

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

DESPACHO ÓTIMO DE REDES INTEGRADAS DE ENERGIA ELÉTRICA E DE GÁS
NATURAL COM RESTRIÇÕES DE SEGURANÇA VIA ÁRVORES DE DECISÃO

DENIS CARLOS LIMA COSTA

TD 20/2016

UFPA / ITEC / PPGEE
Campus Universitário do Guamá
Belém-Pará-Brasil
2016

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

DENIS CARLOS LIMA COSTA

DESPACHO ÓTIMO DE REDES INTEGRADAS DE ENERGIA ELÉTRICA E DE GÁS
NATURAL COM RESTRIÇÕES DE SEGURANÇA VIA ÁRVORES DE DECISÃO

TD 20/2016

UFPA / ITEC / PPGEE
Campus Universitário do Guamá
Belém-Pará-Brasil
2016

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

DESPACHO ÓTIMO DE REDES INTEGRADAS DE ENERGIA ELÉTRICA E DE GÁS
NATURAL COM RESTRIÇÕES DE SEGURANÇA VIA ÁRVORES DE DECISÃO

Tese submetida à Banca Examinadora do
Programa de Pós-Graduação em
Engenharia Elétrica da UFPA para a
obtenção do Grau de Doutor em
Engenharia Elétrica na área de Sistema
de Energia Elétrica.

UFPA / ITEC / PPGEE
Campus Universitário do Guamá
Belém-Pará-Brasil
2016

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

Costa, Denis Carlos Lima, 1970-

Despacho ótimo de redes integradas de energia elétrica e de gás natural com restrições de segurança via árvores de decisão / Denis Carlos Lima Costa. - 2016.

Orientador : Marcus Vinicius Alves Nunes

Tese (Doutorado) - Universidade Federal do Pará, Instituto de Tecnologia, Programa de Pós-Graduação em Engenharia Elétrica, Belém, 2016.

1. Energia elétrica – produção – métodos estatísticos. 2. Gás natural. 3. Energia – fontes alternativas . I. Título.

CDD 23. ed. 621.31

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

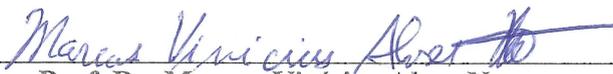
**“DESPACHO ÓTIMO DE REDES INTEGRADAS DE ENERGIA ELÉTRICA E DE GÁS NATURAL COM
RESTRICÇÕES DE SEGURANÇA VIA ÁRVORES DE DECISÃO”**

AUTOR: DENIS CARLOS LIMA COSTA

TESE DE DOUTORADO SUBMETIDA À BANCA EXAMINADORA APROVADA PELO COLEGIADO DO PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA, SENDO JULGADA ADEQUADA PARA A OBTENÇÃO DO GRAU DE DOUTOR EM ENGENHARIA ELÉTRICA NA ÁREA DE SISTEMAS DE ENERGIA ELÉTRICA.

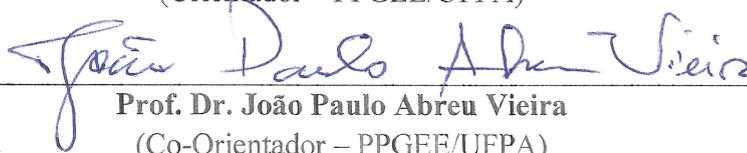
APROVADA EM: 07/12/2016

BANCA EXAMINADORA:



Prof. Dr. Marcus Vinicius Alves Nunes

(Orientador – PPGEE/UFPA)



Prof. Dr. João Paulo Abreu Vieira

(Co-Orientador – PPGEE/UFPA)



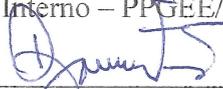
Prof. Dr. Roberto Célio Limão de Oliveira

(Avaliador Interno – PPGEE/UFPA)



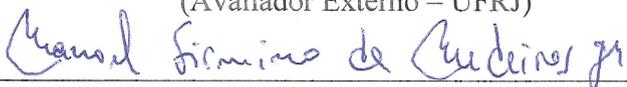
Prof. Dr. Ubiratan Holanda Bezerra

(Avaliador Interno – PPGEE/UFPA)



Prof. Dr. Djalma Mosqueira Falcão

(Avaliador Externo – UFRJ)



Prof. Dr. Manoel Firmino de Medeiros Junior

(Avaliador Externo – UFRN)

VISTO:



(Coordenador do PPGEE/ITEC/UFPA)

DEDICATÓRIA

“... mas quem beber da água que eu darei, nunca mais terá sede”.

(Jo 4, 13-14)

Ao meu pai, Carlos Costa.

Nele eu vejo bem mais que a inteligência:

Eu *bebo* da sua sabedoria.

AGRADECIMENTOS

Agradeço a Deus que me deu inteligência para compreender o Seu plano à minha vida.

Aos meus pais, Carlos e Maria, que sempre foram incentivadores da busca pelo conhecimento.

À minha esposa, Jailsa Costa, pelo seu apoio e sua paciência.

Ao meu filho, Heictor Costa, por sua inspiração e ideias.

Ao meu Orientador, Professor Dr. Marcus Vinícius Alves Nunes, pois nunca deixou de acreditar em mim.

Ao Professor Dr. Ubiratan Holanda Bezerra, por suas contribuições magníficas a esse trabalho.

Aos amigos do CEAMAZON, em especial aos prezados Prof. Me. Bernard Carvalho Bernardes, Eng. Me. Werbeston Douglas e Eng. Me. Dione Vieira com quem vivenciei momentos de trabalho e descontração.

O meu agradecimento em especial ao Professor Dr. João Paulo Abreu Vieira. A sua participação nesse trabalho e no meu crescimento intelectual são incomensuráveis. Professor João Paulo, ou apenas João, o meu muito obrigado por desfrutar da sua competência, da sua inteligência e da sua paciência. Foi uma honra tê-lo como Orientador. Muito obrigado!

EPÍGRAFE

“Se a ÁRVORE não deu FRUTOS, valeu pela beleza das FLORES;
Se não deu FLORES, valeu pela sombra das FOLHAS;
Se não deu FOLHAS, valeu pela intenção da SEMENTE”.

Henfil.

SUMÁRIO

DEDICATÓRIA.....	vi
AGRADECIMENTOS.....	vii
EPÍGRAFE.....	viii
SUMÁRIO.....	ix
LISTA DE ILUSTRAÇÕES.....	xiii
LISTA DE TABELAS.....	xvi
LISTA DE SIGLAS.....	xviii
RESUMO.....	xix
ABSTRACT.....	xx
CAPÍTULO I – INTRODUÇÃO	1
1.1 CONTEXTUALIZAÇÃO.....	1
1.2 REVISÃO DE LITERATURA.....	4
1.3 OBJETIVO DA TESE.....	7
1.4 CONTRIBUIÇÃO DA TESE	7
1.5 PUBLICAÇÃO EM PERIÓDICO.....	8
1.6 ORGANIZAÇÃO DA TESE.....	8
1.7 CONCLUSÃO DO CAPÍTULO I.....	9
CAPÍTULO II – O GÁS NATURAL.....	10
2.1 DEFINIÇÃO.....	10
2.2 ANÁLISE HISTÓRICA.....	10
2.3 PANORAMA NO BRASIL.....	12
2.4 ORIGEM E COMPOSIÇÃO DO GÁS NATURAL.....	13

2.5 GÁS NÃO-ASSOCIADO.....	14
2.6 GÁS ASSOCIADO.....	15
2.7 COMPORTAMENTO DO GÁS NATURAL.....	16
2.8 PARTICIPAÇÃO NA MATRIZ ENERGÉTICA.....	19
2.9 APLICAÇÃO ENERGÉTICA DO GÁS NATURAL.....	22
2.10 APLICAÇÃO DO GÁS NATURAL NA GERAÇÃO DE ENERGIA ELÉTRICA.....	24
2.11 CONCLUSÃO DO CAPÍTULO II.....	28
CAPÍTULO III - ÁRVORE DE DECISÃO.....	29
3.1 INTRODUÇÃO.....	29
3.2 MINERAÇÃO DE DADOS.....	32
3.3 ÁRVORE DE DECISÃO – CONCEITO.....	34
3.4 ESTRUTURAÇÃO DE UMA ÁRVORE DE DECISÃO.....	37
3.5 CRITÉRIOS PARA SELEÇÃO DOS ATRIBUTOS DE UMA ÁRVORE DE DECISÃO.....	38
3.5.1 TOP-DOWN INDUCTION OF DECISION TREE (TDIDT).....	39
3.6 SELEÇÃO DOS ATRIBUTOS PREDITIVOS PARA OS NÓS DA ÁRVORE DE DECISÃO.....	40
3.6.1 GANHO DE INFORMAÇÃO.....	41
3.6.2 RAZÃO DE GANHO.....	41
3.6.3 CRITÉRIO DE GINI.....	42
3.6.4 REPRESENTAÇÃO DOS NÓS PARA ATRIBUTOS CATEGÓRICOS.....	42
3.6.5 REPRESENTAÇÃO DOS NÓS PARA ATRIBUTOS CONTÍNUOS.....	44
3.7 MÉTODOS DE PODA DA ÁRVORE DE DECISÃO.....	44
3.8 ALGORITMOS DE INDUÇÃO DE ÁRVORES DE DECISÃO.....	46

3.8.1 ID3.....	46
3.8.2 C4.5.....	46
3.8.3 CART.....	47
3.9 CONCLUSÃO DO CAPÍTULO III.....	47
CAPÍTULO IV – OTIMIZAÇÃO.....	49
4.1 INTRODUÇÃO.....	49
4.2 MÉTODOS DETERMINÍSTICOS <i>VERSUS</i> HEURÍSTICOS.....	52
4.3 ESTÁGIOS DOS ALGORITMOS GENÉTICOS.....	53
4.4 SISTEMA DE REPRESENTAÇÃO.....	54
4.5 ESQUEMAS DE SELEÇÃO.....	55
4.6 OPERADORES GENÉTICOS.....	56
4.7 PARÂMETROS GENÉTICOS.....	58
4.8 ESTRATÉGIAS ELITISTAS E ADAPTAÇÃO DINÂMICA.....	58
4.9 CONCLUSÃO DO CAPÍTULO IV.....	60
CAPÍTULO V – CONFIABILIDADE DO SIEEGN.....	62
5.1 INTRODUÇÃO.....	62
5.2 A EVOLUÇÃO DO CONCEITO DE SEGURANÇA EM SISTEMAS ELÉTRICOS.....	62
5.3 AVALIAÇÃO DE SEGURANÇA EM SISTEMAS ELÉTRICOS.....	64
5.4 A SEGURANÇA EM SISTEMAS DE GÁS NATURAL.....	65
5.5 CONCLUSÃO DO CAPÍTULO V.....	66
CAPÍTULO VI – MÉTODO DE DESPACHO PROPOSTO.....	67
6.1 INTRODUÇÃO.....	67
6.2 DESPACHO SEGURO - ESQUEMA PROPOSTO.....	67

6.2.1 ANÁLISE DE CONTINGÊNCIAS.....	68
6.2.2 GERAÇÃO DA BASE DE DADOS.....	68
6.3 ÁRVORE DE DECISÃO E REGIÃO DE SEGURANÇA.....	69
6.4 DESENVOLVIMENTO E TREINAMENTO DA ÁRVORE DE DECISÃO.....	73
6.5 FLUXO DE POTÊNCIA E DE GÁS NATURAL OTIMIZADO COM RESTRIÇÕES DE SEGURANÇA.....	73
6.6 CONCLUSÃO DO CAPÍTULO VI.....	76
CAPÍTULO VII – ESTUDO DE CASO.....	77
7.1 INTRODUÇÃO.....	77
7.2 CASO 1: CONTINGÊNCIAS NO GASODUTO 3→4 E NA LINHA DE TRANSMISSÃO 30→38.....	79
7.3 CASO 2: CONTINGÊNCIAS NOS GASODUTOS 3→4 E 13→14 E NA LINHA DE TRANSMISSÃO 30→38.....	84
7.4 CASO 3: CONTINGÊNCIAS NOS GASODUTOS 3→4 E 13→14 E NAS LINHAS DE TRANSMISSÃO 30→38 E 19→34.....	91
7.5 CONCLUSÃO DO CAPÍTULO VII.....	96
CAPÍTULO VIII - CONCLUSÕES.....	98
8.1 CONSIDERAÇÕES FINAIS.....	98
8.2 PROPOSTA DE TRABALHOS FUTUROS.....	99
REFERÊNCIAS.....	100

LISTA DE ILUSTRAÇÕES

CAPÍTULO II

Figura 01: Evolução do consumo de energia.....	12
Figura 02: Formas de ocorrência do gás natural.....	14
Figura 03: Esquema da geologia dos recursos de gás natural.....	15
Figura 04: Representação do comportamento do gás natural.....	17
Figura 05: Participação dos combustíveis na matriz energética mundial.....	20
Figura 06: Gasodutos brasileiros e dos países vizinhos.....	21
Figura 07: Diagrama esquemático do ciclo aberto.....	26
Figura 08: Diagrama esquemático do ciclo combinado.....	27
Figura 09: Rendimento energético do processo de cogeração.....	27

CAPÍTULO III

Figura 10: Diagrama do processo de indução de um classificador e sua utilização na dedução de novos casos.....	31
Figura 11: Uma Árvore de Decisão para classificação de gasodutos.....	35
Figura 12: Conjunto de Regras de uma Árvore de Decisão.....	36
Figura 13: Indução da AD por profundidade.....	38
Figura 14: Indução da AD por largura.....	38

CAPÍTULO IV

Figura 15: Operador de Cruzamento.....	51
Figura 16: Operador de Mutação.....	51
Figura 17: Esquema de Reposição Populacional em um AG.....	54
Figura 18: Exemplificação do Método da Roleta.....	56
Figura 19: Comportamento da População no Decorrer das Gerações.....	60

CAPÍTULO VI

Figura 20: Esquema de geração do Banco de Dados.....	69
Figura 21: Uma típica representação do Cart.....	71
Figura 22: Região de Segurança definida pela AD.....	72

CAPÍTULO VII

Figura 23: Rede de Gás Natural de 15 nós.....	77
Figura 24: Rede de 118 barras – IEEE.....	78
Figura 25: AD para contingências no gasoduto 3→4 e na linha de transmissão 30→38.....	79
Figura 26: Nomograma em 2-D definido pelo conjunto de regras do caso 01 quando a $P_{B010} > 352,395$ MW e a $P_{B103} > 32,633$ MW.....	81
Figura 27: Nomograma em 2-D definido pelo conjunto de regras 01 do caso 01 quando a $P_{B103} > 31,633$ MW e a $58,610 \text{ bar} < Pre_{01} \leq 59,972 \text{ bar}$	81
Figura 28: Nomograma em 2-D definido pelo conjunto de regras 02 do caso 01 quando a $P_{B103} > 31,633$ MW e a $Pre_{01} \leq 58,610 \text{ bar}$	82
Figura 29: Nomograma em 2-D definido pelo conjunto de regras do caso 01 quando a $Pre_{01} \leq 59,972 \text{ bar}$ e a $P_{B010} > 352,395$ MW.....	82
Figura 30: Despacho de geradores a gás natural e a diesel considerando as contingências no duto 3→4 e na linha de transmissão 30→38.....	83
Figura 31: AD para contingências nos gasodutos 3→4 e 13→14 e na linha de transmissão 30→38.....	85
Figura 32: Nomograma em 2-D definido pelo conjunto de regras do caso 02 quando a $Pre_{01} \leq 59,972 \text{ bar}$ e a $Pro_{02} > 5,508 \text{ Mm}^3$	87
Figura 33: Nomograma em 2-D definido pelo conjunto de regras do caso 02 quando a $Pro_{01} > 5.508 \text{ Mm}^3$ e a $0.977 \text{ p.u.} < V_{B10} \leq 1.038 \text{ p.u.}$	87
Figura 34: Nomograma em 2-D definido pelo conjunto de regras do caso 02 quando a $0.977 \text{ p.u.} < V_{B10} \leq 1.006 \text{ p.u.}$ e a $P_{B61} > 155.695 \text{ MW}$, regra 01.....	88
Figura 35: Nomograma em 2-D definido pelo conjunto de regras do caso 02 quando a $0.977 \text{ p.u.} < V_{B10} \leq 1.006 \text{ p.u.}$ e a $P_{B61} > 155.695 \text{ MW}$, regra 02.....	88
Figura 36: Nomograma em 2-D definido pelo conjunto de regras do caso 02 quando a $0.977 \text{ p.u.} < V_{B10} \leq 1.006 \text{ p.u.}$ e a $P_{B61} > 155.695 \text{ MW}$, regra 03.....	89

Figura 37: Despacho de geradores a gás e a diesel considerando as contingências nos dutos 3→4 e 13→14 e na linha de transmissão 30→38.....	89
Figura 38: AD para contingências nos gasodutos 3→4 e 13→14 e nas linhas de transmissão 30→38 e 19→34.....	91
Figura 39: Nomograma em 2-D definido pelo conjunto de regras do caso 03 quando a $Pro_02 > 5,516 \text{ Mm}^3$ e o $F07 \rightarrow 08 > 4,324 \text{ m}^3/\text{h}$	93
Figura 40: Nomograma em 2-D definido pelo conjunto de regras do caso 03 quando o $F07 \rightarrow 08 > 4,324 \text{ m}^3/\text{h}$ e a $P_B49 > 110,335 \text{ MW}$	93
Figura 41: Nomograma em 2-D definido pelo conjunto de regras do caso 03 quando a $P_B49 > 264,790 \text{ MW}$ e a $P_B12 \leq 100,432 \text{ MW}$, regra 01.....	94
Figura 42: Nomograma em 2-D definido pelo conjunto de regras do caso 03 quando a $110,335 \text{ MW} < P_B49 \leq 264,790 \text{ MW}$ e a $P_B12 > 89,517 \text{ MW}$, regra 02.....	94
Figura 43: Despacho de geradores a gás e a diesel considerando as contingências nos dutos 3→4 e 13→14 e nas linhas de transmissão 30→38 e 19→30.....	95

LISTA DE TABELAS

CAPÍTULO II

Tabela 1: Composição do Gás Natural.....	14
Tabela 2: Resolução ANP 16 - 17/06/2008.....	18
Tabela 3: Consumo de Gás Natural por setor em Mtep.....	23
Tabela 4: Centrais Termelétricas a Gás Natural em Operação no Brasil.....	24

CAPÍTULO III

Tabela 5: Conjunto de dados para o diagnóstico de um Sistema de Gasodutos.....	30
Tabela 6: Conjunto de dados para a dedução de diagnóstico de um Sistema de Gasodutos.....	31

CAPÍTULO IV

Tabela 7: Mapeamento de Variáveis Reais para Código Binário.....	54
Tabela 8: Representação de um Indivíduo para um Problema de Três Variáveis.....	55
Tabela 9: Representação de uma População com Quatro Indivíduos.....	55
Tabela 10: Representação do Processo de Seleção.....	56
Tabela 11: Representação do Cruzamento.....	57
Tabela 12: Representação da Mutação.....	58

CAPÍTULO V

Tabela 13: Produção de gás natural para contingências no duto 3→4 e na linha de transmissão 30 → 38.....	83
Tabela 14: Pressão (em bar) dos nós da rede de gás natural para contingências no duto 3→4 e na linha de transmissão 30→38.....	84
Tabela 15: Produção de gás natural para contingências nos dutos 3→4 e 13→14 e na linha de transmissão 30→38.....	89

Tabela 16: Pressão (em bar) dos nós da rede de gás natural para contingências nos dutos 3→4 e 13→14 e na linha de transmissão 30→38.....	90
Tabela 17: Produção de gás natural para contingências nos dutos 3→4 e 13→14 e nas linhas de transmissão 30→38 e 19→34.....	95
Tabela 18: Fluxo de gás (em m ³ /h) na rede de gás natural para contingências nos dutos 3→4 e 13→14 e nas linhas de transmissão 30→38 e 19→30.....	96

LISTA DE SIGLAS

AAM	Algoritmos de Aprendizado de Máquina
AD	Árvore de Decisão
ANAREDE	Análise de Redes Elétricas
ANEEL	Agência Nacional de Energia Elétrica
ANP	Agência Nacional de Petróleo, Gás Natural e Biocombustíveis
BEN	Balanco Energético Nacional
BP	British Petroleum
CART	Classification and Regression Trees
CEPEL	Centro de Pesquisa de Energia Elétrica
EPE	Empresa de Pesquisa Energética
EUA	Estados Unidos da América
GASMIG	Companhia de Gás de Minas Gerais
IA	Inteligência Artificial
IEEE	Institute of Electrical and Electronic Engineers
KDD	Knowledge Discovery in Database
MATLAB	Matrix Laboratory
MME	Ministério das Minas e Energia
Mtep	Mega tonelada equivalente de petróleo
PCS	Poder Calorífico Superior
PIB	Produto Interno Bruto
RIEEGN	Redes Integradas de Energia Elétrica e Gás Natural
RNAs	Redes Neurais Artificiais
SIEEGN	Sistema Integrado de Energia Elétrica e de Gás Natural
TDIDT	Top-Down Induction of Decision Tree
UPGN	Unidades de Processamento de Gás Natural

RESUMO

Esta tese propõe um método de despacho ótimo seguro baseado em Árvore de Decisão (AD) aplicado em Redes Integradas de Energia Elétrica e de Gás Natural (RIEEGNs) considerando contingências críticas que podem causar violações dos limites operacionais. Ajustes preventivos na geração de energia elétrica e na produção de gás natural são realizados com base nas regiões de segurança e nos limites de variáveis de controle determinados pelas ADs. As regras de fácil interpretação da AD que descrevem as regiões de segurança são consideradas como restrições no problema de otimização redespacho da geração de energia elétrica e da produção de gás natural. Algumas contingências críticas específicas foram aplicadas ao sistema teste de 118 barras do IEEE integrado com uma rede teste de gás natural de 15 nós. Tais contingências são tomadas como exemplos para demonstrar uma aplicação promissora do método de despacho seguro proposto para restaurar a segurança das RIEEGNs.

PALAVRAS-CHAVE

Sistemas Integrados de Redes de Energia Elétrica e de Gás Natural, Fluxo de Potência Ótimo, Árvore de Decisão, Despacho Seguro.

ABSTRACT

This thesis proposes the use of a decision tree (DT) based security dispatch method applied to integrated electric power and natural-gas networks (IPGN) against credible contingencies that may cause violations. Preventive adjustments to the optimal electric energy generation and gas production are carried out based on the security regions and boundaries of controllable variables determined by the DTs. The easily interpretable DT's rules that describe the security regions are tractable constraints to be included in the optimization routines of electricity generation and gas production rescheduling. Some specific critical contingencies applied to the IEEE 118-bus test system integrated with the 15-node natural gas network are taken as examples to demonstrate a promising application of the proposed security dispatch method to restore IPGN security.

KEYWORDS

Integrated energy systems, natural gas networks, optimal power flow, decision tree, security dispatch.

CAPÍTULO I – INTRODUÇÃO

1.1 CONTEXTUALIZAÇÃO

As fontes de energia disponíveis no planeta são fundamentais ao fortalecimento de uma sociedade industrializada, uma vez que a economia dos países desenvolvidos e emergentes depende cada vez mais de uma infra-estrutura energética complexa e interdependente.

Tais fontes de energia são divididas em duas categorias principais: convencionais (carvão, petróleo, hidrelétrica, nuclear, gás natural) e emergentes (solar, eólica, biomassa, etc). Nesse contexto, o uso de gás natural como combustível para gerar energia é representado como uma interseção à utilização mais limpa e eficiente entre os meios de geração de energia existentes no mercado atual.

O gás natural (GN) ganhou espaço na matriz energética mundial, especialmente, na substituição ao carvão mineral e ao óleo combustível, uma vez que a sua utilização em processos de geração de energia traz uma série de benefícios operacionais e ambientais. Dentre estes benefícios pode-se destacar: redução de custos de operação, redução de custos de manutenção, redução de emissões de poluentes atmosféricos, principalmente o CO₂.

O crescimento contínuo e rápido de centrais geradoras de energia alimentadas a Gás Natural (GN), está vinculado ao aumento previsto na demanda de GN nas próximas décadas. Esta mudança para o GN é impulsionada por uma maior eficiência, menores custos de implantação, redução no tempo de construção e menor impacto ambiental devido a queima de Gás Natural a ciclo combinado (NATURAL GAS, 2016).

Há várias diferenças entre a eletricidade e o GN. Por exemplo, a eletricidade viaja à velocidade da luz, enquanto que o GN, no gasoduto, desloca-se, em média, à velocidade de 80 km/h. A eletricidade não é armazenável de forma direta. O GN, por sua vez pode ser armazenado em sistema de gasodutos ou em câmaras, a fim de atender a uma demanda durante certo período (RESSEL et al, 2015).

Projetos de transmissão de energia elétrica são muito grandes em termos de escala econômica, assim é muito mais econômico instalar uma linha de transmissão com uma capacidade bem superior à demanda máxima de carga do que instalá-la com capacidade igual à demanda máxima de carga, para posterior ampliação ou instalação de uma nova linha quando a capacidade da linha em uso for ultrapassada. Os gasodutos, por sua vez, são comumente operados a uma pressão mais baixa quando o fluxo de gás é bem inferior a capacidade do gasoduto. A pressão poderá ser elevada na medida que a demanda por gás for

umentando para obtenção de capacidade adicional. E ainda, fluxos de GN em gasodutos podem ser controlados independentes dos componentes da rede de gás.

Mesmo com tantas diferenças, há uma interdependência entre os sistemas de GN e de eletricidade apoiada na sustentabilidade social das infra-estruturas energéticas. A interdependência entre esses sistemas pode, inevitavelmente, resultar em um risco de fornecimento elétrico em uma significativa escala associada com a segurança da infra-estrutura do GN.

A decisão de usar termelétricas a GN normalmente está vinculada à preservação ambiental, à logística, à redução de custos, à disponibilidade e à confiabilidade da fonte de energia (SANTOS, 2000).

Todavia, uma interrupção nos sistemas de gasodutos ou uma diminuição de pressão nos nós produtores de GN pode levar a uma perda de geração de energia elétrica nos turbogeradores, o que poderia reduzir o fornecimento de potência elétrica e colocar em perigo a segurança do sistema integrado de energia elétrica e de gás natural (SIEEGN). Para o caso de certas contingências no sistema de GN, os operadores devem fornecer opções ao fornecimento de GN às unidades termelétricas. Dessa forma, o despacho de energia com base em decisões pertinentes deve ser efetivo e eficiente para contornar problemas de violação de restrições operacionais no SIEEGN.

A segurança da rede de GN é necessária para apoiar o crescimento econômico de uma nação. A diminuição da vulnerabilidade da rede de GN, do ponto de vista de segurança, poderá ser alcançada com o monitoramento, análise e estimativas das variáveis envolvidas no modelo, e conseqüentemente, no controle e adequação às demandas do sistema elétrico de potência. É indispensável incluir no planejamento de um sistema de potência a infra-estrutura do sistema de GN. O planejamento integrado entre esses sistemas poderá identificar as contingências e promover as melhorias necessárias ao modelo integrado.

É importante também considerar que existe uma instabilidade no fluxo de GN em gasodutos devido, principalmente, às variações na demanda do gás. A natureza instável do fluxo de gás indica a necessidade de um modelo que represente, no gasoduto, tais condições. Ou seja, o modelo deverá projetar redes de gasodutos capaz de estimar a demanda do gás, impedindo assim um inadequado dimensionamento (SANTOS, 1997).

Nesse caso, é economicamente viável a instalação de estações de compressores ao longo do gasoduto. Tais compressores ajustam a pressão do gás na rede de modo a satisfazer a demanda de gás, para fins elétricos e não-elétricos.

Quando as Redes Integradas de Energia elétrica e de Gás Natural (RIEEGNs) se encontram vulneráveis, em uma dada condição de operação, a uma provável contingência específica, como por exemplo, vazamento de gás em gasodutos e/ou desligamento de linhas de transmissão, ações de controle preventivo tornam-se indispensáveis para recuperar a segurança de ambos os sistemas de potência e de GN. O redespacho da geração de energia elétrica e da produção de gás, com restrições de segurança, é uma das ações preventivas que podem restaurar as RIEEGNs do estado inseguro para o estado seguro.

Por outro lado, a necessidade de operação coordenada entre os sistemas de energia elétrica e de GN, aumenta a complexidade e incerteza da tarefa de despachar a geração de energia elétrica e a produção do gás, uma vez que as redes de GN não são diretamente supervisionadas e controladas por centros de operação de sistemas de energia. Nesse contexto, a aplicação de técnicas de aprendizado de máquina automática fornece uma abordagem promissora para definir variáveis de controle e seus limites de segurança em operação das RIEEGNs.

Nos últimos anos, a técnica de mineração de dados denominada *Árvore de Decisão* (AD) tem sido muito aplicada na área de sistemas de potência para fins de avaliação da segurança e controle preventivo, como pode-se comprovar nos trabalhos de Diao et al (2010), Genc et al (2010) e Liu et al (2014).

Particularmente, a AD tem como principal vantagem a facilidade de interpretação dos conhecimentos adquiridos a partir do banco de dados. A AD difere de técnicas tradicionais pelo fato de encontrar os atributos críticos e seus limiares diretamente de um banco de dados usando simulações off-line. Os limiares obtidos não só ajudam a construir um modelo preditivo, mas também geram regiões de segurança para os operadores do sistema adotarem ações de controle a fim de garantir a segurança de operação do sistema. Além disso, a AD reduz significativamente o conjunto de variáveis a serem utilizadas nas ações de controle preventivo, permitindo que os operadores fiquem mais focados nas variáveis realmente críticas das RIEEGNs relacionadas com a segurança da rede.

Outro aspecto marcante da AD é o fato de que ela apresenta uma descrição sistêmica em relação às variáveis críticas que afetam a segurança das RIEEGNs. O caráter sistêmico é importante, porque o conjunto de variáveis críticas, para cada configuração topológica das RIEEGNs, pode ser distribuído por vários pontos do sistema elétrico e da rede de GN, muitas vezes em lugares não seria necessariamente tão evidente para o operador.

1.2 REVISÃO DE LITERATURA

Nos últimos anos, a integração entre as redes de gás natural e os sistemas de potência aumentou significativamente em vários países do mundo, devido ao crescimento de instalações de unidades termelétricas a gás natural, principalmente de usinas de ciclo combinado (NATURAL GAS, 2016), (MME: Petróleo, Gás Natural e Combustíveis Renováveis 2016). Como consequência, as Redes Integradas de Energia Elétrica e Gás Natural (RIEEGNs) tornaram-se mais vulneráveis aos problemas de segurança operacional, haja vista as interações cada vez maiores entre o fornecimento de gás e os geradores elétricos (CORREA e SANCHEZ, 2014).

Desde o surgimento dos primeiros SIEEGNs até os dias atuais, o projeto das RIEEGNs foi modificado significativamente como resultado de novos requisitos tais como a segurança do sistema. Os primeiros estudos foram realizados a partir das análises individuais de cada rede. Segundo Goldberg et al (2005), o gás natural tem se mostrado um combustível bastante atrativo uma vez que ele se caracteriza por uma queima eficiente e limpa. Na atualidade os seguintes fatores apontam para o crescimento dos SIEEGNs: o melhor aproveitamento da capacidade de geração de energia pelo usuário, a segurança no fornecimento, maior eficiência na geração de energia, redução de gastos de transporte e distribuição, aproveitamento de um combustível pouco poluente e melhor adequação entre a oferta e demanda por energia.

Castillo e González (1996) propõem a análise de uma rede de distribuição de gás baseada em algoritmo genético, com o objetivo de encontrar a distribuição ótima de GN. Os resultados alcançados mostram uma rede com características seguras considerando os limites de velocidade do gás no interior do gasoduto.

Munoz et al (2003) propõem um modelo que integra as redes de gás natural e de energia elétrica, analisando a confiabilidade do sistema. A análise de confiabilidade foi realizada avaliando a potência máxima fornecida pelas termelétricas em cada período de tempo. Os autores verificaram que a potência máxima depende das condições da rede de gás no que se refere à produção de GN.

Goldberg et al (2005) abordam um caso do problema de sistemas de cogeração de mínimo custo. O objetivo dos autores é modelar e apresentar métodos de solução para o problema descrito e denominado como Problema da Determinação da Configuração de Custo Mínimo de Sistemas de Cogeração de Energia com base em Gás Natural. Nesse trabalho são apresentados algoritmos evolucionários para a solução dos modelos formulados, com abordagens desenvolvidas por meio de algoritmos genéticos, algoritmos meméticos e

algoritmos transgenéticos. Entretanto, não há uma avaliação de segurança dos sistemas.

Devido ao crescimento de termelétricas a gás natural houve um acréscimo no grau de interdependência entre as redes de energia elétrica e de gás natural. Uma vez que essas redes podem apresentar situações de contingências, o sistema de energia tende a exigir mais flexibilidade e confiabilidade do sistema de gás natural.

Alguns trabalhos da literatura apresentam métodos de *Fluxo de Potência e de Gás Natural com Restrições de Segurança* aplicados ao problema do despacho da geração termelétrica e da produção do gás em RIEEGNs. Uma formulação do problema da rede de gás com restrições de segurança com foco na operação curto-prazo de redes de gás-eletricidade é apresentada por Shahidehpour e Wiedman (2005). O impacto de falhas na rede de gás sobre as operações no despacho de energia elétrica é analisado. Entretanto, o modelo da rede de gás não é considerado na formulação.

Mello e Ohishi (2006) propõem um modelo integrado de geração termelétrica a partir do gás natural com restrições ao fornecimento de GN. De acordo com os autores, o despacho das usinas a gás natural afeta os fluxos de gás nos gasodutos, e, por outro lado, os requisitos operacionais da rede de gasoduto podem impor limites a geração termelétrica. Neste contexto, os modelos que integram o funcionamento destes dois sistemas são importantes para uma operação econômica e segura. Este trabalho propõe um modelo que considera um conjunto de usinas de gás natural fornecidos por uma rede de gasoduto em que o objetivo é maximizar a geração de energia. Além disso, os requisitos dos sistemas, como limites de geração de energia elétrica, limites de pressão e do fluxo de gás natural são incorporados ao modelo.

No trabalho apresentado por Tao e Shahidehpour (2008), o modelo da rede de gás é integrado ao modelo tradicional com restrições de segurança para avaliar o impacto da interdependência das redes de gás e de energia elétrica na segurança dos sistemas de potência. Contudo, esforços para restaurar a segurança dos sistemas de potência não são considerados.

Shahidehpour et al (2009), apresentam um modelo integrado, no qual o transporte de gás natural pode afetar a segurança e o custo de energia elétrica. Do ponto de vista da segurança do sistema, as perdas de pressão e as contingências nos gasodutos são as principais causas no abastecimento de GN, produzindo um drástico aumento nos custos operacionais e comprometendo a segurança no sistema elétrico. O modelo integrado proposto por esses autores minimiza os custos operacionais do sistema de energia elétrica, levando em consideração as restrições de transmissão de energia elétrica, as restrições do transporte de GN e as restrições de contratos de GN.

Liu e Li (2009), apresentam uma metodologia para solução do vínculo de segurança com restrições da rede de gás natural. Os resultados mostram que contingências na rede de gás natural podem impactar significativamente na segurança de sistemas de potência. Contudo, nenhum procedimento de ação preventiva para restaurar a segurança é proposto em caso de prováveis contingências.

Martínez e Fuerte (2012) apresentam uma rede integrada de energia elétrica e de gás natural na qual destacam que a operação segura do sistema não depende apenas da disponibilidade do GN e do desempenho das instalações de geração e transmissão de energia elétrica. A segurança, segundo os autores, está vinculada na interdependência das infraestruturas utilizadas para produzir, transportar e armazenar o GN e as maneiras de transformá-lo em energia elétrica. Tradicionalmente, essas estruturas foram projetadas para operar separadamente e a simples integração entre as redes afeta a segurança e os custos de operação. Martínez e Fuerte (2012), aplicam um conjunto de equações algébricas não lineares para representar as redes integradas e solucionam o sistema de equações utilizando o método de Newton sob condições operacionais pré-estabelecidas.

Uma nova formulação de programação linear inteira mista foi utilizada, para a resolução do fluxo ótimo de energia e gás com restrições de segurança, no trabalho de Correa e Sanchez (2014). Para tanto, uma metodologia baseada no cálculo de fatores lineares de sensibilidade é proposta para ajustar as variáveis de controle das redes integradas de gás natural e de energia elétrica, de maneira ótima e rápida, de tal forma que $n-1$ contingências não resultem em violações. Vale ressaltar que o trabalho de Correa e Sanchez (2014) usa uma abordagem matemática clássica para comportar um grande conjunto de variáveis de controle, e assim, restabelecer a segurança das RIEEGNs.

Midthun et al (2015) avaliam a segurança da produção de GN utilizando a geração de cenários ótimos. A principal contribuição deste artigo é uma discussão sobre várias maneiras de utilizar uma abordagem de análise de cenários combinado como uma ferramenta de otimização de redes de GN. O trabalho se beneficiou de uma ferramenta de otimização chamada *GasOpt*, que tem sido muito utilizada em redes de GN na Europa. A ferramenta utiliza modelos de otimização de fluxo de GN que inclui modelagem das relações de pressão e fluxo em gasodutos com a capacidade de produção de gás flexível, com e sem a utilização de compressores.

Ficco et al (2015) publicaram um artigo que destaca a análise de incerteza das medições na rede de transmissão de GN. Segundo os autores, a precisão das medidas é essencial para a segurança e gestão do equilíbrio físico das redes de GN, bem como para a

correção da variação do consumo. Para alcançarem o objetivo, um modelo para a estimar as incertezas nas medições de gás natural foi desenvolvido e validado. O nível de incerteza encontrado foi então aplicado na rede de GN para balancear a precisão dos resultados obtidos.

1.3 OBJETIVO DA TESE

Estabelecer um método de despacho seguro entre Redes Integradas de Energia Elétrica e de Gás Natural, fundamentado em regras instituídas em Árvore de Decisão.

1.4 CONTRIBUIÇÃO DA TESE

A contribuição primária dessa tese de doutorado consiste de um método de despacho com restrições de segurança aplicado às RIEEGNs considerando uma abordagem sistêmica baseada no cálculo de fluxo de carga e de gás natural ótimo e em regiões de segurança definidas por regras de AD. Nesse trabalho, as regiões de segurança obtidas pelas ADs são empregadas para fornecer diretrizes para a tomada de decisão do despacho da geração e da produção de gás natural.

Esta metodologia primeiro escolhe alguns atributos das RIEEGNs controláveis, selecionados de forma off-line e treinados pela AD como os únicos atributos a serem sintonizados durante o processo de controle preventivo. As ADs são então treinadas para identificar o limite de segurança com a finalidade de separar os casos seguros e inseguros. O limite obtido é finalmente utilizado como um guia para desenhar as estratégias de controle preventivo.

Entre as questões mencionadas, este trabalho propõe uma análise da interdependência entre as redes de energia e de gás natural em condições de contingências das unidades termogeradoras movidas a gás natural. O trabalho utiliza as RIEEGNs para expor o impacto das contingências em gasodutos sobre a segurança do sistema de energia elétrica. O trabalho considera que as características desse combustível como produção, fluxo nos gasodutos e pressão nos nós interferem no desempenho das termelétricas e conseqüentemente, nos sistemas de potência.

1.5 PUBLICAÇÃO EM PERIÓDICO

A presente tese de doutorado conta com uma publicação em periódico internacional: “Decision tree-based security dispatch application in integrated electric power and natural gas networks”, *Electric Power System Research*, Vol. 141, pp. 442-449, Dec 2016.

1.6 ORGANIZAÇÃO DA TESE

O desenvolvimento desta tese está estruturado em seis capítulos, distribuídos da seguinte forma:

Capítulo I: Introdução

A introdução contempla uma contextualização e a motivação da pesquisa. A revisão de literatura apresenta um histórico dos trabalhos relevantes a essa tese. E a proposta desse trabalho, com foco em estratégias de controle preventivo.

Capítulo II: O Gás Natural

É apresentado a definição do gás natural no contexto do Brasil. É realizada uma análise histórica que mostra a evolução do gás natural no Brasil e no mundo. Uma seção sobre a origem e composição do gás natural mostra o comportamento desse combustível e como se encontra na natureza. A participação do gás natural na matriz energética realça a importância desse combustível, principalmente, na geração de energia elétrica.

Capítulo III: Árvore de Decisão

Uma introdução sobre Algoritmos de Aprendizado de Máquina mostra as peculiaridades das variáveis que servem como atributos em uma Árvore de Decisão. A seção sobre Mineração de Dados apresenta uma análise dos dados, após a sua extração, com o objetivo de realizar atividades com o menor nível de risco. Em sequência, há uma seção que conceitua a Árvore de Decisão, mostrando os seus elementos, as suas regras, a sua estruturação, os critérios para a seleção de atributos, os métodos de poda e os algoritmos de indução.

Capítulo IV: Otimização

Nesse capítulo são apresentados detalhes sobre o Algoritmo Genético e como ele foi aplicado para encontrar o valor ótimo na região de segurança estimada pela Árvore de Decisão.

Capítulo V: Confiabilidade do SIEEGN

Esse capítulo mostra os elementos que qualificam um SIEEGN quanto ao seu nível de segurança. O capítulo apresenta três quesitos que fundamentam a confiabilidade do sistema integrado: adequação, segurança do sistema e a segurança do combustível. O capítulo também relata, de forma separada, a segurança de cada sistema. A análise separada da segurança dos sistemas torna a sua compreensão mais didática.

Capítulo VI: Método de Despacho Proposto

Esse capítulo apresenta a metodologia utilizada para, a partir da Árvore de Decisão, determinar os limites dos atributos a fim de construir uma região de segurança que define as regiões operacionais do sistema. Essa região de segurança será utilizada para o fluxo ótimo da geração termelétrica e da produção de gás natural de acordo com as regras estimadas pela Árvore de Decisão quando um conjunto de contingências críticas é identificado pelo módulo de avaliação de segurança.

Capítulo VII: Estudo de Caso

Os casos avaliados nesse trabalho utilizam uma rede de gás natural de 15 nós e o sistema teste de 118 barras do IEEE. São apresentados três casos com contingências críticas específicas, tomadas como exemplos para demonstrar o método de despacho seguro proposto para restaurar a segurança do sistema.

Capítulo VIII: Conclusões

São apresentadas as considerações finais e as sugestões para futuros trabalhos.

1.7 CONCLUSÃO DO CAPÍTULO I

Esse capítulo mostrou a importância da inclusão do gás natural no planejamento energético integrado de redes elétricas. A evolução dos trabalhos que avaliam a integração entre a rede de gás natural e os sistemas de potência comprova a relevância do tema desta tese que apresenta um método de despacho seguro, com restrições, considerando as regiões de segurança definidas por regras provenientes da Árvore de Decisão.

CAPÍTULO II – O GÁS NATURAL

2.1 DEFINIÇÃO

A definição formal do gás natural no Brasil é dada pela lei 9.478/97, conhecida como lei do petróleo, em seu Capítulo III, seção I, artigo 6º, item II:

“Gás natural é todo hidrocarboneto que permaneça em estado gasoso nas condições atmosféricas normais, extraído diretamente a partir de reservatórios petrolíferos ou gaseíferos, incluindo gases úmidos, secos, residuais e gases raros” (BRASIL, 1997).

Para Garcia (2002) o gás natural é um combustível fóssil, formado basicamente por metano, em uma proporção de 80 a 90 %, etano (de 5 a 10 %) e outros gases em menores proporções (propano, butano, pentano, hexano, gás carbônico, nitrogênio e gases raros).

Ainda segundo Garcia (2002) o gás natural pode ser encontrado no subsolo ou no fundo do mar, em depósitos naturais de rochas sedimentares, associado ou não ao petróleo. Sua origem é a decomposição de plantas e animais, resultado de um processo de milhões de anos.

2.2 ANÁLISE HISTÓRICA

Com uma existência na natureza há milhões de anos, o gás natural só começou a ser utilizado pelo homem há cerca de mil anos antes de Cristo. O primeiro uso registrado se deu na China, onde o gás era distribuído através de um sistema de bambus e utilizado principalmente na iluminação (GASMIG, 2016).

A descoberta de gás natural está registrada nos tempos antigos no Oriente Médio. Milhares de anos atrás, notou-se labaredas criadas pelo escoamento de gás natural inflamado por relâmpagos. Na Pérsia, Grécia e na Índia, as pessoas construíram templos em torno destas chamas eternas para suas práticas religiosas. Contudo, o valor energético de gás natural não foi reconhecido até aproximadamente 900 a.C., na China. Os chineses foram os primeiros que perfuraram o solo em busca do gás natural, em 211 a.C.. Na Europa, o gás natural era desconhecido, até ser descoberto na Grã-Bretanha em 1659, embora não tenha sido comercializado até cerca de 1790.

Caminhando pelas margens do Lago Maior, no norte da Itália, cientista italiano Alessandro Volta descobriu por acaso que, agitando as águas de um pântano com uma vareta, bolhas que exalavam um gás inflamável eram produzidas. O Ocidente descobria assim o gás natural. No século XVIII, foram definidas pela Ciência as características dos gases

inflamáveis e Alessandro Volta e outros cientistas identificaram o metano, principal elemento da constituição do gás natural.

A exploração comercial iniciou-se de fato em 1820, com a descoberta de uma jazida na Pensilvânia (EUA). Em 1821, em Fredonia, Estados Unidos, os moradores observaram bolhas de gás subindo para a superfície de um riacho. William Hart, considerado na América como o pai de gás natural, cavou o primeiro poço de gás natural na América do Norte (Speight, 1993). Em 1930, foram descobertas outras jazidas, incentivando pesquisas de melhores equipamentos de exploração e melhores tubos para transporte do gás. Outro grande avanço ocorreu como resultado de grandes jazidas descobertas no Mar do Norte em 1965 (GASMIG, 2016).

Historicamente, o gás natural foi descoberto em consequência da prospecção ao petróleo bruto. O gás natural era muitas vezes um subproduto indesejável. Com os reservatórios de gás natural sendo drenados no processo de perfuração, os operários paravam de perfurar para permitir que o gás escapasse livremente no ar. Agora, e particularmente após a escassez de petróleo bruto dos anos de 1970, o gás natural tornou-se uma importante fonte de energia no mundo.

Durante todo o século XIX, o gás natural foi usado quase exclusivamente como um insumo destinado a destinado à iluminação pública. Devido à falta de estruturas de transporte a sua utilização permaneceu local, pois era difícil o transporte de grandes quantidades de gás natural a longas distâncias. Houve uma mudança importante em 1890 com a invenção do gasoduto à prova de vazamentos, mesmo com os avanços tecnológicos em dutos, o transporte de gás natural a longa distância para clientes não se tornou viável até 1920. Além disso, foi somente após a II Guerra Mundial que o uso de gás natural cresceu rapidamente por causa do desenvolvimento das redes de gasodutos e sistemas de armazenamento.

Como a evolução da sociedade é marcada por um consumo crescente de energia, a utilização do gás natural vem se adequando à demanda energética mundial. Cook (1971) apresentou a evolução do consumo de energia por pessoa desde milhares de anos atrás, nas sociedades primitivas, quando a necessidade energética era da ordem de 8.000 kJ/dia (o equivalente a 0,2 litros de petróleo), até chegar próximo a 1.000.000 kJ/dia (equivalente a 25 litros de petróleo), no fim do século XX. A Figura 1 apresenta esta evolução e uma estimativa.

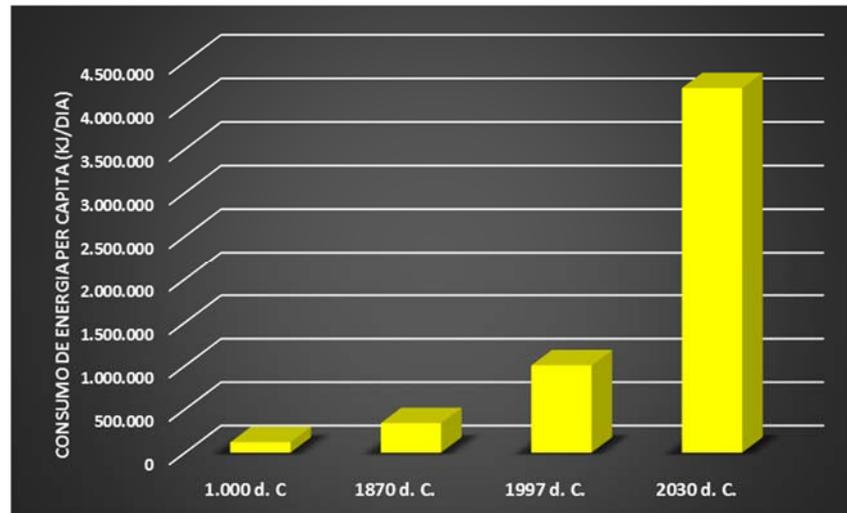


Figura 01: Evolução do consumo de energia.
Fonte: Cook, 1971.

Por volta do ano 1.000 d.C. iniciou-se a utilização da água, do vento, do carvão e do gás natural como fonte de energia. A revolução industrial ocorrida na segunda metade do século XIX ampliou radicalmente o consumo de energia. A utilização de tais fontes de energia, aumentou de maneira significativa e o uso de máquinas a vapor e da eletricidade cresceu rapidamente. Nas sociedades industrializadas atuais a economia depende fortemente do petróleo, de seus derivados e da eletricidade. Grande parte do consumo concentra-se na indústria e nos transportes.

2.3 PANORAMA NO BRASIL

Para Monteiro e Silva (2010) a história do gás canalizado iniciou-se no Brasil no século XIX desenvolvendo-se em ritmo satisfatório até a primeira metade do século XX. Porém, entre as décadas de 1950 e 1990, o setor entrou em estagnação, sendo que em muitos locais houve retrocesso. Apesar de, na primeira metade do século passado, existirem redes de distribuição de gás canalizado nas cidades de Porto Alegre, Salvador, Taubaté, Santos, Belém e Recife, as únicas companhias distribuidoras realmente atuantes localizavam-se nas cidades do Rio de Janeiro e São Paulo.

Ainda de acordo com Monteiro e Silva (2010) a distribuição de gás canalizado no Rio de Janeiro ocorreu por intermédio do Barão de Mauá, que em 1851 assinou contrato para iluminação a gás na cidade. Já em São Paulo a distribuição de gás iniciou-se em 1872 quando a *San Paulo Gas Company* recebeu autorização do Império para a prestação de serviços de distribuição de gás canalizado. Ao longo dos anos diversos tipos de gases foram distribuídos

pelas companhias do Rio de Janeiro e de São Paulo, sendo que o gás de carvão foi distribuído até 1972. De 1972 a 1989 o gás manufaturado de nafta era distribuído, sendo que a distribuição do gás natural teve seu princípio no início dos anos 1990.

Em 1996, a Petrobras assinou contrato de compra e venda de gás com a Bolívia. O volume inicial de importação foi de 4,1 milhões de m³ por dia, sendo que o volume importado atualmente é de 30 milhões de m³ por dia (MONTEIRO e SILVA, 2010).

2.4 ORIGEM E COMPOSIÇÃO DO GÁS NATURAL

O gás natural existe na natureza em reservatórios sob pressão de rochas na crosta Terra, formado por um conjunto de hidrocarbonetos dissolvidos em água ou isolado dela. Este gás é produzido nos reservatório de forma semelhante ao óleo bruto. O gás natural é formado pela degradação de matéria orgânica acumulada nos últimos milhões de anos. Os dois principais mecanismos responsáveis por esta degradação são biogênicos e termogênicos (ROJEV et al., 1997).

O gás biogênico é formado em menores profundidades e baixas temperaturas de decomposição bacteriana anaeróbia de matéria orgânica sedimentar. Em contraste, o gás termogênico é formado em maiores profundidades por o craqueamento térmico da matéria orgânica sedimentar. O gás biogênico é composto quase inteiramente por metano. Em contraste, o gás termogênico contém significativas concentrações de etano, propano, butanos, e outros hidrocarbonetos .

A informação que classifica o tipo de gás natural, entre gás biogênico ou gás termogênico, tem implicações críticas na sua utilização. O principal constituinte do gás natural é o metano, os outros componentes são hidrocarbonetos parafínicos, tais como etano, propano e butanos.

O comportamento de fase do gás natural é analisado pelo conjunto pressão-temperatura, que determina se o fluxo de gás natural, a uma dada pressão e temperatura, consiste de uma única fase gasosa ou de duas fases: gás e líquido. Dessa forma, o reconhecimento do tipo de gás é fundamental à aplicação em que ele será destinado.

O gás natural é considerado "seco" quando é, praticamente, metano puro, pois é possível ter a maioria dos outros hidrocarbonetos associados removidos. Quando outros hidrocarbonetos estão presentes, o gás natural é " molhado ".

A composição do gás natural varia dependendo do campo, formação, ou reservatório a partir do qual ele é extraído. Uma vez que a composição de gás natural nunca é constante,

existem métodos de ensaio padrão pelo qual a composição do gás natural pode ser determinada e, portanto, preparado ao seu uso. Tabela 1 apresenta a composição típica de gás natural antes de ser refinado.

Tabela 1: Composição do Gás Natural

Nome	Fórmula	Volume (%)
Metano	CH ₄	>85
Etano	C ₂ H ₆	3 – 8
Propano	C ₃ H ₈	1 – 2
Butano	C ₄ H ₁₀	<1
Pentano	C ₅ H ₁₂	<1
Dióxido de Carbono	CO ₂	1 – 2
Sulfato de Hidrogênio	H ₂ S	<1
Nitrogênio	N ₂	1 – 5
Hélio	He	<0,5

Fonte: ANP, 2016.

Os reservatórios de gás natural, mostrados na Figura 02, são compostos por rochas porosas que aprisionam gás e petróleo. Dependendo da relação existente entre o gás e o petróleo no poço, classifica-se o gás quanto ao seu estado: gás não-associado ou gás associado.

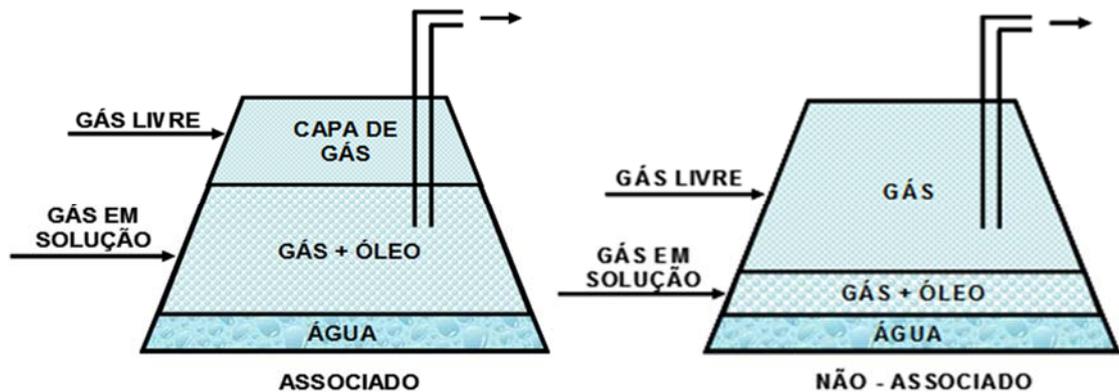


Figura 02: Formas de ocorrência do gás natural
Fonte: Tavares e Mendonça, 2013.

2.5 GÁS NÃO-ASSOCIADO

O gás não-associado é produzido a partir de formações geológicas que tipicamente não contêm muito hidrocarbonetos fervente. O gás não-associado contém gases que não são hidrocarbonetos, tais como o dióxido de carbono e sulfureto de hidrogênio.

O gás não-associado é diretamente monitorado na produção, sendo avaliado durante a sua passagem pelas válvulas de controle. O gás flui para cima do poço sob a sua própria energia, através das válvulas de controle de saída do poço, e ao longo da linha de fluxo, dirige-se à estação de tratamento. O tratamento requer que a temperatura do gás deva ser reduzida a um ponto dependente da pressão. Dentro da tubulação todos os líquidos que existem à temperatura oleoduto condensam e então são removidos.

Segundo Monteiro e Silva (2010), o gás não-associado está, no reservatório, livre ou em presença de quantidades muito pequenas de petróleo.

2.6 GÁS ASSOCIADO

O gás associado é produzido durante a produção de petróleo bruto, sendo que este petróleo não pode ser produzido sem que haja a produção algum gás associado, pois a solução que sai do poço com a pressão é reduzida no caminho para a superfície. O gás natural é concebido de reservatórios de poços de petróleo bruto, e processos de separação são usados para minimizar a produção de gás associado, a fim de reter o máximo energia no reservatório e, assim, aumentar a produção do petróleo em bruto.

Para Monteiro e Silva (2010) o gás associado é aquele que, no reservatório, está dissolvido no óleo ou sobre o óleo na forma de uma capa de gás. Nesse caso, a produção do gás é vinculada a produção de petróleo, sendo o gás utilizado para ampliação da produção de petróleo ou consumida na geração de energia na unidade de produção. A Figura 03 apresenta um esquema da localização do gás natural, associado e não-associado, na natureza. A capa que aprisiona o gás, chamada de lacre, é fundamental para detectá-lo.

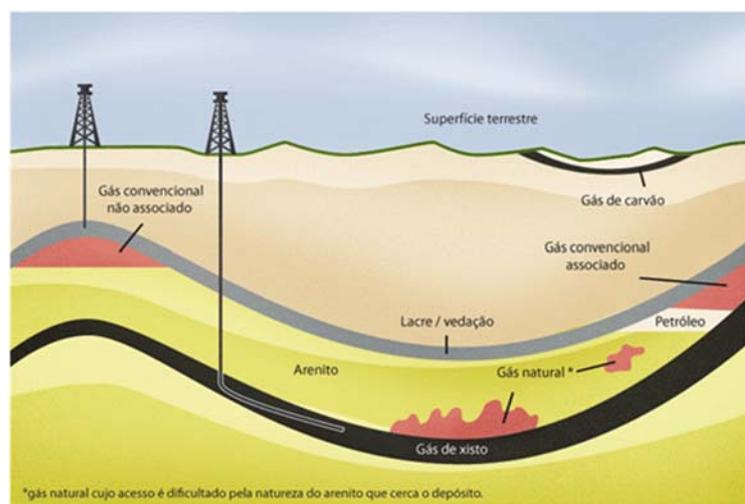


Figura 03: Esquema da geologia dos recursos de gás natural
Fonte: Tavares e Mendonça, 2013.

2.7 COMPORTAMENTO DO GÁS NATURAL

Na exploração de petróleo e gás o interesse econômico é somente na produção de hidrocarbonetos. Porém, na extração dos produtos dos poços, sejam eles terrestres ou marítimos, além dos hidrocarbonetos (gás e óleo), também são obtidas impurezas (água, CO, compostos ácidos, areia, dentre outros) que devem ser retirados (THOMAS, 2001). Sendo assim, após a etapa de exploração e produção o gás natural passa por diversos processos com o objetivo de separar o óleo, do gás e da água e condicioná-lo aos parâmetros de qualidade exigidos. Estas etapas são denominadas de condicionamento e processamento e são realizadas nas Unidades de Processamento de Gás Natural (UPGN) (MONTEIRO e SILVA, 2010).

Segundo Thomas (2001) o condicionamento é um conjunto de processos físicos e químicos que visam remover ou reduzir impurezas e contaminantes presentes no gás, de modo a atender as especificações de mercado, segurança, transporte ou posterior processamento. As principais especificações que são observadas no condicionamento do gás natural são: teores de compostos de enxofre, de dióxido de carbono e de água, ponto de bolha e poder calorífico. Nesta etapa dois processos são fundamentais:

a) Desidratação: são processos de absorção ou adsorção para a retirada de água presente no gás;

b) Dessulfurização: são processos de absorção química ou física para a retirada de gases ácidos (CO e compostos de enxofre).

Conforme Mokhatab et al (2006), o comportamento de fase de gás natural é um gráfico da *Pressão X Temperatura* que determina se o fluxo de gás natural, a uma dada pressão e temperatura, consiste de uma única fase gasosa ou de duas fases: gás e líquido. O comportamento do gás natural é exibido em um diagrama de fase, conforme mostrado na Figura 04.

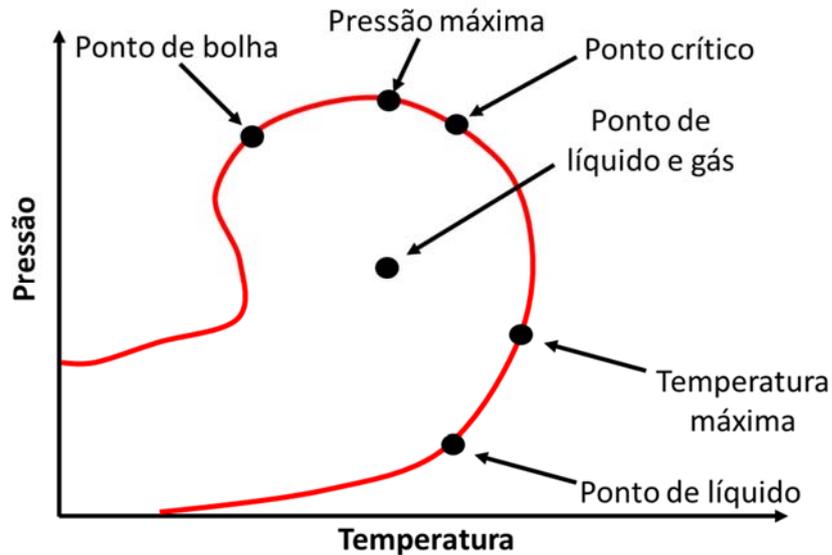


Figura 04: Representação do comportamento do gás natural
 Fonte: Mokhatab et al, 2006.

O comportamento de fase de gás natural é uma função da composição da mistura de gás e é fortemente influenciada pela concentração dos mais pesados hidrocarbonetos, especialmente C_6^+ . A presença de hidrocarbonetos mais pesados irá aumentar a estimativa do laço da fase do gás e falta desses hidrocarbonetos torna mais difícil essa predição.

A qualidade do gás natural comercializado no Brasil, seja de origem nacional ou de origem importada, é regulamentada pela Agência Nacional de Petróleo, Gás Natural e Biocombustíveis, através da resolução ANP nº 16, de 17 de junho de 2008. Os limites impostos pela resolução da ANP estão apresentados na Tabela 2.

De acordo com a Gasmig (2016) o gás natural possui seis características importantes:

- O gás natural não é tóxico e sua inalação acidental pode causar asfixia;
- Sua densidade, em relação ao ar, é igual a 0,6. Isto significa que ele é mais leve que o ar e em caso de vazamento se dissipa na atmosfera naturalmente;
- O gás é incolor e inodoro em seu estado natural. O cheiro característico (mercaptana) é inserido em sua composição para identificar eventuais vazamentos;
- O gás natural é inflamável apenas quando atinge temperaturas superiores a 622°C , o que dificulta muito qualquer tipo de queima acidental;
- O gás natural não é explosivo, apesar de ser inflamável. Se houver combustão, ele manterá a chama sem explosões até que seu suprimento se esgote;
- De todos os combustíveis utilizados atualmente, o gás natural é um dos menos poluentes, dispensando por isso equipamentos de filtragem dos gases de combustão.

Tabela 2 – Resolução ANP 16 - 17/06/2008

CARACTERÍSTICA	UNIDADE	LIMITE		
		Norte	Nordeste	Centro-Oeste, Sudeste e Sul
Poder calorífico superior	kJ/ m ³	34.000 a 38.400	35.000 a 43.000	
	kWh/m ³	9,47 a 10,67	9,72 a 11,94	
Índice de Wobbe	kJ/m ³	40.500 a 45.000	46.500 a 53.500	
Número de metano, mín.		60	65	
Metano, mín.	% mol.	68	85	
Etano, máx.	% mol.	12	12	
Propano, máx.	% mol.	3	6	
Butanos e mais pesados, máx.	% mol.	1,5	3	
Oxigênio, máx.	% mol.	0,8	0,5	
Inertes (N ₂ +CO ₂), máx.	% mol.	18	8	6
CO ₂ , máx.	% mol.	3		
Enxofre Total, máx.	mg/m ³	70		
Gás Sulfídrico (H ₂ S), máx.	mg/m ³	10	13	10
Ponto de bolha de água a 1 atm, máx.	°C	-39	-39	-45
Ponto de bolha de hidrocarbonetos a 4,5 MPa, máx.	°C	15	15	0
Mercúrio, média.	µg/m ³	38,5 kg/ano para cada 10 ⁶ MWh		

Fonte: ANP, 2016.

Ainda de acordo com a Gasmig (2016) as principais características físicas e químicas do gás natural que terão consequências sobre sua potencialidade energética são:

- a) Poder calorífico superior (PCS): quantidade de energia liberada na forma de calor, na combustão completa de uma quantidade definida de gás com o ar à pressão constante. Os produtos de combustão são considerados na mesma temperatura dos reagentes. O valor típico do PCS, para o gás natural, é de 9.400 kcal/m³ (VAZ et al, 2008);

- b) Limite de inflamabilidade: proporção entre combustível e ar na qual a combustão é possível. No caso do gás natural o limite de inflamabilidade varia entre 5 – 15 % em volume (VAZ et al, 2008);
- c) Temperatura de ignição espontânea: 622 °C;
- d) Velocidade de chama: é definida como a velocidade, em relação ao gás não queimado, na qual uma chama adiabática normal propaga-se através de uma mistura homogênea de gases. A velocidade de chama do gás natural varia entre 35 a 50 cm/s (GREEN e PERRY, 2008);
- e) Temperatura de chama: é a temperatura da zona onde ocorre a reação entre o combustível e o oxidante, geralmente o ar. A temperatura da chama do gás natural é de 1.945°C quando a reação ocorre com ar e 2.810 °C quando ocorre com oxigênio (GREEN e PERRY, 2008);
- f) Ponto de ebulição: -162 °C;
- g) Ponto de Fulgor: é definido como a menor temperatura na qual o combustível libera vapor em quantidade suficiente para formar uma mistura inflamável por uma fonte externa de calor. Esta temperatura para o gás natural é de – 189 °C (GARCIA, 2002);
- h) Massa específica: 0,766 kg/m³ considerando temperatura de 20 °C e pressão de 1 atm.

2.8 PARTICIPAÇÃO NA MATRIZ ENERGÉTICA

Várias organizações elaboram modelos sobre a utilização de combustíveis e fazem estimativas para os próximos anos do mercado energético. Dentre estas organizações, podem-se destacar como referências no assunto as publicações da *British Petroleum – BP*, e da Empresa de Pesquisa Energética – EPE. A BP elabora anualmente o *BP Statistical Review of World Energy* que apresenta diversas projeções sobre o consumo de combustíveis no mundo.

A EPE, órgão ligado ao Ministério de Minas e Energia, publica anualmente o Balanço Energético Nacional – BEN. Estas organizações estimam um grande aumento na participação do gás natural na matriz energética, crescimento superior ao observado no petróleo e no carvão mineral, que são, no momento, as duas principais fontes de energia primária.

O Balanço Energético Nacional – BEN 2015 (EPE, 2015) demonstra que no ano base de 2014 o gás natural representava cerca de 10% da matriz energética, sendo que a produção média diária, em 2014, foi da ordem de 87,1 milhões de m³, um aumento de 13% em relação a 2013, enquanto que, em média, 46,8 milhões de m³ foram importados diariamente, sobretudo

da Bolívia. O consumo total brasileiro, portanto, aproxima-se dos 100 milhões de m³ diários, com aumento do consumo termelétrico. Neste valor está incluído tanto o gás fornecido para o mercado consumidor (indústrias, geração elétrica, veículos e mercado residencial) quanto o gás utilizado pela Petrobras na exploração e produção de petróleo e derivados.

Dados publicados no BEN 2015 (EPE, 2015), ano base 2014, mostram um crescimento superior a 8,3% no consumo de gás natural, frente a um crescimento de 0,1% do PIB e 3,1% na oferta interna de energia. Demonstrando, desta forma, um significativo crescimento da participação do gás natural na matriz energética nacional, principalmente pelo aumento da geração termelétrica.

Para que haja crescimento da economia mundial significa que é necessário o aumento na oferta de mais energia. O consumo de energia tenderá a aumentar entorno de 34% entre 2014 e 2030. As previsões da British Petroleum - BP (2016) sobre o consumo mundial de energia, para o período, indicam que o consumo de gás natural será o que apresentará maior crescimento dentre os combustíveis fósseis. Enquanto a demanda mundial por gás natural crescerá na ordem de 2% ao ano, a procura por petróleo e seus derivados deverá crescer menos de 1% ao ano.

A Figura 05 apresenta a evolução da participação dos combustíveis na matriz energética mundial. O gás natural, em conjunto com as energias renováveis, deverá apresentar incremento na participação, ao passo que o petróleo e o carvão deverão sofrer grandes reduções.

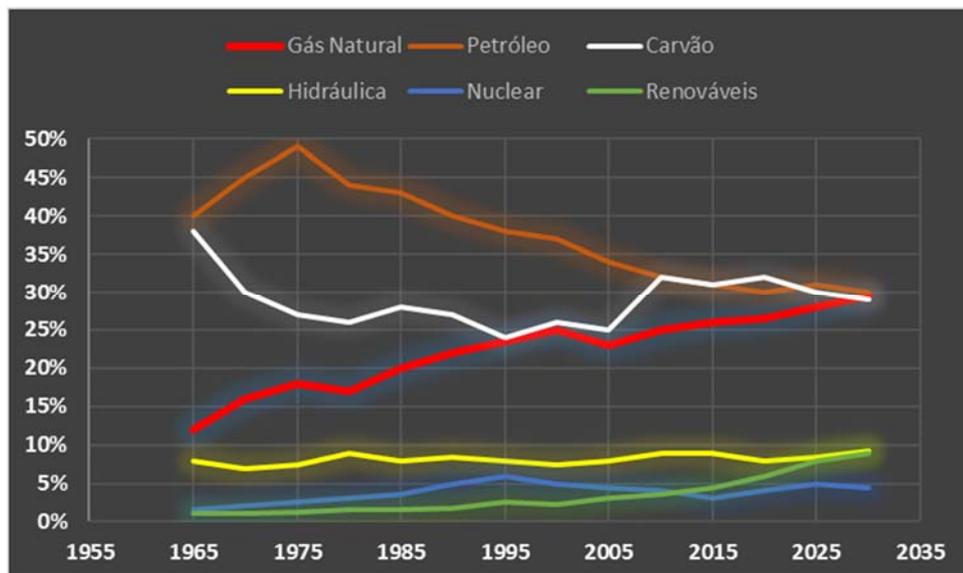


Figura 05: Participação dos combustíveis na matriz energética mundial
Fonte: BP, 2016.

Para o Brasil as expectativas de crescimento do consumo de gás natural são mais fortes. Ainda de acordo com a BP (2016) o crescimento na produção de todos os tipos de energia, até

2030, deverá ser superior a 80%. Por sua vez, a produção de gás natural deverá crescer mais de 140%, sendo superior ao crescimento da produção de petróleo, previsto em 87%. Já a demanda pelo gás natural deverá ter um incremento superior a 130%, sendo a maior alta dentre os combustíveis fósseis. A expansão da infraestrutura de transporte e distribuição de gás será o principal entrave para que o país possa concretizar as previsões.

A legislação brasileira, desde a década de 1990, aponta para o incremento da participação do gás natural na matriz energética nacional. A Lei nº 9478, de 06 de agosto de 1997, que dispõe sobre a Política Energética Nacional, as atividades relativas ao monopólio do petróleo, institui o Conselho Nacional de Política Energética e a Agência Nacional do Petróleo, prevê logo no Art. 1º que as políticas nacionais de utilização de energia deverão incrementar, em bases econômicas, a utilização do gás natural (BRASIL, 1997).

O Conselho Nacional de Política Energética, criado através do Decreto nº 3520, de 21 de junho de 2000, também prevê em seu Art. 1º que a Política Energética Brasileira deverá, dentre outros princípios, incrementar a utilização do gás natural (BRASIL, 2000).

O crescimento da oferta e da demanda do gás natural no Brasil deverá passar, necessariamente, pelo incremento da infraestrutura de transporte e distribuição deste combustível. A Figura 06 apresenta os gasodutos de transporte em operação, em implantação e em estudo no Brasil e nos países vizinhos.



Figura 06: Gasodutos brasileiros e dos países vizinhos

Fonte: Gasnet, 2016.

Observa-se a grande concentração dos gasodutos de transporte, atualmente em operação, na zona litorânea brasileira, enquanto que boa parte do interior do Brasil, com grandes polos econômicos e industriais, não possui infraestrutura de transporte, e conseqüentemente de distribuição, deste combustível.

Uma das principais etapas da cadeia de suprimento do gás natural é, justamente, a distribuição, função que no caso brasileiro, é efetuada pelas companhias distribuidoras de cada estado, que possuem concessão para a realização de tais serviços. A transferência de propriedade das empresas transportadoras para as companhias distribuidoras ocorre nas estações de transferência de custódia, denominadas *City Gates*. No *City Gate* equipamentos específicos realizam medições e regulam a pressão de entrega do gás. Neste ponto também é realizada a odorização do gás, conforme exigência da legislação (MONTEIRO e SILVA, 2010).

Dentre os vários aspectos positivos da utilização do gás natural destaca-se que entre os combustíveis fósseis ele é o que apresenta menor potencial de poluição. Reduzindo a quantidade de CO₂, materiais particulados e compostos de enxofre emitidos para a atmosfera.

2.9 APLICAÇÃO ENERGÉTICA DO GÁS NATURAL

De acordo com Esteves e Matz (2013) o gás natural possui uma ampla gama de aplicações podendo ser utilizado para fins energéticos, como combustível, ou não energético, sendo utilizado em indústrias químicas ou na geração de energia elétrica. Para Shreve e Brink Jr. (1997) o gás natural, além de ser o gás combustível predominante mundialmente, também é utilizado como matéria-prima para diversas sínteses.

O gás de síntese obtido da reforma do gás natural pode ser utilizado para a produção de metanol, hidrogênio e outros compostos líquidos, que por sua vez produzem ácidos, combustíveis e amônia. A amônia é um componente fundamental para o mercado brasileiro, uma vez que serve de matéria-prima para a produção de fertilizantes nitrogenados, produto muito utilizado pela indústria agropecuária, sendo que mais de 70% da demanda nacional é importada (ESTEVEES e MATZ, 2013).

Como fonte de energia térmica, ou como combustível, o gás natural é utilizado nos setores energético, industrial, comercial, institucional (repartições públicas), residencial e automotivo (ESTEVEES e MATZ, 2013). A Tabela 3 apresenta o consumo do gás natural (em Mega tonelada equivalente de petróleo – Mtep) como combustível em diversos setores econômicos.

Tabela 3: Consumo de Gás Natural por setor em Mtep

Setor	Consumo em 2012 (Mtep)	Consumo em 2015 (Mtep)
Energético	5700	6112
Comercial/Público	270	158
Residencial	336	312
Agropecuário	0	0
Transporte Rodoviário	1942	1553
Industrial	11192	9947
Cimento	34	12
Ferro-gusa e Aço	1072	1223
Ferroligas	3	6
Mineração e Pelotização	765	657
Química	2740	2222
Não ferroso	873	593
Têxtil	360	215
Alimentos e Bebidas	738	834
Papel e Celulose	812	805
Cerâmica	1493	1324
Outras Industrias	2300	2057

Fonte: EPE, 2016.

No setor energético, o gás natural é utilizado em motores e turbinas, para geração de energia elétrica. Em processos de geração distribuída ou cogeração, o gás natural é fonte energética para a produção de energia elétrica e vapor (ESTEVEES e MATZ, 2013). Nos dados da Tabela 3 não está contabilizado o volume de gás fornecido para usinas termelétricas. O BEN (EPE, 2016) apresenta o gás natural como a segunda principal fonte de combustível para o setor energético, ficando atrás apenas do bagaço de cana.

Na indústria o gás natural possui inúmeras aplicações, podendo ser utilizado em motores, turbinas e caldeiras, para geração de eletricidade e vapor; em fornos, de cerâmica, de tratamento térmico; ou em aquecedores, de fluidos térmicos, por exemplo. A utilização do gás natural na indústria pode ser dividida através da temperatura de trabalho. A maior aplicação do gás natural na indústria ocorre nos equipamentos de alta temperatura, nesta subdivisão está concentrado o maior consumo de gás natural industrial.

2.10 APLICAÇÃO DO GÁS NATURAL NA GERAÇÃO DE ENERGIA ELÉTRICA

Conforme as informações do BEN (EPE, 2016) o consumo de gás natural para geração e energia elétrica atingiu uma média diária de 45,90 milhões de m³ no ano de 2015, ao passo que em 2011 a média diária foi de 10,39 milhões de m³. Pode-se observar assim, que o volume de gás natural destinado à geração de energia elétrica teve um expressivo aumento, de 341,77 %, nesse período.

O Boletim Mensal de Acompanhamento da Indústria do Gás Natural (MME, 2016) mostra uma tendência de aumento no consumo de gás natural para geração termelétrica. A média diária observada em 2013 foi de 40,08 milhões de m³. No ano de 2015 a média diária cresceu para 45,90 milhões de m³. A escassez de chuvas registrada nos anos de 2014 e 2015 reduziu o nível dos reservatórios das usinas hidrelétricas, o que exigiu uma maior operação das usinas termelétricas, que entram assim como um fator de segurança energética.

A aplicação do gás natural para a geração de energia elétrica pode ser dividida em duas modalidades: geração termelétrica e cogeração.

A principal modalidade é a utilização do gás natural para geração exclusiva de eletricidade. Neste processo, o gás, em geral, serve como combustível para turbinas a gás. As turbinas a gás são máquinas complexas que comprimem o ar, misturando-o com o combustível com o objetivo de obter-se a combustão. Os gases gerados no processo de combustão possuem elevada temperatura e junto com o ar a alta pressão, provocam o movimento de turbinas conectadas a geradores elétricos. Desta forma, a energia térmica presente no combustível transforma-se em energia mecânica que logo a seguir é convertida em energia elétrica (ANEEL, 2008).

Ainda segundo a ANEEL (2008) o desenvolvimento das turbinas a gás é relativamente recente, após a segunda guerra mundial. Porém, o uso mais intenso destes equipamentos tem ocorrido somente nos últimos 20 a 30 anos, devido à maior oferta de gás natural e também a melhoria dos rendimentos termodinâmicos das máquinas (ANEEL, 2008). A Tabela 4 apresenta as principais usinas a gás natural em operação no Brasil.

Tabela 4: Centrais Termelétricas a Gás Natural em Operação no Brasil

Usina	Potência (MW)	Proprietário	Estado
Santa Cruz	1.000	Furnas (Eletrobras)	RJ
Campos	30	Furnas (Eletrobras)	RJ
Camaçari	131	Braskem (47% Petrobras)	BA
Modular de Campo Grande	206	Tractebel	MS
Uruguaiana	640	AES	RS

Cuiabá	529	Petrobras	MT
Sepé Tiaraju	249	Petrobras	RS
Barbosa Lima Sobrinho	386	Petrobras	RJ
Mário Lago	923	Petrobras	RJ
Termo Norte II	427	50% CS, 50% Termogas	RO
Juiz de Fora	87	Petrobras	MG
Termo Ceará	220	Petrobras	CE
Aureliano Chaves	226	Petrobras	MG
Termocabo	50	Eenergia	PE
Araucária	484	80% Copel; 20% Petrobras	PR
EnergyWorks Mogi	31	Energyworks	SP
Camaçari	360	CHESF (Eletrobras)	BA
Celpav IV	139	Fibria Celulose	SP
Rômulo Almeida	138	Petrobras	BA
Fortaleza	347	Endesa	CE
Luiz Carlos Prestes	386	Petrobras	MS
Celso Furtado	186	Petrobras	BA
Norte Fluminense	827	EDF	RJ
Termopernambuco	533	Neoenergia (49% Previ, 39% Iberdrola, 12% Banco do Brasil)	PE
Gov. Leonel Brizola	1.058	Petrobras	RJ
Fernando Gasparian	386	Petrobras	SP
Jaraquib	157	Breitener (94% Petrobras)	AM
Manauarab	85	Cia En. Manauara	AM
Jesus Soares Pereirac	323	Petrobras	RN
Euzébio Rochac	250	Petrobras	SP
Luiz Oscar Rodrigues de Melo	204	Linhares Geração	ES
Maranhão IV	338	70% ENEVA, 30% Petra	MA
Maranhão V	338	70% ENEVA, 30% Petra	MA
MC2 Nova Venécia 2	176	35% ENEVA (EoN), 35% JV, 30% Petra	MA
Parnaíba IV	56	35% ENEVA (EoN), 35% JV, 30% Petra	MA
Baixada Fluminense	530	Petrobras	RJ
Total	12.433		

Fonte: ANEEL, 2015.

A geração de energia elétrica, a partir do gás natural, em usinas termelétricas pode ocorrer de duas formas. A forma mais simples e de implementação mais barata denomina-se ciclo aberto (ou ciclo simples). Neste sistema os gases que saem da turbina são resfriados e liberados para a atmosfera por meio de uma chaminé (ANEEL, 2008). De acordo com o Plano Decenal de Expansão de Energia 2021 (EPE, 2012), no ciclo simples, a eficiência termodinâmica do processo é de 38,7 %, ou seja, de cada 100 unidades de combustível que entram na câmara de combustão aproximadamente 39 se transformam em energia elétrica, sendo o restante perdido.

De acordo com a MME (2016) a produção média das usinas termelétricas movidas a gás natural em operação no Brasil é de 12.439 MW.

A Figura 07 ilustra o diagrama de uma central termelétrica de ciclo aberto, com os seus principais componentes.

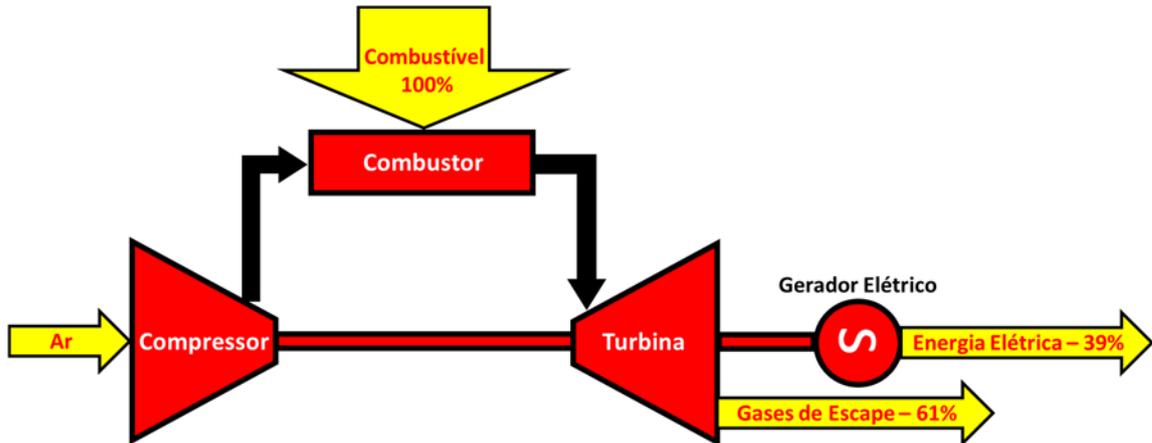


Figura 07: Diagrama esquemático do ciclo aberto
Fonte: Esteves e Matz, 2013.

No ciclo combinado, os gases que saem da turbina e que ainda estão em alta temperatura, são encaminhados para caldeiras onde produzem vapor em alta pressão. Este vapor é então direcionado para turbinas onde é expandido, provocando o movimento da turbina. Assim, a característica básica de termelétricas a ciclo combinado é a operação conjunta de turbinas movidas a gás e a vapor.

Esta tecnologia é recente, sendo aplicada principalmente a partir da década de 1980, e passa por processo de expansão em todo o mundo, inclusive no Brasil. Este sistema requer maiores investimentos, porém a eficiência do processo de geração é maior, chegando ao nível de 50% (ANEEL, 2008). A Figura 08 apresenta um diagrama com os principais componentes de uma central termelétrica com ciclo combinado.

Ainda de acordo com a ANEEL (2008), existe o processo de geração de energia elétrica, a partir do gás natural, simultaneamente com a geração de vapor e de energia térmica. Este processo, conhecido como cogeração, pode ser realizado com todos os combustíveis usados em usinas termelétricas – por exemplo, óleos, biomassa e carvão, além do gás natural.

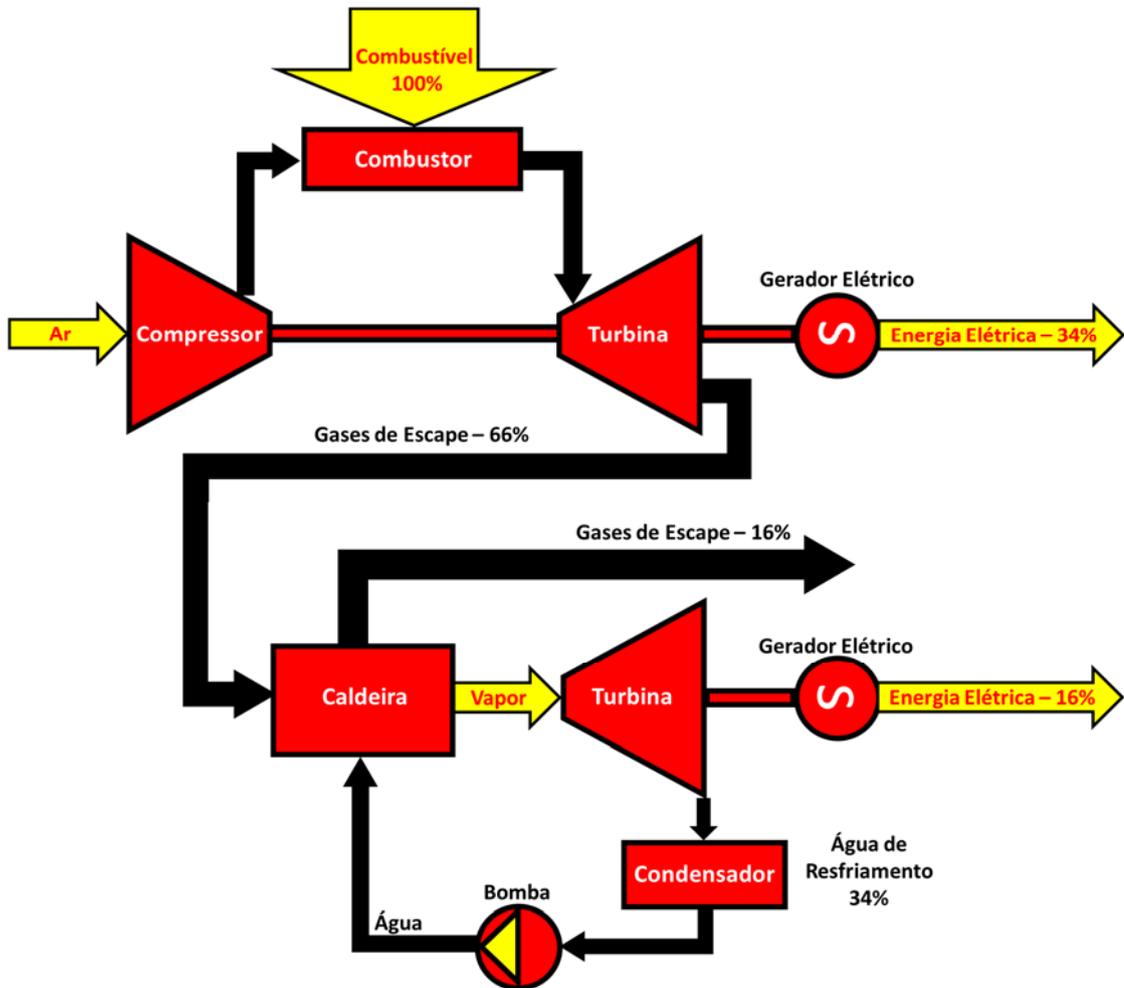


Figura 08: Diagrama esquemático do ciclo combinado
 Fonte: Esteves e Matz, 2013.

A Figura 09 apresenta os percentuais de aproveitamento energético em cada uma das etapas da cogeração.

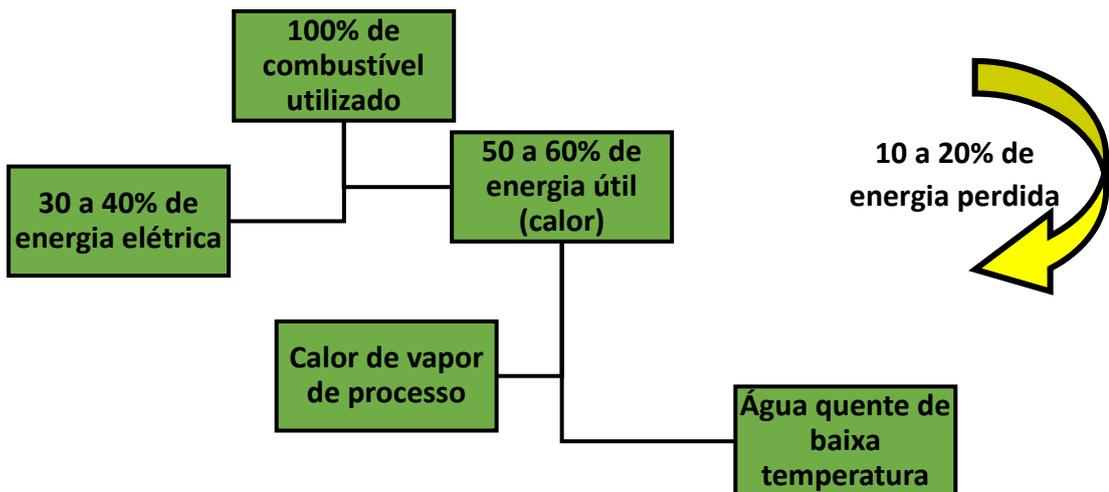


Figura 09: Rendimento energético do processo de cogeração
 Fonte: Monteiro e Silva, 2010

Na cogeração o calor gerado na produção da eletricidade por usinas em ciclo simples e que, se não utilizado, seria liberado na atmosfera é recuperado e destinado à produção de vapor, do ar quente ou da refrigeração.

Três fatores tornam o processo da cogeração bastante interessante para as aplicações industriais. O primeiro é a possibilidade de utilização da energia que naturalmente se perde no processo de geração da eletricidade. O segundo fator é a possibilidade de independência em relação ao suprimento fornecido pelas distribuidoras ou comercializadoras de energia elétrica. Finalmente, um terceiro é a redução do volume de gases lançados na atmosfera, o que pode ser um fator de competitividade no momento atual, em que os consumidores estão cada vez mais exigentes com relação ao impacto ambiental provocado pelos produtos que adquirem.

2.11 CONCLUSÃO DO CAPÍTULO II

O capítulo II apresentou a definição do gás natural e analisou historicamente a evolução da sua utilização. A origem e a composição desse combustível também foram estudadas, apresentando-se os tipos de gás natural avaliados pela geofísica e pela geoquímica. A participação na matriz energética, especialmente, na geração elétrica, são os itens que finalizam o capítulo II. Esses itens apresentam dados otimistas quanto à importância do gás natural na segurança do sistema elétrico.

CAPÍTULO III - ÁRVORE DE DECISÃO

3.1 INTRODUÇÃO

Algoritmos de Aprendizado de Máquina (AAMs) são muito utilizados na resolução de problemas de um universo de aplicações (MITCHELL, 1997). Esses algoritmos são principalmente utilizados em três casos:

- i. problemas de mineração de dados em que grandes conjuntos de dados podem apresentar inestimáveis harmonias implícitas que poderão ser revelados automaticamente, por exemplo, fornecer as variáveis críticas que afetam a segurança operacional de um sistema de potência;
- ii. controles mal assimilados em que humanos não possuem o conhecimento necessário para desenvolver algoritmos efetivos, por exemplo, diagnóstico de regiões de segurança apenas por imagens;
- iii. propriedade em que o programa deve se adaptar dinamicamente à mudanças do sistema, como ocorre nos centros de operação de sistemas elétricos de potência. Os AAMs envolvem diferentes áreas de conhecimento, tais como Inteligência Artificial (IA), Probabilidade e Estatística, Complexidade Computacional, Teoria da Informação, Psicologia, Neurobiologia e Teoria de Controle.

Diversos AAMs utilizam o conhecimento de IA e da Estatística para estruturação de modelos capazes de obter conhecimento a partir de um conjunto de dados. Esse conjunto de dados é formado por exemplos (ou casos) que são descritos por várias peculiaridades (atributos), e a cada um desses casos pode ser atribuída uma classe dentre um conjunto finito de classes possíveis.

As peculiaridades são variáveis observáveis e independentes, e seus domínios podem ser tanto contínuos (numéricos) como discretos (categóricos). A classe é uma variável dependente, e seu valor é determinado a partir dos valores das respectivas variáveis independentes.

Os casos desses conjuntos são chamados de dados de treinamento e a partir desses faz-se a inferência indutiva, que poderá gerar hipóteses verdadeiras ou não (ALPAYDIN, 2004). Todo algoritmo de aprendizado, ou algoritmo de indução, possui *bias* indutivo que é a preferência de uma hipótese sobre outra dentre as hipóteses que são igualmente consistente com os dados.

O aprendizado indutivo pode ser dividido em supervisionado e não-supervisionado. Atualmente, o aprendizado semi-supervisionado também atrai a atenção de grande parte da comunidade de AAMs (CHAPELLE et al., 2006).

No aprendizado supervisionado, os valores da classe (atributo classe) dos casos do conjunto de treinamento são conhecidos. Se o domínio do atributo classe for discreto, tem-se um problema de classificação. Caso o domínio do atributo classe for numérico, tem-se um problema de regressão ou aproximação de funções.

Classificação é a tarefa de aprender uma função alvo f que mapeia cada conjunto de atributos x , ou seja, cada caso, a uma das y classes pré-difinidas (TAN et al, 2005). Em um problema de classificação, os dados de entrada podem ser descritos pelo par ordenado (\bar{x}, y) , em que \bar{x} é um valor que representa os atributos preditivos, $\bar{x} = (x_1, x_2, x_3, \dots, x_n)$, e y é rótulo da classe à qual esse caso pertence.

A Tabela 5 apresenta um conjunto de dados para classificação do estado de segurança de um sistema de gasodutos. Nessa tabela, cada linha representa um caso do conjunto de dados e cada coluna um atributo desse caso. O atributo *Diagnóstico* é especial, pois apresenta o rótulo da classe para cada caso, ou seja, *Seguro* ou *Inseguro*.

Tabela 5: Conjunto de dados para o diagnóstico de um Sistema de Gasodutos

Caso	Produção de Gás Natural	Pressão nos nós produtores	Pressão nos nós não produtores	Fluxo de Gás Natural nos dutos	Diagnóstico do Sistema
C1	Baixa	Baixa	Baixa	Lento	Inseguro
C2	Normal	Normal	Baixa	Normal	Seguro
C3	Normal	Normal	Baixa	Lento	Seguro
C4	Elevada	Elevada	Normal	Intenso	Inseguro
C5	Normal	Normal	Normal	Normal	Seguro
C6	Elevada	Baixa	Normal	Lento	Inseguro

Fonte: Autor.

Técnicas de AAM têm sido utilizadas com êxito em problemas de classificação e cada uma faz uso de um algoritmo de aprendizado para construir o classificador (modelo) capaz de inferir a classe baseado nos atributos preditivos. Um desafio aos algoritmos de aprendizado é construir modelos que apresentem relevante capacidade de generalização, ou seja, consigam prever, com alta taxa de acerto, rótulos das classes para casos que não foram utilizados durante a construção do classificador (TAN et al, 2005).

Os classificadores possuem outra propriedade a ser observada: a interpretabilidade do conhecimento adquirido. Alguns classificadores têm uma representação interna que, normalmente, não se consegue interpretar facilmente, ou seja, torna-se difícil reconhecer as características do problema que o levaram a uma certa dedução de classe. Um exemplo desse tipo de classificador são as Redes Neurais Artificiais (RNAs). Os classificadores orientados a conhecimento, como a Árvore de Decisão (AD), criam estruturas simbólicas que são mais compreensíveis do que as RNAs (MONARD e BARANAUSKAS, 2003).

A Tabela 6 ilustra casos, cujos diagnósticos deverão ser estimados pelo classificador a partir dos casos fornecidos. A Tabela 5 foi utilizada como um conjunto de treinamento, no qual os rótulos de classes dos casos conhecidos serviram de treinamento para o algoritmo de aprendizado construir um modelo.

Tabela 6: Conjunto de dados para a dedução de diagnóstico de um Sistema de Gasodutos

Caso	Produção de Gás Natural	Pressão nos nós produtores	Pressão nos nós não produtores	Fluxo de Gás Natural nos dutos	Diagnóstico do Sistema
C7	Baixa	Elevada	Normal	Intenso	?
C8	Normal	Elevada	Baixa	Normal	?

Fonte: Autor.

A Figura 10 ilustra um diagrama do processo de indução de um classificador e posteriormente a sua utilização. Inicialmente o conjunto de treinamento, no qual os rótulos das classes dos exemplos são conhecidos, é utilizado por um algoritmo de aprendizado na construção do modelo. Conhecido o classificador ele pode ser aplicado para prever os rótulos das classes dos casos do conjunto teste, ou seja, dos casos cujas classes ainda não se conhece.

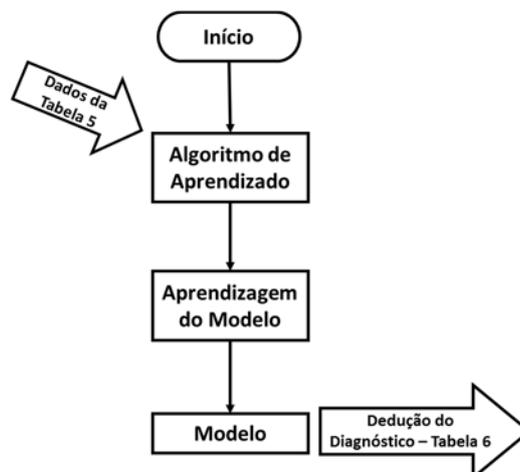


Figura 10: Diagrama do processo de indução de um classificador e sua utilização na dedução de novos casos.

Segundo Tan et al (2005), a classificação pode ser utilizada para as seguintes finalidades: modelagem descritiva e modelagem preditiva. Na descritiva, o modelo de classificação é utilizado como uma ferramenta para distinguir casos de diferentes classes. Como por exemplo, um operador do sistema de gás natural poderá aplicar o modelo de classificação a fim de identificar quais as principais causas de uma determinada contingência da rede de gasodutos. A partir disso, é possível chegar a conclusão que a rede de gás natural está com um nível de produção de gás superior ao nível de segurança exigido.

Quando há o interesse em análise descritiva é importante que o modelo de classificação seja de fácil interpretação, ou seja, que fique claro para o operador do sistema o porquê de um determinado caso pertencer à determinada classe.

Na modelagem preditiva, o modelo de classificação é utilizado para classificar casos cujas classes são desconhecidas, ou seja, casos que não foram utilizados na construção do modelo. Por exemplo: um operador de um sistema de gás natural poderá utilizar um conjunto de dados de um outro sistema de gasodutos, em que as classes (diagnósticos) já são conhecidas, a fim de construir um modelo de classificação a ser utilizado para diagnosticar contingências do seu sistema de gás.

Um modelo de classificação pode apresentar o seu conhecimento adquirido de várias formas, tais como, Regras de Decisão, Redes Bayesianas, Máquinas de Vetores de Suporte, Redes Neurais Artificiais e Árvores de Decisão.

As Árvores de Decisão (ADs) constituem uma técnica muito poderosa e amplamente utilizada em problemas de classificação. Uma das razões dessa técnica ser muito aplicada é pelo fato do conhecimento adquirido ser representado por meio de regras. Essas regras são expressas em linguagem natural, o que facilita o entendimento por parte de especialista humano.

3.2 MINERAÇÃO DE DADOS

Mineração de Dados (*Data Mining*) é um ramo da computação que teve início nos anos 80, quando os profissionais das empresas e organizações começaram a se preocupar com os grandes volumes de dados computacionais estocados e inutilizados dentro da empresa. Nesta época, *Data Mining* consistia essencialmente em extrair informação de gigantescas bases de dados da maneira mais automatizada possível.

Atualmente, *Data Mining* consiste, sobretudo na análise dos dados após a extração, buscando-se, por exemplo, levantar as necessidades reais e hipotéticas de cada especialista

realizar atividades com o menor nível de risco. Assim, uma empresa distribuidora de gás natural, por exemplo, tem um conjunto valioso de informações, pois ela conhece as atividades de cada um dos seus clientes: o quanto costuma consumir, qual o seu padrão de pressão, qual fluxo de gás requerido, etc. Para a empresa essas informações são extremamente úteis na estimação do limite demandado para cada cliente.

Os seguintes pontos são algumas das razões pelas quais o *Data Mining* vem se tornando necessário para uma boa gestão empresarial: (i) os volumes de dados são muito importantes para um tratamento que utilizaria somente técnicas clássicas de análise, (ii) o especialista final não precisar ser necessariamente um estatístico, (iii) a intensificação do tráfego de dados aumenta a possibilidade de acesso aos dados.

A Mineração de Dados ou *Data Mining* é o núcleo do processo de Descoberta de Conhecimento em Banco de Dados ou *Knowledge Discovery in Database* (KDD). A Descoberta de conhecimento em bancos de dados é o processo não trivial de identificar em dados padrões que sejam válidos, novos (previamente desconhecidos), potencialmente úteis e compreensíveis, visando melhorar o entendimento de um problema ou um procedimento de tomada de decisão.

Segundo Fayyad (1996), o processo de KDD é interativo, iterativo, cognitivo e exploratório, envolvendo vários passos com muitas decisões sendo tomadas pelo analista. KDD é um processo que consiste das seguintes etapas:

- i. Definição do tipo de conhecimento a descobrir, o que pressupõe uma compreensão do domínio da aplicação, bem como do tipo de decisão que tal conhecimento pode contribuir para melhorar.
- ii. Criação de um conjunto de dados alvo (*Selection*): selecionar um conjunto de dados, ou focar num subconjunto, onde a descoberta deve ser realizada.
- iii. Limpeza de dados e pré-processamento (*Preprocessing*): operações básicas, tais como, remoção de ruídos quando necessário, coleta da informação necessária para modelar ou estimar ruído, escolha de estratégias para manipular campos de dados ausentes, e formatação de dados de forma a adequá-los à ferramenta de mineração.
- iv. Redução de dados e projeção (*Transformation*): localização de características úteis para representar os dados dependendo do objetivo da tarefa, visando a redução do número de variáveis e/ou instâncias a serem consideradas para o conjunto de dados, bem como o enriquecimento semântico das informações.
- v. Mineração de dados (*Data Mining*): selecionar os métodos a serem utilizados para localizar padrões nos dados, seguida da efetiva busca por padrões de interesse numa

forma particular de representação ou conjunto de representações; busca pelo melhor ajuste dos parâmetros do algoritmo para a tarefa em questão.

- vi. Interpretação dos padrões minerados (*Interpretation/Evaluation*), com um possível retorno aos passos (i-vi) para posterior iteração.
- vii. Implantação do conhecimento descoberto (*Knowledge*): incorporar este conhecimento ao desempenho do sistema, ou documentá-lo e reportá-lo às partes interessadas.

3.3 ÁRVORE DE DECISÃO - CONCEITO

Árvore de Decisão (AD) é essencialmente um gráfico em forma de árvore, contendo as sentenças a serem tomadas e suas possíveis consequências (riscos). Uma AD é utilizada para criar um conjunto de estratégias para se alcançar determinado objetivo. Uma AD é um modelo preditivo, isto é, um mapeamento de observações sobre um item para conclusões sobre o seu valor-alvo. Cada nó interno corresponde uma variável; um arco para um nó-filho representar um possível valor daquela variável. Uma *folha* representa o valor previsto da variável-alvo, dadas as variáveis representadas no caminho até ela desde a *raiz*.

A Árvore de Decisão é um algoritmo de classificação de dados no ambiente chamado Mineração de Dados (*Data Mining*). Possui como uma importante característica a função de particionar de modo recursivo um conjunto complexo de dados, até que cada subconjunto, assim obtido, contenha apenas os casos de uma única classe possibilitando uma melhor análise. Após a construção da AD, os resultados obtidos, são formados por dados organizados de maneira simples e de fácil entendimento que servirão como importante ferramenta para tomada de decisão.

A Árvore de Decisão é uma estrutura que pode ser utilizada para dividir sucessivamente uma grande coleção de registros em conjuntos menores, por meio de uma simples regra de decisão. A cada divisão realizada, os dados são separados de acordo com características em comum até chegar a pontos indivisíveis, que representam as classes.

Cada nó da árvore representa um teste a ser realizado e os ramos definem um caminho para cada resposta desses testes. O nó *raiz* é aquele que não possui nenhum ramo de entrada, havendo zero ou mais ramos na saída. Os nós internos possuem exatamente um ramo de entrada e dois ou mais ramos de saída. As *folhas* da árvore são os pontos indivisíveis, as quais representam as classes. A Figura 11 ilustra um exemplo de AD para classificar gasodutos como seguros ou inseguros.

Para classificar um registro utilizando uma AD, basta começar pelo nó *raiz* da árvore, em que é aplicado o primeiro teste com o atributo referente a este nó. O processo se repete até ser encontrado uma *folha*, a qual representa o valor associado pela árvore ao atributo classe do registro em questão.

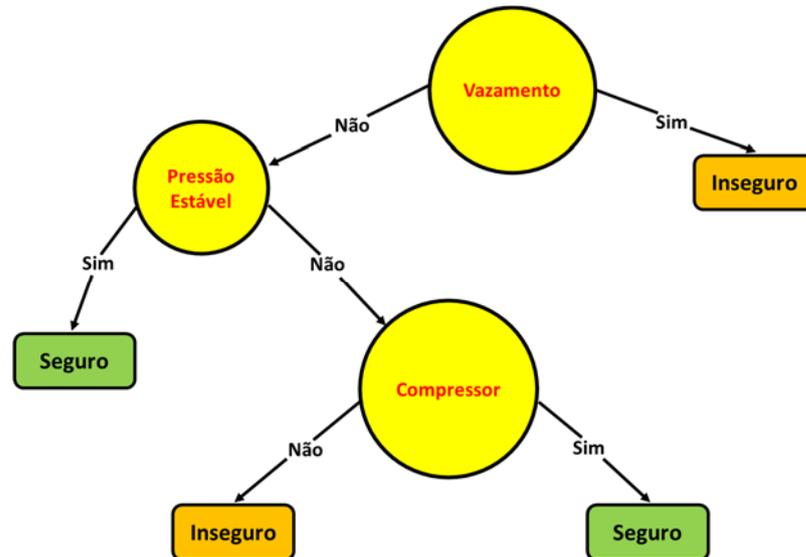


Figura 11: Uma Árvore de Decisão para classificação da segurança gasodutos.

As Árvores de Decisão são construídas baseadas no modelo hierárquico *top-down*, isto é, do nó *raiz* em direção às *folhas*. Embora haja diferenças na forma de realizar os passos, os algoritmos utilizam a técnica dividir para conquistar. Esta técnica está baseada nas sucessivas divisões do problema estudado em vários subproblemas de menores dimensões, até que seja encontrada uma solução mais simples para cada um dos problemas.

Uma AD é essencialmente uma série de declarações *if-then*, que quando aplicados a um registro de uma base de dados, resultam na classificação daquele registro. A Figura 12 representa um exemplo de uma árvore hipotética, e é constituída dos seguintes elementos:

- Raiz: é o nó do topo da árvore.
- Nós: são todos os elementos que estão conectados por ramos;
- Ramos: Ligações entre nós;
- Folhas: são os últimos nós da árvore e representam as classes do conjunto de dados.

Em resumo pode-se afirmar que, sobre uma árvore de decisão:

- Cada nó interno representa um teste em um atributo;
- O ramo que sai de um nó interno representa o resultado do teste no atributo;
- A folha da árvore representa um rótulo de classe (*Seguro/Inseguro*);

- Uma nova observação é classificada seguindo um caminho na árvore, da raiz até a folha.

Analisando a estrutura da Figura 12, nota-se que é possível extrair um conjunto de regras do tipo “*if-then*” para uma melhor compreensão e interpretação dos resultados. As regras de decisão seguem de acordo com o trajeto dos nós sucessivos até um nó *folha* com a sua respectiva classe atribuída. Geralmente essas regras são utilizadas juntamente com a árvore. Em alguns casos as regras substituem as ADs. Assim, podem ser derivados os seguintes conjuntos de regras para o exemplo em questão:

- Conjunto de Regras 1 → *If Fluxo de Gás Natural* < 5 *Then Inseguro*.
- Conjunto de Regras 2 → *If Fluxo de Gás Natural* ≥ 5 e *Pressão do nó* > 45 *Then Seguro*.
- Conjunto de Regras 3 → *If Fluxo de Gás Natural* ≥ 5, *Pressão do nó* ≤ 45 e *Produção de Gás Natural* < 20 *Then Inseguro*.
- Conjunto de Regras 4 → *If Fluxo de Gás Natural* ≥ 5, *Pressão do nó* ≤ 45 e *Produção de Gás Natural* ≥ 20 *Then Seguro*.

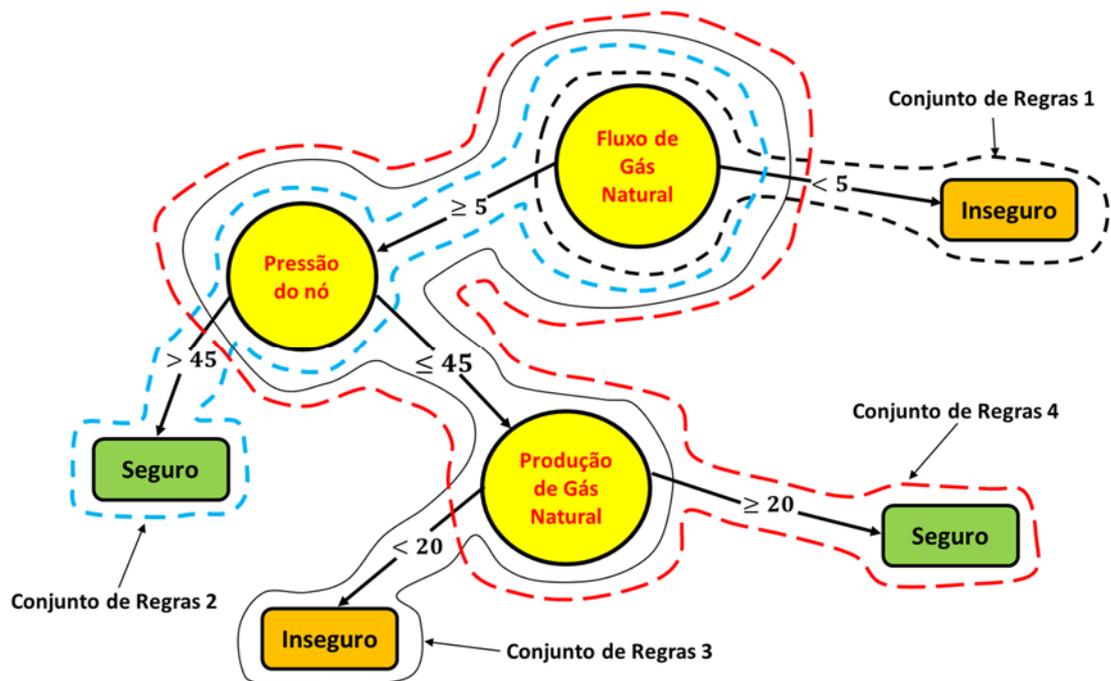


Figura 12: Conjunto de Regras de uma Árvore de Decisão.

Devem ser aplicados testes em todos os nós definidos e um conjunto de exemplos, para que seja escolhido o caminho a percorrer na árvore durante o processo de classificação. Os resultados destes testes dependem da natureza dos atributos, que podem ser numéricos,

categoricos ou simbolicos. Deve-se utilizar apenas um nó para cada teste, pois, isso torna a árvore bem mais simples.

Desta forma, é possível utilizar uma AD para classificar um sistema de gasodutos como SEGURO ou INSEGURO. Para isso, basta a partir do nó *raiz* da árvore e ir percorrendo-a, através das respostas aos testes dos nós internos, até chegar em um nó *folha*, o qual indica a classe correspondente ao estado do gasoduto. Além da obtenção da classe, a grande vantagem é que, a trajetória percorrida até o nó *folha* representa um conjunto de regras que facilita a interpretabilidade do modelo pelo operador do sistema.

Um mesmo conjunto de dados pode gerar várias ADs distintas. Assim, usando o exemplo da Figura 12, o nó raiz poderia ser **Pressão do Nó** em vez de **Fluxo de Gás Natural**, fazendo com que o nó **Fluxo de Gás Natural** passe a ocupar uma outra posição na árvore. Essa troca de nós faz com que seja necessário percorrer um caminho maior ou menor para se chegar a uma decisão.

3.4 ESTRUTURAÇÃO DE UMA ÁRVORE DE DECISÃO

Na estruturação de uma AD, procura-se associar a cada nó de decisão o atributo “mais informativo” entre aqueles que ainda não foram utilizados no caminho desde a *raiz* até a *folha* da árvore. No entanto, cada algoritmo tem a sua própria metodologia para distinguir o atributo mais informativo, fazendo com que a topologia da árvore e a qualidade da árvore variem em função do algoritmo utilizado.

O processo de estruturação das ADs é chamado de indução. A abordagem mais comum para indução é particionar recursivamente, conjuntos de exemplos, classificados até que seja encontrado um critério de parada. A partição é definida relacionando um teste que tenha um pequeno conjunto de saídas, criando uma ligação para cada uma possível, passando cada exemplo para sua correspondente ligação e tratando cada bloco de partição como um subproblema para o qual uma subárvore é construída recursivamente. Como critério de parada é comum adotar quando todos os exemplos de uma ligação pertencem a uma mesma classe.

Para realizar o processo de indução da AD é utilizada a estratégia *top-down*, em que se inicia a geração da árvore a partir do nó raiz, percorrendo os ramos até o próximo nó interno. Este processo de indução pode ser realizado de duas disposições distintas: por profundidade ou por largura.

A Figura 13 apresenta a disposição por profundidade. Nota-se que, a partir de um nó inicial, é feita ligação até um novo nó, caso este nó não seja uma *folha*, serão acrescentados

uma nova ligação e um novo nó, e assim sucessivamente. Quando o processo chegar a uma *folha* ele retorna até o nó que gerou a *folha* e acrescenta um novo ramo e um novo nó. Caso novamente não seja possível adicionar um novo nó, o processo retorna até o nó que seja possível incluir uma nova ligação. O processo finalizará somente quando todos os caminhos terminarem por uma *folha* e nenhum nó for capaz de ser encontrado.

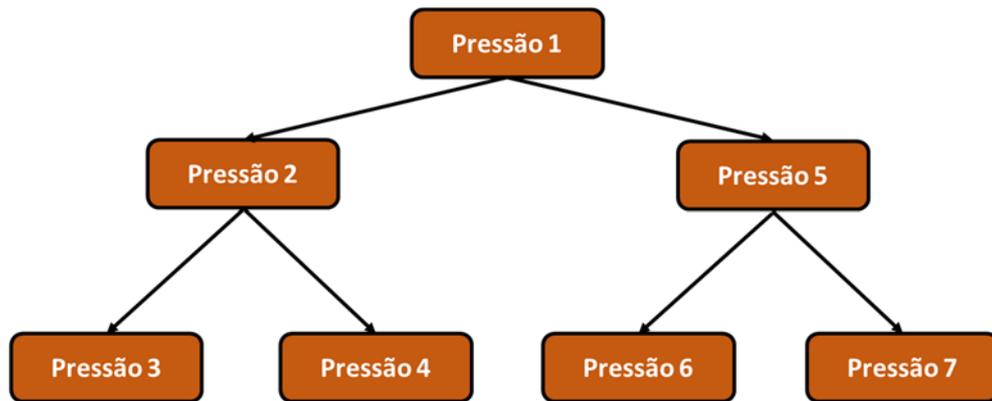


Figura 13: Indução da AD por profundidade.

A Figura 14 representa a disposição por largura. A estruturação por largura realiza todas as ligações de um nó chegando a novos nós. Para cada nó resultante, o critério é testado e é adicionada uma *folha* (ou novos nós), e assim sucessivamente, até que não haja nenhum nó a ser testado. Este método de indução não retorna ao nó *raiz*.

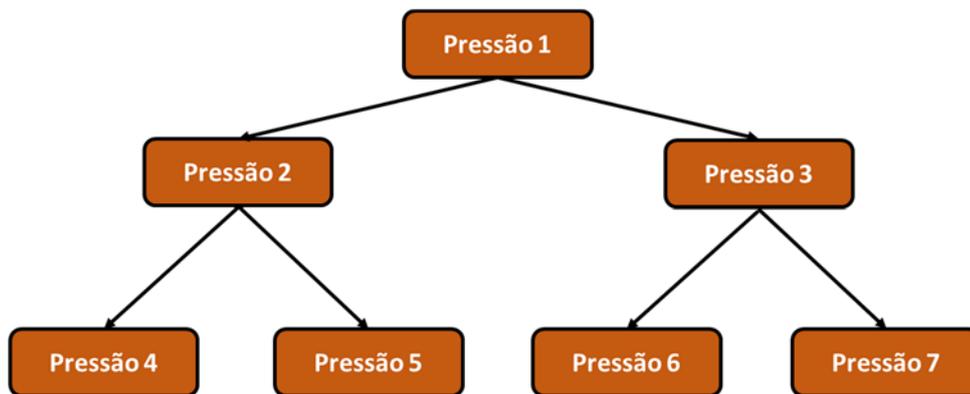


Figura 14: Indução da AD por largura.

3.5 CRITÉRIOS PARA SELEÇÃO DOS ATRIBUTOS DE UMA ÁRVORE DE DECISÃO

A estruturação de uma AD, chamado de processo de indução, pode ser uma tarefa de alta demanda computacional. Embora a indução de uma AD possa ser realizada de forma manual, em uma abordagem *top-down*, as principais demandas por AD geralmente recorrem a processos automáticos de indução, numa abordagem *bottom-up*, a partir de dados disponíveis.

Há muitas maneiras de uma AD ser estruturada a partir de um conjunto de atributos. De forma exaustiva, o número de ADs possíveis cresce fatorialmente à medida que o número de atributos aumenta. Logo, torna-se impraticável definir a estrutura da AD ótima para um determinado problema, devido ao elevado custo computacional envolvido nessa busca.

Nesse sentido, algoritmos baseados em heurísticas têm sido desenvolvidos para a indução de ADs. Mesmo que eles não garantam uma solução ótima, apresentam resultados satisfatórios em tempo factível.

No processo de indução de uma AD, a escolha correta do atributo definirá o sucesso do algoritmo de indução. Vários critérios foram desenvolvidos para escolha do melhor atributo que poderá ser utilizado em um nó. Este critério verifica cada atributo candidato e seleciona aquele que maximiza (ou minimiza) alguma função heurística sobre os subconjuntos.

3.5.1 TOP-DOWN INDUCTION OF DECISION TREE (TDIDT)

O Top-Down Induction of Decision Tree (TDIDT) é um algoritmo bem conhecido e é utilizado como base para muitos algoritmos de indução de árvores de decisão, dentre eles os mais conhecidos como ID3 (QUINLAN, 1986), C4.5 (QUINLAN, 1993) e CART (BREIMAN et al., 1984).

O TDIDT produz regras de decisão de forma implícita em uma AD, a qual é construída por sucessivas divisões dos exemplos de acordo com os valores de seus atributos preditivos. De acordo com BRAMER (2007), esse processo é conhecido como particionamento recursivo.

A essência do algoritmo de TDIDT é baseado em três possibilidades sobre um conjunto de treinamento T contendo classes $C_1, C_2, C_3, \dots, C_K$:

1. T contém um ou mais objetos, sendo todos da classe C_j . Assim, a AD para T é um nó folha que identifica a classe C_j .
2. T não contém objetos. A AD também é um nó folha, mas a classe associada deve ser determinada por uma informação externa. Por exemplo, pode-se utilizar o conhecimento do domínio do problema.
3. T contém exemplos pertencentes a mais de uma classe. Neste caso, a ideia é dividir T em subconjuntos que são, ou tendem a dirigir-se para, coleções de exemplos com classes únicas. Para isso, é escolhido um atributo preditivo A , que possui um ou mais possíveis resultados $O_1, O_2, O_3, \dots, O_n$. T é particionado em subconjuntos $T_1, T_2, T_3, \dots, T_n$, em que T_i contém todos os exemplos de T que têm resultado O_i para o

atributo A . A AD para T consiste de um nó de decisão identificando o teste sobre o atributo A , e uma aresta para cada possível resultado, ou seja, n arestas. No lugar de um único atributo A , pode também ser considerado um subconjunto de atributos.

O mesmo algoritmo de indução das ADs (passos 1, 2 e 3) é aplicado recursivamente para cada subconjunto de exemplos T_i , com i variando de 1 até n .

Fundamentalmente, o TDIDT é um algoritmo recursivo de busca ávida que procura, em um conjunto de atributos, aqueles que “melhor” dividem o conjunto de exemplos em subconjuntos. Inicialmente, todos os exemplos são colocados em um único nó, chamado de raiz.

Em seguida, um atributo preditivo é escolhido para representar o teste desse nó e, assim, dividir os exemplos em subconjuntos de exemplos. Esse processo se repete recursivamente até que todos os exemplos já estejam classificados ou então até que todos os atributos preditivos já tenham sido utilizados.

3.6 SELEÇÃO DOS ATRIBUTOS PREDITIVOS PARA OS NÓS DA ÁRVORE DE DECISÃO

O critério de seleção define qual atributo preditivo é utilizado em cada nó da árvore. Existem diferentes tipos de critérios de seleção, sendo esta uma das variações entre os diversos algoritmos de indução de AD. Esses critérios são definidos em termos da distribuição de classe dos exemplos antes e depois da divisão (TAN *et al.*, 2005).

A maioria dos algoritmos de indução de AD trabalha com funções de divisão univariável, ou seja, cada nó interno da árvore é dividido de acordo com um único atributo. Nesse caso, o algoritmo tenta encontrar o melhor atributo para realizar essa divisão.

Os critérios de seleção para a melhor divisão são baseados em diferentes medidas, tais como impureza, distância e dependência. A maior parte dos algoritmos de indução busca dividir os dados de um nó-pai de forma a minimizar o grau de impureza dos nós-filhos.

Quanto menor o grau de impureza, mais desbalanceada é a distribuição de classes. Em um determinado nó, a impureza é nula se todos os exemplos nele pertencerem à mesma classe. Analogamente, o grau de impureza é máximo no nó se houver o mesmo número de exemplos para cada classe possível. A seguir, serão apresentadas as medidas mais utilizadas para a seleção da melhor divisão.

3.6.1 GANHO DE INFORMAÇÃO

Uma das medidas baseadas em impureza é o **Ganho de Informação**, o qual usa a entropia como medida de impureza. O algoritmo ID3 (QUINLAN, 1986), pioneiro em indução de árvores de decisão, utiliza essa medida. Para determinar o quanto uma condição de teste realizada é adequada, é necessário comparar o grau de entropia do nó-pai (antes da divisão) com o grau de entropia dos nós-filhos (após a divisão). O atributo que gerar uma maior diferença é escolhido como condição de teste. O ganho é dado pela Equação (1), na forma:

$$Ganho = Entropia(pai) - \sum_{j=1}^n \left[\frac{N(v_j)}{N} Entropia(v_j) \right] \quad (1)$$

Em que n é o número de valores do atributo, ou seja, o número de nós-filhos, N é o número total de objetos do nó-pai e $N(v_j)$ é o número de exemplos associados ao nó filho v_j .

O grau de entropia é definido pela Equação (2), em que $p(i/nó)$ é a fração dos registros pertencentes à classe i no nó, e c é o número de classes.

$$Entropia(nó) = - \sum_{i=1}^c p(i/nó) \cdot \log_2 [p(i/nó)] \quad (2)$$

O critério de ganho seleciona como atributo-teste aquele que maximiza o ganho de informação. A grande contrariedade ao se utilizar o ganho de informação é que ele dá preferência a atributos com muitos valores possíveis (número de arestas).

Um exemplo claro dessa contrariedade ocorreria ao utilizar um atributo totalmente irrelevante (por exemplo, um identificador único). Nesse caso, seria criado um nó para cada valor possível, e o número de nós seria igual ao número de identificadores. Cada um desses nós teria apenas um exemplo, o qual pertence a uma única classe, ou seja, os exemplos seriam totalmente discriminados. Assim, o valor da entropia seria mínimo porque, em cada nó, todos os exemplos (no caso um só) pertencem à mesma classe. Essa divisão geraria um ganho máximo, embora seja totalmente inútil.

3.6.2 RAZÃO DE GANHO

Para resolver o problema do ganho de informação, foi proposto em Quinlan (1993) a Razão de Ganho (*Gain Ratio*), que nada mais é do que o ganho de informação relativo (ponderado) como critério de avaliação. A razão de ganho é definida pela Equação (3).

$$\text{Razão de Ganho}(\text{nó}) = \frac{\text{Ganho}}{\text{Entropia}(\text{nó})} \quad (3)$$

Pela Equação (3), é possível perceber que a razão não é definida quando o denominador é igual a zero. Além disso, a razão de ganho favorece atributos cujo denominador, ou seja, a entropia, possui valor pequeno.

Em Quinlan (1988), é sugerido que a razão de ganho seja realizada em duas etapas: na primeira etapa, é calculado o ganho de informação para todos os atributos. Na segunda etapa, considera-se apenas aqueles atributos que obtiveram um ganho de informação acima da média, e então escolher aquele que apresentar a melhor razão de ganho. Dessa forma, Quinlan mostrou que a razão de ganho supera o ganho de informação tanto em termos de acurácia quanto em termos de complexidade das árvores de decisão geradas.

3.6.3 CRITÉRIO DE GINI

Outra medida bastante conhecida é o *Gini*, a qual emprega um índice de dispersão estatística proposto em 1912 pelo estatístico italiano Corrado Gini. Ele é utilizado no algoritmo CART (BREIMAN et al., 1984). Para um problema de c classes, o $gini_{index}$ é definido pela Equação (4).

$$gini_{index}(\text{nó}) = 1 - \sum_{i=1}^c p(i/\text{nó}) \quad (4)$$

Da mesma forma como no cálculo do ganho de informação, basta calcular a diferença entre o $gini_{index}$ antes e após a divisão. Essa diferença, *Gini*, é representada pela Equação (5).

$$Gini = gini_{index}(pai) - \sum_{j=1}^n \left[\frac{N(v_j)}{N} gini_{index}(v_j) \right] \quad (5)$$

Em que n é o número de valores do atributo, ou seja, o número de nós-filhos, N é o número total de objetos do nó-pai e $N(v_j)$ é o número de exemplos associados ao v_j , nó-filho. Dessa forma, é selecionado o atributo que gerar o maior valor para *Gini*.

3.6.4 REPRESENTAÇÃO DOS NÓS PARA ATRIBUTOS CATEGÓRICOS

A forma de representação dos nós pode influenciar de maneira decisiva no desempenho das árvores de decisão induzidas. Dependendo do tipo de atributo, existem diferentes tipos de

representação dos nós para o particionamento dos dados. A seguir, são apresentadas algumas formas de representação considerando atributos categóricos ordinais e não-ordinais.

1. Um ramo para cada valor de atributo: É a partição mais comum, na qual é criada uma aresta para cada valor do atributo usado como condição de teste. Embora esse tipo de partição permita extrair do atributo todo o seu conteúdo informativo, possui a desvantagem de tornar a árvore de decisão mais complexa. O algoritmo C4.5 (QUINLAN, 1993) utiliza esse tipo de divisão para atributos categóricos não-ordinais.
2. Solução de Hunt: A partição utilizada pelo algoritmo ID3, sugere uma partição binária. Nesse caso, um dos valores é atribuído a uma das arestas e todos os outros valores à outra aresta. A desvantagem desse tipo de partição é não aproveitar todo o poder de discriminação do atributo em questão.
3. Atributos categóricos ordinais: Como já definido, um atributo é ordinal quando há uma relação de ordem entre os seus possíveis valores. Por exemplo, tem-se um atributo **pressão** que pode possuir os valores (baixa), (média) e (alta). Com atributos desse tipo, é possível realizar uma partição binária do tipo **pressão** < (média), em que todos os exemplos cujo atributo **pressão** tem valor (baixa) seguem por uma aresta e os outros seguem por outra aresta. Esse tipo de partição é uma das que foram implementadas para o algoritmo CART (BREIMAN *et al.*, 1984).
4. Agrupamento de valores em dois conjuntos: De acordo com Breiman et al. (1984), a divisão binária também pode ser realizada de uma forma mais complexa, em que cada um dos dois subconjuntos pode ser formado por registros com mais de um valor para o atributo utilizado como condição de teste. O grande desafio desse tipo de partição é o elevado custo computacional para encontrar a melhor divisão, pois o número de combinações possíveis é $2^n - 1$, em que n é o número de valores possíveis para o atributo em questão.
5. Agrupamento de valores em vários conjuntos: Visando permitir o agrupamento de valores em vários conjuntos com uma complexidade de cálculo razoável, o algoritmo C4.5 (QUINLAN, 1993) permite encontrar uma solução de boa qualidade. Para isso, inicia criando uma aresta para cada valor do atributo em questão. Em seguida são testadas todas as combinações possíveis de dois valores e, caso nenhuma dessas combinações produza um ganho maior que a divisão anterior, o processo é interrompido e a divisão anterior é adotada como divisão final. Caso contrário, é repetido o processo tendo como base a melhor das soluções anteriores. Nota-se que

não se pode garantir que a divisão encontrada seja a melhor possível, pois é verificado se houve melhoria apenas um passo à frente. Esse é o custo de se produzir um algoritmo mais simples.

3.6.5 REPRESENTAÇÃO DOS NÓS PARA ATRIBUTOS CONTÍNUOS

Os atributos contínuos permitem uma maior variedade de testes e, conseqüentemente, implicam uma maior complexidade de cálculo. Segundo Fonseca (1994), alguns dos testes mais usados para partição de atributos contínuos são: testes simples ou pesquisa exaustiva, testes múltiplos (segmentação global e segmentação ao nível do nó) e combinação linear de características.

O teste simples, também conhecido como pesquisa exaustiva, é o mais utilizado. Um dos algoritmos que o utiliza é o C4.5, e a divisão é sempre binária. Supondo um atributo contínuo X a ser utilizado como nó teste, mesmo que seu domínio seja infinito, o número de exemplos num conjunto de treinamento T é finito e, portanto, o número de valores diferentes para esse atributo também é finito.

Assim, os exemplos do conjunto T são ordenados de acordo com seus valores para o atributo X . Supondo que os diferentes valores de X sejam, em ordem crescente, $\{a_1, a_2, a_3, \dots, a_m\}$, T é dividido em duas partes. São elas: T_1 , cujos exemplos possuem valores $\{a_1, a_2, a_3, \dots, a_i\}$ e T_2 , com valores $\{a_{i+1}, a_{i+2}, a_{i+3}, \dots, a_m\}$ para o atributo X . Para cada $a_i, i = 1, 2, 3, \dots, m-1$, é calculado o ganho (independente do critério utilizado) para a respectiva divisão. Após avaliar todas as divisões possíveis, é escolhida aquela que fornecer o maior ganho.

Finalmente, é necessário definir o valor que será usado como limiar (valor usado para dividir os exemplos no nó). Tendo posse do a_i que produziu o melhor ganho, o valor mais utilizado como limiar é $\frac{a_i + a_{i+1}}{2}$, pois assim espera-se que a árvore resultante apresente melhores resultados para exemplos que não participaram do conjunto de treinamento.

3.7 MÉTODOS DE PODA DA ÁRVORE DE DECISÃO

Quando as ADs são construídas, muitas das arestas ou subárvores podem refletir ruídos ou erros. Isso possibilita, ao problema conhecido como sobreajuste, um aprendizado muito específico do conjunto de treinamento, não permitindo ao modelo generalizar.

Para detectar e excluir essas arestas e subárvores, são utilizados métodos de poda (*pruning*) da árvore, cujo objetivo é melhorar a taxa de acerto do modelo para novos exemplos, os quais não foram utilizados no conjunto de treinamento (HAN, 2001).

Consequentemente, a árvore podada se torna mais simples, facilitando a sua interpretabilidade por parte do usuário. Junto ao método de seleção, o método de poda também varia de acordo com os diferentes algoritmos de indução da AD.

Existem diversas formas de realizar poda em uma AD, e todas elas são classificadas como pré-poda ou pós-poda. O método pré-poda é realizado durante o processo de construção da árvore, em que o processo pode simplesmente parar de dividir o conjunto de elementos e transformar o nó corrente em um nó folha da árvore.

O ganho de informação, por exemplo, pode ser utilizado como critério de poda. Caso todas as divisões possíveis utilizando um atributo A gerem ganhos menores que um valor pré-estabelecido, então esse nó vira folha, representando a classe mais frequente no conjunto de exemplos.

A dificuldade é encontrar um valor adequado para x , visto que um valor muito alto pode gerar uma árvore muito simplificada, enquanto que um valor muito baixo pode produzir uma ínfima simplificação na árvore.

O pós-poda é realizado após a estruturação da AD, removendo ramos completos, onde tudo que está abaixo de um nó interno é excluído e esse nó é transformado em folha, representando a classe mais frequente no ramo.

Para cada nó interno da árvore, o algoritmo calcula a taxa de erro caso a subárvore abaixo desse nó seja podada. Em seguida, é calculada a taxa de erro caso não haja a poda. Se a diferença entre essas duas taxas de erro for menor que um valor preestabelecido, a árvore é podada. Caso contrário, não ocorre a poda.

Esse processo se repete progressivamente, gerando um conjunto de árvores podadas. Finalmente, para cada uma delas é calculada a acurácia na classificação de um conjunto de dados independente dos dados de treinamento (o conjunto de validação), e a árvore que obtiver a melhor acurácia será a escolhida.

A poda é um método bastante utilizado e eficaz na solução do problema de sobreajuste, todavia, deve-se ter cuidado para não podar demais a árvore. Quando isso ocorre, tem-se o problema conhecido como subajuste, em que o modelo de classificação não aprendeu o suficiente sobre os dados de treinamento.

Entre os métodos de poda existentes, destacam-se: *Cost Complexity Pruning*, *Reduced Error Pruning*, *Minimum Error Pruning* (MEP), *Pessimistic Pruning*, *Error-Based Pruning*

(EBP), *Minimum Description Length* (MDL) *Pruning*, *Minimum Message Length* (MML) *Pruning*, *Critical Value Pruning* (CVP), OPT e OPT-2 (ROKACH e MAIMON, 2008).

3.8 ALGORITMOS DE INDUÇÃO DE ÁRVORES DE DECISÃO

Os três principais algoritmos para indução de ADs são: ID3 (QUINLAN, 1986), C4.5 (QUINLAN, 1993) e CART (BREIMAN *et al.*, 1984). Esses algoritmos têm como fundamentação o TDIDT.

3.8.1 ID3

O ID3 é o algoritmo pioneiro em indução das ADs. Ele é um algoritmo recursivo e baseado em busca voraz, que procura em um conjunto de atributos, aqueles que “melhor” dividem os exemplos, gerando subárvores.

A principal limitação do ID3 é que ele só lida com atributos categóricos não-ordinais, não sendo possível apresentar a ele conjuntos de dados com atributos contínuos. Nesse caso, os atributos contínuos devem ser previamente discretizados. Outra limitação do ID3 é de não apresentar uma forma para tratar valores desconhecidos, ou seja, todos os exemplos do conjunto de treinamento devem ter valores conhecidos para todos os seus atributos.

O ID3 utiliza o ganho de informação para selecionar a melhor divisão. Contudo, esse critério não considera o número de divisões (número de arestas), e isso pode acarretar em árvores mais complexas. Somado a isso, o ID3 também não apresenta nenhum método de pós-poda, o que poderia amenizar esse problema de árvores mais complexas.

3.8.2 C4.5

O algoritmo C4.5 representa uma significativa evolução do ID3. O C4.5 trabalha tanto com atributos categóricos (ordinais ou não-ordinais) como com atributos contínuos. Para lidar com atributos contínuos, o algoritmo C4.5 define um limiar e então divide os exemplos de forma binária: aqueles cujo valor do atributo é maior que o limiar e aqueles cujo valor do atributo é menor ou igual ao limiar.

O algoritmo C4.5 permite que os valores desconhecidos para um determinado atributo sejam representados como ‘?’, e o algoritmo trata esses valores de forma especial. Esses valores não são utilizados nos cálculos de ganho e entropia.

Outra vantagem em relação ao ID3 é que o C4.5 utiliza a medida de razão de ganho para selecionar o atributo que melhor divide os exemplos. Essa medida se mostrou superior ao

ganho de informação, gerando árvores mais precisas e menos complexas. O C4.5 também trata de problemas em que os atributos possuem custos diferenciados.

O algoritmo C4.5 faz uma busca na árvore, de baixo para cima, e transforma em nós folha aqueles ramos que não apresentam nenhum ganho significativo. Dessa forma, diferentemente do ID3, ele apresenta um método de pós-poda das árvores geradas.

3.8.3 CART

O algoritmo CART (*Classification and Regression Trees*) consiste de uma técnica não-paramétrica que induz tanto árvores de classificação quanto árvores de regressão, dependendo se o atributo é nominal (classificação) ou contínuo (regressão).

Entre as principais virtudes do CART está a grande capacidade de pesquisa de relações entre os dados, mesmo quando elas não são evidentes, bem como a produção de resultados sob a forma de ADs de grande simplicidade e legibilidade (FONSECA, 1994).

As árvores geradas pelo algoritmo CART são sempre binárias, as quais podem ser percorridas da sua raiz até as folhas respondendo apenas a questões simples do tipo “sim” ou “não”.

Os nós que correspondem a atributos contínuos são representados por agrupamento de valores em dois conjuntos. Da mesma forma que no algoritmo C4.5, o CART utiliza a técnica de pesquisa exaustiva para definir os limiares a serem utilizados nos nós para dividir os atributos contínuos.

Adicionalmente, o CART dispõe de um tratamento especial para atributos ordenados e também permite a utilização de combinações lineares entre atributos (agrupamento de valores em vários conjuntos).

Diferente das abordagens adotadas por outros algoritmos, os quais utilizam pré-poda, o CART expande a árvore exaustivamente, realizando pós-poda por meio da redução do fator custo-complexidade (BREIMAN *et al.*, 1984). Segundo esses autores, a técnica de poda utilizada é muito eficiente e produz árvores mais simples, precisas e com boa capacidade de generalização.

3.9 CONCLUSÃO DO CAPÍTULO III

Esse capítulo apresentou o Algoritmo de Aprendizado de Máquina (AAM) como sendo o início do modelo da Árvore de Decisão (AD). Foi verificado que, na modelagem preditiva, o modelo de classificação é utilizado para classificar casos cujas classes são desconhecidas. O

presente capítulo também mostrou que a Mineração de Dados ou *Data Mining* corresponde ao núcleo do processo de Descoberta de Conhecimento em Banco de Dados ou *Knowledge Discovery in Database* (KDD) e que, a Árvore de Decisão é um algoritmo de classificação de dados nesse ambiente. O capítulo III exibiu, também, os critérios para a seleção de atributos de uma AD; como é avaliado o nível de impureza das informações; os métodos de poda da árvore e os algoritmos de indução mais importantes.

CAPÍTULO IV – OTIMIZAÇÃO

4.1 INTRODUÇÃO

A primeira teoria sobre evolução das espécies foi proposta em 1809, pelo naturalista francês Jean Baptiste Pierre Antoine de Monet, conhecido como Lamarck. Para Lamarck as características que um animal adquire durante sua vida podem ser transmitidas hereditariamente, este estudo ficou conhecido pela ciência como a “lei do uso e desuso” (DARWIN, 2004).

Charles Darwin vem debater a teoria de Lamarck, de forma agressiva tentando de forma científica explicar como as espécies evoluem. A seleção natural é um processo de evolução, geralmente aceito pela comunidade científica como a melhor explicação para a adaptação. O meio ambiente seleciona os seres mais aptos, em geral, só estes conseguem reproduzir-se e os menos adaptados são eliminados ou pelo menos reduzidos em um primeiro momento a uma minoria. Assim, só as diferenças que facilitam a sobrevivência são transmitidas à geração seguinte (STEARNS, 2003).

A computação evolucionária se inspira na teoria evolutiva para o desenvolvimento de métodos computacionais (SRINIVAS, 1994). Em meados da década de 1970, John H. Holland propõe a técnica de Algoritmo Genético (AG), inspirada nas teorias darwinianas. Deste estudo surgiu a publicação “*Adaptation in Natural and Artificial Systems*” (DIAS et al, 1998).

Os Algoritmos Genéticos (AGs) são algoritmos de otimização e busca baseados nos mecanismos de seleção natural e genética. A teoria emula o processo da natureza em que os mais aptos vencem e se reproduzem e, conseqüentemente, os mais fracos se extinguem. O método consiste inicialmente de uma população formada por uma série de bits (*string* – representando os cromossomos), que é transformada por três operadores genéticos: seleção, reprodução e mutação (GOLDBERG, 1989).

Cada *string* (cromossomo) representa uma possível solução do problema a ser otimizado, e cada bit (ou grupo de bits), representa o valor associado a determinadas variáveis do problema (gene). As soluções são classificadas por uma função de aptidão ou de avaliação (*fitness*) que desempenha o papel do ambiente. O par cromossomo e função de aptidão representam o indivíduo (GOLDBERG, 1989).

O interesse por este algoritmo tem aumentado de forma significativa, uma vez que o mesmo fornece um mecanismo de busca bastante robusto, além de se basear em importantes conceitos biológicos (MONTROYA et al, 2000).

Geralmente, os métodos convencionais de otimização e busca trabalham de forma sequencial, avaliando a cada instante uma possível solução, enquanto que os AGs trabalham com um conjunto de possíveis soluções simultaneamente. As principais diferenças entre os AGs e outros métodos são: trabalham com a codificação do conjunto de parâmetros e não com os próprios parâmetros; operam com uma população de soluções candidatas simultaneamente, e não com uma única solução; utilizam funções de custo ou recompensa, e não derivadas de funções; e utilizam regras de transição probabilística, e não determinística (GOLDEBERG, 1989). Os algoritmos genéticos utilizam os operadores, tais como:

- O operador de seleção, responsável por selecionar indivíduos da população anterior que vão participar das etapas de cruzamento e mutação, formando assim uma nova geração. Tal seleção é influenciada, no sentido de escolher os indivíduos com aptidão acima da média como matrizes para os membros da nova população. O processo de seleção pode ser implementado de vários modos, sendo o sorteio por meio de uma roleta, o mais utilizado (MONTROYA et al, 2000);
- A combinação ou cruzamento (*crossover*) é o principal operador genético de reprodução. Sua função é trocar porções de cromossomo entre dois indivíduos, de modo a explorar novas partes do espaço de busca. Existem diferentes tipos de operadores de cruzamento, tais como: combinação em um ponto, dois pontos e uniforme. Nem todos os indivíduos realizam a combinação, sendo esta frequentemente controlada por uma probabilidade de reprodução.
- A mutação também é um operador importante e consiste na troca aleatória de um ou dois *bits* da *string* utilizada na representação do indivíduo pelo seu complemento (na codificação binária). Este operador deve ser utilizado com um pouco de cuidado, com baixa probabilidade, tipicamente, 0.0001, ou entre 0.001 e 0.1 (GOLDEBERG, 1989). A mutação serve como uma espécie de “garantia de vida”. Algum gene (*bit*) importante pode ser perdido durante a seleção, cabendo à mutação, a possibilidade em recuperá-lo de volta. Além disto, em alguns casos, as *strings* de uma população podem apresentar o mesmo valor em cada bit, sendo a reprodução incapaz de introduzir novos indivíduos.

Em (MONTROYA et al, 2000) a operação de *crossover* é descrita em passos. Primeiro, dois indivíduos são escolhidos aleatoriamente da população de indivíduos ancestrais gerados pelo operador seleção. Segundo, uma ou mais posições do indivíduo são escolhidas como sendo o(s) ponto(s) em que haverá o cruzamento que resultará em segmentos dos indivíduos.

E por último, os segmentos são recombinados para resultar em dois indivíduos descendentes. Estes passos podem ser visualizados na Figura 15 a seguir.

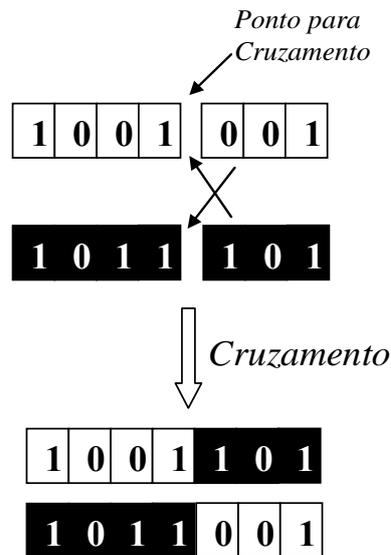


Figura 15: Operador de Cruzamento

O operador mutação procede determinando-se aleatoriamente a posição que sofrerá mutação, sendo que cada posição possui a mesma probabilidade de ocorrer mutação. Na Figura 16, fica claro que a mutação é simplesmente a mudança de um bit de 0 para 1 ou vice-versa; neste caso da Figura 16, apenas uma posição sofreu a mutação, apesar da mutação ser bastante ocasional (com baixa probabilidade), mais de uma posição pode sofrer a mutação.

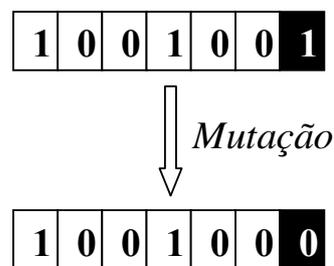


Figura 16: Operador de Mutaçãõ

Com base nos conceitos expostos, são apresentados a seguir os passos a serem seguidos por um algoritmo genético típico, ou seja:

Início

1. Gerar aleatoriamente uma população inicial (população anterior)
2. Repetir até que o critério de convergência seja atingido
3. Avaliação da população
 - 3.1. Avaliar (população anterior)
 - 3.2. Nova população = Selecionar (população anterior)

- 3.3. Reprodução (nova população)
 - 3.4. Mutação (nova população)
 - 3.5. População anterior = população nova
4. Fim.

O processo termina quando for atingido o critério de convergência definido pelo usuário ($X\%$ dos genes possuem o mesmo valor em pelo menos $Y\%$ dos indivíduos), ou quando o indivíduo adequado tiver sido obtido, ou ainda, um determinado número de gerações tiver sido alcançado (ou o valor ótimo não se alterou por Z gerações).

Por fim, pode-se resumir os algoritmos genéticos como sendo um processo onde se gera uma população inicial aleatoriamente, avalia-se esta, e criam-se novas populações por meio de operadores. Os indivíduos com um bom desempenho são selecionados e partes destes são combinadas, criando cópias mais fortes, pelo uso de seleção e reprodução. Apesar desta aparente simplicidade, cabe ressaltar que o processo de busca exige uma substancial força computacional que podem inviabilizar sua utilização em aplicações on-line.

4.2 MÉTODOS DETERMINÍSTICOS *VERSUS* HEURÍSTICOS

Nos métodos determinísticos, a solução pode sempre ser encontrada, dependendo somente do ponto de partida. O conhecimento prévio deste ponto de partida é o grande problema destes métodos. Por outro lado, os métodos heurísticos, que são baseados na experiência, podem levar a uma solução boa, mas nem sempre ótima. Quase todas as decisões tomadas diariamente envolvem a solução heurística de um problema.

Métodos determinísticos foram amplamente empregados para resolver problemas de despacho econômico, formulados através da Função Lagrangeana ou da Função Penalidade. Tanto a Função Lagrangeana quanto a Função Penalidade são empregadas para tornar o problema restrito em um problema irrestrito, cujas soluções devem conter todas as variáveis mais o multiplicador de Lagrange, no primeiro caso e, as variáveis e um fator de penalidade, no segundo caso. Métodos do tipo Gradiente são uma alternativa de buscar o ótimo destas funções, sempre que a função objetivo for contínua e diferenciável. O Método de Newton também pode ser usado, desde que as duas primeiras derivadas da função objetivo sejam contínuas. Este último método converge em apenas uma iteração, dependendo do ponto de partida, mas requer um esforço computacional maior para calcular as derivadas de primeira e segunda ordem (MONTROYA et al, 2000).

A partir da década de 50, métodos heurísticos foram propostos para simular fenômenos biológicos, entre estes, os Algoritmos Genéticos. Na analogia matemática, os indivíduos da população são os pontos de partida ou soluções iniciais. É importante destacar, que os AGs trabalham sempre com um conjunto de indivíduos. A determinação do número de indivíduos é um parâmetro que deve ser ajustado para cada caso e não existe um parâmetro ótimo, porém, há como se determinar uma faixa de trabalho (MICHALEWICZ,1994). Os operadores básicos dos AGs são a criação da população inicial aleatoriamente, o cruzamento, a mutação e a seleção. Inicialmente, a população inicial possui indivíduos de características boas e ruins, devido ao processo de criação que é aleatório. Estas características podem ser trocadas com outros indivíduos da população e transmitidas para as gerações posteriores através do operador cruzamento. No operador mutação, os indivíduos são “modificados” na tentativa de torná-los mais aptos a permanecerem nas próximas gerações. Na analogia matemática, a aptidão está relacionada à função objetivo. No operador seleção, os indivíduos mais aptos serão escolhidos para fazerem parte das gerações posteriores (MICHALEWICZ,1994).

4.3 ESTÁGIOS DOS ALGORITMOS GENÉTICOS

Um ciclo dos AGs consiste dos seguintes estágios: (a) geração aleatória de uma população de soluções potenciais e codificadas (indivíduos); (b) avaliação dos indivíduos; (c) seleção dos indivíduos mais capazes; (d) geração da nova população através de manipulações genéticas.

A princípio, os AGs utilizam uma população de soluções potenciais criada por um gerador aleatório de valores lógicos: verdadeiro ou falso (codificação binária). No estágio (b) os indivíduos são avaliados para, em (c), se fazer a seleção daqueles que vão participar das próximas etapas. O desempenho de cada indivíduo é determinado pelo valor da função aptidão, relacionada com a função objetivo do problema. Esta seleção visa determinar os indivíduos genitores para serem utilizados no estágio (d). Finalmente, no último estágio, pode-se realçar o papel dos operadores genéticos como fundamental para a obtenção de novos pontos, em busca da solução ótima. Os novos indivíduos criados substituem os anteriores, terminando assim, um ciclo dos AGs.

Os algoritmos então prosseguem ciclicamente a partir dessa nova população e só terminam quando algum critério de convergência é alcançado. Nos AGs geralmente mantém-se o tamanho da população com o mesmo número n_{pop} de indivíduos. O esquema de reposição populacional em um AG pode ser visto na Figura 17.

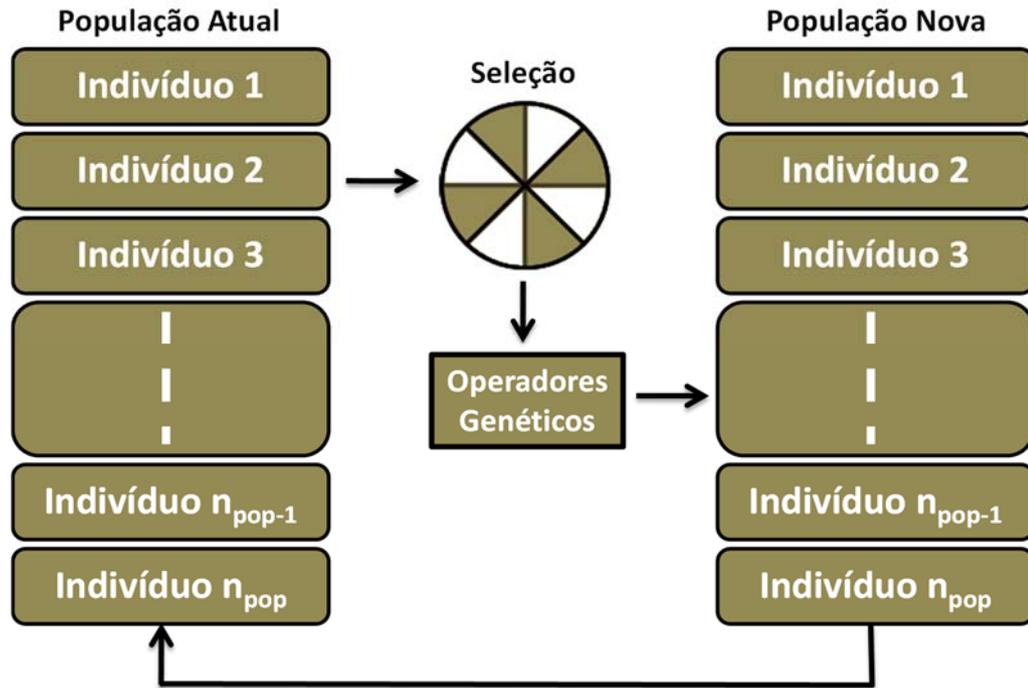


Figura 17: Esquema de Reposição Populacional em um AG.

4.4 SISTEMA DE REPRESENTAÇÃO

O problema de controle do fluxo de gás natural e a minimização do custo serão modelados através de variáveis reais. Então, deve-se fazer o mapeamento das variáveis reais para um código com o qual os AGs trabalham.

Neste trabalho, a representação cromossômica das variáveis de otimização, utilizadas no Algoritmo Genético implementado no MATLAB, são modeladas através do código binário.

A relação existente entre o código binário, supondo que o cromossomo tenha comprimento $l = 5$, e as variáveis reais limitadas por x_{\min} e x_{\max} é representada na Tabela 7.

Tabela 7: Mapeamento de Variáveis Reais para Código Binário

Variável Real	Código Binário
x_{\min}	0 0 0 0
x_{\max}	1 1 1 1

O mapeamento das variáveis reais entre o intervalo limitado por $[x_{\min}, x_{\max}]$ é feito linearmente, ou seja, um valor real intermediário entre x_{\min} e x_{\max} , é calculado através da relação:

$$\text{Valor real} = x_{\min} + \left(\frac{\text{valor binário}}{2^l - 1} \right) (x_{\max} - x_{\min}) \quad (6)$$

Nos AGs cada cromossomo (ou cadeia de caracteres) representa uma variável. Em problemas multi-variáveis, deve-se avaliar a ação conjunta do grupo de cromossomos que constituem o indivíduo. Por exemplo, seja um problema multi-variável, cuja solução é um vetor $X = [x_1 \ x_2 \ x_3]$. Cada x_i é um cromossomo e seu agrupamento representa um indivíduo. Para ilustrá-lo, suponha que cada cromossomo tenha comprimento $l = 5$. Dessa forma, um possível indivíduo para um problema de 3 variáveis poderia ser visto na Tabela 8:

Tabela 8: Representação de um Indivíduo para um Problema de Três Variáveis

X1	X2	X3
1 1 0 0 1	0 1 0 1 1	0 0 0 1 1

Conforme foi dito, os AGs trabalham com grupos de indivíduos (população). Então uma típica população com quatro indivíduos, todos com comprimento $l = 5$, pode ser vista na Tabela 9.

Tabela 9: Representação de uma População com Quatro Indivíduos

Número do Indivíduo	Indivíduo
1	1 1 0 0 0
2	1 0 0 1 0
3	0 1 1 0 1
4	0 0 1 0 1

4.5 ESQUEMAS DE SELEÇÃO

O Algoritmo Genético Simples trabalha com um número fixo de indivíduos na população ao longo das gerações. Então, a cada geração, deve-se selecionar quais indivíduos terão cópias e quais desaparecerão. Este processo é denominado de reprodução. Há diversas possibilidades de implementação deste processo de seleção. Neste trabalho, optou-se pelo método da roleta, no qual um indivíduo com valor de aptidão ff_i tem uma probabilidade de seleção dada pela expressão:

$$\frac{ff_i}{\sum ff_i} \quad (7)$$

em que $\sum ff_i$ é a soma dos valores da função de adaptação da população.

Na implementação, pelo método da roleta, cada indivíduo corresponde a um setor circular de ângulo $2\pi \frac{ff_i}{\sum ff_i}$. Um número gerado aleatoriamente entre 0 e 2π determina qual o setor escolhido e o indivíduo correspondente.

Suponha que os quatro indivíduos apresentados na Tabela 9 possuam respectivamente probabilidades de seleção iguais a 40, 30, 20 e 10%, conforme apresentado na Figura 18. Girando a roleta quatro vezes, pode-se ter, por exemplo, a escolha representada na Tabela 10, onde o indivíduo 1 obteve duas cópias, os indivíduos 2 e 3 uma, enquanto que o indivíduo 4 não obteve nenhuma cópia.

Tabela 10: Representação do Processo de Seleção

Número do Sorteio	Probabilidade de Seleção	Indivíduo Escolhido
1	40%	1
2	30%	2
3	40%	1
4	20%	3

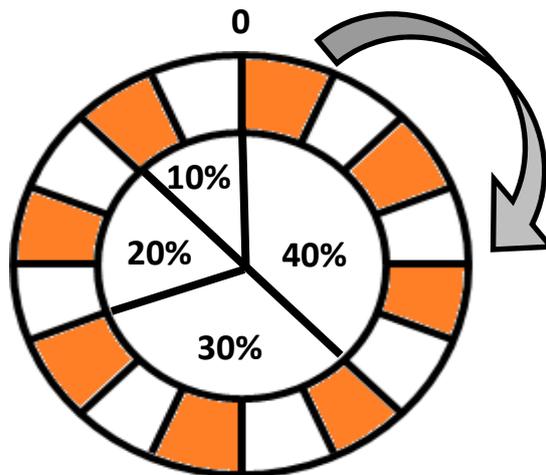


Figura 18: Exemplificação do Método da Roleta

4.6 OPERADORES GENÉTICOS

Entre os principais operadores genéticos existentes, pode-se citar três: cruzamento, mutação e inversão. Dos três operadores, o cruzamento e a mutação têm uma importância significativa na evolução dos AGs. A inversão, por sua vez, é considerada como secundária para os AGs.

a) Cruzamento:

Mantendo a analogia com sistemas naturais, entre os indivíduos que foram selecionados no processo de reprodução, dois são escolhidos para se cruzarem. O cruzamento nada mais é do que um processo que possibilita a troca de material genético entre os indivíduos participantes, dessa forma, fica fácil entender que ele é uma importante operação para gerar novos pontos, que são possíveis soluções.

Entre os diversos tipos possíveis, pode-se citar o cruzamento com um único ponto de corte, com múltiplos pontos de corte, cruzamento uniforme, cruzamento entre vários indivíduos simultaneamente, cruzamento por variável, dentre outros.

O cruzamento só se efetiva com uma probabilidade p_c . A decisão para executá-lo, é tomada usando um gerador de números aleatórios para gerar um número entre 0 e 1. Se o número gerado é inferior a p_c , a decisão é positiva. Escolhe-se então, o parceiro para o indivíduo em questão gerando um número aleatório entre 1 e n_{pop} . Um terceiro número aleatório j ($1 < j < L-1$), no caso do cruzamento com um único ponto de corte, é necessário para determinar a posição de corte na cadeia de *bits*. Finalmente, o material genético do lado direito dessa posição de corte é permutada entre eles.

Esse processo é ilustrado na Tabela 11 para dois indivíduos apresentados em código binário com comprimento $l = 5$ e posição de corte $j=2$. O resultado do processo são dois novos indivíduos, gerando-se duas novas possíveis soluções. Este é o principal mecanismo na produção de novos pontos a serem testados.

Tabela 11: Representação do Cruzamento

Indivíduo	Indivíduo após Cruzamento
0 1 1 0 1	0 1 0 0 0
1 1 0 0 0	1 1 1 0 1

b) Mutação:

A mutação altera *bits* em uma cadeia de caracteres com a probabilidade de mutação p_m . Esta operação protege o processo de busca da solução contra perdas de valiosas características genéticas, que possam ter ocorrido durante as operações de reprodução e cruzamento. A mutação, introduz novas informações, no âmbito da população, permitindo que novos pontos sejam testados, aumentando assim, a probabilidade de se encontrar o ótimo global. A implementação da mutação consiste apenas na mudança do valor do *bit* escolhido.

Considere o indivíduo apresentado na Tabela 12. Supondo que o quarto *bit* ($j = 4$) esteja sob consideração, gera-se um número aleatório entre 0 e 1. Se esse número é inferior à

probabilidade p_m , a operação é executada e o valor do *bit* é modificado para se obter o novo indivíduo.

Tabela 12: Representação da Mutaç o

Indiv�duo	Indiv�duo ap�s Mutaç�o
0 1 1 0 1	0 1 1 1 1

4.7 PAR METROS GEN TICOS

Os par metros gen ticos s o as entidades que determinam o desempenho dos AGs, que, geraç o ap s geraç o, tentam encontrar indiv duos com melhor capacidade que os anteriores.

- a) *Tamanho da Populaç o*: O tamanho da populaç o afeta a efici ncia do algoritmo. Uma populaç o pequena acarreta um fraco desempenho, podendo conduzir o algoritmo na direç o de um m nimo local. Em contrapartida, uma populaç o grande desencoraja a converg ncia prematura para soluç es locais, por m, pode resultar em um tempo computacional inaceit vel, tendo em vista que a mesma exige um maior n mero de avaliaç es da funç o objetivo. Os AGs usam tipicamente um tamanho de populaç o entre 10 e 200.
- b) *Probabilidade de cruzamento (p_c)*: A probabilidade de cruzamento,   um valor percentual do n mero de indiv duos que experimentam cruzamento em relaç o ao n mero total de indiv duos de uma populaç o. Uma alta probabilidade de cruzamento, em um GA sem elitismo, faz com que indiv duos, com alto desempenho, sejam eliminados antes mesmo que a seleç o possa produzir aperfeiçoamentos. Por outro lado, uma baixa probabilidade de cruzamento pode convergir lentamente devido   baixa taxa de exploraç o das caracter sticas gen ticas. A probabilidade de cruzamento, sem adaptaç o din mica, utilizada pelo GA varia entre 0,5 e 1,0.
- c) *Probabilidade de mutaç o (p_m)*: A probabilidade de mutaç o,   um valor percentual do n mero de *bits* escolhidos aleatoriamente para serem alterados em relaç o ao n mero total de *bits* da populaç o. Uma baixa probabilidade de mutaç o n o permite a restauraç o da diversidade gen tica da populaç o, ao passo que uma alta probabilidade de mutaç o resulta em uma busca essencialmente aleat ria. A probabilidade de mutaç o, sem adaptaç o din mica, utilizada pelo AG varia entre 0,005 e 0,05.

4.8 ESTRAT GIAS ELITISTAS E ADAPTAÇ O DIN MICA

Ap s a aplicaç o dos operadores gen ticos, pode acontecer que os indiv duos da nova geraç o n o apresentem melhores resultados que o melhor da geraç o anterior. Se isso

ocorrer, o valor da função aptidão cai do patamar que já havia alcançado. Outro problema surge em gerações mais avançadas, onde a população contém várias cópias de bons indivíduos. Se por acaso surgir um indivíduo melhor que os atuais, este corre o risco de não ser selecionado, justamente pela grande probabilidade proporcionada aos indivíduos com mais cópias. Neste caso, para a composição da nova geração, pode-se aplicar estratégias elitistas para evitar estes problemas.

As estratégias elitistas, implementadas neste trabalho, visam substituir os dois piores indivíduos da geração nova por dois outros, que são o melhor da geração anterior e um dos filhos do cruzamento deste último com o melhor da nova geração.

Os AGs podem ficar presos numa região de ótimo local, independente de se ter aplicado estratégias elitistas. Essa possibilidade aumenta à medida em que a diversidade genética no meio da população reduz. Se não houver nenhum mecanismo que restaure essa diversidade, a convergência para o ótimo global pode ficar comprometida. Para se conseguir resultados satisfatórios, deve haver diversidade genética suficiente no âmbito da população, durante as gerações, de tal modo a permitir que outras regiões, além da definida pelo indivíduo de melhor *performance*, sejam representadas na população. Com isso, problemas de convergência prematura podem ser evitados.

Assim, com o objetivo de manter um nível suficiente de diversidade genética, aplica-se a adaptação dinâmica às probabilidades de cruzamento e mutação. Os valores das probabilidades de cruzamento e mutação são modificados observando a medida da diversidade genética m_{dg} da população.

$$m_{dg} = \frac{ff_{med}}{ff_{max}} \quad (8)$$

em que ff_{med} é o desempenho médio da população e ff_{max} o melhor resultado da população.

Se m_{dg} é próximo da unidade, significa que há pouca diversidade, e muita, quando ele se aproxima de zero. A Figura 19 mostra que, na medida em que o número de gerações cresce, o valor médio da função desempenho se aproxima do valor máximo encontrado, o que dá valores de m_{dg} próximos da unidade, indicando escassez de material genético dentro da população.

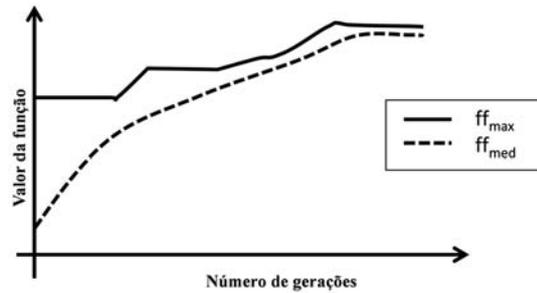


Figura 19: Comportamento da População no Decorrer das Gerações

Procura-se, portanto, encontrar valores ideais para V_{dg}^{\min} e V_{dg}^{\max} tais que, na faixa compreendida entre eles, o nível de diversidade genética seja aceitável. Ou seja, se o valor de m_{dg} estiver abaixo de um valor V_{dg}^{\min} , considera-se que há grande diversidade genética. Nesse caso, deve-se diminuir a taxa de introdução de novas características genéticas no seio da população. Por outro lado, se m_{dg} for maior que um valor V_{dg}^{\max} , considera-se que há pouca diversidade genética, o que pode proporcionar uma convergência prematura. Para contornar essa situação, aumenta-se a taxa de introdução de novas características, aumentando a possibilidade de se encontrar o ótimo global. Para manter m_{dg} dentro do intervalo entre V_{dg}^{\min} e V_{dg}^{\max} , será utilizado o critério de adaptação *Fora da Faixa* - FF.

A adaptação FF pode ser implementada da seguinte forma. Quando $m_{dg} > V_{dg}^{\max}$, aumenta-se p_m e simultaneamente diminui-se p_c , pois, com p_m maior, há mais inserção de material genético novo na população e com p_c baixo, há pouca troca de material genético. Por outro lado, quando $m_{dg} < V_{dg}^{\min}$, faz-se o contrário. A alteração dos valores das probabilidades é percentual, agindo quando os valores de m_{dg} são respectivamente superiores e inferiores a V_{dg}^{\max} e V_{dg}^{\min} .

4.9 CONCLUSÃO DO CAPÍTULO IV

Esse capítulo mostrou a importância do Algoritmo Genético comparando-o com as metodologias convencionais. Em analogia com os modelos matemáticos, os indivíduos da população são os pontos de partida ou soluções iniciais. É importante destacar, que o AG trabalha sempre com um conjunto de indivíduos. A determinação do número de indivíduos é um parâmetro que deve ser ajustado para cada caso e não existe um parâmetro ótimo, porém, há como se determinar uma faixa de trabalho. Nessa tese, a representação

cromossômica das variáveis de otimização, utilizadas no Algoritmo Genético implementado no MATLAB, são modeladas através do código binário.

CAPÍTULO V – CONFIABILIDADE DO SIEEGN

5.1 INTRODUÇÃO

De acordo com a *International Energy Agency* (2013), a confiabilidade do SIEEGN é uma área muito ampla e está construída em torno da produção do gás e da geração termelétrica. Essa confiança pode ser definida como "manter, constantemente, as luzes acesas". No entanto, esta definição relativamente simples fornece pouco do conhecimento sobre a sua natureza complexa que envolve a integração de tais sistemas.

O conceito da confiabilidade do SIEEGN pode ser avaliado separadamente para ser melhor compreendido e gerido. A confiabilidade neste contexto, inclui a capacidade de produção de gás responsável pela geração termelétrica necessária para atender todos os usuários conectados dentro dos padrões aceitáveis e nos montantes desejados. Dessa forma, três requisitos são fundamentais:

- **Adequação:** a capacidade do SIEEGN em utilizar os recursos existentes para atender as mudanças e para agregar requisitos de modo eventual;
- **Segurança do Sistema:** capacidade do SIEEGN em utilizar os recursos existentes, para manter o abastecimento em um nível seguro em face das restrições, por exemplo, a perda de pressão em componentes da rede de gasodutos ou rápidas mudanças na demanda por energia elétrica;
- **Segurança do Combustível:** Refere-se às questões relacionadas com a manutenção do acesso a fontes de combustível para geração de energia. Um gasoduto inoperante pode impactar no fluxo de gás e, conseqüentemente, na geração termelétrica.

5.2 A EVOLUÇÃO DO CONCEITO DE SEGURANÇA EM SISTEMAS ELÉTRICOS

A partir do blackout na cidade de Nova York em 1965, o conceito de segurança nos sistemas elétricos passou a ser estabelecido no sentido de representar a habilidade dos sistemas em continuar sua operação, respeitando seus limites operacionais de segurança, mesmo após a ocorrência de falhas em alguns de seus componentes (BALU et al, 1992). Segundo LIMMER (1966), após este evento, foi mencionado, pela primeira vez na literatura, o termo “segurança”.

Também nesta época foi estabelecida a estrutura básica da classificação de estado de operação dos Sistemas Elétricos de Potência (DY LIACCO, 1967). Dy Liacco definiu o sistema capaz de operar em três condições: Normal, Emergência e Restaurativo. Em 1969 foi acrescentado a essa classificação o estado Alerta (CIHLAR, 1969). Para a análise de

segurança em regime permanente, o estado Alerta é definido quando o sistema está operando dentro de seus limites de segurança, mas uma contingência pode levá-lo a uma condição de operação violada.

Monticelli et. al. (1987) enriquecem o conceito de segurança com as definições de modo de segurança preventivo e modo de segurança corretivo, quando é levado em consideração o fato do sistema elétrico suportar e operar com limites de operação violados durante certo tempo, antes da atuação dos equipamentos de proteção.

Atualmente, a análise de segurança em sistemas elétricos de potência está dividida em duas condições de análise:

- Análise Estática (regime permanente);
- Análise Dinâmica;

Na avaliação de segurança estática, as perturbações, como perda de unidades de geração e de transmissão, são processadas levando em conta que a transição entre os estados atinge sempre um equilíbrio estável. Este estado pode, contudo, nunca ser alcançado, devido aos efeitos de saída em cascata associado a problemas de instabilidade, onde uma análise dinâmica é recomendada. A avaliação de segurança dinâmica visa examinar o comportamento dinâmico do sistema logo após um distúrbio, para identificar casos que possam provocar efeitos em cascata que levem o sistema a não atingir um estado estável (MENDES,1999).

Um sistema de energia elétrica, quando operando em regime estacionário, está sujeito a três tipos de restrições: de carga, de operação e de segurança.

- *As restrições de carga* representam as injeções de potência ativa e reativa nas barras de carga e as injeções de potência ativa nas barras de geração.
- *As restrições de operação* representam os limites impostos às magnitudes das tensões nas barras, aos fluxos de potência aparente nas linhas de transmissão e transformadores, e às injeções de potência reativa nas barras de geração.
- *As restrições de segurança* estão associadas a um conjunto preestabelecido de contingências possíveis no sistema como: perda de linha, geradores, capacitores ou reatores em derivação, entre outros equipamentos.

Esta lista de contingências, em geral, inclui contingências simples, ou seja, perda de um equipamento por vez, porém podendo conter algumas contingências múltiplas. Não existem critérios bem estabelecidos para a elaboração destas listas. Pode-se usar alguma ferramenta computacional, como a seleção automática de contingências, mas também costuma-se levar em consideração a experiência do operador e as probabilidades de ocorrências de falhas. Para a Análise de Contingências, a lista de contingências é

normalmente formulada levando-se em consideração a probabilidade de ocorrência de um determinado evento (MONTICELLI, 1983).

Uma excelente abordagem para a seleção de contingências é a utilização de métodos automáticos de seleção de contingência, ou também chamados de *Contingency Screening*. Neste caso, é realizada uma pré-análise do efeito da contingência, sendo atribuído um índice de desempenho que será utilizado para ordenação da lista a partir dos casos mais críticos para os menos críticos (WOOD e WOLLENBERG, 1996).

5.3 AVALIAÇÃO DE SEGURANÇA EM SISTEMAS ELÉTRICOS

O objetivo básico de um sistema de energia elétrica é atender a todas as cargas existentes numa dada área de serviço. O sistema deve ser projetado e operado de forma a atender alguns requisitos de qualidade. Deve disponibilizar energia em todos os pontos exigidos pelos consumidores, estando apto para atender à demanda variável. A energia fornecida deve possuir frequência constante, tensão constante e alta confiabilidade. Além disso, o sistema deve operar com custos mínimos, tanto econômicos como ambientais (ELGERD, 1977).

A constante busca pelo aumento da confiabilidade, redução dos custos e uma operação mais otimizada deu origem às interligações entre sistemas. Foram interligados os pequenos sistemas isolados através de linhas de transmissão mais longas, formando assim os sistemas interligados. As interligações possibilitaram otimização dos investimentos e um melhor aproveitamento eletroenergético.

A segurança de um sistema elétrico pode ser definida, de forma geral, como a característica deste sistema que lhe permite permanecer em estado de equilíbrio, em condições normais de funcionamento, e retornar a um estado aceitável após ter sido submetido a uma determinada perturbação (KUNDUR, 1994).

A avaliação de segurança em sistemas elétricos de potência trata da análise dos impactos no sistema, quando da ocorrência das possíveis contingências a que ele está sujeito. Os principais benefícios da avaliação da segurança são: o aumento da confiabilidade operativa e a melhor utilização dos recursos eletroenergéticos do sistema (MONTICELLI, 1983). Dessa forma, pode-se concluir que a **avaliação da segurança** de um sistema elétrico consiste na determinação de uma região que apresente uma lista de contingências, consideradas mais prováveis ou de maior importância, e da consequente verificação das restrições a que o sistema está sujeito.

Sendo assim, o sistema é considerado em um estado seguro de operação quando todas

as restrições são respeitadas. Dessa maneira, ocorre o suprimento total da demanda com todos os limites operativos dentro da faixa permitida. A verificação da restrição de segurança indica que as restrições de carga e operação continuam sendo respeitadas mesmo na ocorrência das contingências definidas (MONTICELLI, 1983).

5.4 SEGURANÇA EM SISTEMAS DE GÁS NATURAL

O gás natural é frequentemente encontrado em lugares onde não há mercado local. Para que o GN esteja disponível ao mercado, ele deve ser explorado, processado e transportado. Em muitos casos, o GN é transportado ao longo de uma substancial distância em tubulações de diferentes tamanhos, sofrendo, dessa forma, uma mudança de pressão durante o seu percurso.

Atualmente, a maior parte do GN transportado tem como destino a geração elétrica. Essa geração de eletricidade geralmente é alimentada por redes de gasodutos. Contudo, há considerações práticas que devem ser observadas da produção até a entrega do GN (GUDMUNDSSON, 1997).

O volume de um gás real é, normalmente, menor do que o volume de um gás ideal. A proporção entre o volume real e o volume ideal, é a medida chamadada de fator de compressibilidade e é comumente abreviado para o símbolo Z . Logo, o fator Z de um gás é, por definição, a razão entre o volume efetivo ocupado por um gás, a uma dada pressão e temperatura, e o volume ocupado para que se tenha o comportamento ideal (Wichert e Aziz, 1972).

A equação do estado do gás real é escrita como

$$PV = ZnRT \quad (9)$$

Onde, P é a pressão, V é o volume, T é a Temperatura absoluta, Z é o fator de compressibilidade, n é o número de moles e R é a constante dos gases.

O valor crítico de pressão produzirá no gás uma temperatura crítica e esse conjunto de variáveis em níveis críticos, provocará um pseudofluxo de gás nos gasodutos, comprometendo todo o sistema de carga de gás (ECONOMIDES et al, 2005). É possível prever esse pseudofluxo estimando o fator de compressibilidade do gás. Quando o fator Z , for menor que 1, há uma baixa pressão e uma elevada temperatura, o que significa que o gás comporta-se como um gás ideal, nestas condições. Em condições normais ou atmosféricas o fator Z do gás é sempre igual a 1. Segundo a teoria do estado do gás, o fator Z pode ser definido em função da pressão reduzida, portanto, o monitoramento e o controle da pressão de

um gás em uma rede de gasodutos são imprescindíveis à segurança desse sistema (WICHERT e AZIZ, 1972).

São muitas as vantagens da geração termelétrica a gás natural, como por exemplo, o prazo relativamente curto de maturação do empreendimento e a flexibilidade para o atendimento de cargas de ponta. Por outro lado, as turbinas a gás são máquinas extremamente sensíveis às condições de operação, principalmente em relação à quantidade GN utilizada, pois apresentam alterações substanciais de rendimento térmico no caso de operação em cargas parciais (NRCan, 2004).

Apesar dos ganhos alcançados no rendimento térmico das turbinas a gás operando em ciclo simples, seu desempenho tem sido prejudicado pela perda de energia nos gases de exaustão. Entre outras tecnologias empregadas na recuperação dessa energia, destaca-se a de ciclo combinado, por meio da geração de vapor e da produção de potência adicional. Tem-se, assim, uma combinação dos ciclos de turbinas a gás e turbinas a vapor, por meio de trocadores de calor, nos quais ocorre a geração de vapor, aproveitando-se a energia dos gases de exaustão da turbina a gás. Esse processo ainda pode ser melhorado com a queima de combustível suplementar, principalmente quando há disponibilidade de combustíveis residuais. Todavia, caso haja uma variação na pressão nos nós produtores de gás capaz de provocar um fluxo de gás inferior ao mínimo exigido, o rendimento das termelétricas ficará comprometido em razão da menor quantidade de GN enviada (NRCan, 2004).

Ainda de acordo com NRCan (2004), existe um risco de condensação do GN nas paredes internas do gasoduto devido ao fluxo de gás. Assim, as pressões nos nós produtores, dentro das características do melhor escoamento, são necessárias para uma operação segura no sistema de gasodutos. A pressão e o fluxo do gás em gasodutos estão diretamente relacionados com a produção do GN. Os níveis de pressão e as características do fluxo de GN têm uma natureza não linear, e uma abordagem que identifique níveis seguros de produção do gás evita uma queda, ou variação da pressão e uma oscilação muito acentuada no seu fluxo.

5.5 CONCLUSÃO DO CAPÍTULO V

O presente capítulo mostrou que o SIEEGN é projetado para trabalhar sob determinados limites operacionais, e dispositivos de controle automáticos podem retirá-lo de operação, caso haja uma violação nesses limites. Um evento, como a interrupção de um ramo no gasoduto, pode provocar uma série de ações de retirada de outros equipamentos, e esse processo de falhas em cascata, pode vir a produzir um colapso em todo o SIEEGN.

CAPÍTULO VI – MÉTODO DE DESPACHO PROPOSTO

6.1 INTRODUÇÃO

Pelo fato de não se conseguir precisar ou estabelecer o momento exato em que ocorrerá uma falha no SIEEGN, o mesmo deve ser operado de modo a que nunca passe por uma condição insegura, caso algum provável evento venha a ocorrer. Como os equipamentos de um SIEEGN são projetados para trabalhar sob determinados limites, estes são protegidos por dispositivos automáticos que podem retirá-los de operação, caso haja uma violação de limites operativos. Este evento pode ser seguido de uma série de ações de retirada de outros equipamentos, e se esse processo de falhas em cascata continua é possível que todo o sistema possa entrar em colapso ou *blackout*.

Um exemplo clássico desse tipo de evento que pode ocasionar um *blackout* é dado por uma simples abertura de linha devido a uma falha nos isoladores. Nesta condição, os circuitos remanescentes do sistema vão servir de caminho para a energia que estava fluindo pela linha agora aberta. Se uma das linhas remanescentes se tornar muito carregada, ela pode abrir pela ação de um relé de proteção, causando um carregamento maior nas linhas remanescentes e assim por diante. Esse tipo de processo é chamado de “desligamentos em cascata” (WOOD e WOLLENBERG, 1996).

Na rede gás, por exemplo, a simples diminuição da pressão em um nó ativo, comprometerá toda a distribuição de GN, para fins elétricos e não elétricos. Da mesma forma que a interrupção em um dos ramos na rede de gasodutos poderá alterar todo o despacho em um sistema de gás natural (MANSELL, 1991).

6.2 DESPACHO SEGURO - ESQUEMA PROPOSTO

O despacho coordenado da geração termelétrica e da produção de gás natural consiste em ajustar tanto a potência ativa entre os geradores térmicos, quanto os níveis de produção de gás natural entre os poços de gás, a fim de proporcionar a segurança das Redes Integradas de Energia Elétrica e Gás Natural (RIEEGNs), quando um conjunto de contingências críticas é identificado pelo módulo de avaliação de segurança.

Neste trabalho, as regiões de segurança, relativas às variáveis de controle das RIEEGNs, proporcionam margens de segurança, que servem para orientar a operação e garantir a segurança do sistema. Desta forma, os conhecimentos obtidos pela AD podem ser diretamente utilizados no processo de despacho ótimo, avaliando de forma mais segura, as contingências no sistema das RIEEGNs.

As áreas de segurança são calculadas por meio da formação da AD, a partir de uma base de dados que considera vários pontos de operação das RIEEGNs e o estado do sistema (seguro/inseguro) para estas condições de operação. O esquema proposto neste trabalho é resumido da seguinte forma, como exposto nos itens a seguir:

6.2.1 ANÁLISE DE CONTINGÊNCIAS

A maioria dos operadores das RIEEGNs exige que alguns critérios específicos de segurança, como $(n-1)e(n-2)$, sejam satisfeitos. Para conhecer quais as principais contingências de segurança são prejudiciais ao sistema, foram realizadas análises de contingências no sistema elétrico de potência. Neste trabalho foi utilizado o software de fluxo de carga ANAREDE (CEPEL, 1999).

As técnicas para a análise de contingência baseadas na análise de fluxo de gás podem também ser aplicadas para detectar possíveis problemas na rede de gás. Neste trabalho, a análise de contingência foi considerada apenas para a rede elétrica.

6.2.2 GERAÇÃO DA BASE DE DADOS

A aplicação dos algoritmos da AD, fundamenta-se na base do conhecimento anterior, que descreve as características das RIEEGNs. Para o despacho de segurança aplicado nas RIEEGNs, o banco de dados deve cobrir pontos de operação que sejam suficientes para aproximar o espaço do ponto de operação das RIEEGNs, podendo resistir às mudanças incertas da condição operacional.

Para ser efetivo, o banco de dados deve refletir a região de operação possível e representativa no qual as RIEEGNs deverão experimentar. Na prática, o banco de dados pode ser adquirido a partir de arquivos históricos e/ou ser gerados por meio de exaustivas simulações off-line. Neste trabalho, o banco de dados utilizado foi gerado por simulação, conforme mostrado na Figura 20.

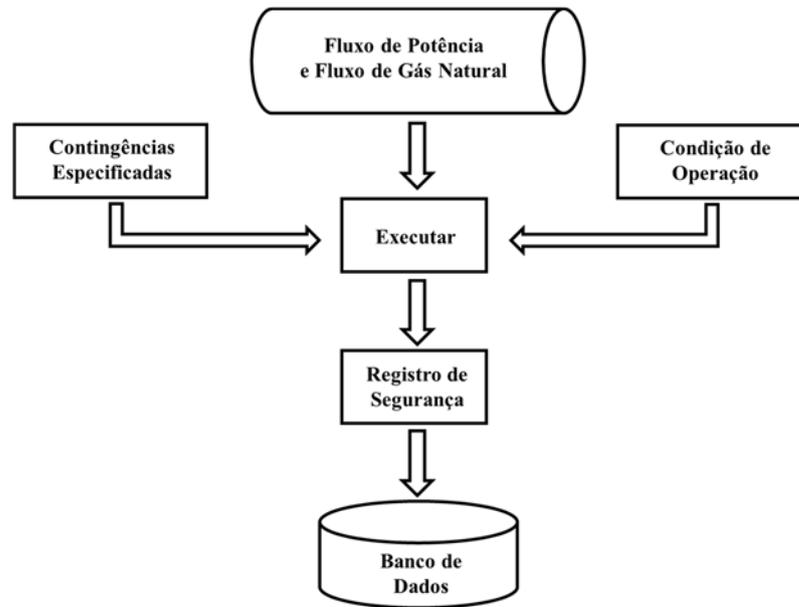


Figura 20: Esquema de geração do Banco de Dados.

O banco de dados é gerado pela execução do fluxo de carga e rotinas de fluxo de gás, considerando-se uma gama de pontos de operação e contingências especificadas nas RIEEGNs. As potências ativa e reativa demandadas em cada barra de carga e o gás natural demandado em cada nó estão uniformemente distribuídos dentro de uma faixa de 3% do ponto de operação inicial.

A fim de adquirir um banco de dados abundante, os perfis de demanda de carga do sistema e de alimentação de gás natural, podem ser estendidos com a finalidade de enriquecer a base de dados. Para cada contingência definida nas RIEEGNs é gerado um banco de dados que consiste de valores numéricos das variáveis de controle (atributos candidatos), e os resultados da classificação, que representam a segurança do sistema, são utilizados para obter a AD.

Os atributos candidatos são: produção de gás natural em poços de gás, pressões nos nós de produção, os fluxos de gás em dutos que possuem compressor, geradores de potência ativa e magnitudes de tensão. O rótulo de segurança para as variáveis de controle é introduzido na base de dados: seguro ou inseguro. Esse rótulo depende, respectivamente, se as restrições são satisfeitas ou não.

6.3 ÁRVORE DE DECISÃO E REGIÃO DE SEGURANÇA

Para fins de classificação, a AD é uma ferramenta de aprendizagem de máquina, supervisionado, destinado a resolver problemas com dados de alta dimensionalidade. O princípio fundamental é a obtenção de um modelo preditivo para classificar uma meta usando

os atributos que contribuem diretamente para este objetivo. A AD converte um processo de classificação complexa em algumas declarações lógicas *If-Then*, de acordo com os limites da entrada de atributos ou as suas combinações lineares.

Para que o treinamento de uma AD resulte em uma ferramenta com bom desempenho, é necessário, em primeiro lugar, construir um banco de dados constituído por um número suficientemente grande de casos. Cada caso é representado por um objetivo alvo (por exemplo, seguro ou inseguro), e atributos, tais como níveis de geração de potência ativa e potência reativa, os volumes de produção de gás em poços, etc. A AD é então projetada para representar um modelo que identifique os atributos críticos que afetam o objetivo alvo de forma mais eficaz e direta. O modelo da AD tem uma estrutura binária com dois tipos de nós: o nó interno com dois sucessores e o nó terminal sem qualquer sucessor.

Para cada nó terminal, também chamado de *folha*, um resultado de classificação será atribuído de acordo com o objetivo da classe majoritária, por exemplo, seguro ou não seguro. O processo de classificação começa a partir do nó raiz e termina em um nó terminal, no qual é alcançado o resultado de triagem.

Após a criação da AD, um processo de poda é realizado para a remoção dos nós desnecessários e, finalmente, reduzir o tamanho final da AD. O algoritmo de classificação e regressão da AD usado neste trabalho é conhecido como CART (BREIMAN *et al.*, 1984).

Um típico exemplo do modelo CART é apresentado na Figura 21, mostrando o resultado de uma série de separação de nós. Dado um caso representado pelo conjunto de variáveis (x_1, x_2, \dots) para uma condição operacional em particular, a classe (seguro ou inseguro) do processo pode ser prevista, deixando as variáveis do caso se dirigirem, a partir do nó raiz em direção a um nó terminal de uma AD.

A AD cresce e é dividida recursivamente a partir dos casos de aprendizagem em seus nós. A ideia fundamental da seleção de cada divisão é, de tal modo, que os processos de aprendizagem em cada nó descendente são mais puros do que o nó *pai*. A seleção ideal de regras de decomposição pode ser calculada por repetidas tentativas de minimizar o índice de impureza GINI index (BREIMAN *et al.*, 1984).

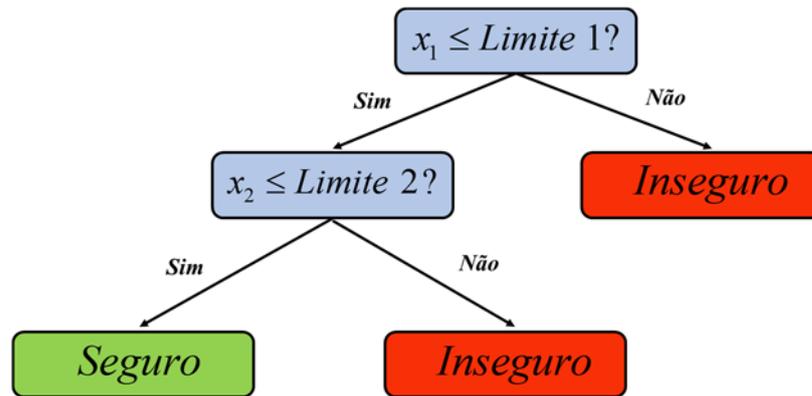


Figura 21: Uma típica representação do Cart.

Em que a impureza do nó é máxima quando todas as classes possuem igual distribuição e mínima quando existe apenas uma classe.

Considerando um conjunto de dados S , que contém n registros, cada um com uma classe A , o índice *Gini* de S é dado pela equação (10) adaptada da equação (4),

$$Gini(S) = 1 - \sum_{i=1}^m p_i \left[\frac{A}{n} \right]^2 \quad (10)$$

Sendo:

p_i a probabilidade relativa da classe A em S .

n o número de registros em S .

m o número de classes.

Se S for particionado em dois subconjuntos S_1 e S_2 , um para cada ligação, o índice *Gini* dos dados particionados será dado pela equação (11),

$$Gini(S | A) = \frac{n_1}{n} Gini(S_1) + \frac{n_2}{n} Gini(S_2) \quad (11)$$

Sendo:

n_1 o número de exemplos de S_1 .

n_2 o número de exemplos de S_2 .

O Treinamento *off-line* de uma AD seleciona atributos críticos do sistema como bons indicadores de segurança e ajuda a construir as regiões que realçam o aprimoramento da situação do sistema. Os limites da AD fornecem as regiões de segurança que definem as regiões operacionais, conforme mostra a Figura 22.

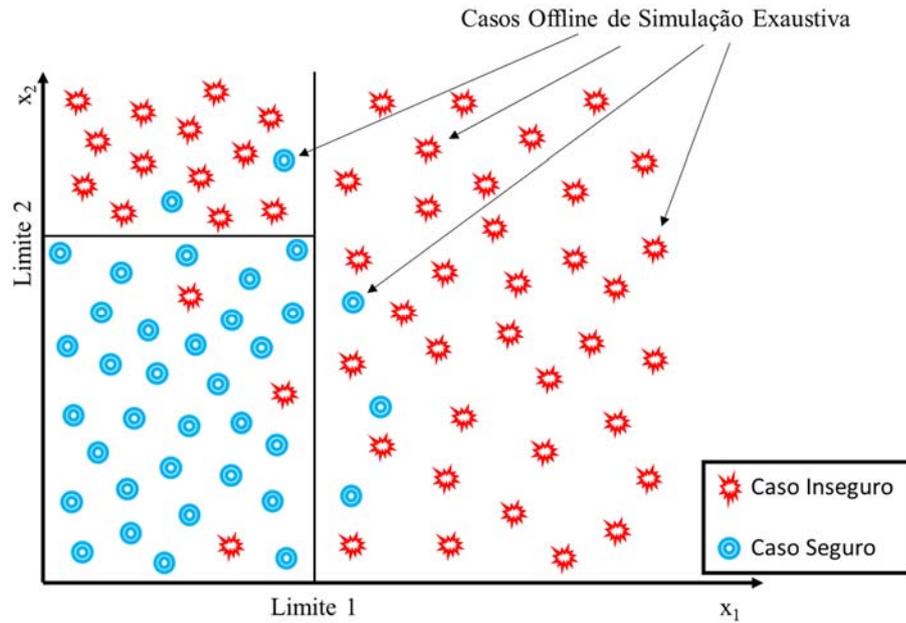


Figura 22: Região de Segurança definida pela AD.

A região de segurança estabelecida pela AD é definida pela interseção das regiões de segurança declaradas. Com regiões de segurança e os seus limites definidos pela AD, a margem de segurança pode ser definida pela menor distância entre o ponto de funcionamento e o limite de segurança estabelecido.

Dois tipos de Árvores de Decisão podem ser usados para identificar as regiões de segurança e os seus contornos: a ortogonal e a oblíqua. As Árvores de Decisão ortogonais dividem as regras em regiões retangulares que são chamados de hiperplanos e são ortogonais ao eixo para que possam associar cada região a uma classe; as Árvores de Decisão oblíquas dividem as regras, que são definidas por combinações lineares de atributos e seus limites, em regiões não retangulares, reduzindo drasticamente o tamanho da AD.

As Árvores de Decisão que cobrem o critério de partição, dividem o espaço P de parâmetros relacionados em atributos críticos. Para uma AD ortogonal, cada regra de separação k define uma divisão R_k de parâmetros pertencentes a P tais como:

$$R_k = \{P \in P : P_i^k \leq P_i \leq \bar{P}_i^k, i \in A\} \quad (12)$$

Sendo A um conjunto de atributos críticos. Para uma AD oblíqua, cada regra de divisão é definida como uma partição.

$$R_k = \left\{ P \in P : \sum_{i \in A} \alpha_i^k P_i \leq \beta_k, \alpha_i^k \in A, i \in A \right\} \quad (13)$$

Sendo P_i^k , \bar{P}_i^k e β_k os limites ótimos dos valores calculados.

Cada nó terminal de uma AD determina uma sub-região segura ou insegura de P , que é uma interseção das divisões previamente definidas. Dessa forma, cada AD divide o espaço de parâmetros em duas regiões, uma segura e outra insegura, que são a união das sub-regiões definidas pelos nós terminais correspondentes.

6.4 DESENVOLVIMENTO E TREINAMENTO DA ÁRVORE DE DECISÃO

A AD é treinada a partir de um banco de dados de casos correspondentes a um conjunto de contingências críticas das RIEEGNs. A AD identifica os atributos (variáveis de controle) que são mais relevantes e que afetam a segurança do sistema e também fornece regras que descrevem, de forma transparente, a relação entre as variáveis de controle e a segurança das RIEEGNs no que diz respeito a maioria dos conjuntos de prováveis contingências. Neste trabalho foi utilizado k – vezes o algoritmo de validação cruzada na formação da AD.

6.5 FLUXO DE POTÊNCIA E DE GÁS NATURAL OTIMIZADO COM RESTRIÇÕES DE SEGURANÇA

Se as RIEEGNs estão operando em uma região insegura ou com uma pequena margem de segurança, os ajustes preventivos para a produção de gás natural e geração de energia elétrica devem ser aplicados de tal forma que o ponto de operação será movido para uma região segura, minimizando os custos operacionais totais para o novo ponto de operação. Na operação de otimização é utilizado o método do Algoritmo Genético.

O problema de despacho ótimo de energia elétrica e fluxo de gás natural, via AG, com a geração de energia e minimização de custos de produção de gás natural sob restrições operacionais e de rede pode ser declarado como

$$\min f(x_1, x_2, u_1, u_2) \quad (14)$$

$$\text{sujeito a } g_1(x_1, u_1) = 0 \text{ e } g_2(x_2, u_2) = 0 \quad (15)$$

$$h_1(x_1, u_1) \leq 0 \text{ e } h_2(x_2, u_2) \leq 0 \quad (16)$$

Sendo $f(x, u)$ a função do custo de geração de energia e produção de gás natural, $g_1(x_1, u_1) = 0$ e $g_2(x_2, u_2) = 0$ são as restrições de igualdade que representam as equações de energia e fluxo de gás natural, respectivamente. $h_1(x_1, u_1) \leq 0$ e $h_2(x_2, u_2) \leq 0$ são as restrições de desigualdade que descrevem as restrições operacionais das redes de eletricidade e de gás natural, respectivamente. As variáveis dependentes e de controle são indicados por x

e u , respectivamente (O subscrito 1 denota a rede de energia elétrica e o subscrito 2 denota a rede de gás natural).

a) Restrições operacionais nas redes de energia elétrica e de gás natural:

A geração de energia elétrica em cada termelétrica é delimitada por limites máximos e mínimos como:

$$P_{gi \min} \leq P_{gi} \leq P_{gi \max} \quad \forall i \quad (17)$$

Do mesmo modo, a produção de gás é delimitada por qualquer das características físicas ou montante contraído como:

$$Pw_j \min \leq Pw_j \leq Pw_j \max \quad \forall j \quad (18)$$

b) Restrições nas redes de energia elétrica e gás natural:

Ambas as redes, de energia elétrica e de gás natural, devem promover o equilíbrio entre oferta e demanda em cada nó. A equação (19) inclui o gás natural, na integração das redes de energia elétrica e gás natural, consumido por cada unidade de energia.

$$\sum fg_p + \sum Pw_j + \sum nd_j - \sum P_{gi} = GL_j \quad (19)$$

O balanço de potência elétrica em cada nó é dado por:

$$P_{gi} - V_i \sum_k V_k (G_{ik} \cos \theta_{ik} + B_{ik} \text{sen} \theta_{ik}) = P_{Li} \quad (20)$$

$$Q_{gi} - V_i \sum_k V_k (G_{ik} \text{sen} \theta_{ik} - B_{ik} \cos \theta_{ik}) = Q_{Li} \quad (21)$$

$$V_i \min \leq V_i \leq V_i \max \quad \forall i \quad (22)$$

A modelagem das variações de pressão em gasodutos é baseada na relação não-linear de Weymouth, dada por:

$$\text{sign}(fg_p) \cdot fg_p = C_p^2 (ps_i^2 - ps_j^2) \quad (23)$$

Sendo:

$P_g \rightarrow$ Geração de Potência Elétrica (MW);

$P_w \rightarrow$ Produção de Gás Natural (Mm³);

$fg \rightarrow$ Fluxo de Gás Natural (m³/h);

$nd \rightarrow$ Gás Natural não-entregue (m³);

$GL \rightarrow$ Demanda de Gás Natural (Mm^3);

$V \rightarrow$ Magnitude da tensão (p.u.);

$\theta \rightarrow$ Diferença de fase da tensão ($^\circ$);

$P_L \rightarrow$ Carga elétrica (MW);

$G \rightarrow$ Condutância (Ω^{-1});

$B \rightarrow$ Susceptância (Ω^{-1});

$C_p \rightarrow$ Constante de Weymouth;

$ps \rightarrow$ Pressão (bar);

$i, k \rightarrow$ Nós de potência;

$j, l \rightarrow$ Nós do gasoduto;

$p \rightarrow$ Gasoduto;

$w \rightarrow$ Poço produtor de Gás Natural.

Uma vez que as regiões de segurança são determinados pelas ADs, as restrições de segurança podem ser facilmente integradas no problema de fluxo ótimo de energia e de gás natural na forma algébrica de restrições de desigualdade, via AG.

Na abordagem de reescalonamento de geração de energia e produção de gás natural, as variáveis de controle são: geração de potência ativa, a magnitude de tensão em barras de geração, a produção de gás natural, o fluxo de gás em dutos que possuem compressor e a pressão em poços de gás natural. Assim, para demonstrar a abordagem proposta, um fluxo ótimo de energia elétrica e de gás natural é usado para resolver a função objetivo

$$f(u_1, u_2) = \sum_{i=1}^{ng} (a_i P_{gi}^2 + b_i P_{gi} + c_i) + \sum_{j=1}^{mw} (a_j P_{wj}) \quad (24)$$

Sendo:

a_i, b_i e $c_i \rightarrow$ Constantes;

$P_{gi} \rightarrow$ Geração de Potência Elétrica (MW);

$ng \rightarrow$ Número de Geradores;

$a_j \rightarrow$ Constante;

$P_{wj} \rightarrow$ Produção de Gás Natural (Mm^3)

$nw \rightarrow$ Número de poços de Gás Natural.

Supondo que a região de segurança S é representada pela união das regiões convexas definidas por S_i em cada nó terminal então, cada S_i pode ser derivada das regras de decomposição da AD indicando que o sistema é seguro. Portanto, o problema de fluxo de energia e de gás natural ótimo, sobre o conjunto viável de S_i , pode ser declarado como

$$\min f(x_1, x_2, u_1, u_2) \quad (25)$$

$$\text{sujeito a } g_1(x_1, u_1) = 0 \text{ e } g_2(x_2, u_2) = 0 \quad (26)$$

$$h_1'(x_1, u_1) \leq 0 \text{ e } h_2'(x_2, u_2) \leq 0 \quad (27)$$

$$u_1 \text{ e } u_2 \in S \quad (28)$$

6.6 CONCLUSÃO DO CAPÍTULO VI

Esse capítulo identificou que o algoritmo de classificação e regressão da Árvore de Decisão usado neste trabalho é o CART e que, a região de segurança estabelecida pela AD é definida pela interseção das regiões de segurança declaradas. Também foi mostrado que dois tipos de Árvores de Decisão podem ser usados para identificar as regiões de segurança e os seus contornos: a ortogonal e a oblíqua. Contudo, a ortogonal define de forma mais específica os limites das regiões de segurança, reduzindo substancialmente o tamanho da árvore. Ainda no presente capítulo verificou-se que, os conhecimentos obtidos pela AD podem ser diretamente utilizados no processo de despacho ótimo seguro considerando as contingências no sistema das Redes Integradas de Energia Elétrica e de Gás Natural (RIEEGNs), via AG.

CAPÍTULO VII – ESTUDO DE CASO

7.1 INTRODUÇÃO

A metodologia proposta neste trabalho foi aplicada para obtenção do despacho de energia de um sistema integrado de energia composto por uma rede de gás natural de 15 nós, mostrada na Figura 23 e pelo sistema teste de 118 barras do IEEE, com unidades de gás natural (realçadas na cor vermelha), apresentado na Figura 24 a seguir.

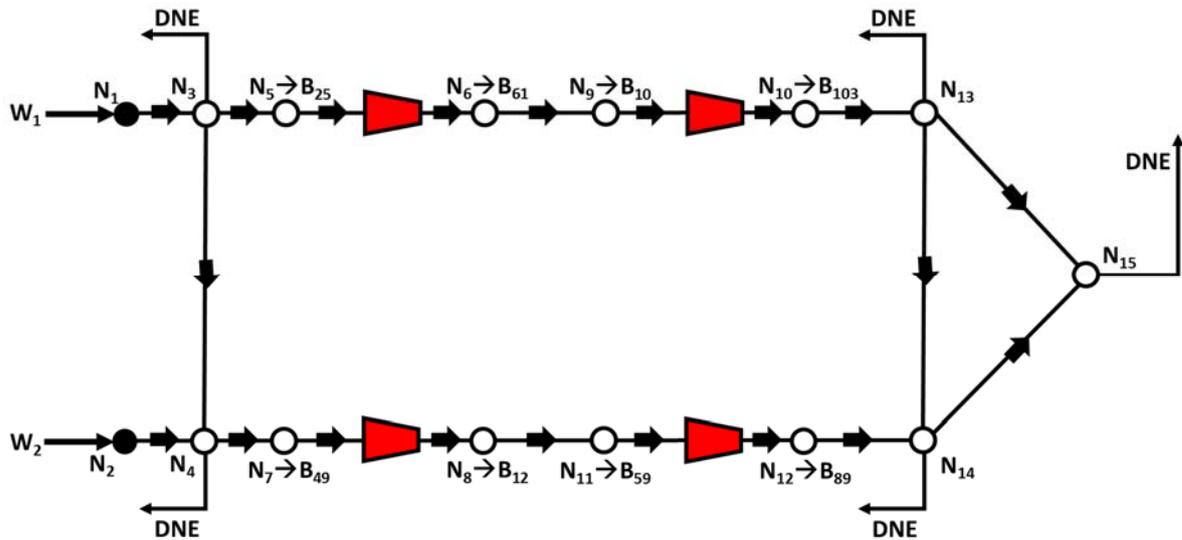


Figura 23: Rede de Gás Natural de 15 nós.

Sendo adotada a seguinte notação:

$W \rightarrow$ Poço produtor de gás natural;

$N \rightarrow$ Nó da rede de gasodutos;

$B \rightarrow$ Barra geradora associada ao nó do gasoduto;

$DNE \rightarrow$ Demanda não-elétrica de gás natural;

 \rightarrow Compressor de gás.

Com base na análise de contingência com soluções exaustivas dos fluxos de gás natural e de energia, algumas contingências críticas específicas são tomadas como exemplos para demonstrar a eficácia do método de despacho seguro proposto para restaurar a segurança do sistema integrado.

No estudo de caso apresentado nesta tese, os pontos de operação da geração de energia e de análise de fluxo de gás natural são definidos com base em um programa desenvolvido em ambiente MATLAB. As ADs foram treinadas e testadas usando o software RapidMiner.

As ADs foram desenvolvidas para identificar os limites de segurança com maior precisão e com estruturas mais simples, facilitando assim a tomada de decisão na implementação do despacho seguro.

A escolha das várias contingências a seguir tanto nos gasodutos quanto no sistema elétrico pretende, conforme descrito anteriormente, demonstrar a consistência da metodologia baseada em AD implementada na presente tese.

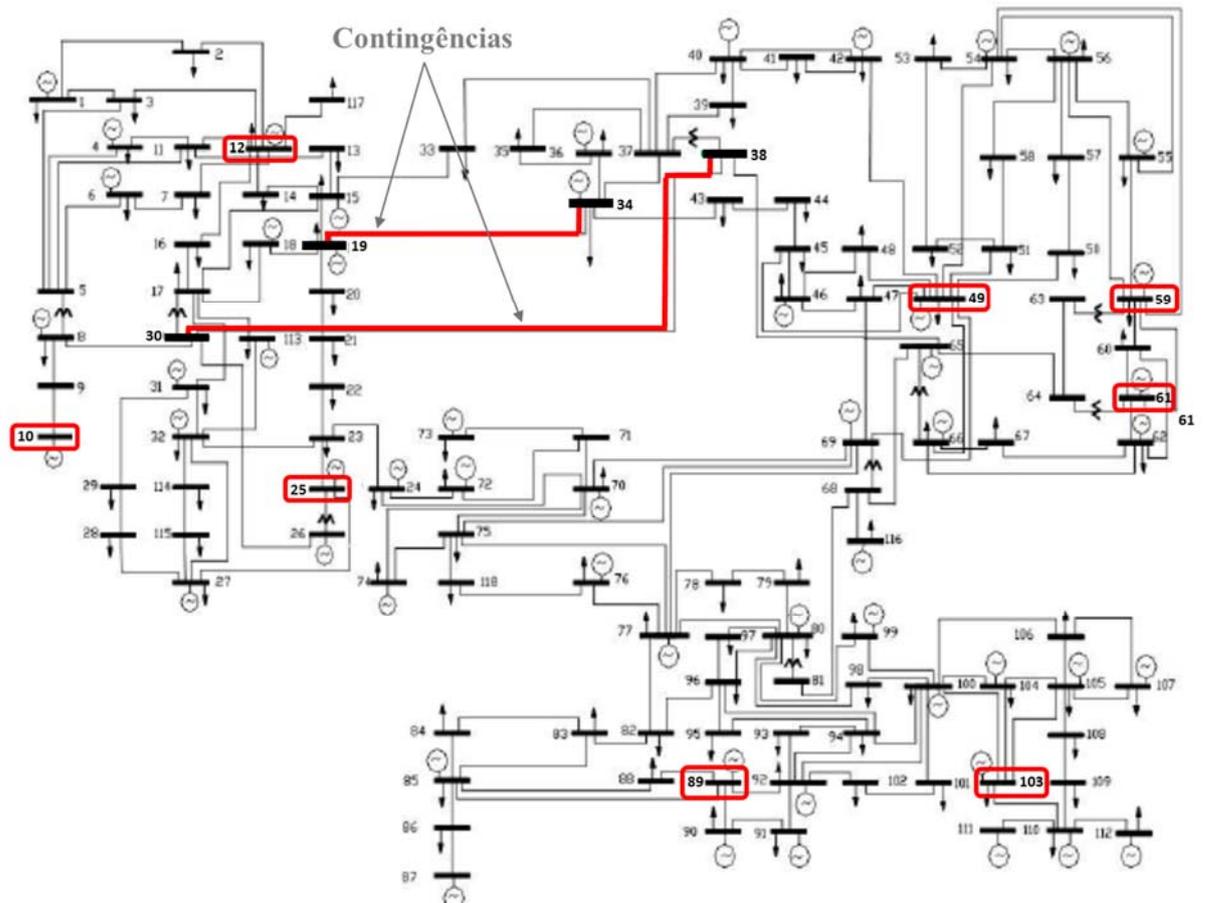


Figura 24: Rede de 118 barras – IEEE.

7.2 CASO 1: CONTINGÊNCIAS NO GASODUTO 3→4 E NA LINHA DE TRANSMISSÃO 30→38

O desligamento do gasoduto 3→4 da rede de gás natural de 15 nós e a desconexão da linha de transmissão 30→38 da rede elétrica de 118 barras são as contingências simultâneas consideradas no estudo do caso 1. A Figura 25, a seguir, ilustra a estrutura da Árvore de Decisão com dois conjuntos de regras que determinam as restrições de segurança da rede de gás e rede elétrica, de acordo com as contingências do presente estudo de caso.

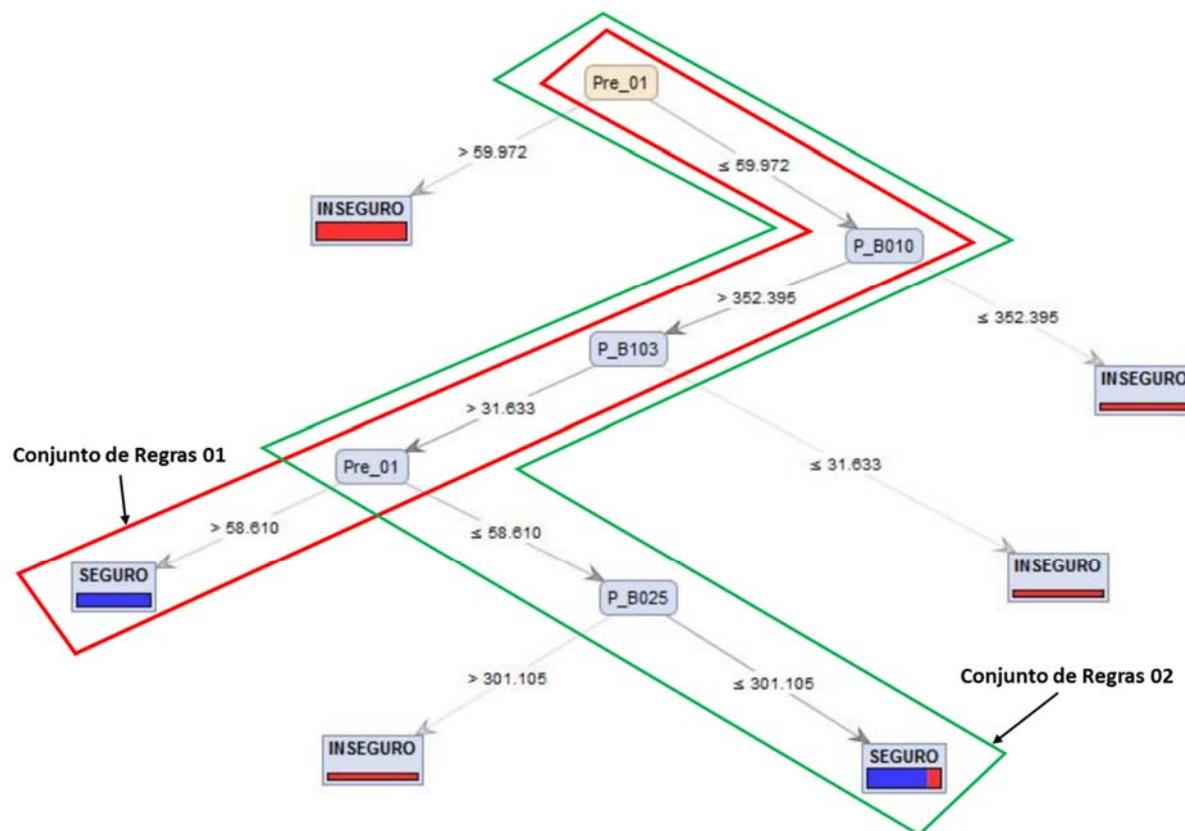


Figura 25: AD para contingências no gasoduto 3→4 e na linha de transmissão 30→38.

As regras da AD da Figura 25 são consideradas como restrições de segurança sobre o problema do despacho ótimo entre as redes integradas de gás e eletricidade. Cada conjunto de regras determina a região de segurança com sub-regiões definidas pelos nós terminais seguro da AD:

Conjunto de Regras 01:

Variável de Controle	Conjunto de Regras
Pressão no nó 01	$Pre_01 \leq 59,972 \text{ bar}$
Potência Ativa na barra 10	$P_B010 > 352,395 \text{ MW}$
Potência Ativa na barra 103	$P_B103 > 31,633 \text{ MW}$
Pressão no nó 01	$Pre_01 > 58,610 \text{ bar}$

Conjunto de Regras 02:

Variável de Controle	Conjunto de Regras
Pressão no nó 01	$Pre_01 \leq 59,972 \text{ bar}$
Potência Ativa na barra 10	$P_B010 > 352,395 \text{ MW}$
Potência Ativa na barra 103	$P_B103 > 31,633 \text{ MW}$
Pressão no nó 01	$Pre_01 \leq 58,610 \text{ bar}$
Potência Ativa na barra 25	$P_B025 \leq 301,105 \text{ MW}$

As regras da AD podem ser reproduzidas por nomogramas (como, por exemplo, a potência ativa na barra 10 X potência ativa na barra 103) para descrever as regiões de segurança, como mostrado nas Figuras 26, 27, 28 e 29.

O asterisco em vermelho em cada nomograma representa um caso inseguro, o retângulo em azul representa a região de segurança e a circunferência em azul representa um caso ótimo dentro da região de segurança. Pode-se observar que cada trajetória de um asterisco para uma circunferência é uma indicação de controle confiável que leva o sistema de um região insegura para uma região segura de operação.

O despacho ótimo coordenado entre a geração termoeletrica e produção de gás natural pode ser realizado como um controle preventivo, que por sua vez leva o sistema a funcionar na região segura antes de ocorrer a contingência.

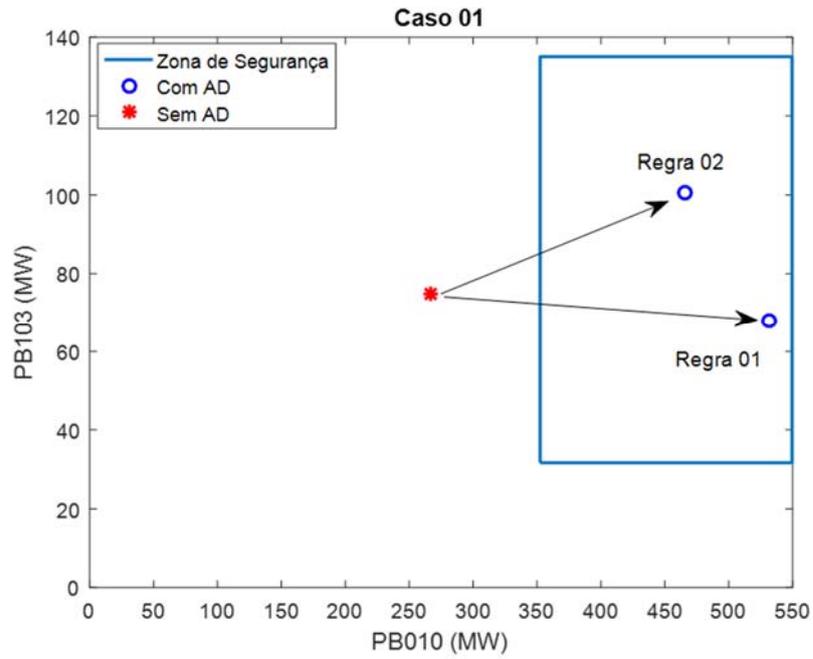


Figura 26: Nomograma em 2-D definido pelo conjunto de regras do caso 01 quando a $P_B10 > 352,395$ MW e a $P_B103 > 32,633$ MW.

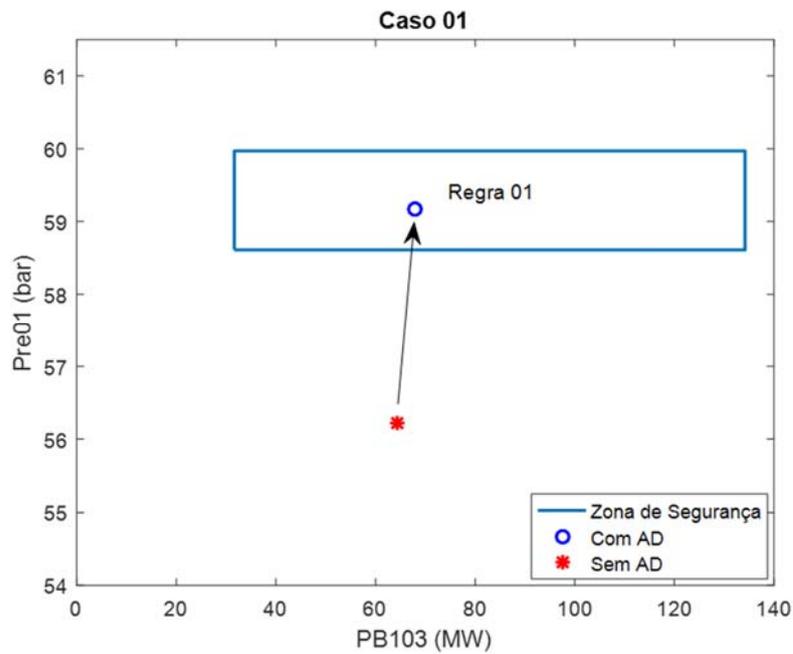


Figura 27: Nomograma em 2-D definido pelo conjunto de regras 01 do caso 01 quando a $P_B103 > 31,633$ MW e a $58,610 \text{ bar} < Pre_01 \leq 59,972 \text{ bar}$.

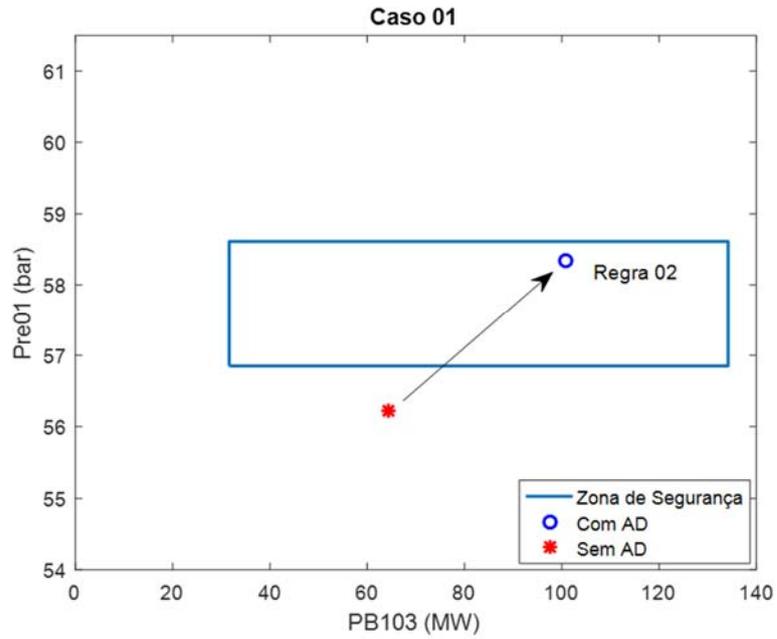


Figura 28: Nomograma em 2-D definido pelo conjunto de regras 02 do caso 01 quando a $P_{B103} > 31,633$ MW e a $Pre_{01} \leq 58,610$ bar.

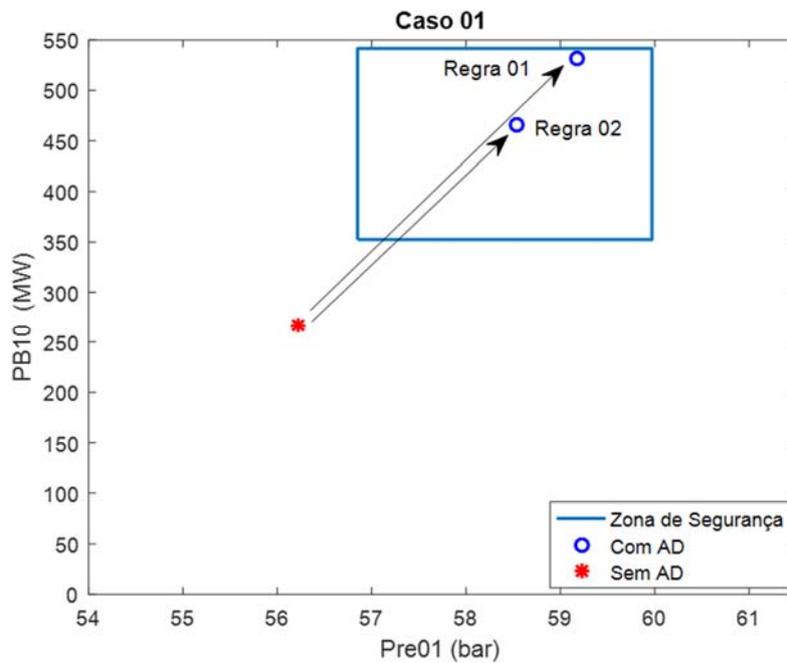


Figura 29: Nomograma em 2-D definido pelo conjunto de regras do caso 01 quando a $Pre_{01} \leq 59,972$ bar e a $P_{B010} > 352,395$ MW.

A Tabela 13 e a Figura 30 mostram as produções de gás nos nós 1 e 2 da rede de gás natural e a potência ativa dos geradores térmicos, respectivamente.

Tabela 13: Produção de gás natural para contingências no duto 3 → 4 e na linha de transmissão 30 → 38

Nós	Sem AD	Regra 01	Regra 02
Nó 1	9,5444966	9,2756797	8,0800112
Nó 2	3,1571202	4,1160594	4,7327209

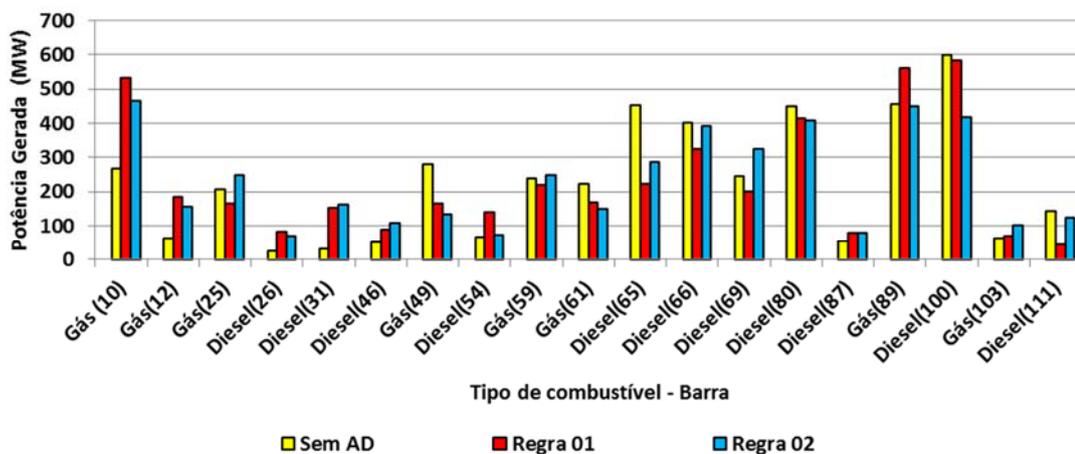


Figura 30: Despacho de geradores a gás natural e a diesel considerando as contingências no duto 3→4 e na linha de transmissão 30→38

Note que a produção de gás natural no nó 2 é violada quando se utiliza um método convencional de fluxo de carga e fluxo de gás, isto é, sem considerar as regras da Árvore de Decisão. Neste caso, os geradores a gás natural despacham menos potência ativa, e, como consequência, a produção de gás natural no nó 2 é reduzida para 3,1571202 Mm³.

Por outro lado, quando o método de despacho utiliza as restrições da AD, os geradores a gás entregam mais potência elétrica, a fim de reduzir o custo da produção de energia elétrica, em vista do maior custo de geração dos motores a diesel.

A Tabela 14 mostra as pressões em todos os nós da rede de gás natural. Verifica-se que sem as regras da AD a pressão no nó 1, produtor de gás, é violada. O valor da pressão de 56,2211581 bar calculado para esse nó, utilizando um método convencional de fluxo de gás natural e fluxo de carga, altera o fluxo de gás na rede de gasodutos e, conseqüentemente, o despacho seguro de gás natural para as demandas não-elétrica e elétrica.

Tabela 14: Pressão (em bar) dos nós da rede de gás natural para contingências no duto 3 → 4 e na linha de transmissão 30 → 38

Nós	Sem AD	Regra 01	Regra 02
Nó 1	56,2211581	59,1732668	58,3480625
Nó 2	53,9721754	53,8578956	53,7811671
Nó 3	57,1240478	57,1262473	57,7826408
Nó 4	51,2783578	51,2828783	51,2871310
Nó 5	56,7783113	56,7508832	56,5029878
Nó 6	43,4965569	43,4974538	43,4769701
Nó 7	51,1922335	51,1972267	51,2000060
Nó 8	41,0075202	40,9873833	40,9916068
Nó 9	41,0930100	41,071900	41,0793212
Nó 10	34,6734042	34,670797	34,6663407
Nó 11	39,0730185	39,0686535	39,0662491
Nó 12	32,9711596	32,9710958	32,9711422
Nó 13	31,1419106	31,1418471	31,1417273
Nó 14	30,0285461	31,0286506	31,0284911
Nó 15	29,3768778	30,3768291	30,3768208

Neste caso, devido a redução da pressão no nó produtor 1 para um valor inferior ao estipulado pela AD, o fluxo de gás natural é reduzido, fazendo com que os geradores a gás natural despachem menos potência ativa, aumentando o custo da produção de energia elétrica, em virtude do maior custo de geração dos motores a diesel.

7.3 CASO 2: CONTINGÊNCIAS NOS GASODUTOS 3→4 E 13→14 E NA LINHA DE TRANSMISSÃO 30→38

O desligamento dos gasodutos 3→4 e 13→14 da rede de gás natural de 15 nós e a desconexão da linha de transmissão 30→38 da rede elétrica de 118 barras são as contingências simultâneas consideradas no estudo do caso 2. A Figura 31 ilustra a estrutura da Árvore de Decisão com três conjuntos de regras que determinam as restrições de segurança das redes de gás e elétrica, de acordo com as contingências do presente estudo de caso.

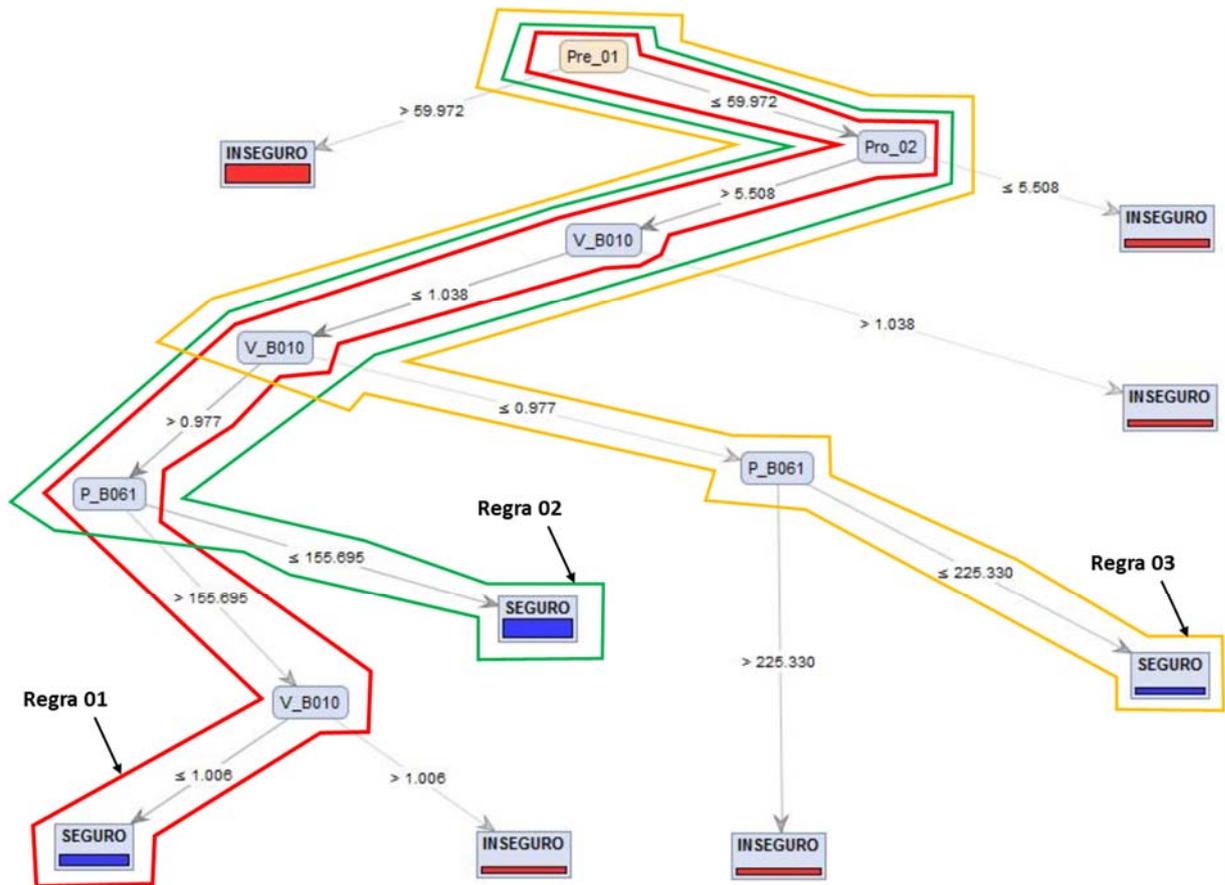


Figura 31: AD para contingências nos gasodutos 3→4 e 13→14 e na linha de transmissão 30→38

As regras da AD da Figura 31 são consideradas como restrições de segurança sobre o problema do despacho entre as redes integradas de gás e eletricidade. Cada conjunto de regras determina a região de segurança com sub-regiões definidas pelos nós terminais seguro da AD:

Conjunto de Regras 01:

Variáveis de Controle	Conjunto de Regras
Pressão no nó 01	$Pre_01 \leq 59,972$ bar
Produção de Gás Natural no nó 02	$Pro_02 > 5,508$ Mm ³
Tensão na barra 10	$0,977 \text{ p.u.} < V_B10 \leq 1,038 \text{ p.u.}$
Potência ativa na barra 61	$P_B061 > 155,695$ MW
Tensão na barra 10	$V_B10 \leq 1,006 \text{ p.u.}$

Conjunto de Regras 02:

Variáveis de Controle	Conjunto de Regras
Pressão no nó 01	$Pre_01 \leq 59,972 \text{ bar}$
Produção de Gás Natural no nó 02	$Pro_02 > 5,508 \text{ Mm}^3$
Tensão na barra 10	$0,977 \text{ p.u.} < V_B10 \leq 1,038 \text{ p.u.}$
Potência ativa na barra 61	$P_B061 > 155,695 \text{ MW}$
Tensão na barra 10	$V_B10 \leq 1,006 \text{ p.u.}$

Conjunto de Regras 03:

Variáveis de Controle	Conjunto de Regras
Pressão no nó 01	$Pre_01 \leq 59,972 \text{ bar}$
Produção de Gás Natural no nó 02	$Pro_02 > 5,508 \text{ Mm}^3$
Tensão na barra 10	$V_B10 \leq 1,038 \text{ p.u.}$
Tensão na barra 10	$V_B10 \leq 0,977 \text{ p.u.}$
Potência ativa na barra 61	$P_B061 \leq 225,330 \text{ MW}$

As regras da AD podem ser reproduzidas por nomogramas (como, por exemplo, a pressão no nó 01 X produção de gás no nó 02) para descrever as regiões de segurança, como mostrado nas Figuras 32, 33, 34, 35 e 36.

O asterisco em vermelho em cada nomograma representa um caso inseguro, o retângulo em azul representa a região de segurança e a circunferência em azul representa um caso ótimo dentro da região de segurança. Pode-se observar que cada trajetória de um asterisco para uma circunferência é uma indicação de controle confiável que leva o sistema de um região insegura para uma região segura de operação.

O despacho ótimo coordenado entre a geração termoeletrica e produção de gás natural pode ser realizado como um controle preventivo, que por sua vez leva o sistema a funcionar na região segura antes de ocorrer a contingência.

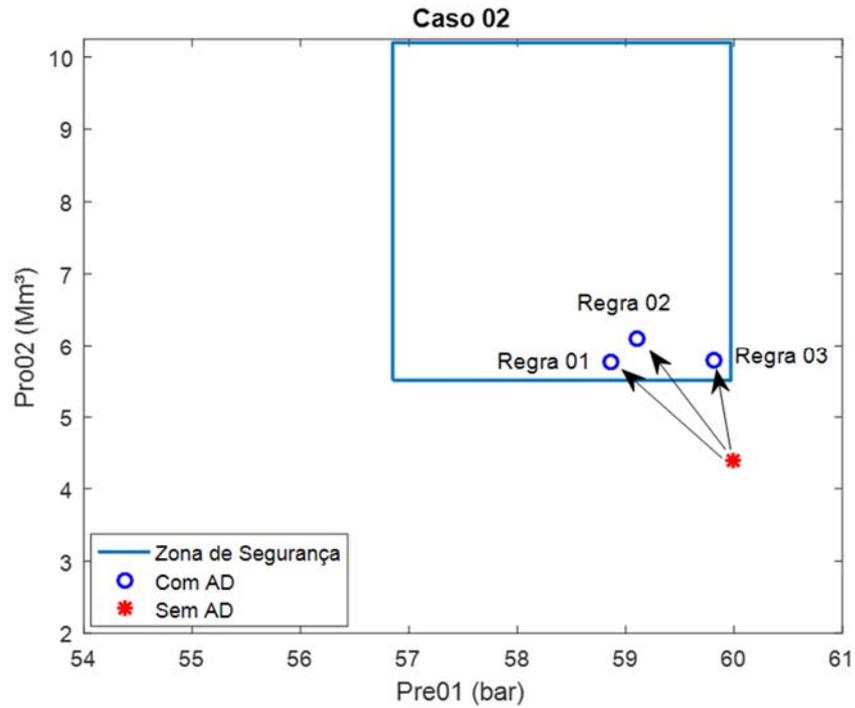


Figura 32: Nomograma em 2-D definido pelo conjunto de regras do caso 02 quando a $Pre_{01} \leq 59,972$ bar e a $Pro_{02} > 5,508$ Mm^3

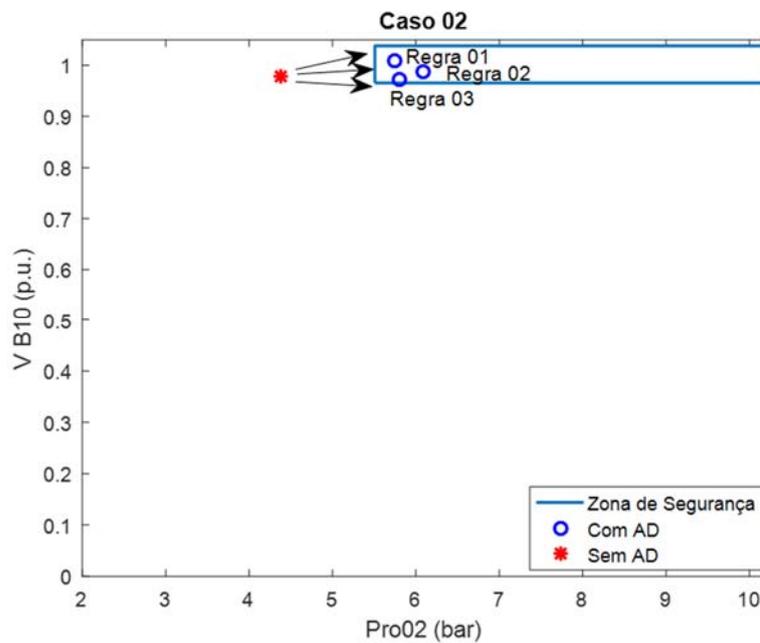


Figura 33: Nomograma em 2-D definido pelo conjunto de regras do caso 02 quando a $Pro_{01} > 5.508$ Mm^3 e a 0.977 p.u. $< V_{B10} \leq 1.038$ p.u.

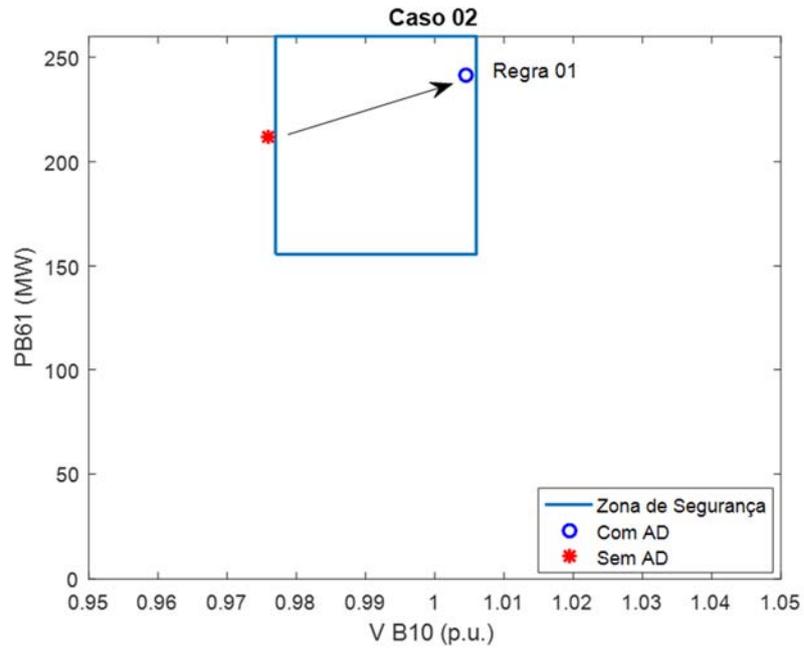


Figura 34: Nomograma em 2-D definido pelo conjunto de regras do caso 02 quando a $0.977 \text{ p.u.} < V_{B10} \leq 1.006 \text{ p.u.}$ e a $P_{B61} > 155.695 \text{ MW}$, regra 01.

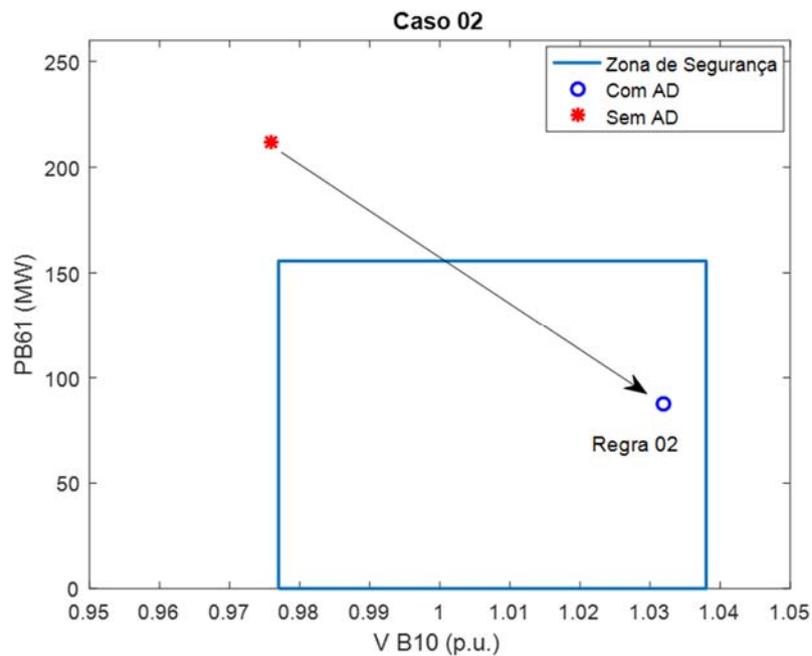


Figura 35: Nomograma em 2-D definido pelo conjunto de regras do caso 02 quando a $0.977 \text{ p.u.} < V_{B10} \leq 1.006 \text{ p.u.}$ e a $P_{B61} > 155.695 \text{ MW}$, regra 02.

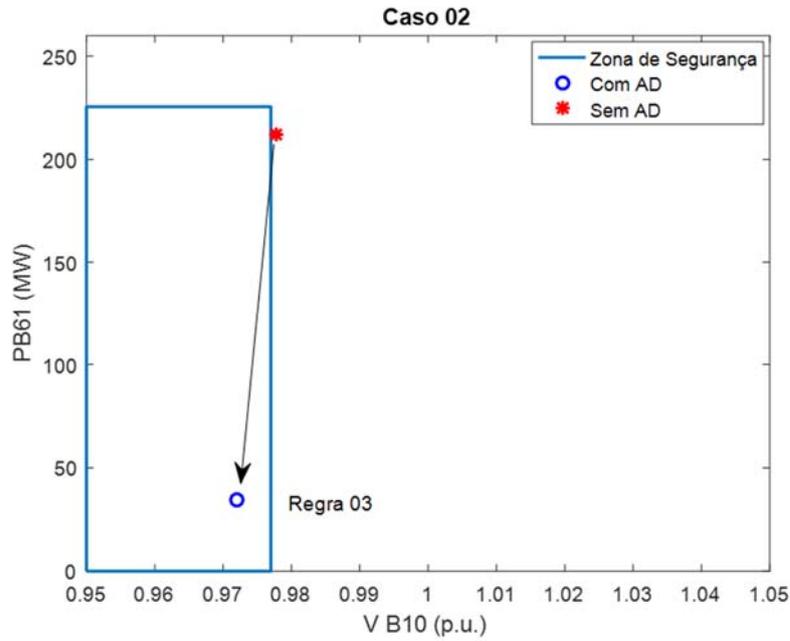


Figura 36: Nomograma em 2-D definido pelo conjunto de regras do caso 02 quando a $0.977 \text{ p.u.} < V_{B10} \leq 1.006 \text{ p.u.}$ e a $P_{B61} > 155.695 \text{ MW}$, regra 03.

A Tabela 15 e a Figura 37 mostram as produções de gás nos nós 1 e 2 da rede de gás natural e a potência ativa dos geradores térmicos, respectivamente.

Tabela 15: Produção de gás natural para contingências nos dutos 3 → 4 e 13 → 14 e na linha de transmissão 30 → 38

Nós	Sem AD	Regra 01	Regra 02	Regra 03
Nó 1	7,3124450	7,1380933	6,1017917	7,9254161
Nó 2	4,386605	5,7547093	6,0905476	5,7975963

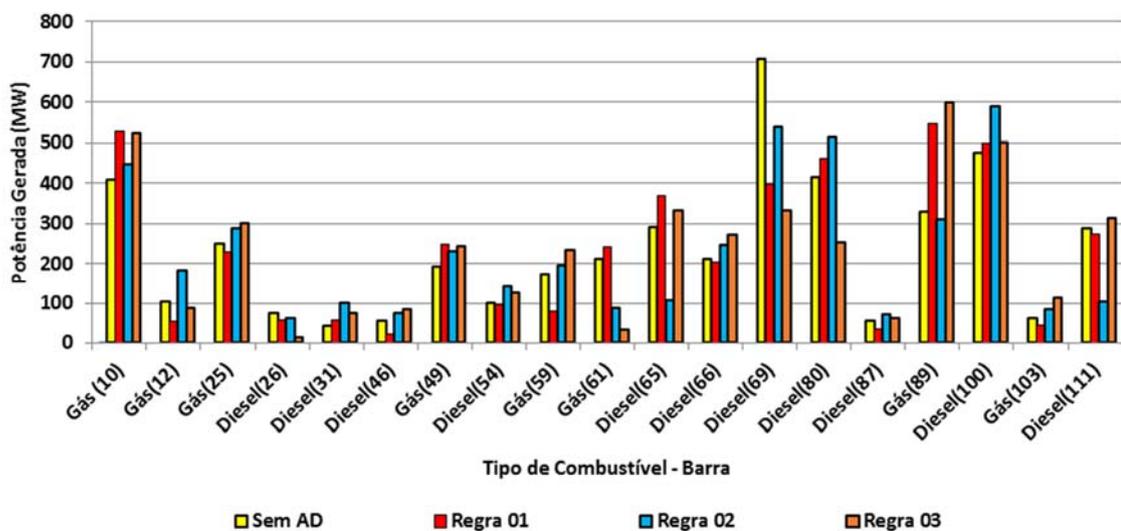


Figura 37: Despacho de geradores a gás e a diesel considerando as contingências nos dutos 3→4 e 13→14 e na linha de transmissão 30→38

Note que a produção de gás natural no nó 2 é violada quando se utiliza um método convencional de fluxo de carga e fluxo de gás, isto é, sem considerar as regras da Árvore de Decisão. Neste caso, os geradores a gás despacham menos potência ativa, e, como consequência, a produção de gás natural no nó 2 é reduzida para 4,386605 Mm³, valor inferior ao definido pela AD, 5,508 Mm³.

Por outro lado, quando o método de despacho utiliza as restrições da AD, os geradores a gás entregam mais potência elétrica, a fim de reduzir o custo da produção de energia elétrica, em vista do maior custo de geração dos motores a diesel.

A Tabela 16 mostra as pressões em todos os nós da rede de gás natural. Verifica-se que sem as regras da AD a pressão no nó 1, produtor de gás, é violada. O valor da pressão de 59,9972039 bar calculado para esse nó, utilizando um método convencional de fluxo de gás natural e fluxo de carga, supera o valor máximo estimado pela AD.

Tabela 16: Pressão (em bar) dos nós da rede de gás natural para contingências nos dutos 3 → 4 e 13 → 14 e na linha de transmissão 30 → 38

Nós	Sem AD	Regra 01	Regra 02	Regra 03
Nó 1	59,9972039	58,8770597	59,1183161	59,8185593
Nó 2	51,8237884	53,6588724	53,6186521	53,6541566
Nó 3	53,1622492	57,1250670	57,1242944	57,1235182
Nó 4	48,2834950	51,2879170	51,2893423	51,2881734
Nó 5	50,6455083	56,7623014	56,7729905	56,7999869
Nó 6	47,4835220	43,4850306	43,5189153	43,5281671
Nó 7	42,1972335	51,1999170	51,2010611	51,2001852
Nó 8	41,0041164	41,0173009	40,9893134	41,0040999
Nó 9	44,0887588	41,0692400	41,0905387	41,0801466
Nó 10	38,6749204	34,6746373	34,6710376	34,6642397
Nó 11	35,0811633	39,0982610	39,0729669	39,0725182
Nó 12	31,9712292	32,9711238	37,9712230	35,9710606
Nó 13	35,1419006	31,1419391	33,1327682	32,2616801
Nó 14	29,6283835	31,0286324	33,0283304	32,0146475
Nó 15	28,0369122	30,3768685	31,3768742	31,2267659

Neste caso, o aumento da pressão no nó produtor 1 para um valor superior ao estipulado pela AD, compromete o desempenho da rede de gasodutos e por consequência, a segurança de todo o sistema das RIEEGNs. Com uma pressão acima do valor seguro, a diferença de pressão nos dutos é alterada. Dessa forma, o fluxo de gás natural no sistema alcança valores que alteram o funcionamento do despacho de gás natural impactando na geração segura de energia elétrica.

7.4 CASO 3: CONTINGÊNCIAS NOS GASODUTOS 3→4 E 13→14 E NAS LINHAS DE TRANSMISSÃO 30→38 E 19→34

O desligamento dos gasodutos 3→4 e 13→14 da rede de gás natural de 15 nós e a desconexão das linhas de transmissão 30→38 e 19→34 da rede elétrica de 118 barras são as contingências simultâneas consideradas no estudo do caso 3. A Figura 38 ilustra a estrutura da Árvore de Decisão com dois conjuntos de regras que determinam as restrições de segurança das redes de gás e elétrica, de acordo com as contingências do presente estudo de caso.

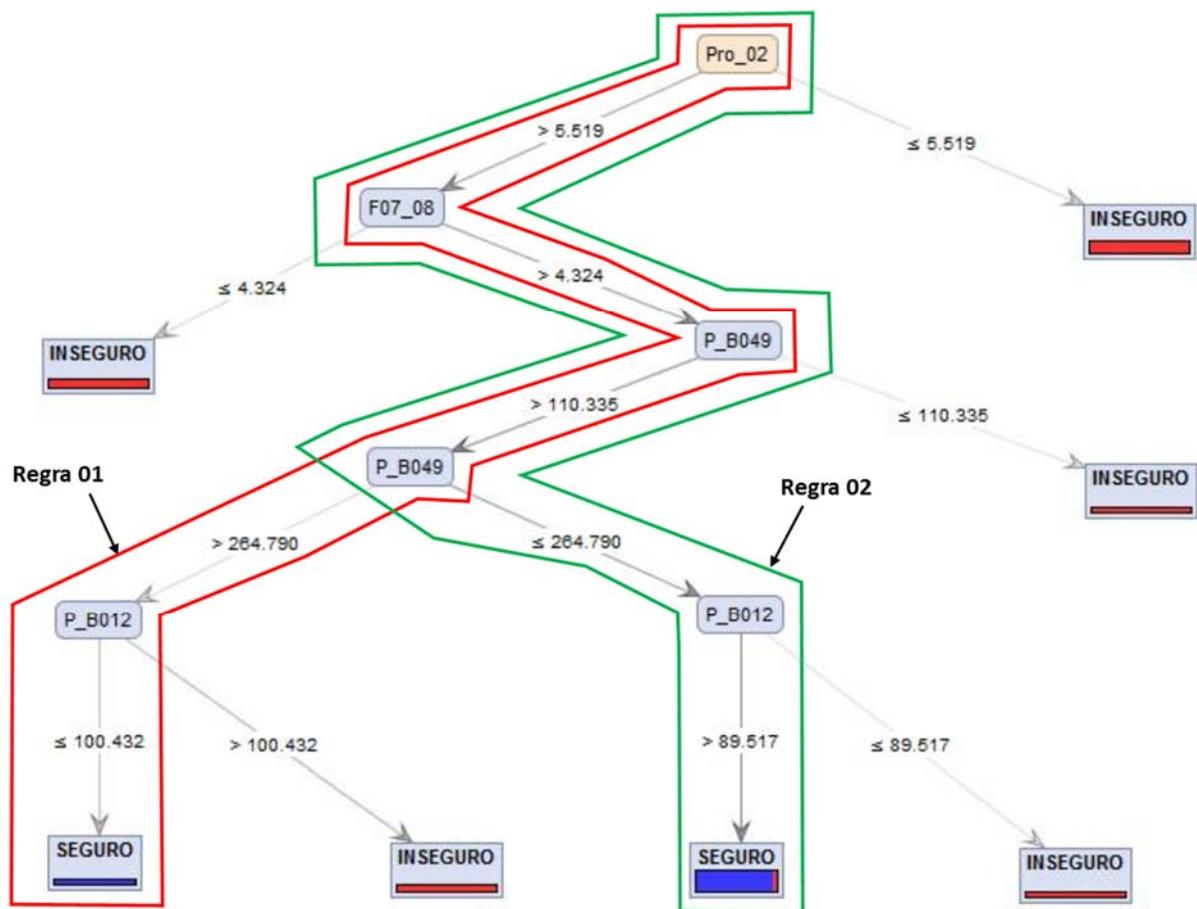


Figura 38: AD para contingências nos gasodutos 3→4 e 13→14 e nas linhas de transmissão 30→38 e 19→34.

As regras da AD da Figura 38 são consideradas como restrições de segurança sobre o problema do despacho entre as redes integradas de gás e eletricidade. Cada conjunto de regras determina a região de segurança com sub-regiões definidas pelos nós terminais seguro da AD:

Conjunto de Regras 01:

Variáveis de Controle	Conjunto de Regras
Produção de Gás Natural no nó 02	$Pro_02 > 5,519 \text{ Mm}^3$
Fluxo de gás no duto 07→08	$F07_08 > 4,324 \text{ m}^3/\text{h}$
Potência ativa na barra 49	$P_B049 > 110,335 \text{ MW}$
Potência ativa na barra 49	$P_B049 > 264,790 \text{ MW}$
Potência ativa na barra 12	$P_B012 \leq 100,432 \text{ MW}$

Conjunto de Regras 02:

Variáveis de Controle	Conjunto de Regras
Produção de Gás Natural no nó 02	$Pro_02 > 5,519 \text{ Mm}^3$
Fluxo de gás no duto 07→08	$F07_08 > 4,324 \text{ m}^3/\text{h}$
Potência ativa na barra 49	$P_B049 > 110,335 \text{ MW}$
Potência ativa na barra 49	$P_B049 \leq 264,790 \text{ MW}$
Potência ativa na barra 12	$P_B012 > 89,517 \text{ MW}$

As regras da AD podem ser reproduzidas por nomogramas (como, por exemplo, a produção no nó 02 X fluxo de gás no duto 07→08) para descrever as regiões de segurança, como mostrado nas Figuras 39, 40, 41 e 42.

O asterisco em vermelho em cada nomograma representa um caso inseguro, o retângulo em azul representa a região de segurança e a circunferência em azul representa um caso ótimo dentro da região de segurança. Pode-se observar que cada trajetória de um asterisco para uma circunferência é uma indicação de controle confiável que leva o sistema de um região insegura para uma região segura de operação.

O despacho ótimo coordenado entre a geração termoeletrica e produção de gás natural pode ser realizado como um controle preventivo, que por sua vez leva o sistema a funcionar na região segura antes de ocorrer a contingência.

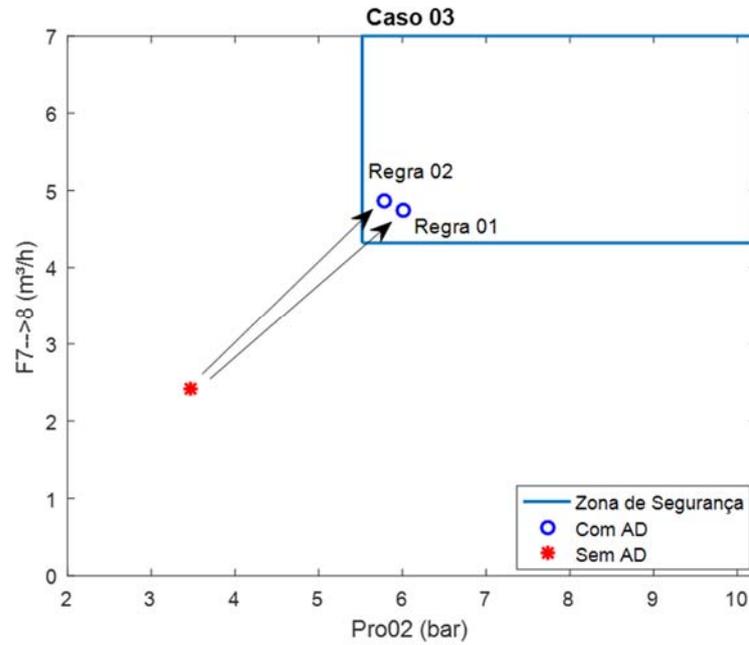


Figura 39: Nomograma em 2-D definido pelo conjunto de regras do caso 03 quando a $Pro_02 > 5,516 \text{ Mm}^3$ e o $F07 \rightarrow 08 > 4,324 \text{ m}^3/\text{h}$.

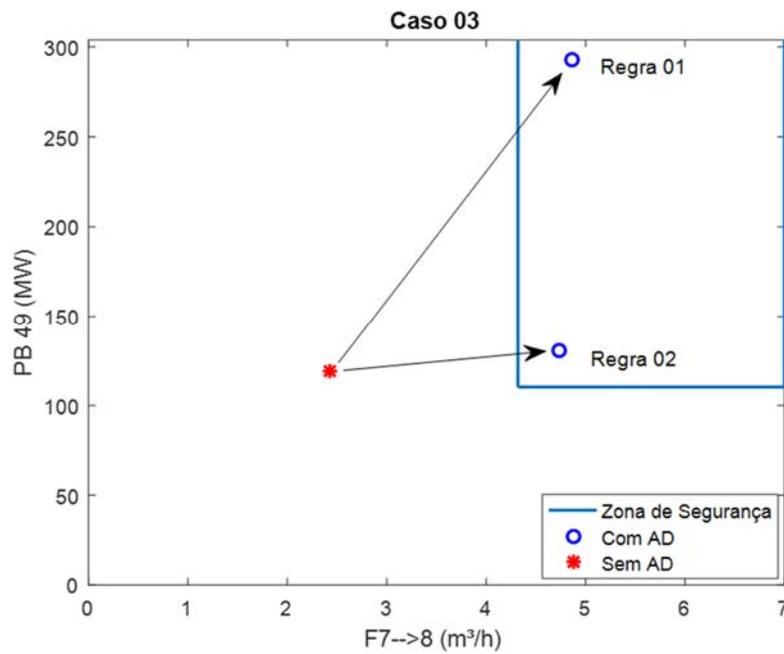


Figura 40: Nomograma em 2-D definido pelo conjunto de regras do caso 03 quando o $F07 \rightarrow 08 > 4,324 \text{ m}^3/\text{h}$ e a $P_B49 > 110,335 \text{ MW}$.

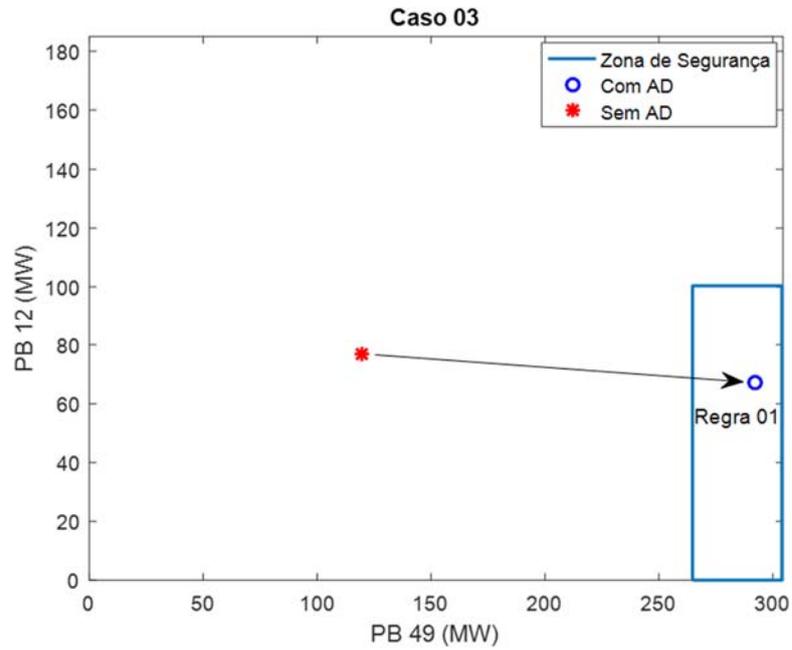


Figura 41: Nomograma em 2-D definido pelo conjunto de regras do caso 03 quando a $P_{B49} > 264,790$ MW e a $P_{B12} \leq 100,432$ MW, regra 01.

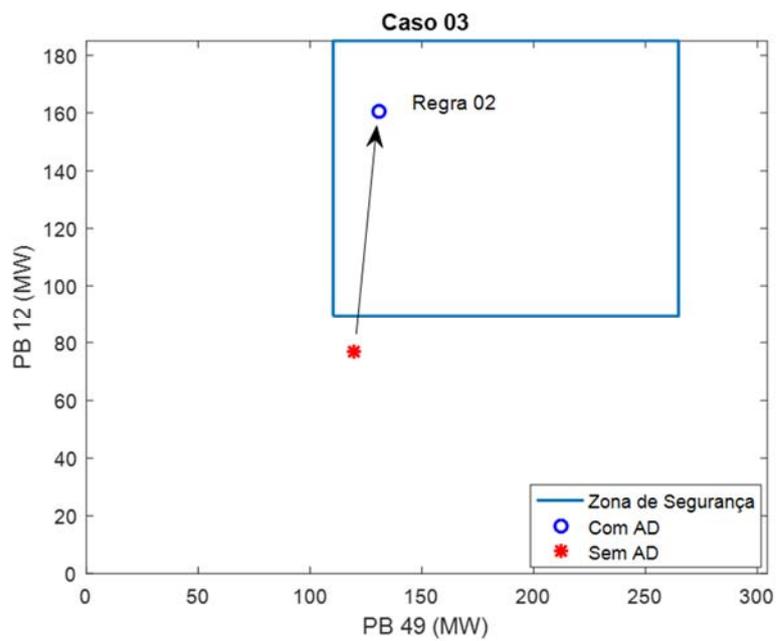


Figura 42: Nomograma em 2-D definido pelo conjunto de regras do caso 03 quando a $110,335$ MW $< P_{B49} \leq 264,790$ MW e a $P_{B12} > 89,517$ MW, regra 02.

A Tabela 17 e a Figura 43 mostram as produções de gás nos nós 1 e 2 da rede de gás natural e a potência ativa dos geradores térmicos, respectivamente.

Tabela 17: Produção de gás natural para contingências nos dutos 3 → 4 e 13 → 14 e nas linhas de transmissão 30→38 e 19→34

Nós	Sem AD	Regra 01	Regra 02
Nó 1	7,3823296	7,5082638	7,9638867
Nó 2	3,475682	5,7884872	6,0103381

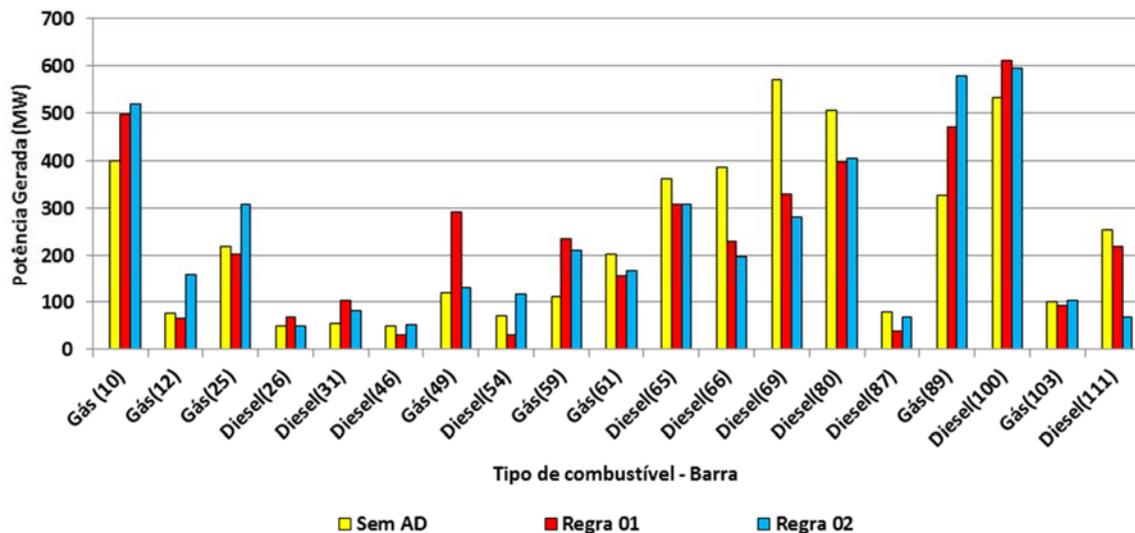


Figura 43: Despacho de geradores a gás e a diesel considerando as contingências nos dutos 3→4 e 13→14 e nas linhas de transmissão 30→38 e 19→30

Note que a produção de gás natural no nó 2 é violada quando se utiliza um método convencional de fluxo de carga e fluxo de gás, isto é, sem considerar as regras da Árvore de Decisão. Neste caso, os geradores a gás despacham menos potência ativa, e, como consequência, a produção de gás natural no nó 2 é reduzida para 3,475682 Mm³, valor inferior ao definido pela AD, 5,519 Mm³.

Por outro lado, quando o método de despacho utiliza as restrições da AD, os geradores a gás entregam mais potência elétrica, a fim de reduzir o custo da produção de energia elétrica, em vista do maior custo de geração dos motores a diesel.

A Tabela 18 mostra os fluxos de gás natural em todos os dutos da rede. Verifica-se que sem as regras da AD o fluxo de gás no duto 07→08 é violado. O valor do fluxo de gás natural de 2,424018 m³/h calculado para esse ramo, utilizando um método convencional de fluxo de gás natural e fluxo de carga, é inferior ao valor mínimo estimado pela AD, 4,324 m³/h.

Tabela 18: Fluxo de gás (em m³/h) na rede de gás natural para contingências nos dutos 3→4 e 13→14 e nas linhas de transmissão 30→38 e 19→30

Gasodutos	Sem AD	Regra 01	Regra 02
F1→3	7,3823296	7,5082638	7,9638867
F2→4	3,475682	5,7884872	6,0103381
F3→4	0,000000	0,0000000	0,0000000
F3→5	6,5443296	6,6702638	7,1258867
F4→7	3,257682	5,5704872	5,7923381
F5→6	5,049418	5,6571103	5,5759898
F7→8	2,424018	4,8724868	4,7382841
F6→9	2,659882	4,9078041	5,1364789
F8→11	2,274682	3,6222877	4,3332636
F9→10	2,423418	2,3746943	2,1424165
F11→12	1,922368	1,9110966	1,6219462
F10→13	1,713982	2,4409386	3,2802400
F12→14	0,878368	1,1370966	1,1520538
F13→14	0,000000	0,0000000	0,0000000
F13→15	1,199368	1,6880966	1,3989462
F14→15	0,126325	0,4309034	0,3200538

Neste caso, com o fluxo de gás no duto 07→08 menor que o valor estimado pela AD, provocará uma instabilidade no abastecimento de gás nos geradores termelétricos alimentados por esses nós (nó 07 e nó 08), respectivamente nas barras B49 e B12.

A ausência das regras determinadas pela AD, compromete o desempenho da rede de gasodutos e por consequência, a segurança de todo o sistema das RIEEGNs. Aplicando-se o conjunto de regras da AD, o fluxo de gás natural no sistema alcança valores que asseguram o bom funcionamento do despacho de gás natural e, conseqüentemente, em uma geração segura de energia elétrica.

7.5 CONCLUSÃO DO CAPÍTULO VII

Este capítulo apresentou estudos de casos aplicados ao despacho de um sistema integrado de energia composto pela rede de gás natural de 15 nós e pelo sistema de 118 barras do IEEE. Essa análise foi dividida em três situações: caso 01, caso 02 e caso 03 que consideram diferentes níveis de contingências. Nas três situações verificou-se que, sem a utilização das regras estimadas pela AD declaradas aos atributos das RIEEGNs, as regiões segura e insegura não apresentam limites bem definidos para o sistema, comprometendo a segurança das redes integradas. E ainda, quando se utiliza um método convencional de fluxo

de carga ou fluxo de potência e fluxo de gás natural, isto é, sem considerar as regras da Árvore de Decisão, os limites de segurança dos atributos, como pressão e fluxo de gás, são violados, alterando o despacho seguro ótimo e prejudicando a tomada de decisão que manterá o sistema das Redes Integradas de Energia Elétrica e Gás Natural em seu estado seguro de operação.

CAPÍTULO VIII - CONCLUSÕES

8.1 CONSIDERAÇÕES FINAIS

O monitoramento das Redes Integradas de Energia Elétrica e Gás Natural (RIEEGNs) é uma das funções mais importantes dos Sistemas Elétricos de Potência que visa aumentar a segurança do Sistema Integrado de Energia Elétrica e de Gás Natural (SIEEGN). No entanto, a complexidade inerente da operação de um grande sistema como este torna necessária a utilização de funções sofisticadas de diagnóstico, análise e aconselhamento. Ações preventivas disponíveis nos sistemas de gerenciamento de energia objetivam a melhora na qualidade da informação disponível para o operador, na tomada de decisão por parte deste.

Existem várias formas de avaliar a segurança de um SIEEGN. Cada maneira está associada ao tipo de ferramenta computacional, metodologia e aos critérios utilizados para análise. Nessa tese, a avaliação foi realizada pelo Método de Mineração de Dados, denominado, *Árvore de Decisão (AD)*.

Com o algoritmo da *Árvore de Decisão* desenvolvido para análise de segurança, a partir de uma lista de contingências prováveis de ocorrência na rede, conforme literatura vigente foi possível simular as ocorrências das RIEEGNs e ao final do processamento fornecer ao operador quais medidas devem ser tomadas para aumentar, ou manter, o nível de segurança do sistema, baseados nas instruções estabelecidas pela *Árvore*. As ADs demonstraram assim serem ferramentas eficazes e efetivas na análise da segurança das RIEEGNs.

As ADs foram aplicadas na determinação de regiões de segurança sistêmica e as suas fronteiras, foram avaliadas pela metodologia proposta neste trabalho, com o objetivo de fornecer diretrizes para a tomada de decisões relativas às ações de controle preventivas necessárias contra violações que possam ocorrer na rede elétrica.

Os conjuntos de regras de fácil interpretação que descrevem os limites de segurança, que são definidas pelas ADs, foram utilizados como restrições tratáveis na otimização da geração de energia elétrica e do reescalonamento da produção de gás natural, via Algoritmo Genético.

Os resultados apresentados mostram que o despacho ótimo seguro guiado pelas *Árvores de Decisão* é capaz de estabelecer ações de controle que trazem o sistema integrado de energia de volta para a região de segurança, conseqüentemente, impedindo que as Redes Integradas de Energia Elétrica e Gás Natural se tornem inseguras mediante a ocorrência de uma ou mais contingências no sistema.

Em todos os casos simulados, os resultados são satisfatórios e o método inteligente se mostrou eficaz na determinação das regiões de segurança, considerando as regiões de fronteiras aplicadas, descrevendo assim critérios para os limites operacionais nos quais as RIEEGNs poderão atuar.

Pode-se assim afirmar que a aplicação do método proposto é uma boa opção para se obter despacho seguro entre as redes elétricas e de gás natural, pois a AD explorou e identificou as principais variáveis de controle existentes nesse tipo de abordagem.

8.2 PROPOSTA DE TRABALHOS FUTUROS

Com as análises apresentadas no estudo de caso, verificou-se a viabilidade e a capacidade da técnica de Árvore de Decisão para orientar as ações de despacho seguro e ações corretivas nas Redes Integradas de Energia Elétrica e Gás Natural. Contudo, algumas ações serão necessárias realizar no futuro, visando o aperfeiçoamento da metodologia para aplicação na função de avaliação de segurança de um sistema formado pelas redes elétrica e de gás natural:

- Adicionar um número maior de contingências agregando novos casos simulados em torno do caso base da topologia considerada.
- Análise de contingências poderão ser realizadas para detectar possíveis problemas na rede de gás natural.
- Promover a integração da metodologia na função de avaliação de segurança do centro de operação do sistema de Redes Integradas de Energia Elétrica e Gás Natural, disponibilizando-a como ferramenta de suporte à operação de tempo real.

REFERÊNCIAS BIBLIOGRÁFICAS

ALPAYDIN, E. Introduction to Machine Learning: Adaptive Computation and Machine Learning, The MIT Press. 2004.

AN, Seungwon., LI, Qing., GEDRA, T. W. Natural Gas and Electricity Optimal Power Flow. Transmission and Distribution Conference and Exposition, IEEE PES. V.1 p, 138-143. 2003

ANEEL – AGÊNCIA NACIONAL DE ENERGIA ELÉTRICA. Atlas de energia elétrica no Brasil. 3 ed. Brasília: Aneel, 2008.

Disponível em <http://www.aneel.gov.br/visualizar_texto.cfm?idtxt=1687>. Acesso 26 maio 2016.

ANEEL – AGÊNCIA NACIONAL DE ENERGIA ELÉTRICA. Atlas de energia elétrica no Brasil. 3 ed. Brasília: Aneel, 2015. Disponível em <

<http://www.aneel.gov.br/documents/655816/14805196/Usinas+Termel%C3%A9tricas+3/e728cf8e-1743-419b-bfb8-8b1346344277>>. Acesso 23 de junho 2016.

ANP – AGÊNCIA NACIONAL DE PETRÓLEO, GÁS E BIOCOMBUSTÍVEIS. Resolução Nº 16, de 17 jun. 2008. Dispõe sua a qualidade do gás natural distribuído no Brasil. Disponível em <<http://www.anp.gov.br>>. Acesso em 10 de julho 2016.

BALU, N. J., CAULEY, G., LAUBY, M. G., BERTRAM, T. BOSE, A., BRANDWAJN, V., CURTICE, D., FOUAD, A., FINK, L. H., WOLLENBERG, B. F. and WRUBEL, J. N., “On-line Power System Security Analysis”. In: Proceedings of the IEEE, Vol. 80, No 2, pp. 262-280, Fevereiro 1992.

BRAMER, M. Principles of data mining. Springer, London. 2007.

BRASIL. Lei no 9.478, de 06 ago. 1997. Dispõe sobre a política energética nacional, as atividades relativas ao monopólio do petróleo, institui o Conselho Nacional de Política. Disponível em: < http://www.planalto.gov.br/ccivil_03/leis/L9478.htm>. Acesso em 15 maio de 2016.

BRASIL. Decreto 3520, de 21 jun. 2000. Dispõe sobre a estrutura e o funcionamento do Conselho Nacional de Política Energética - CNPE e dá outras providências. Disponível em <http://www.planalto.gov.br/ccivil_03/decreto/D3520.htm>. Acesso em 15 maio de 2016.

BREIMAN, L., Friedman, J. H., OLSHEN, R. A. & STONE, C. J. Classification and Regression Trees. Belmont, Wadsworth. 1984.

BRITISH PETROLEUM - BP. BP energy Outlook 2035. BP, 2016. Disponível em: <<https://www.bp.com/content/dam/bp/pdf/energy-economics/energy-outlook-2016/bp-energy-outlook-2016.pdf>>. Acesso em: 26 de junho 2016.

CASTILLO, L., GONZALEZ, A. Distribution network optimization: Finding the most economic solution by using genetic algorithms. European Journal of Operational Research 108, 527-537. 1998.

CEPEL, Centro de Pesquisas de Energia Elétrica, ANAREDE: Programa de Análise de Redes, Manual Guide, 07-08 Version, Aug.1999.

CHAELE, O., SCHÖLKOPF, B. & ZIEN, A. Semi-Supervised Learning. MIT Press. 2006.

CIHLAR, T. C., WEAR, J. H., EWART, D. N., KIRCHMAYER, L. K., “Electric Utility System Security”. In: Proceedings of the American Power Conference, Vol. 31, pp. 891-908, 1969.

COOK, Earl. The flow of energy in industrial society. New York: Scientific American. Vol. 225. Pag. 134-144. 1971.

CORREA-POSADA, C. & SANCHEZ-MARTIN, P. Security-Constrained Optimal Power and Natural-Gas Flow, IEEE Trans. Power Syst., vol. 29, no. 4, pp. 1780–1787, Jul. 2014.

CORREA-POSADA, C. & SANCHEZ-MARTIN, P. Integrated Power and Natural Gas Model for Energy Adequacy in Short-Term Operation, IEEE Trans. Power Syst., vol. 30, no. 6, pp. 3347–3355, Nov. 2015.

DARWIN, C., “A origem das espécies”. Rio de Janeiro: Ediouro, 2004.

DIAO, R., VITTAL, V. & LOGIC, N. Design of a Real-Time Security Assessment Tool for Situational Awareness Enhancement in Modern Power Systems, IEEE Trans. PowerSyst., vol. 25, no. 2, pp. 957–965, May 2010.

DIAS, J.S., BARRETO, J.M., “Algoritmo genético: inspiração biológica na solução de problemas - uma introdução”, Revista Marítima Brasileira - Suplemento Especial, Pesquisa Naval, nº 11, p. 105-128, 1998.

DY LIACCO, T. E., “The Adaptative Reliability Control System”, IEEE Trans. Parallel Dist. System, Vol. 86, pp 517-531, 1967.

ECONOMIDES, M.J., SUN, K., and SUBERO, G., “Compressed Natural Gas (CNG): An Alternative to Liquid Natural Gas (LNG),” SPE 92047, Paper presented at the Asia Pacific Oil and Gas Conference and Exhibition, Jakarta, Indonesia, April 5–7, 2005.

ELGERD, O. I. Introdução à Teoria de Sistemas de Energia Elétrica. São Paulo, Brasil: McGraw-Hill, 1977.

EPE – EMPRESA DE PESQUISA ENERGÉTICA. Plano decenal de expansão de energia 2021. EPE, 2012. Disponível em:
<<http://www.epe.gov.br/pdee/forms/epeestudo.aspx>>. Acesso em 02 julho 2016.

EPE – EMPRESA DE PESQUISA ENERGÉTICA. BEN - Balanço Energético Nacional 2015 ano base 2014. Disponível em <<https://ben.epe.gov.br/>>. Acesso em 08 de julho de 2016.

EPE – EMPRESA DE PESQUISA ENERGÉTICA. Matriz energética Nacional 2016. EPE 2016. Disponível em:

< <https://ben.epe.gov.br/BENRelatorioInicial.aspx?anoColeta=2016&anoFimColeta=2015>>. Acesso em 01 julho 2016.

ESTEVES, Marcos Cramer. & MATZ, Marcello. Fundamentos da cadeia do gás natural. Rio de Janeiro: Petrobras. 2013.

FAYYAD, U. PIATETSKI-SHAPIRO, Gregory & SMYTH, Padhraic. The KDD Process for Extracting Useful Knowledge from Volumes of Data. In: Communications of the ACM, 1996.

FICCO, G., DELL'ISOLA, M., VIGO, P., CELENZA, L. Uncertainty analysis of energy measurements in natural gas transmission networks. ELSEVIER. Flow Measurement and Instrumentation. v. 42, p. 58–68. April 2015.

FONSECA, J. Indução de Árvores de Decisão. Tese de Mestrado, Lisboa. 1994.

GARCIA, Roberto. Combustíveis e combustão industrial. Rio de Janeiro: Interciência. 202 p, 2002.

GASMIG – COMPANHIA DE GÁS DE MINAS GERAIS. Gás Natural em Minas Gerais. Belo Horizonte, 2016.

GASNET. Gasodutos de transporte da América do Sul. GASNET, 2016. Disponível em: < <http://www.gasnet.com.br/gasodutos/projetados.asp> >. Acesso em 09 de maio de 2016.

GENC, I. DIAO, R. VITTAL, V. KOLLURI, S. & MANDAL, S. Decision Tree-Based Preventive and Corrective Control Applications for Dynamic Security Enhancement in Power Systems IEEE Trans. Power Systems, vol. 25, no. 3, pp. 1611–1619, Aug 2010.

GOLDBARG, M.C., GOLDBARG., E.F.G., MEDEIROS NETO., F. D. de. Algoritmos Evolucionários na Determinação da Configuração de Custo Mínimo de Sistemas de Cogeração de Energia com Base no Gás Natural. *Pesquisa Operacional*, V.25 n.2, p. 231-259. 2005.

GOLDEBERG, D. Computer-aided gás pipeline operation using genetic algorithms and rule learning. Dissertation (Ph.D.) – University of Michigan. 1983.

GOLDEBERG, D. Genetic algorithms in search, optimization, and machine learning. Addison-Wesley Professional. 423 p, 1989.

GREEN, Don W. & PERRY, Robert H. Perry's chemical engineers handbook. 8 ed. New York: McGraw-Hill, 2008.

GUDMUNDSSON, J.S., ANDERSSON, V., and LEVIK, O.I., "Gas Storage and Transport Using Hydrates," paper presented at the Offshore Mediterranean Conference, Ravenna, March 19–21, 1997.

HAN, J. Data Mining: Concepts and Techniques. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA. 2001.

INTERNATIONAL ENERGY AGENCY. Secure and Efficient Electricity Supply: During the Transition to Low Carbo Power Systems. Disponível em: www.iea.org. 2013.

KUNDUR, P. Power System Stability and Control. New York, USA: McGraw-Hill, 1994.

LIMMER, H. D., “Security Application of On-line Digital Computers”. In: Proceedingd of the Second PSCC, Julho 1966.

LIU, C. SHAHIDEHPOUR, Fu Y. & LI, Z. Security-Constrained Unit Commitment with Natural Gas Transmission Constraints, IEEE Trans. Power Syst., vol. 24, no. 3, pp. 1523–1536, Aug. 2009.

LIU, C. SUN, K. RATHER, Z. H. CHEN, Z. BAK, C. L. THOGERSEN, P. & LUND, P. A Systematic Approach for Dynamic Security Assessment and the Corresponding Preventive Control Scheme Based on Decision Trees,” IEEE Trans. Power Systems, vol. 29, no. 2, pp. 717–730, Mar 2014.

MANSELL, W. Research Ltd., “Methodology for evaluating natural gas transmission system reliability levels and alternatives,” for the Canadian Petroleum Association, Sept. 1991.

MARTINES-MARES, Alberto. FUERTE-ESQUIVEL, Claudio R. A Unified Gas and Power Flow Analysis in Natural Gas and Electricity Coupled Networks. IEEE Transactions on Power Systems, v. 27. n. 4. Nov. 2012.

MATLAB: Matrix Laboratory. Disponível em: <https://www.mathworks.com>

MELLO, O. D., OHISHI., T. An Integrate Dispatch Model of Gas Supply and Thermolectric Generation With Constraints on the Gas Suplly. Journal Proc. of X Symposium of Specialists in Electric Operation and Expantion (X SEPOPE), Florianópolis (SC), Brazil. 2006.

MENDES, J. C., Avaliação de Segurança Estática. Tese de M.Sc., UFMA, São Luiz, MA, Brasil, 1999.

MICHALEWICZ, Z. Genetic Algorithms + Data Structure = Evolution Programs. 3 rd rev. ed. Spring, 1996. 387 p.

MIDTHUN, Kjetil. HELLEMO, Lars. KAUT, Michal & WERNER, Adrian. Assessing production assurance in a natural gas network by using scenario generation and optimization. Elsevier ScienceDirect Journal: Energy Procedia - Volume 64, 2015, Pages 120-129. 2015.

MITCHELL, T. M. Machine Learning. McGraw-Hill, New York. 1997.

MME – MINISTÉRIO DE MINAS E ENERGIA. Boletim mensal de acompanhamento da indústria do gás natural.

Disponível em: <<http://www.mme.gov.br/web/guest/secretarias/petroleo-gas-natural-e-combustiveis-renovaveis/publicacoes/boletim-mensal-de-acompanhamento-da-industria-de-gas-natural>>. MME, 2016. Acesso em 25 junho 2016.

MME – MINISTÉRIO DE MINAS E ENERGIA. Petróleo, Gás Natural e Combustíveis Renováveis. Disponível em:

<<http://www.mme.gov.br/web/guest/secretarias/petroleo-gas-natural-e-combustiveis-renovaveis/>>. MME, 2016. Acesso em 25 junho 2016.

MOKHATAB, Saeid. POE, William A. SPEIGHT, James G. Handbook of Natural Gas Transmission and Processing. United States of America. Elsevier. 2006.

MONARD, M. C. & BARANAUSKAS, J. A. Conceito sobre aprendizado de máquina. In Rezende, S. O., editor, Sistemas Inteligentes: Fundamentos e Aplicações. Editora Manoele Ltda.

MONTEIRO, José Venâncio de Freitas. SILVA, José Roberto Nunes M. da. Gás natural aplicado à indústria e ao grande comércio. São Paulo: Blucher. 182 p. 2010.

MONTICELLI, A., Fluxo de Carga em Redes de Energia Elétrica. 1 ed. São Paulo, Edgard Blücher Ltda, 1983.

MONTICELLI, A., PEREIRA, M.V. F., GRANVILLE, S. "Security-Constrained Optimal Power Flow With Post-Contingency Corrective Rescheduling", IEEE Transactions on Power System, Vol. PWRs-2, No 1, pp. 175-182, Fevereiro 1987.

MONTOYA-O., S. J. et al. Genetic algorithms applied to the optimum design of gas transmission networks. In: SPE International Petroleum Conference and Exhibition, 2000, Villahermosa.

MUNOZ, Jorge., REDONDO, N. Jimenez., RUIZ, J. Perez., BARQUIN, Julian. Natural Gas Network Modeling for Power Systems Reliability Studies. IEEE Bologna Power Tech Conference, Jun. 2003.

NATURAL GAS, U.S. Energy Information Administration (EIA) [Online]. Available: <http://www.eia.gov/naturalgas/>

NRCan (Natural Resources Canada) Report, “Canadian Natural Gas: Review of 2003 and Outlook to 2020,” Ottawa. Energy Policy Sector – Petroleum Resources Branch – Natural Gas Division, Ottawa, Canada, Dec. 2004.

POWER SYSTEMS: Test Case Archive. 188 Bus Power Flow Test Case. Disponível em: <https://www.ee.washington.edu/research/pstca/pf118/pg_tca118bus.htm>

QUINLAN, J. R. Induction of decision trees. Machine Learning, 1(1):81-106. 1986

QUINLAN, J. R. Decision trees and multivalued attributes. Machine Intelligence, 11:305-318. 1988.

QUINLAN, J. R. C4.5: programs for machine learning. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA. 1993.

RAPIDMINER for Windows V6.0—User’s Guide. USA. Disponível em: <http://rapid-i.com/content/view/181/190/lang,en/>

RESSEL, Hans-William. ALKEFJÄRD, Tobias. MÖRÉE, Gustav. MORÉN, Göran. BRODIN, Elin. BJÖRKSTRÖM, Thomas. The Swedish electricity and natural gas market 2014-Ei-R2015:16. Printed by Elanders Sverige A B, 2015.

ROJEY, A. JAFFRET, C. CORNOT-GANDOLPH, S. DURAND, B. JULLIN, S. & VALAIS, M. Natural Gas Production, Processing, Transport. Editions Technip, Paris, France. 1997.

ROKACH, L. & MAIMON, O. Data mining with decision trees. Theory and applications. World Scientific Publishing. 2008.

SANTOS, S.P., “Transient Analysis: A Must in Gas Pipeline Design.” Paper presented at 29th PSIG Annual Meeting, Arizona, Oct. 15–17, 1997.

SANTOS, S.P., “Series or Parallel: Tailor Made Design or a General Rule for a Compressor Station Arrangement?” Paper presented at 32nd PSIG Annual Meeting, Savannah, GA, Oct. 28–30, 2000.

SHAHIDEHPOUR, M. Fu Y. & WIEDMAN, T. Impact of Natural Gas Infrastructure on Electric Power Systems. Proceedings of the IEEE, vol. 93, no. 5, pp. 1042–1056, May. 2005.

SHAHIDEHPOUR, M., Fu Y. & LI, Zuyi. Security-Constrained Unit Commitment With Natural Gas Transmission Constraints. IEEE Transactions on Power Systems, vol. 24, no. 3, p. 1523–1536, Aug. 2009.

SHREVE, R. Norris; BRINK JR., Joseph A. Indústrias de processos químicos. 4^a Ed. Rio de Janeiro: Guanabara, 1997. 717 p.

SPEIGHT, J. G. Gas Processing: Environmental Aspects and Methods. Butterworth Heinemann, Oxford, England, 1993.

SRINIVAS, M., PATNAIK, L.M., “Genetic algorithms: A survey”, IEEE, 1994.

STANDING, M. B., and KATZ, D. L., Density of natural gases, *Trans. AIME*, 140–149, 1942.

STEARNS, S.C., “Evolução: uma introdução”, Atheneu: São Paulo, 2003.

TAO, Li. EREMIA, M. & SHAHIDEHPOUR, M. Interdependency of Natural Gas Network and Power System Security. IEEE Trans. Power Syst., vol. 23, no. 4, pp. 1817–1824, Nov. 2008.

TAN, P. STEINBACH, M. & KUMAR, V. Introduction to Data Mining, first edition. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA. 2005.

THOMAS, José Eduardo. Fundamentos de engenharia de petróleo. 2^a ed. Rio de Janeiro: Interciência, 2001. 271 p.

VAZ, Célio Eduardo M. MAIA, João Luiz P. & SANTOS, Waldir Gomes dos. Tecnologia da indústria do gás natural. 1 ed. São Paulo: Blucher. 416 p, 2008.

WICHERTt, E., and AZIZ, K., Calculate Z's for sour gases, *Hydrocarbon Proc.* 51, 119–122. May 1972.

WOOD, A. J., WOLLENBERG, B.F., Power Generation, Operation, and Control. 2 ed. New York, John Wiley & Sons, 1996.