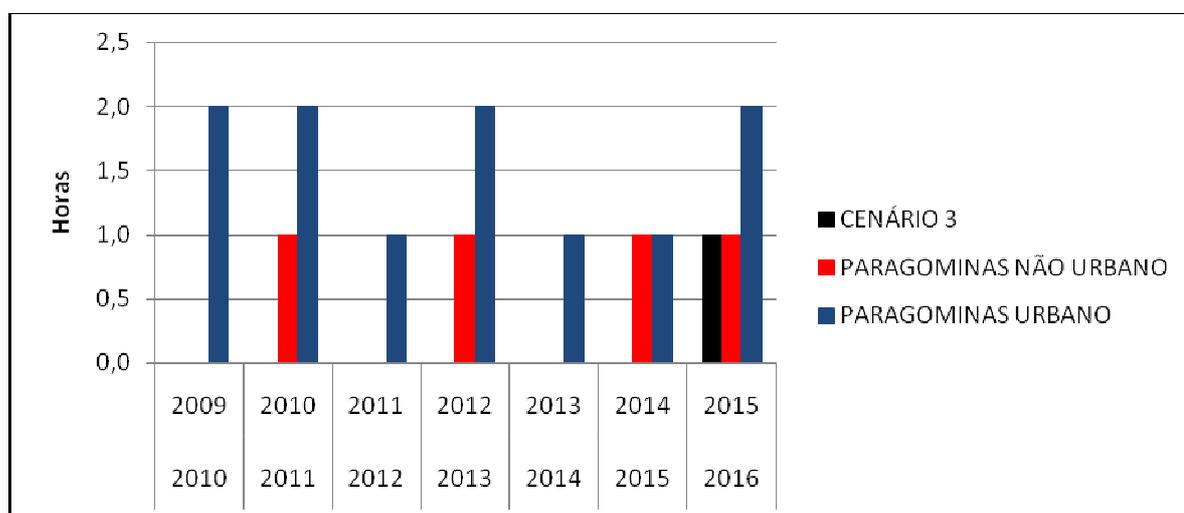


Figura 5.13: Taxa de redução anual para o cenário 3 para o indicador DEC

A Tabela 5.18 mostra a comparação das metas obtidas pela Metodologia Proposta com o DEC agregado dos conjuntos participantes da formação do cenário 3 para o primeiro e oitavo anos.

Tabela 5.18: Comparação do DEC agregado com o obtido pela Metodologia Proposta para o cenário 3.

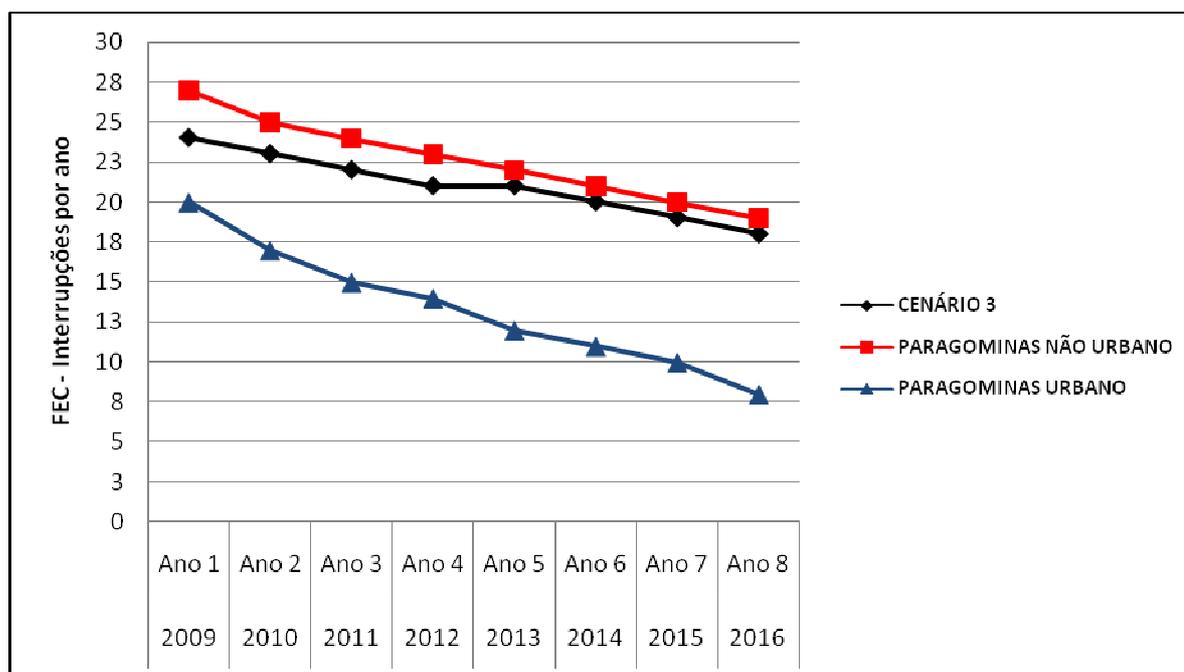
DEC	2009	2016
Agregado	24,21	14,62
Metodologia proposta	29	28

Comparando-se os valores das metas obtidas pela aplicação da Metodologia com o DEC agregado dos conjuntos Paragominas Urbano e Paragominas não-urbano, nota-se que os valores das metas do indicador DEC obtidas pela Metodologia são superiores aos valores de DEC agregado para os anos de 2009 e 2016. Portanto, a união desses conjuntos é satisfatória para o indicador DEC.

A Tabela 5.19 e a Figura 5.14 mostram os valores das metas e o gráfico de decaimento das metas ao longo de oito anos para o cenário 3 e para os conjuntos formadores deste cenário para o indicador de continuidade FEC.

Tabela 5.19: Metas para o cenário 3 ao longo dos 8 anos para o indicador FEC

Conjunto	Metas Anuais							
	2009	2010	2011	2012	2013	2014	2015	2016
Cenário 3	24	23	22	21	21	20	19	18
Paragominas não urbano	27	25	24	23	22	21	20	19
Paragominas urbano	20	17	15	14	12	11	10	8

**Figura 5.14: Metas do cenário 3 ao longo de 8 anos para o indicado FEC**

A comparação dos gráficos de metas do indicador FEC do conjunto Paragominas urbano com o conjunto Paragominas não urbano mostra que o conjunto Paragominas urbano apresenta elevada taxa de redução anual e possui as metas mais rígidas se comparadas ao conjunto Paragominas não urbano.

A união dos conjuntos proporcionou folgas nas metas do conjunto Paragominas urbano que é o conjunto que apresenta maior número de unidades consumidoras, maior potência instalada e maior consumo médio mensal. As folgas nas metas de FEC no conjunto Paragominas urbano são de 4 interrupções para o primeiro ano, 6 interrupção para o segundo ano, 7 interrupções no terceiro e quarto anos, 9 interrupções para o quinto, sexto e sétimo anos e 10 interrupções para o último ano. Observa-se diminuição da taxa de redução anual para o conjunto Paragominas urbano e Paragominas não urbano. A taxa de redução anual do cenário 3 ficou constante exceto para o período entre 2012 a 2013. A Figura 5.15 ilustra a diminuição da taxa de redução anual ao se efetuar a união dos conjuntos para o indicador FEC.

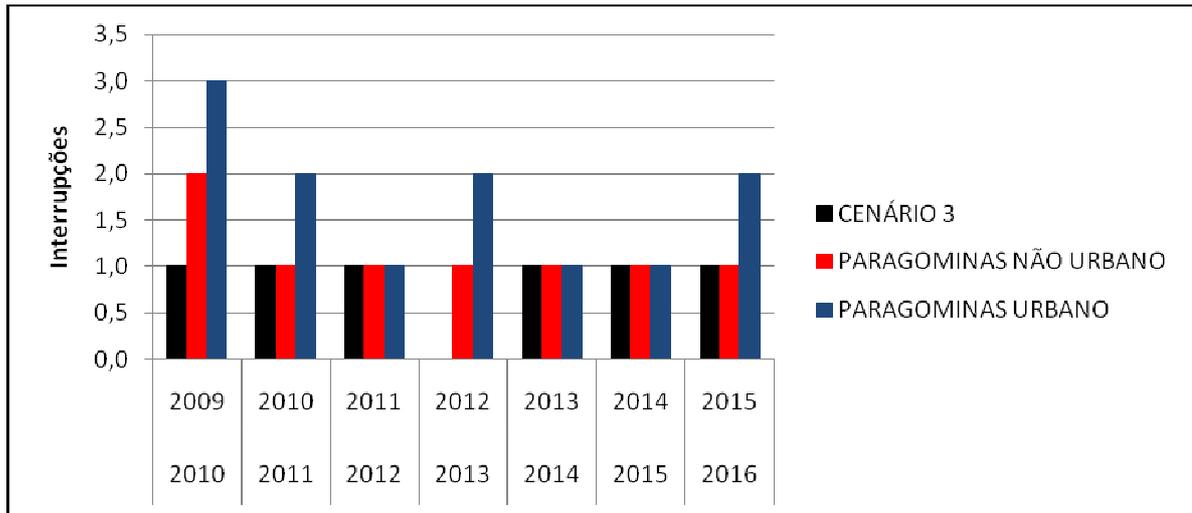


Figura 5.15: Taxa de redução anual para o cenário 3 para o indicador FEC

A Tabela 5.20 mostra a comparação das metas obtidas pela Metodologia Proposta com o FEC agregado dos conjuntos Paragominas Urbano e Paragominas não-urbano para os anos de 2009 e 2016.

Tabela 5.20: Comparação do FEC agregado com o obtido pela Metodologia Proposta para o cenário 3

FEC	2009	2016
Agregado	24,21	14,62
Metodologia proposta	29	28

A comparação dos valores das metas obtidas pela aplicação da Metodologia com o FEC agregado dos conjuntos participantes da formação do cenário 3 mostrou que os valores das metas do indicador FEC obtidas pela Metodologia são superiores aos valores de FEC agregado para o primeiro e último anos. Portanto, a união desses conjuntos é satisfatória para o indicador FEC.

5.4.4 Cenário 4: Água Azul do Norte e Canaã dos Carajás

O quarto cenário a ser analisado é a união dos conjuntos Água Azul do Norte e Canaã dos Carajás que são dois conjuntos do tipo interligado e possuem as seguintes características resumidas na Tabela 5.21.

Tabela 5.21: Atributos dos conjuntos Água Azul do Norte e Canaã dos Carajás

CONJUNTO	ERAP (km)	AREA (km ²)	PNI (kVA)	CMM (MWh)	NUC
Água Azul do Norte	1085,63	8004,06	16370	1033,91	2365
Canaã dos Carajás	414,66	1483,68	17915	1202,11	5182

O conjunto Canaã dos Carajás apresenta aproximadamente o dobro de unidades consumidoras do que o conjunto Água Azul do Norte, além de apresentar o maior consumo médio mensal. Na classificação feita pelo software ANABENCH o conjunto Água Azul do Norte foi alocado no *cluster* 3 cujas metas finais são, para conjuntos do tipo interligado, 40 horas/ano e 23 interrupções/ano para os indicadores DEC e FEC, respectivamente. O conjunto Canaã dos Carajás foi alocado no *cluster* 30 cujas metas finais são, para conjuntos do tipo interligado, 15 horas/ano e 8 interrupções/ano para os indicadores DEC e FEC, respectivamente.

Ao se apresentar os novos cinco atributos resultantes da união destes dois conjuntos, o classificador alocou o cenário 4 no *cluster* 11 cujas metas finais são 33 horas/ano e 19 interrupções/ano para os indicadores DEC e FEC, respectivamente. A partir destas informações mais a informação do valor dos indicadores no ano zero utilizam-se as equações (2.6) a (2.9) para determinação das metas no período compreendido entre duas revisões tarifárias (8 anos). A Tabela 5.22 resume as principais informações necessárias à definição das metas.

Tabela 5.22: Ano zero e valor da meta final dos indicadores DEC e FEC para o cenário 4

Conjunto	DEC		FEC	
	Ano 0	Meta	Ano 0	Meta
Cenário 4	106,30	33	67,78	19
Água Azul do Norte	153,49	40	102,89	23
Canaã dos Carajás	88,28	15	54,37	8

A Tabela 5.23 e a Figura 5.16 mostram os valores das metas e o gráfico de decaimento das metas ao longo de duas revisões tarifárias para o cenário 4 e para os conjuntos formadores deste cenário para o indicador de continuidade DEC.

Tabela 5.23: Metas para o cenário 4 ao longo dos 8 anos para o indicador DEC

Conjunto	Metas Anuais							
	2009	2010	2011	2012	2013	2014	2015	2016
Cenário 4	92	80	69	60	52	45	39	33
Água Azul do Norte	130	110	93	79	67	56	48	40
Canaã dos Carajás	71	57	46	37	30	24	19	15

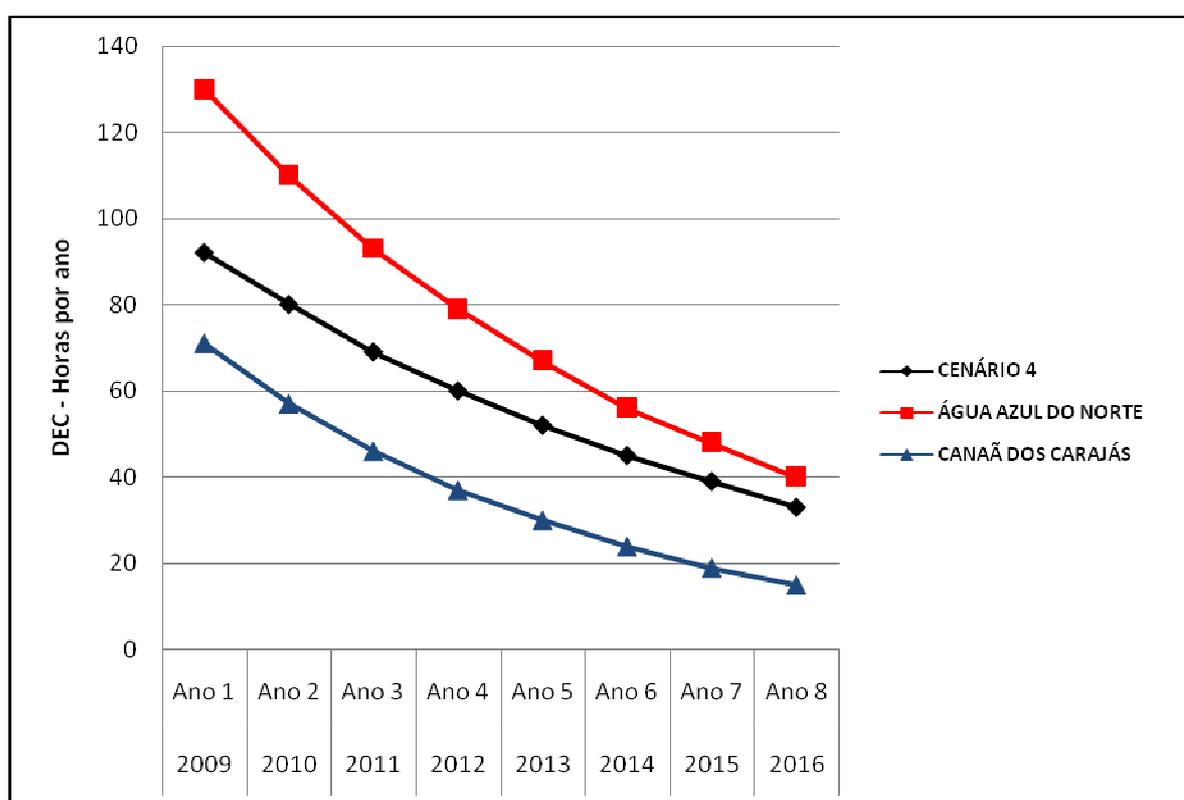


Figura 5.16: Metas do cenário 4 ao longo de 8 anos para o indicado DEC

Os gráficos de metas do indicador DEC mostram que o conjunto Canaã dos Carajás possui as metas mais rígidas se comparadas ao conjunto Água Azul do Norte. O conjunto Água Azul do Norte apresenta elevada taxa de redução anual.

A união dos conjuntos proporcionou folgas nas metas do conjunto Canaã dos Carajás que é o conjunto que apresenta maior número de unidades consumidoras e maior consumo médio mensal. As folgas observadas foram de 21 horas no primeiro ano, 23 horas no segundo, terceiro e quarto anos, 22 horas no quinto ano, 21 horas no sexto ano, 20 horas no sétimo ano

e 18 horas no oitavo ano. Notou-se diminuição da taxa de redução anual para o conjunto Água Azul do Norte. A Figura 5.17 ilustra a diminuição da taxa de redução anual ao se efetuar a união dos conjuntos para o indicador DEC.

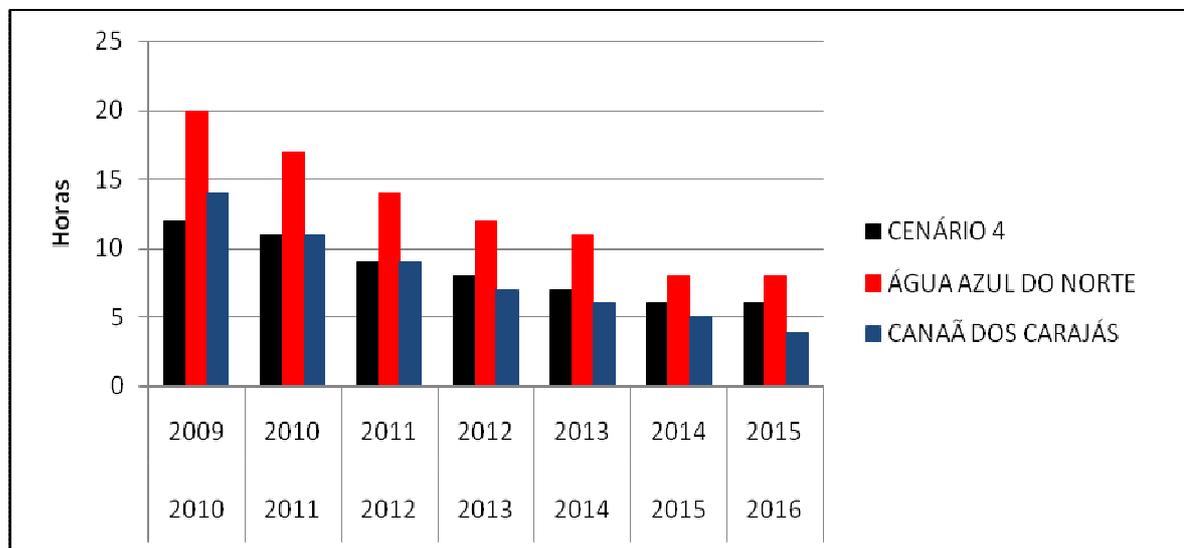


Figura 5.17: Taxa de redução anual para o cenário 4 para o indicador DEC

A Tabela 5.24 apresenta os resultados das metas obtidas pela aplicação da Metodologia Proposta e do DEC agregado dos conjuntos formadores do cenário 4.

Tabela 5.24: Comparação do DEC agregado com o obtido pela Metodologia Proposta para o cenário 4

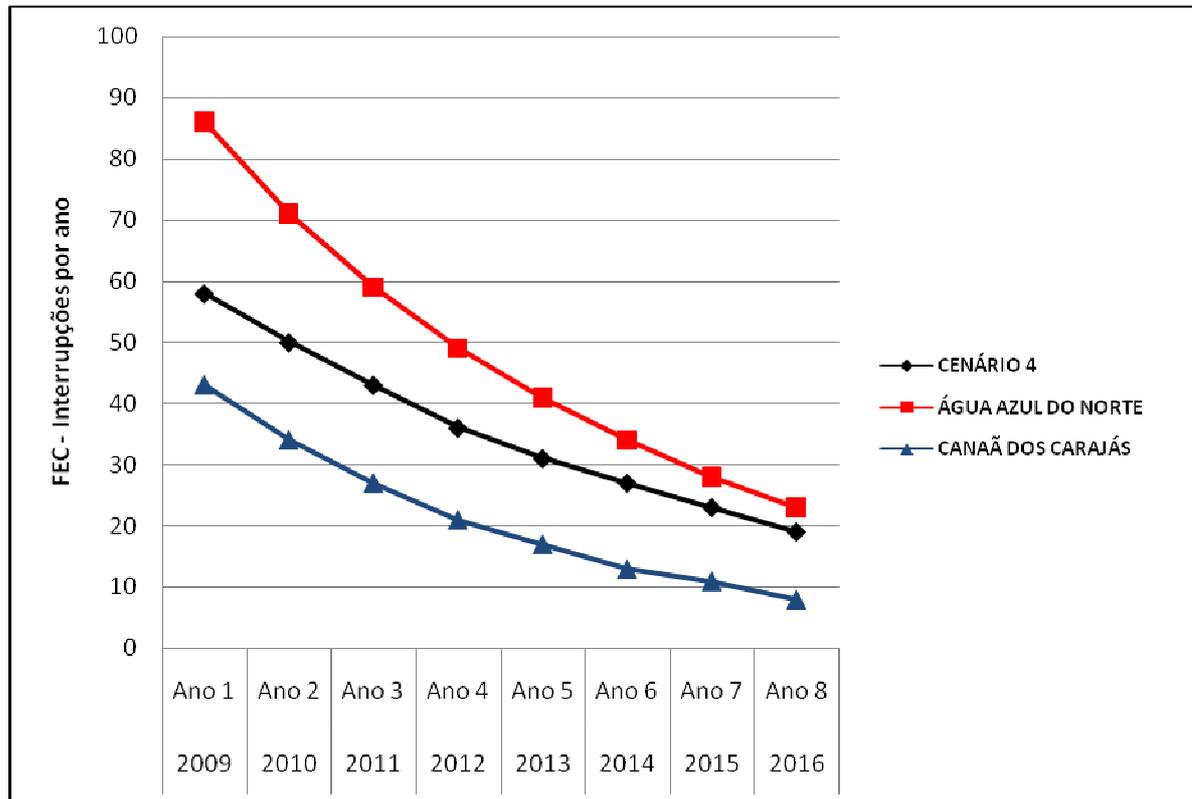
DEC	2009	2016
Agregado	89,49	22,83
Metodologia proposta	92	33

A comparação dos valores obtidos pela aplicação da Metodologia com os valores do DEC agregado dos conjuntos Água Azul do Norte e Canaã dos Carajás mostra que esta é uma união favorável para o indicador DEC, pois os valores de metas obtidos pela Metodologia são superiores aos do DEC agregado para os anos de 2009 e 2016.

A Tabela 5.25 e a Figura 5.18 mostram os valores das metas e o gráfico de decaimento das metas ao longo de 8 anos para o cenário 4 e para os conjuntos formadores deste cenário para o indicador de continuidade FEC.

Tabela 5.25: Metas para o cenário 4 ao longo dos 8 anos para o indicador FEC

Conjunto	Metas Anuais							
	2009	2010	2011	2012	2013	2014	2015	2016
Cenário 4	58	50	43	36	31	27	23	19
Água Azul do Norte	86	71	59	49	41	34	28	23
Canaã dos Carajás	43	34	27	21	17	13	11	8

**Figura 5.18: Metas do cenário 4 ao longo de 8 anos para o indicado FEC**

Os gráficos de metas do indicador FEC mostram que o conjunto Canaã dos Carajás possui as metas mais rígidas se comparadas ao conjunto Água Azul do Norte. O conjunto Água Azul do Norte apresenta elevada taxa de redução anual. A Figura 5.19 mostra o gráfico de diminuição da taxa de redução anual ao se efetuar a união dos conjuntos para o indicador FEC.

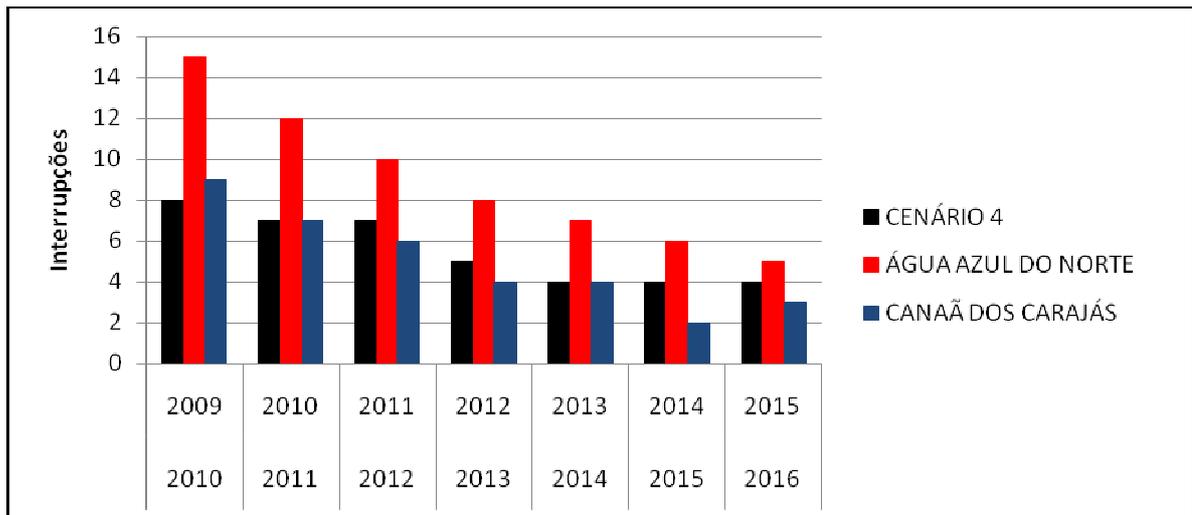


Figura 5.19: Taxa de redução anual para o cenário 4 para o indicador FEC

A união dos conjuntos proporcionou folgas nas metas do conjunto Canaã dos Carajás que é o conjunto que apresenta maior número de unidades consumidoras e maior consumo médio mensal. As folgas nas metas de FEC no conjunto Canaã dos Carajás são de 15 interrupções para o primeiro ano, 16 interrupções para o segundo e terceiro anos, 15 interrupções no quarto ano, 14 interrupções no quinto e sexto anos, 12 interrupções para o sétimo ano e 11 no último ano. Observou-se diminuição da taxa de redução anual para o conjunto Água Azul do Norte.

A Tabela 5.26 apresenta os resultados das metas obtidas pela aplicação da Metodologia Proposta e do DEC agregado dos conjuntos formadores do cenário 4.

Tabela 5.26: Comparação do FEC agregado com o obtido pela Metodologia Proposta para o cenário 4

FEC	2009	2016
Agregado	56,47	12,70
Metodologia proposta	58	19

A comparação dos valores obtidos pela aplicação da Metodologia com os valores do FEC agregado dos conjuntos Água Azul do Norte e Canaã dos Carajás mostra que esta é uma união favorável para o indicador FEC, pois os valores de metas obtidos pela Metodologia são superiores aos do FEC agregado para os anos de 2009 e 2016.

5.4.5 Cenário 5: Santa Maria não urbano e São Miguel do Guamá

O quinto cenário a ser analisado é a união dos conjuntos Santa Maria não urbano e São Miguel do Guamá que são dois conjuntos do tipo interligado e possuem as seguintes características resumidas na Tabela 5.27.

Tabela 5.27: Atributos dos conjuntos Santa Maria não urbano e São Miguel do Guamá

CONJUNTO	ERAP (km)	AREA (km ²)	PNI (kVA)	CMM (MWh)	NUC
Santa Maria não urbano	484,45	653	9397,50	544,01	3734
São Miguel do Guamá	102,93	265	17952,5	2282,49	8388

O conjunto São Miguel do Guamá apresenta cerca de o dobro de unidades consumidoras do que o conjunto Santa Maria não urbano, além de possuir maior potência instalada e consumo médio mensal. Na classificação feita pelo software ANABENCH o conjunto Santa Maria não urbano foi alocado no *cluster* 4 cujas metas finais são, para conjuntos do tipo interligado, 14 horas/ano e 10 interrupções/ano para os indicadores DEC e FEC, respectivamente. O conjunto São Miguel do Guamá foi alocado no *cluster* 20 cujas metas finais são, para conjuntos do tipo interligado, 10 horas/ano e 8 interrupções/ano para os indicadores DEC e FEC, respectivamente.

Ao se apresentar os novos cinco atributos resultantes da união destes dois conjuntos, o classificador alocou o cenário 5 no *cluster* 17 cujas metas finais são 12 horas/ano e 9 interrupções/ano para os indicadores DEC e FEC, respectivamente. A partir destas informações mais a informação do valor dos indicadores no ano zero utilizam-se as equações (2.6) a (2.9) para a determinação das metas no período compreendido entre duas revisões tarifárias (8 anos). A Tabela 5.28 resume as principais informações necessárias à definição das metas anuais.

Tabela 5.28: Ano zero e valor da meta final dos indicadores DEC e FEC para o cenário 5

Conjunto	DEC		FEC	
	Ano 0	Meta	Ano 0	Meta
Cenário 5	35,65	12	21,21	9
Santa Maria não urbano	50,9	14	25,11	10
São Miguel do Guamá	29,73	10	19,7	8

A Tabela 5.29 e a Figura 5.20 mostram os valores das metas e o gráfico de decaimento das metas ao longo de duas revisões tarifárias para o cenário 5 e para os conjuntos formadores deste cenário para o indicador de continuidade DEC.

Tabela 5.29: Metas para o cenário 5 ao longo dos 8 anos para o indicador DEC

Conjunto	Metas Anuais							
	2009	2010	2011	2012	2013	2014	2015	2016
Cenário 5	32	28	24	21	19	16	14	12
Santa Maria não urbano	44	37	32	27	23	20	17	14
São Miguel do Guamá	26	23	20	18	16	14	12	10

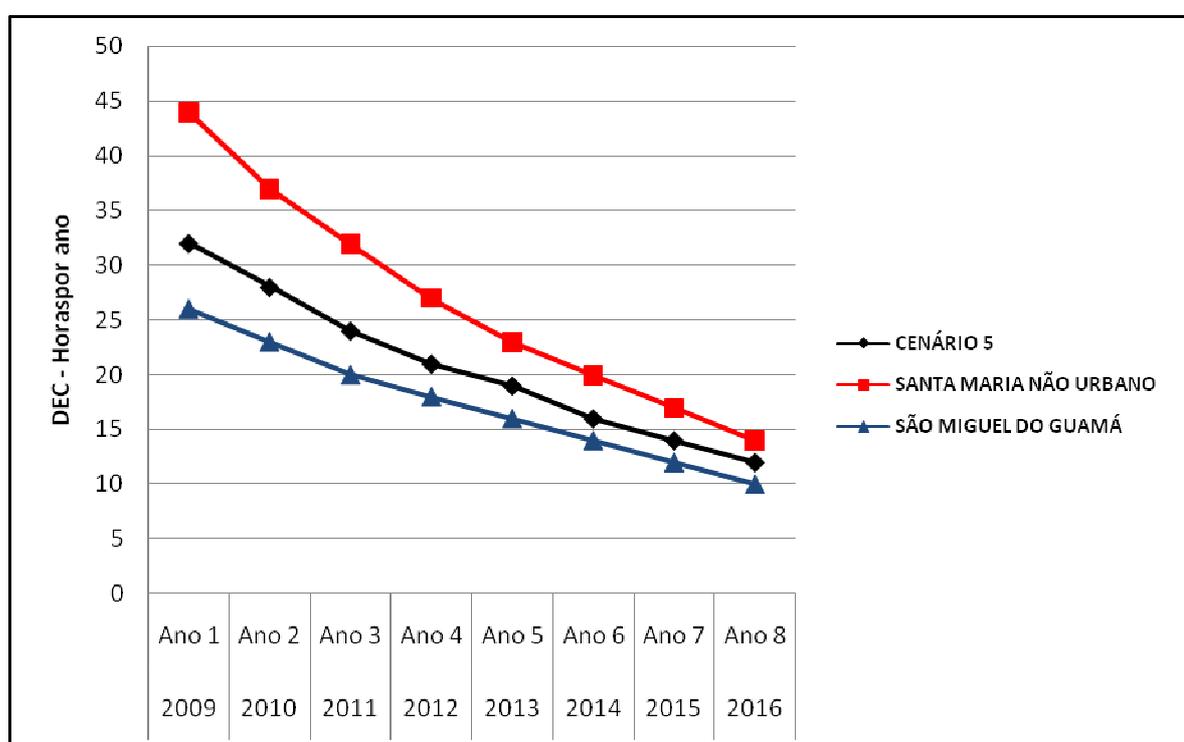


Figura 5.20: Metas do cenário 5 ao longo de 8 anos para o indicado DEC

Os gráficos de metas do indicador DEC mostram que o conjunto São Miguel do Guamá possui as metas mais rígidas se comparadas ao conjunto Santa Maria não urbano. O conjunto Santa Maria não urbano apresenta elevada taxa de redução anual se comparada ao conjunto São Miguel do Guamá.

A união dos conjuntos proporcionou folgas nas metas do conjunto São Miguel do Guamá que é o conjunto que apresenta maior número de unidades consumidoras, maior consumo médio mensal e potência instalada e uma diminuição da taxa de redução anual para o conjunto Santa Maria não urbano. As folgas observadas no conjunto São Miguel do Guamá

foram de 6 horas no primeiro ano, 5 horas no segundo ano, 4 horas no terceiro ano, 3 horas no quarto e quinto anos e 2 horas nos três últimos anos.

A Figura 5.21 ilustra a diminuição da taxa de redução anual ao se efetuar a união dos conjuntos para o indicador DEC.

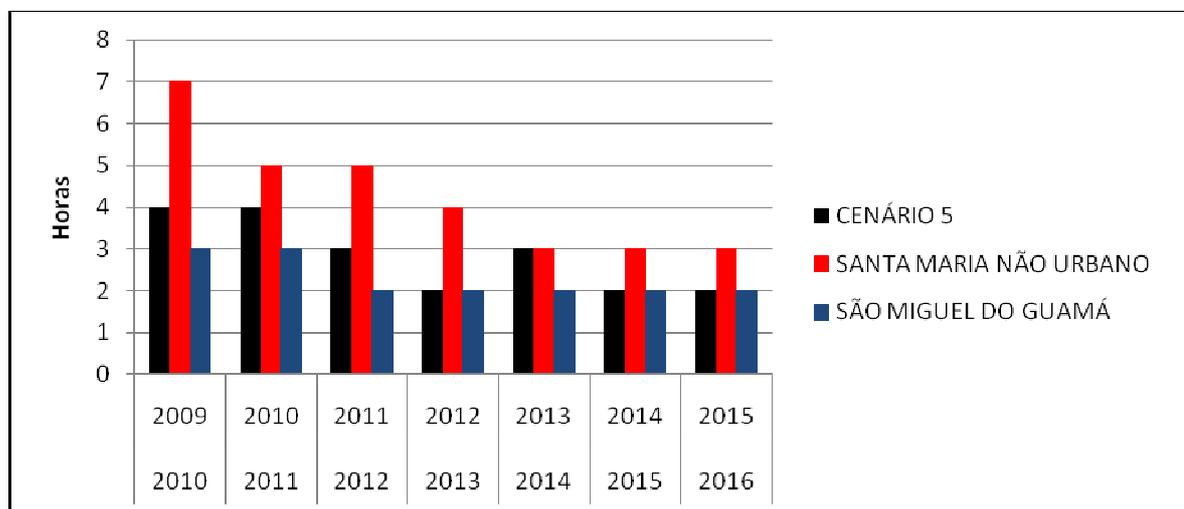


Figura 5.21: Taxa de redução anual para o cenário 5 para o indicador DEC

A Tabela 5.30 apresenta uma comparação entre as metas obtidas por meio da aplicação da Metodologia Proposta para o indicador DEC e o valor do DEC agregado para o primeiro e oitavo anos.

Tabela 5.30: Comparação do DEC agregado com o obtido pela Metodologia Proposta para o cenário 5

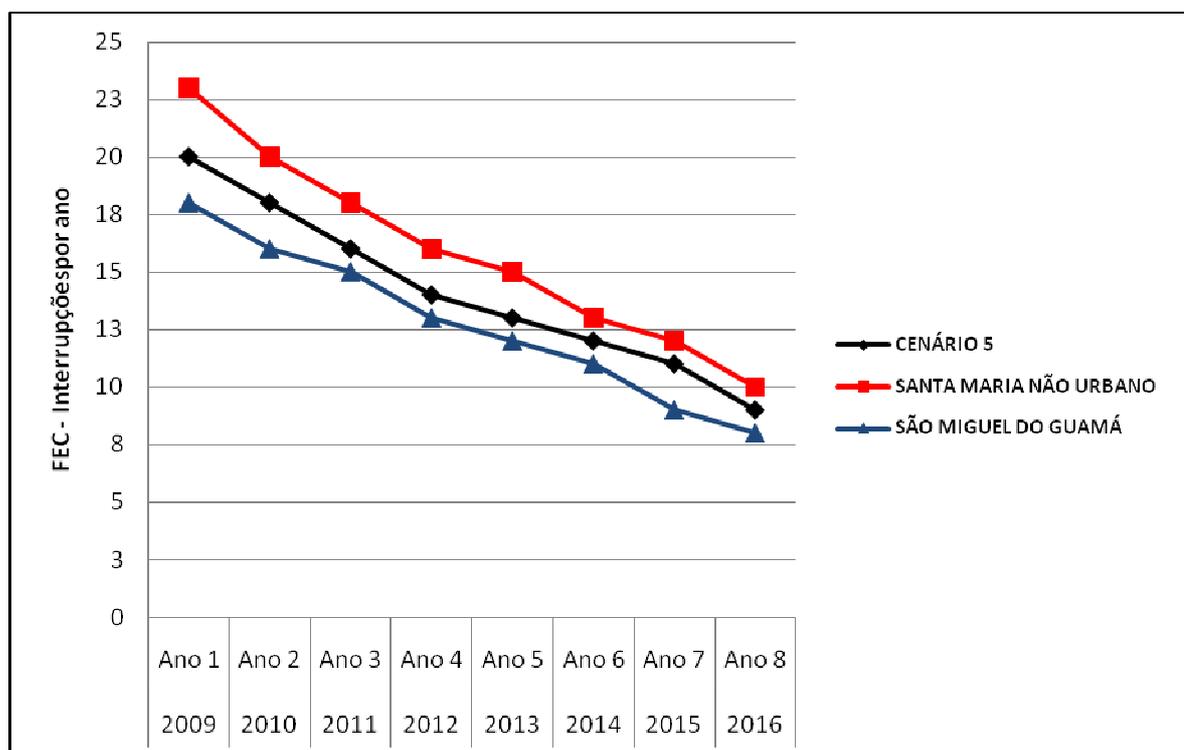
DEC	2009	2016
Agregado	31,54	11,23
Metodologia proposta	32	12

A Tabela 5.30 mostra que apesar da união ter proporcionado folgas nas metas para o conjunto São Miguel do Guamá que é o conjunto com maior número de unidades consumidoras a união não é satisfatória para o indicador DEC, pois os valores de metas obtidos pela Metodologia estão próximos aos obtidos pelo DEC agregado.

A Tabela 5.31 e a Figura 5.22 mostram os valores das metas e o gráfico de decaimento das metas ao longo de oito anos para o cenário 5 e para os conjuntos formadores deste cenário para o indicador de continuidade FEC.

Tabela 5.31: Metas para o cenário 5 ao longo dos 8 anos para o indicador FEC

Conjunto	Metas Anuais							
	2009	2010	2011	2012	2013	2014	2015	2016
Cenário 5	22	22	22	22	22	22	22	22
Santa Maria não urbano	23	20	18	16	15	13	12	10
São Miguel do Guamá	18	16	15	13	12	11	9	8

**Figura 5.22: Metas do cenário 5 ao longo de 8 anos para o indicado FEC**

Os gráficos de metas do indicador FEC mostram que o conjunto São Miguel do Guamá possui as metas mais rígidas se comparadas ao conjunto Santa Maria não urbano. O conjunto Santa Maria não urbano apresenta taxa de redução anual similar ao do conjunto São Miguel do Guamá.

A união dos conjuntos proporcionou folgas pouco significativas nas metas do conjunto São Miguel do Guamá que é o conjunto que apresenta maior número de unidades consumidoras, maior consumo médio mensal e potência instalada. As folgas observadas no conjunto São Miguel do Guamá foram de 2 interrupções no primeiro e segundo anos, 1 interrupção do terceiro ao sexto ano, 2 interrupções no sétimo ano e 1 interrupção no último ano. A Figura 5.21 mostra a taxa de redução anual do cenário 5 e dos conjuntos formadores deste cenário para o indicador FEC.

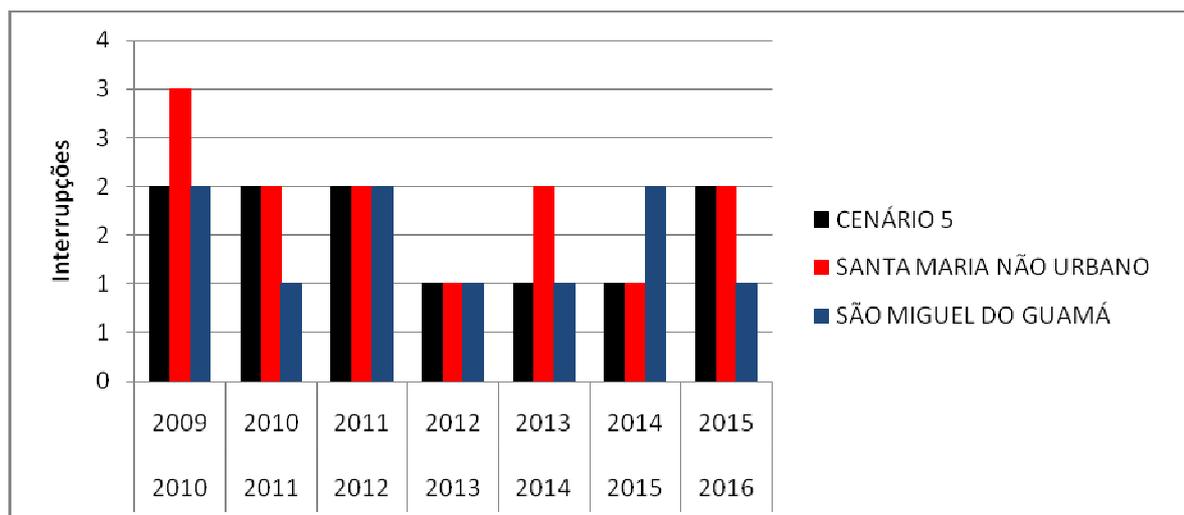


Figura 5.23: Taxa de redução anual para o cenário 5 para o indicador FEC

A Figura 5.23 mostra que não houve ganhos significativos na redução da taxa de redução anual que apresentou um comportamento bastante similar para o cenário 5 e para o conjuntos formadores deste cenário.

A comparação das metas obtidas pela Metodologia Proposta e o FEC agregado é apresentada na Tabela 5.32.

Tabela 5.32: Comparação do FEC agregado com o obtido pela Metodologia Proposta para o cenário 5

FEC	2009	2016
Agregado	19,54	8,62
Metodologia proposta	20	9

A Tabela 5.8 que a união não é satisfatória devido aos valores de FEC obtidos por meio da Metodologia serem similares aos obtidos por meio do FEC agregado. Portanto, a união dos conjuntos Santa Maria não urbano com o conjunto São Miguel do Guamá não é favorável para o indicador FEC, pois apresentou poucas folgas nas metas do cenário 5 se comparadas com as metas do conjunto com maior número de unidades consumidoras, não diminuição da taxa de redução anual e valores de metas muito similares à obtidas por meio do FEC agregado.

5.4.6 Cenário 6: Curionópolis e Eldorado dos Carajás

O sexto cenário a ser analisado é a união dos conjuntos Curionópolis e Eldorado dos Carajás que são dois conjuntos do tipo interligado e possuem as seguintes características resumidas na Tabela 5.33.

Tabela 5.33: Atributos dos conjuntos Curionópolis e Eldorado dos Carajás

CONJUNTO	ERAP (km)	AREA (km ²)	PNI (kVA)	CMM (MWh)	NUC
Curionópolis	532,02	2420,45	7467,50	441,95	3696
Eldorado dos Carajás	440,57	1339,28	15032,50	1191,88	5320

O conjunto Eldorado dos Carajás apresenta maior número de unidades consumidoras, consumo médio mensal e potência instalada do que o conjunto Curionópolis. Na classificação feita pelo software ANABENCH o conjunto Curionópolis foi alocado no *cluster 27* cujas metas finais são, para conjuntos do tipo interligado, 22 horas/ano e 11 interrupções/ano para os indicadores DEC e FEC, respectivamente. O conjunto Eldorado dos Carajás foi alocado no *cluster 30* cujas metas finais são, para conjuntos do tipo interligado, 15 horas/ano e 8 interrupções/ano para os indicadores DEC e FEC, respectivamente.

Ao se apresentar os novos cinco atributos resultantes da união destes dois conjuntos, o classificador alocou o cenário 6 no *cluster 11* cujas metas finais são 33 horas/ano e 19 interrupções/ano para os indicadores DEC e FEC, respectivamente. A partir destas informações mais a informação do valor dos indicadores no ano zero utilizam-se as equações (2.6) a (2.9) para a determinação das metas no período compreendido entre duas revisões tarifárias (8 anos). A Tabela 5.34 resume as principais informações necessárias à definição das metas anuais.

Tabela 5.34: Ano zero e valor da meta final dos indicadores DEC e FEC para o cenário 6

Conjunto	DEC		FEC	
	Ano 0	Meta	Ano 0	Meta
Cenário 6	54,00	33	52,13	19
Curionópolis	57,48	22	51,64	11
Eldorado dos Carajás	50,39	15	52,17	8

A Tabela 5.35 e a Figura 5.24 mostram os valores das metas e o gráfico de decaimento das metas ao longo de duas revisões tarifárias para o cenário 6 e para os conjuntos formadores deste cenário para o indicador de continuidade DEC.

Tabela 5.35: Metas para o cenário 6 ao longo dos 8 anos para o indicador DEC

Conjunto	Metas Anuais							
	2009	2010	2011	2012	2013	2014	2015	2016
Cenário 6	51	48	45	43	40	38	36	33
Curionópolis	51	46	41	36	32	28	25	22
Eldorado dos Carajás	44	38	32	28	24	21	18	15

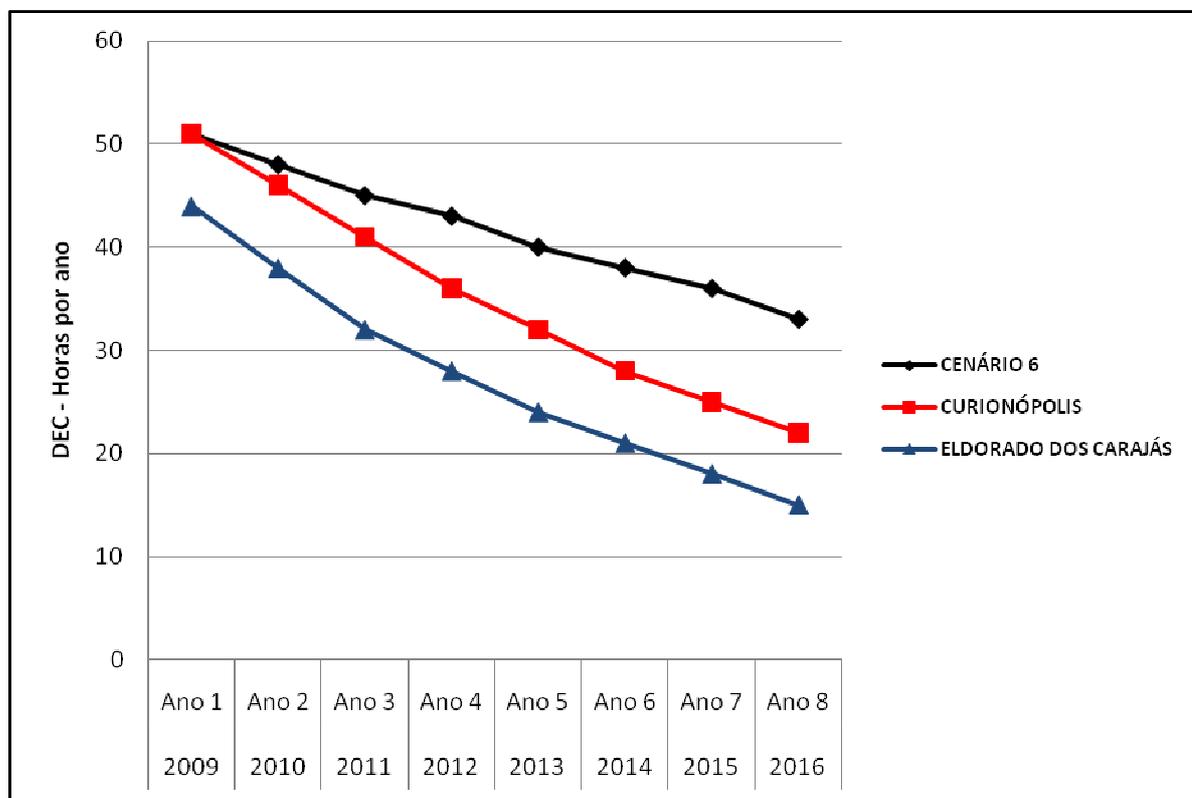


Figura 5.24: Metas do cenário 6 ao longo de 8 anos para o indicado DEC

Os gráficos de metas do indicador DEC mostram que o conjunto Eldorado dos Carajás possui as metas mais rígidas se comparadas ao conjunto Curionópolis. A taxa de redução anual é aproximadamente igual nos dois conjuntos. A Figura 5.25 ilustra as taxas de redução anual do cenário 6 e dos conjuntos formadores deste cenário para o indicador DEC.

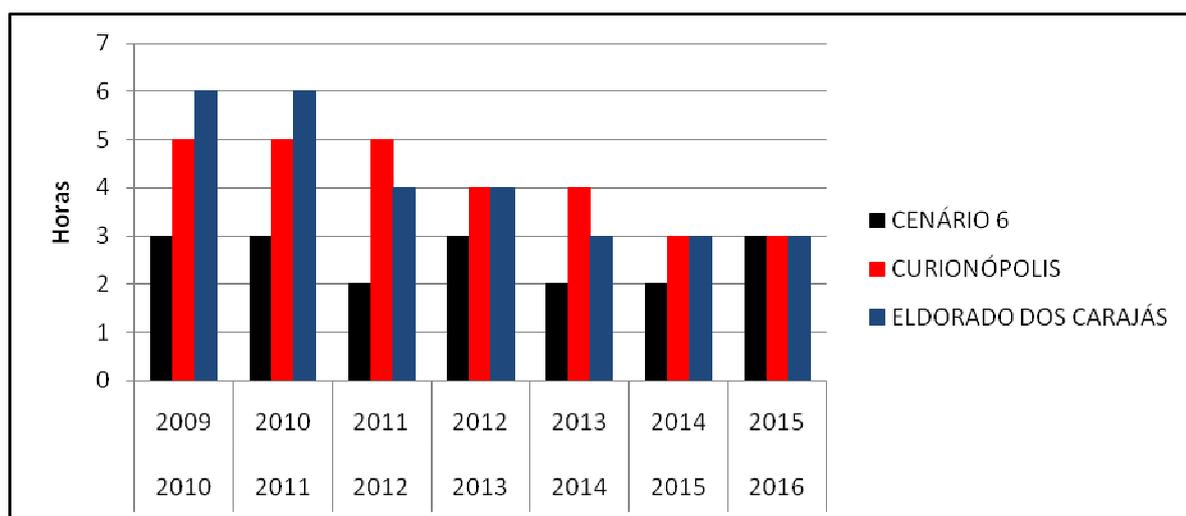


Figura 5.25: Taxa de redução anual para o cenário 6 para o indicador DEC

A união dos conjuntos proporcionou folgas nas metas do indicador DEC tanto para o conjunto Eldorado dos Carajás que é o conjunto com o maior número de unidades consumidoras, maior consumo médio mensal e potência instalada como para o conjunto Curionópolis. As folgas das metas do indicador DEC do cenário 6 em relação ao conjunto Eldorado dos Carajás são de 7 horas no primeiro ano, 10 horas no segundo ano, 13 horas no terceiro ano, 15 horas no quarto ano, 16 horas no quinto ano, 17 horas no sexto ano e 18 horas para os dois últimos anos. As folgas das metas do indicador DEC do cenário 6 em relação ao conjunto Curionópolis são de 2 horas no segundo ano, 4 horas no terceiro ano, 7 horas no quarto ano, 8 horas no quinto ano, 10 horas no sexto ano e 11 horas para os demais anos. A taxa de redução do cenário 6 é menor durante os sete primeiros anos se comparada com a taxa de redução dos conjuntos formadores deste cenário.

A Tabela 5.36 apresenta os resultados das metas obtidas pela aplicação da Metodologia Proposta e do DEC agregado dos conjuntos formadores do cenário 6.

Tabela 5.36: Comparação do DEC agregado com o obtido pela Metodologia Proposta para o cenário 6

DEC	2009	2016
Agregado	46,87	17,87
Metodologia proposta	51	33

A comparação dos valores obtidos pela aplicação da Metodologia com os valores do DEC agregado dos conjuntos Curionópolis e Eldorado dos Carajás mostra que esta é uma

união favorável para o indicador DEC, pois os valores de metas obtidos pela Metodologia são superiores aos do DEC agregado para os anos de 2009 e 2016.

A Tabela 5.37 e a Figura 5.26 mostram os valores das metas e o gráfico de decaimento das metas ao longo de oito anos para o cenário 6 e para os conjuntos formadores deste cenário para o indicador de continuidade FEC.

Tabela 5.37: Metas para o cenário 6 ao longo dos 8 anos para o indicador FEC

Conjunto	Metas Anuais							
	2009	2010	2011	2012	2013	2014	2015	2016
Cenário 6	46	41	36	32	28	25	22	19
Curionópolis	43	36	29	24	20	17	14	11
Eldorado dos Carajás	42	33	26	21	17	13	11	8

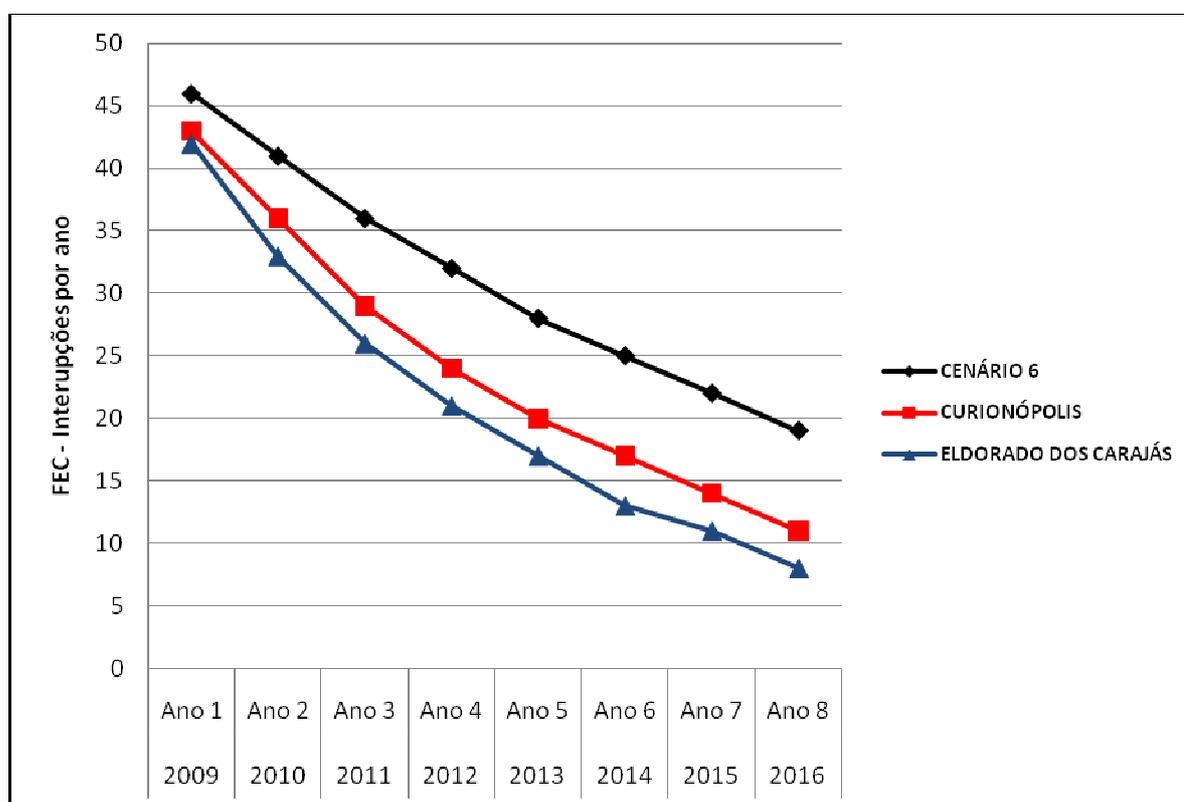


Figura 5.26: Metas do cenário 6 ao longo de 8 anos para o indicado FEC

Os gráficos de metas do indicador FEC mostram que o conjunto Eldorado dos Carajás as metas mais rígidas se comparadas ao conjunto Curionópolis. A taxa de redução anual é aproximadamente igual nos dois conjunto com máxima diferença de 2 interrupções. A Figura 5.27 ilustra as taxas de redução anual do cenário 6 e dos conjuntos formadores deste cenário para o indicador FEC.

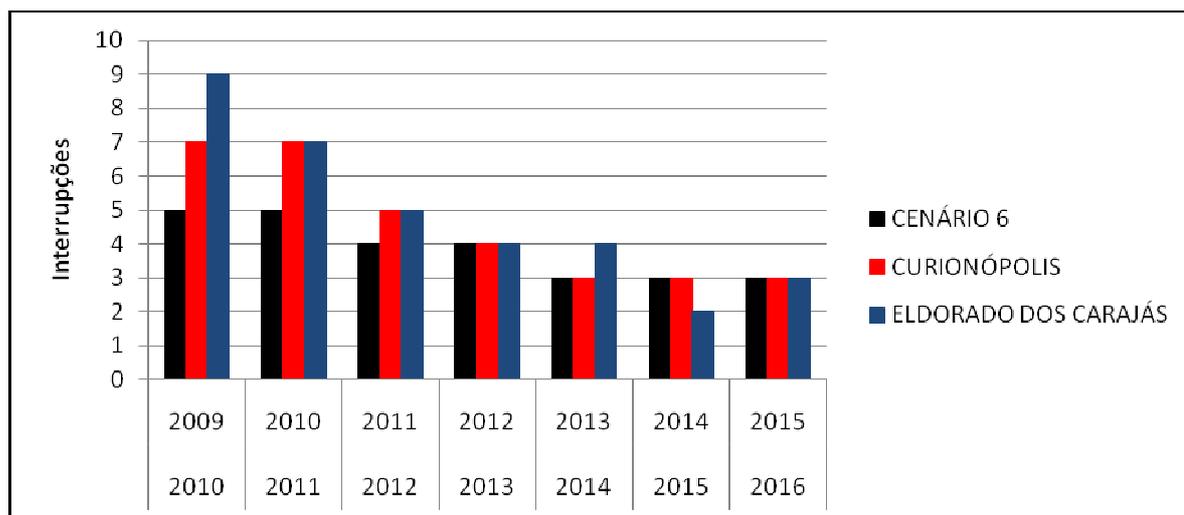


Figura 5.27: Taxa de redução anual para o cenário 6 para o indicador FEC

A união dos conjuntos proporcionou folgas nas metas do indicador FEC tanto para o conjunto Eldorado dos Carajás que é o conjunto com o maior número de unidades consumidoras, maior consumo médio mensal e potência instalada como para o conjunto Curionópolis. As folgas das metas do indicador FEC do cenário 6 em relação ao conjunto Eldorado dos Carajás são de 4 interrupções no primeiro ano, 8 interrupções no segundo ano, 10 interrupções no terceiro ano, 11 interrupções no quarto e quinto anos, 12 interrupções no sexto ano e 11 interrupções para os dois últimos anos. As folgas das metas do indicador FEC do cenário 6 em relação ao conjunto Curionópolis são de 3 interrupções no primeiro ano, 5 interrupções no segundo ano, 7 interrupções no terceiro ano e 8 interrupções para os demais anos. A taxa de redução do cenário 6 é menor durante os dois primeiros anos se comparada com a taxa de redução dos conjuntos formadores deste cenário. A Tabela 5.38 apresenta os resultados das metas obtidas pela aplicação da Metodologia e do FEC agregado dos conjuntos formadores do cenário 6.

Tabela 5.38: Comparação do FEC agregado com o obtido pela Metodologia Proposta para o cenário 6

FEC	2009	2016
Agregado	42,41	9,23
Metodologia proposta	46	19

A comparação dos valores obtidos pela aplicação da Metodologia com os valores do DEC agregado dos conjuntos Curionópolis e Eldorado dos Carajás mostra que esta é uma união favorável para o indicador FEC, pois os valores de metas obtidos pela Metodologia são superiores aos do FEC agregado para os anos de 2009 e 2016.

5.4.7 Cenário 7: Abaetetuba e Igarapé Mirim

O último cenário a ser analisado é a união dos conjuntos Abaetetuba e Igarapé Mirim que são dois conjuntos do tipo interligado e possuem as seguintes características resumidas na Tabela 5.39.

Tabela 5.39: Atributos dos conjuntos Abaetetuba e Igarapé Mirim

CONJUNTO	ERAP (km)	AREA (km ²)	PNI (kVA)	CMM (MWh)	NUC
Abaetetuba	351,03	1171,54	21625	3135,59	17136
Igarapé Mirim	148,24	1728,33	7865	581,09	4122

O conjunto Abaetetuba apresenta aproximadamente 4 vezes mais número de unidades consumidoras do que o conjunto Igarapé Mirim, além de possuir também maior consumo médio mensal e potência instalada. Na classificação feita pelo software ANABENCH o conjunto Abaetetuba foi alocado no *cluster* 29 cujas metas finais são, para conjuntos do tipo interligado, 10 horas/ano e 8 interrupções/ano para os indicadores DEC e FEC, respectivamente. O conjunto Igarapé Mirim foi alocado no *cluster* 15 cujas metas finais são, para conjuntos do tipo interligado, 18 horas/ano e 10 interrupções/ano para os indicadores DEC e FEC, respectivamente.

Ao se apresentar os novos cinco atributos resultantes da união destes dois conjuntos, o classificador alocou o cenário 7 no *cluster* 17 cujas metas finais são 12 horas/ano e 9 interrupções/ano para os indicadores DEC e FEC, respectivamente. A partir destas informações mais a informação do valor dos indicadores no ano zero utilizam-se as equações (2.6) a (2.9) para a determinação das metas no período compreendido entre duas revisões tarifárias (8 anos). A Tabela 5.40 resume as principais informações necessárias à definição das metas.

Tabela 5.40: Ano zero e valor da meta final dos indicadores DEC e FEC para o cenário 7

Conjunto	DEC		FEC	
	Ano 0	Meta	Ano 0	Meta
Cenário 7	32,07	12	26,34	9
Abaetetuba	24,07	10	22,91	8
Igarapé Mirim	61,89	18	38,86	10

A Tabela 5.41 e a Figura 5.28 mostram os valores das metas e o gráfico de decaimento das metas ao longo de duas revisões tarifárias para o cenário 7 e para os conjuntos formadores deste cenário para o indicador de continuidade DEC.

Tabela 5.41: Metas para o cenário 7 ao longo dos 8 anos para o indicador DEC

Conjunto	Metas Anuais							
	2009	2010	2011	2012	2013	2014	2015	2016
Cenário 7	29	26	23	20	18	16	14	12
Abaetetuba	22	20	18	16	14	13	12	10
Igarapé Mirim	54	46	39	34	29	25	22	18

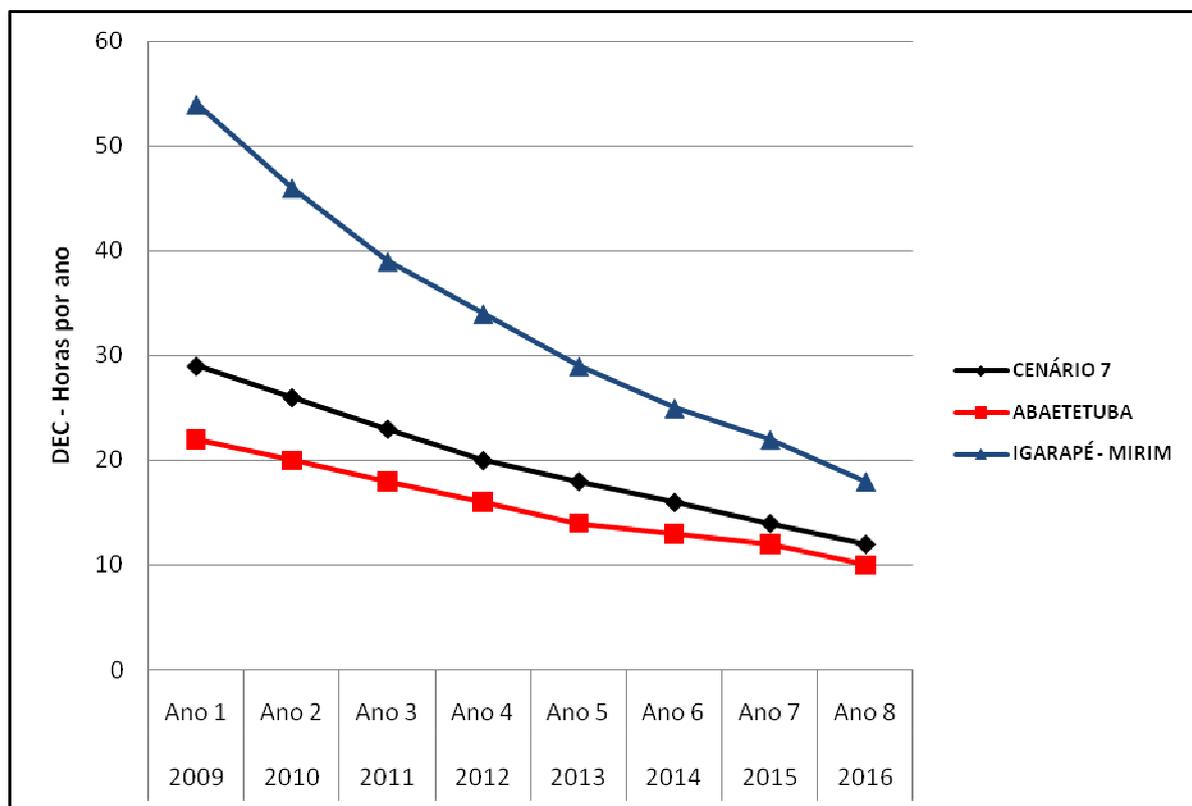


Figura 5.28: Metas do cenário 7 ao longo de 8 anos para o indicado DEC

Os gráficos de metas do indicador DEC mostram que o conjunto Abaetetuba possui as metas mais rígidas se comparadas ao conjunto Igarapé Mirim. O conjunto Igarapé Mirim apresenta elevada taxa de redução anual.

A união dos conjuntos proporcionou folgas nas metas do conjunto Abaetetuba que é o conjunto que apresenta maior número de unidades consumidoras, maior consumo médio mensal e potência instalada. As folgas nas metas foram de 7 horas no primeiro ano, 6 horas no segundo ano, 5 horas no terceiro ano, 4 horas no quarto e quinto anos, 3 horas no sexto ano e

2 horas nos dois últimos anos. A Figura 5.29 ilustra a diminuição da taxa de redução anual para o conjunto Igarapé Mirim ao se efetuar a união dos conjuntos para o indicador DEC.

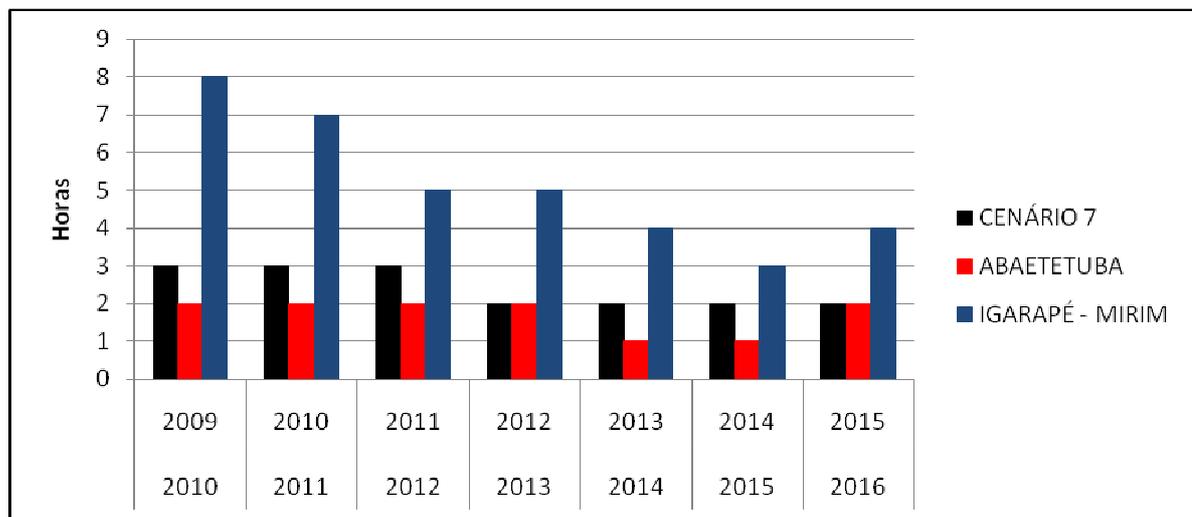


Figura 5.29: Taxa de redução anual para o cenário 7 para o indicador DEC

A Tabela 5.42 apresenta os resultados das metas obtidas pela aplicação da Metodologia Proposta e do FEC agregado dos conjuntos formadores do cenário 7.

Tabela 5.42: Comparação do DEC agregado com o obtido pela Metodologia Proposta para o cenário 7

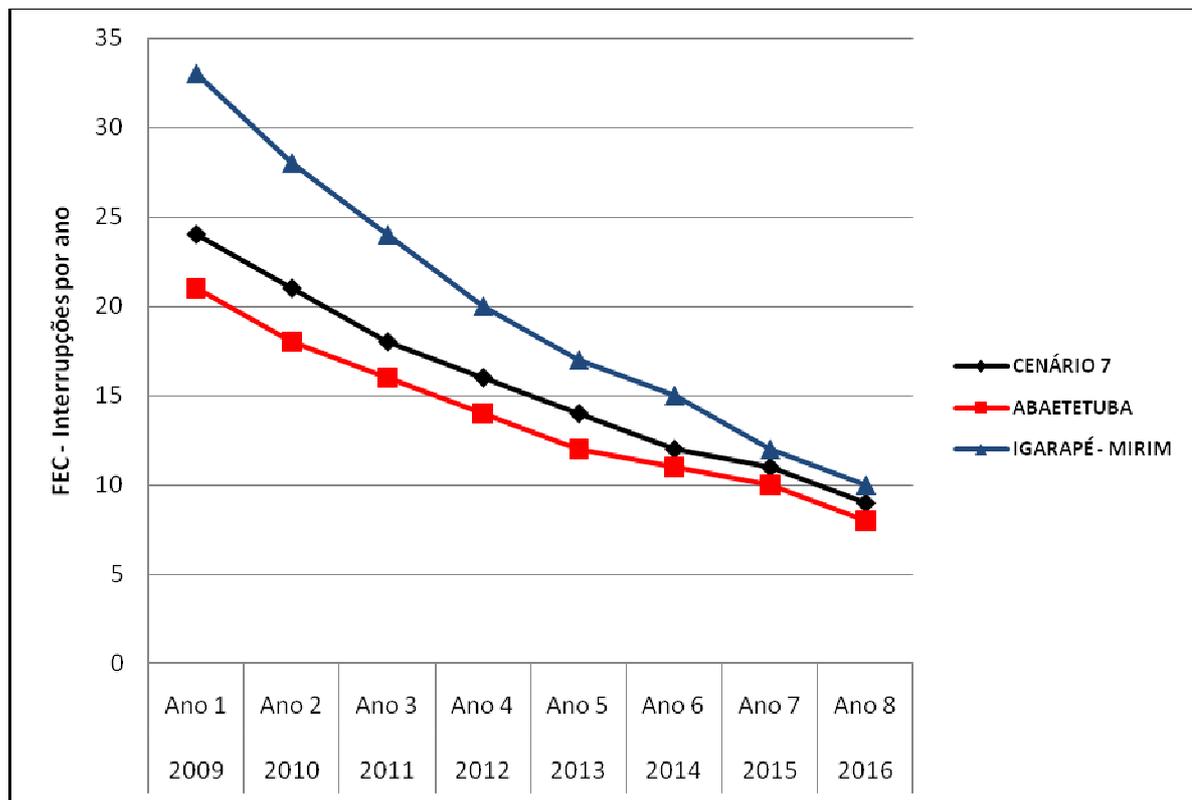
DEC	2009	2016
Agregado	28,20	11,55
Metodologia proposta	29	12

A comparação dos valores obtidos pela aplicação da Metodologia com os valores do DEC agregado dos conjuntos envolvidos na união dos conjuntos formadores do cenário 7 mostra que esta não é uma união favorável para o indicador DEC, pois os valores de metas obtidos pela Metodologia são semelhantes aos valores obtidos pelo FEC agregado para os anos de 2009 e 2016.

A Tabela 5.43 e a Figura 5.30 mostram os valores das metas e o gráfico de decaimento das metas ao longo de oito anos para o cenário 7 e para os conjuntos formadores deste cenário para o indicador de continuidade FEC.

Tabela 5.43: Metas para o cenário 7 ao longo dos 8 anos para o indicador FEC

Conjunto	Metas Anuais							
	2009	2010	2011	2012	2013	2014	2015	2016
Cenário 7	24	21	18	16	14	12	11	9
Abaetetuba	21	18	16	14	12	11	10	8
Igarapé Mirim	33	28	24	20	17	15	12	10

**Figura 5.30: Metas do cenário 7 ao longo de 8 anos para o indicado FEC**

Os gráficos de metas do indicador FEC mostram que o conjunto Abaetetuba possui as metas mais rígidas se comparadas ao conjunto Igarapé Mirim. O conjunto Igarapé Mirim apresenta elevada taxa de redução anual. A Figura 5.31 ilustra as taxas de redução anual do cenário 6 e dos conjuntos formadores deste cenário para o indicador FEC.

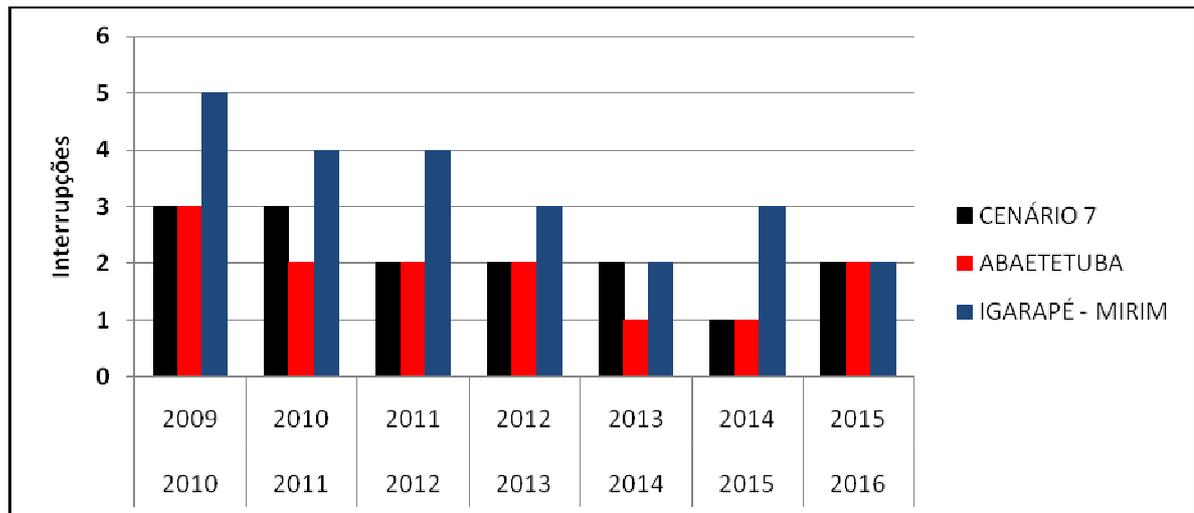


Figura 5.31: Taxa de redução anual para o cenário 7 para o indicador FEC

A união dos conjuntos proporcionou folgas nas metas do conjunto Abaetetuba que é o conjunto que apresenta maior número de unidades consumidoras, maior consumo médio mensal e potência instalada. As folgas nas metas foram de 3 interrupções no primeiro e segundo anos, 2 interrupções no terceiro, quarto e quinto anos e 1 interrupção nos 3 últimos anos. Notou-se diminuição da taxa de redução anual para o conjunto Igarapé Mirim.

A Tabela 5.44 apresenta os resultados das metas obtidas pela aplicação da Metodologia Proposta e do FEC agregado dos conjuntos formadores do cenário 7.

Tabela 5.44: Comparação do FEC agregado com o obtido pela Metodologia Proposta para o cenário 7

FEC	2009	2016
Agregado	23,33	8,39
Metodologia Proposta	24	9

A comparação dos valores obtidos pela aplicação da Metodologia com os valores do FEC agregado dos conjuntos Abaetetuba e Igarapé Mirim mostra que esta não é uma união favorável para o indicador FEC, pois os valores de metas obtidos pela Metodologia são similares aos valores obtidos pelo FEC agregado para os anos de 2009 e 2016.

5.5 CONSIDERAÇÕES FINAIS

Este capítulo apresentou um estudo de caso realizado em uma empresa distribuidora de energia elétrica, a CELPA, atuante no estado do Pará. A CELPA possui uma área de concessão muito atípica onde os cinco atributos utilizados pela ANEEL na caracterização dos conjuntos de unidades consumidoras não são suficientes para uma correta caracterização de seus conjuntos. Muitas características que influenciam nos indicadores de continuidade DEC e FEC não são consideradas na análise realizada pela ANEEL provocando a definição de metas não condizentes com a realidade dos conjuntos da CELPA. Neste sentido, estudos de configurações de conjuntos são uma alternativa para se obter metas menos rígidas nos conjuntos de unidades consumidoras. A metodologia proposta no capítulo anterior foi utilizada para realizar a classificação de cenários definidos pela união de dois conjuntos de unidades consumidoras pertencentes à área de concessão da CELPA. Foram feitas análises que avaliavam se a união dos conjuntos proporcionava melhoria nas metas dos indicadores de continuidade para o cenário em questão. Foram analisadas as trajetórias das metas ao longo dos dois períodos de revisão tarifária onde era realizada a comparação entre a trajetória de metas do cenário e as trajetórias dos conjuntos formadores do cenário, a taxa de redução anual e a comparação com o indicador de continuidade agregado. As análises dos cenários mostraram que em alguns casos é vantajosa a união de conjuntos proporcionando folgas satisfatórias nas metas dos indicadores de continuidade dos cenários se comparadas aos conjuntos com maior número de unidades consumidoras e redução da taxa de redução anual. A comparação das metas obtidas pela Metodologia Proposta com o indicador agregado mostrou ser um importante indicador da qualidade da união do conjunto. A Metodologia Proposta provou ser adequada como auxiliadora no processo de tomada de decisão na definição dos conjuntos de unidades consumidoras da área de concessão das empresas de distribuidoras de energia elétrica.

CAPÍTULO 6

CONCLUSÃO

6.1 REALIZAÇÕES E CONTRIBUIÇÕES

Este trabalho aborda o aspecto de serviço da energia elétrica onde é explorada a continuidade do fornecimento de energia elétrica aos consumidores finais e apresentou o desenvolvimento de uma metodologia de auxílio no processo de tomada de decisão da definição dos conjuntos de unidades consumidoras da área de concessão de empresas de distribuição de energia elétrica.

O Módulo 8: Qualidade da Energia Elétrica do PRODIST confirma a análise comparativa entre empresas distribuidoras de energia elétrica estabelecida pela Resolução nº 024/2000 como base para o estabelecimento de metas dos indicadores de continuidade coletivos DEC e FEC. A metodologia adotada atualmente pela ANEEL consiste na caracterização dos conjuntos de unidades consumidoras de todo o Brasil por meio de cinco atributos. Os atributos são submetidos ao processo de agrupamento via algoritmo *k means*. Os conjuntos de todo o Brasil são agrupados em 30 famílias, ou *clusters*. Para cada um dos *clusters* é aplicada a técnica *Yardstick Competition* que identifica por meio de diversas métricas (percentis, mediana, etc) o conjunto de referência. Os demais conjuntos de uma mesma família deverão alcançar o mesmo desempenho do conjunto de referência em um determinado espaço de tempo que foi definido como oito anos (duas revisões tarifárias). A definição dos conjuntos de unidades consumidoras exerce forte influência na definição das metas anuais dos indicadores de continuidade coletivo.

A metodologia da ANEEL não permite que sejam feitas análises de configurações de conjuntos de unidades consumidoras sem que se utilizem as características de todos os conjuntos existentes no Brasil. Este fato dificulta estudos que indiquem melhores configurações de conjuntos de modo a diminuir a violação das metas dos indicadores de continuidade. Logo, fica evidente o desenvolvimento de ferramentas que auxiliem as empresas de distribuição de energia elétrica na definição de melhores configurações de conjuntos de unidades consumidoras nas suas respectivas áreas de concessão.

A metodologia proposta por este trabalho consiste no emprego de técnicas de aprendizado de máquina com o objetivo de construir um classificador que capture as

características intrínsecas de cada um dos 30 *clusters* definidos pela metodologia empregada pela ANEEL. Uma vez extraídas essas características não é necessário mais utilizar toda a base de dados de todos os conjuntos existentes no Brasil e é possível realizar análises de configurações de conjuntos visando a diminuição das violações das metas dos indicadores de continuidades.

As técnicas avaliadas na construção do classificador foram duas: as redes neurais do tipo *perceptron* de múltiplas camadas (MLP) e o algoritmo kNN. Para descobrir qual das técnicas é mais apropriada na construção do classificador foi realizada a análise de desempenho utilizando a validação cruzada com 10-*folds*. Os resultados obtidos na análise de desempenho mostraram que as redes neurais tipo MLP são a técnica mais apropriada como classificador não somente pelo fato de uma menor taxa de erro na classificação de um novo exemplo, mas também por não necessitar relembrar todos os dados do conjunto de treinamento. Para melhorar ainda mais o desempenho do classificador optou-se por utilizar não apenas uma rede neural, mas um conjunto de 10 redes neurais que apresentariam o resultado de qual *cluster* deveria ser alocado o novo conjunto. A escolha final de em qual *cluster* o novo conjunto será alocado é feita por eleição simples das respostas das 10 redes neurais. Conhecendo-se o *cluster* do novo conjunto e o ano zero pode-se determinar as metas dos indicadores de continuidade ao longo de 8 anos. Com isso, é possível a realização de estudos de cenários explorando diversas situações que são de interesse para as empresas distribuidoras, tais que: influência da variação de atributos, estudos em sistemas malhados para alocação de blocos de consumidores e construção de novas subestações. Este trabalho explorou um cenário como sendo a união de dois conjuntos já existentes na área de concessão da empresa de distribuição de energia elétrica.

A metodologia proposta foi aplicada em um estudo de caso de uma empresa de distribuição de energia elétrica atuante na região norte do Brasil, a CELPA, cuja área de concessão é todo o estado do Pará. As cinco variáveis empregadas pela ANEEL não conseguem caracterizar as peculiaridades existentes no estado do Pará que exercem influência no resultado dos indicadores de continuidade realizados pela CELPA. Devido essa má caracterização, as metas definidas para os conjuntos da CELPA são muito rígidas não refletindo a realidade dos conjuntos e comprometendo a qualidade da energia no que tangem a continuidade do fornecimento de energia elétrica.

O estudo de caso apresentou as análise de 7 cenários constituídos pela união de dois conjuntos de unidades consumidoras de energia elétrica pertencentes à área de concessão da empresa CELPA. As análises tinham objetivo encontrar conjuntos que se unidos

possibilitassem a definição de metas mais condizentes com a realidade da empresa CELPA. As análises dos cenários mostraram que em alguns casos é vantajoso realizar a reconfiguração dos conjuntos de unidades consumidoras por se obter folgas nas metas dos conjuntos com maior número de unidades consumidas, diminuição da taxa de redução anual do novo conjunto se comparados com os conjuntos formadores deste cenário e nos melhores casos a união proporcionou folgas nas metas para ambos os conjuntos participantes da união como foi o caso do cenário 6: Curionópolis e Eldorado do Carajás. Para os cenários onde a união foi considerada não satisfatória, a diferença entre os valores obtidos pela metodologia proposta e o cálculo do indicador agregado não eram muito grandes. A comparação das metas obtidas pela metodologia com o cálculo do indicador agregado mostrou-se uma importante métrica na qualidade da união dos conjuntos. Notou-se que a melhoria de um determinado indicador não implica necessariamente na melhoria do outro. Para estes casos, a empresa de distribuição de energia elétrica poderá favorecer um dos indicadores por meio das folgas nas metas e implantar medidas concretas para a melhoria do outro indicador, tais que: melhor distribuição ou aumento de equipes de reparo para o caso do indicador DEC ou investimentos em manutenção para o caso do indicador FEC.

Portanto, este trabalho cumpriu o objetivo proposto que era o desenvolvimento de uma metodologia que auxiliasse a tomada de decisão da formação dos conjuntos de unidades consumidoras de uma determinada empresa distribuidora que fosse baseada na atual metodologia empregada pelo órgão regulador do sistema elétrico nacional.

6.2 SUGESTÃO DE TRABALHOS FUTUROS

Citam-se alguns pontos que poderão ser explorados em trabalhos futuros:

- Construção de modelos de redes neurais utilizando a base de dados onde os conjuntos são definidos por subestação: Os modelos de redes neurais utilizados neste trabalho foram construídos com base no banco de dados fornecido pela ANEEL onde ainda não existia o critério de formação definido. Com a publicação da primeira revisão do PRODIST, que definiu o critério de formação foi estabelecido como a área de abrangência das subestações, um novo banco de dados será construído pela ANEEL para realizar a definição das metas dos indicadores de continuidade. Ao se utilizar um novo banco de dados as características intrínsecas de cada um dos 30 *clusters* sofrerão

alterações. Por esta razão é necessário realizar todo o trabalho de construção das redes neurais descrito no Capítulo 4 deste trabalho;

- Melhoria do processo de escolha do *cluster* no arranjo de redes neurais: como descrito no Capítulo 4 o processo de escolha do *cluster* é feito por eleição por maioria das respostas dos 10 classificadores neurais. Este processo de eleição poderá ser melhorado se ocorresse uma atribuição de pesos a cada uma das respostas das redes neurais integrantes do arranjo. Maiores pesos seriam atribuídos às redes que obtivessem menor taxa de erro;
- Adicionar uma análise comparativa da trajetória das metas com os indicadores realizados pela empresa distribuidora de energia elétrica. Com a adição dos valores realizados dos indicadores será possível a construção de cenários que levem em consideração fatores que influenciam na diminuição dos indicadores realizados tais como: planos de obras, investimentos na rede de distribuição, números de equipes de manutenção, entre outras ações. Com isso, é possível a realização de estudos de projeção dos indicadores DEC e FEC realizados pela empresa de distribuição que indiquem futuras transgressões nas metas dos indicadores de continuidade. Também nestes cenários poderão ser analisadas variáveis exógenas que em certas áreas de concessões possuem forte influência no resultado dos indicadores realizados e que não são levadas em consideração na análise feita pela ANEEL;
- Implementação via software da metodologia proposta: como demonstrado no Capítulo 5 deste trabalho a Metodologia Proposta é uma importante ferramenta na tomada de decisão da definição dos conjuntos de unidades consumidoras da área de concessão da empresa de distribuição de energia elétrica, por tanto fica evidente que a implementação via software desta metodologia facilitará o processo de definição dos conjuntos de unidades consumidoras.

REFERÊNCIAS BIBLIOGRÁFICAS

AGÊNCIA NACIONAL DE ENERGIA ELÉTRICA. Atlas de Energia Elétrica do Brasil. 3 ed. Brasília, 2008.

_____. Resolução normativa nº 024, de 27 de janeiro de 2000. Estabelece as disposições relativas à continuidade da distribuição de energia elétrica às unidades consumidoras.

_____. Resolução normativa nº 345, de 16 de dezembro de 2008. Aprova os Procedimentos de Distribuição de Energia Elétrica no Sistema Elétrico Nacional – PRODIST, e dá outras providências.

_____. Resolução normativa nº 395, de 15 de dezembro de 2009. Aprova a Revisão 1 dos Procedimentos de Distribuição de Energia Elétrica no Sistema Elétrico Nacional – PRODIST, e dá outras providências.

_____. Procedimentos de Distribuição de Energia Elétrica no Sistema Elétrico Nacional – PRODIST. Módulo 8 – Qualidade da Energia Elétrica. Revisão 1.

AZEVEDO, F. M., BRASIL, L. M., OLIVEIRA, R. C. L. Redes Neurais com Aplicação em Controle e em Sistemas Especialistas. 1.ed. Visual Books, 2000. 401p.

BRAGA, A. P.; CARVALHO, A. C. P. L. F; LUDEMIR, T. B. Redes Neurais Artificiais: teoria e aplicações. 1.ed. Rio de Janeiro. LTC, 2000.

CALDEIRA, A. M., AGUIAR, H., JUNIOR, O., MACHADO, M. A. S., SOUZA, R. C., TANSHELL, R. Inteligência Computacional aplicada à Administração, Economia e Engenharia em Matlab®. 1. ed. São Paulo: Thosom Learning, 2007. 370p.

DEPARTAMENTO NACIONAL DE ÁGUA E ENERGIA ELÉTRICA. Portaria nº 046, de 17 de abril de 1978.

FERRERO, C. A. Algoritmo kNN para previsão de dados temporais: função de previsão e critérios de seleção de vizinhos próximos aplicados a variáveis ambientais em limnologia. 2009. 129f. Dissertação (Mestrado em Ciência da Computação e Matemática Computacional) – Universidade de São Paulo, São Carlos, 2009. [Orientador: Prof^a Dr^a Maria Carolina Monard].

HAYKIN, S. Redes Neurais: Princípios e Prática. 2.ed. Porto Alegre: Bookman, 2001. 900p.

HUNZILER, M. A., NASCIMENTO, M. H. L. Determinação de Metas de Qualidade de Serviço. Congresso LatinoAmericano de Distribuion Electrica, 2008.

KAGAN, N., ROBBA, E. J., SCHMIDT, H. P. Estimação de Indicadores de Qualidade de Energia Elétrica. 1 ed. São Paulo: Editora Blucher, 2009. 230p.

KOTLER, Philip Administração de Marketing: Análise, Planejamento, Implementação e Controle. 4.ed. São Paulo: Atlas, 1994. 676p.

MACIEL, J. Análise de um Sistema de Crédito Cooperativo através de Redes Neurais (MLP) com a utilização do Algoritmo Levenberg Marquardt. 2005. 99f. Dissertação (Mestrado em Ciências) – Universidade Federal do Paraná, Curitiba, 2005. [Orientador: Prof^a Dr^a Andrea Sell Dyminski].

MICHIE, D., SPIEGELHALTER, D. J., TAYLOR, C. C. Machine Learning, Neural and Statistical Classification. 1.ed. Prentice Hall, 1994. 289p.

MINGOTI, S. A. Análise de Dados Através de Métodos de Estatística Multivariada. 1. ed. Belo Horizonte: Editora UFMG, 2005. 297 p.

MITCHELL, T. M. Machine Learning. 1.ed. McGraw-Hill, 1997. 432p.

NAUCK, D., KLAWONN, F., KRUSE, R. Foundations of Neuro-Fuzzy Systems. 1.ed. New York: John Wiley & Sons, 1997. 305p.

NETO, A. S., LONGUE, C. A., SANTOS, E. C. Metodologia para Reavaliação dos limites de Indicadores de Performances dos Conjuntos Elétricos - ANEEL. Anais do V Seminário Brasileiro sobre Qualidade da Energia Elétrica (SBQEE), vol 2, pp: 463 – 467, Aracajú, Brasil, 2003.

PESSANHA, J. F. M., CASTTELLANI, V. L. O., HASSIM, E. S., CHEBERLE, L. A. D. ANABENCH – Sistema Computacional para Estabelecimento de Metas de Continuidade. XVI Seminário Nacional de Distribuição de Energia Elétrica, 2004.

PESSANHA, J. F. M. Um Modelo de Análise Envoltória de Dados para Estabelecimento das Metas de Continuidade do Fornecimento de Energia Elétrica. 2006. 161f. Tese (Doutorado em Engenharia Elétrica) – Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2006. [Orientador: Prof. Reinaldo Castro Souza].

QUEIROZ, H. L. Metodologia para Definição de Agrupamentos de Consumidores e os Requisitos de Qualidade do Fornecimento. XV Seminário Nacional de Distribuição de Energia Elétrica (SENDI), Salvador, 2002.

ROSA, F. A. F., QUEIROZ, H. L., LIMA, J. W. M. Aprimoramento da Regulação da Qualidade Utilizando Atributos de Clima e Envelhecimento da Rede. XI Simpósio de Especialistas em Planejamento da Operação e Expansão Elétrica (SEPOPE), Belém, 2009.

SANTOS, J. G. M. Indicadores e Indutores da Qualidade do Serviço de Distribuição de Energia Elétrica. VI Congresso Brasileiro de Regulação, Rio de Janeiro, 2009.

SPERANDIO, M., COELHO, J., QUEIROZ, H. L. Identificação de Agrupamentos de Consumidores de Energia Elétrica através de Mapas Auto-Organizáveis. Anais do V Seminário Brasileiro sobre Qualidade da Energia Elétrica (SBQEE), vol. 2, pp 439-443, Aracajú, 2003.

SPERANDIO, M. Classificação de Conjuntos Consumidores de Energia Elétrica via Mapas Auto-Organizáveis e Estatística Multivariada. 2004. 92f. Dissertação (Mestrado em Engenharia Elétrica) – Universidade Federal de Santa Catarina, Florianópolis, 2004. [Orientador: Prof. Jorge Coelho].

TANURE, J. E. P. S. Análise Comparativa de Empresas de Distribuição para o Estabelecimento de Metas de Desempenho para Indicadores de Continuidade do Serviço de Distribuição. 2000. 157f. Dissertação (Mestrado em Ciências em Engenharia) – Escola Federal de Engenharia de Itajubá, Itajubá, 2000. [Orientador: Prof. José Wanderley Marangon Lima].

TANURE, J. E. P. S. Proposta de Procedimentos e Metodologia para Estabelecimento de Metas de Qualidade (DEC e FEC) para Concessionárias de Distribuição de Energia Elétrica através da Análise Comparativa. 2004. 228f. Tese (Doutorado em Engenharia) – Universidade de São Paulo, São Paulo, 2004. [Orientador: Prof. Dr. Carlos Marcio Vieira Tahan].

TOSTES, W. S., TOSTES, M. E. L., SILVA, R. D. S., OLIVEIRA, T. F., SOUSA, S. C. P., MÜLLER, E. Análise de Novos Atributos para Determinação de Padrões de Continuidade no Fornecimento de Energia Elétrica – Um Estudo de Caso no Pará. VIII Congresso sobre Qualidade da Energia Elétrica (CBQEE), Blumenau, 2009.

WITTEN, I. H., FRANK, E. Data Mining: Practical Machine Learning Tools and Techniques. 2.ed. Elsevier, 2005. 525p.

WU, X., KUMAR, V. The Top Ten Algorithms in Data Mining. 1.ed. CRC Press, 2009. 215p.

ANEXO 1
Metas dos Indicadores DEC e FEC para conjuntos do tipo interligado e isolado.

METAS PARA DEC E FEC				
CLUSTER	Interligado		Isolado	
	DEC	FEC	DEC	FEC
1	15	10	21	16
2	16	11	30	18
3	40	23	63	34
4	14	10	23	14
5	22	10	43	25
6	23	27	41	43
7	6	5	10	8
8	8	8	15	12
9	28	18	43	28
10	17	11	22	15
11	33	19	53	33
12	10	8	14	11
13	10	8	16	14
14	11	8	17	13
15	18	10	33	19
16	14	9	22	13
17	12	9	28	14
18	9	7	15	11
19	17	10	26	14
20	10	8	16	12
21	22	12	32	18
22	15	21	40	36
23	6	5	11	8
24	9	8	15	13
25	11	8	20	12
26	7	6	10	7
27	22	11	33	15
28	32	26	60	40
29	10	8	17	13
30	15	8	26	13