Escola de Química
UFRJ

Carlos André Vaz Junior

DETECÇÃO E DIAGNÓSTICO DE FALHAS COM BASE EM DADOS HISTÓRICOS DE PROCESSO: APLICAÇÃO EM DUTOVIAS

Rio de Janeiro

2010

Carlos André Vaz Junior

DETECÇÃO E DIAGNÓSTICO DE FALHAS COM BASE EM DADOS HISTÓRICOS DE PROCESSO: APLICAÇÃO EM DUTOVIAS

Tese de Doutorado apresentada ao Programa de Pós Graduação em Tecnologia de Processos Químicos e Bioquímicos, Escola de Química, Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do título de Doutor em Ciências.

Orientadores
Prof. José Luiz de Medeiros, D.Sc.
Profa. Ofélia de Queiroz Fernandes Araújo, Ph.D.

Rio de Janeiro
2010

V393 - Vaz Junior, Carlos André

Detecção e diagnóstico de falhas com base em dados históricos de processo: aplicação em dutovias / Carlos André Vaz Junior – 2010.

194f.: il.

Tese de Doutorado (Programa de Tecnologia de Processos Químicos e Bioquímicos)

– Universidade Federal do Rio de Janeiro, Escola de Química, 2010.

Orientadores: José Luiz de Medeiros
Ofélia de Queiroz Fernandes Araújo

1. Redes de Dutos, 2. Falhas. 3. Detecção. 4. Séries temporais

I. Universidade Federal do Rio de Janeiro. Escola de Química. II. Título.

Carlos André Vaz Junior

Detecção e diagnóstico de falhas com base em dados históricos de processo: aplicação em dutovias

Tese de Doutorado submetida ao Programa de Pós Graduação em Tecnologia de Processos Químicos e Bioquímicos, Escola de Química, Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do título de Doutor em Ciências.

Aprovado por:

_____
Prof. José Luiz de Medeiros, D.Sc.
(orientador – presidente da banca)

_____
Profa. Ofélia de Queiroz Fernandes Araújo, Ph.D.
(orientadora)

_____
Claudia Sayer, D.Sc.

_____
Beatriz de Souza Leite Pires de Lima, D.Sc.

_____
Cláudia do Rosário Vaz Morgado, D.Sc.

_____
Maurício Bezerra de Souza Jr, D.Sc.

_____
Luiz Fernando Lopes R. Silva, D.Sc.

Rio de Janeiro, RJ – Brasil
Abril de 2010

Agradecimentos

A BRASKEM pela utilização de informações e dados gerados no âmbito do Projeto TECDUTOS com respeito à rede industrial estudada

Aos meus orientadores

A Escola de Química da Universidade Federal do Rio de Janeiro

Ao CNPq pelo auxílio financeiro por meio de bolsa de doutorado

A todos aqueles que direta ou indiretamente colaboraram no desenvolvimento desse estudo.

RESUMO

VAZ JUNIOR, Carlos André. Detecção e diagnóstico de falhas com base em dados históricos de processo: aplicação em dutovias.Tese de Doutorado - Escola de Química, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2010.

Extrair informações relevantes a partir do grande volume de dados atualmente disponível em sistemas de controle de plantas industriais é um desafio. Devido às dimensões das bases de dados, ferramentas estatísticas específicas são necessárias para obter informações a respeito das condições operacionais do processo. Metodologias tais como “redes neurais”, “análise de componentes principais”, “método de protótipos”, “classificação hierárquica” e “modelos auto-regressivos” (AR) são algumas das diversas técnicas atualmente empregadas. Tais ferramentas passaram a ser vastamente aplicadas em plantas químicas e petroquímicas, inclusive em dutovias. Dentre os diversos recursos e possibilidades criados a partir do uso de ferramentas de mineração de dados, tornou-se possível o desenvolvimento de sistemas de detecção de falhas baseados em séries temporais. As abordagens baseadas em séries temporais tendem a apresentar processamento matemático rápido e elevada capacidade adaptativa. Nesse trabalho, são desenvolvidas metodologias baseadas em análise de componentes principais (PCA), classificação hierárquica e teoria dos protótipos com a finalidade de detectar, identificar, e quantificar a ocorrência de anomalias operacionais em rede de dutos. Também são desenvolvidos recursos de detecção de vazamentos baseados na utilização de preditores auto-regressivos com entrada externa (ARX) em contexto multi-entrada e multi-saída (MIMO). São estudas as respostas do sistema frente a duas espécies distintas de problemas operacionais: falhas em sensores e ocorrências de vazamento na tubulação. Diferentes severidades de erro no sinal gerado pelos sensores e diversos diâmetros de vazamento são testados. Após estar devidamente treinado, o detector ARX é lançado em sessão de vigilância do processo fazendo predições do mesmo um instante adiante. Admitindo-se que não ocorreram outros tipos de anormalidades, estabelece-se a detecção de uma ruptura quando há desvio estatisticamente significativo entre predições do detector e valores amostrados das variáveis de saída do processo. Estando definida a detecção de um vazamento, é necessário proceder a sua localização e quantificação. Entre as várias abordagens aqui propostas está a utilização de preditores ARX também para tal tarefa. Por outro lado modelos PCA também são usados para diagnosticar mudanças no comportamento do processo. Modelos que descrevem o comportamento normal das variáveis podem indicar quando o sinal gerado pelos sensores desvia-se do padrão. Com objetivo semelhante, a classificação hierárquica é aplicada sobre os dados visando separar as amostras em agrupamentos: “com” e “sem” falha. Quatro métricas de similaridade são aplicadas: Euclidiana, City Block e Minkowski de ordens 3 e 4. Melhores resultados são obtidos com uso da métrica Minkowski de ordem 4. A ferramenta baseada em PCA revela maior capacidade de detecção quando comparada aos classificadores hierárquicos. Porém, os classificadores revelaram-se fundamentais para identificar com precisão o instante de ocorrência da falha. Uma vez as amostras classificadas quanto a presença ou ausência de falha, o método de classificação é empregado na determinação do tipo de anomalia encontrada. O classificador alcança grau de acerto superior a 95%. A ferramenta de classificação também estima a severidade da anomalia. Para localizar a origem do problema usa-se abordagem híbrida, combinando classificação hierárquica e método dos protótipos. Amostras representando condições anormais de operação previamente estabelecidas são denominadas “protótipos”. Os protótipos são comparados com os dados a serem classificados, permitindo identificar a origem do vazamento.

ABSTRACT

VAZ JUNIOR, Carlos André. Detecção e diagnóstico de falhas com base em dados históricos de processo: aplicação em dutovias. Tese de Doutorado - Escola de Química, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2008.

Extracting relevant information from large operational database available in chemical industrial plants control systems is a real challenge. Due to the size of those databases, statistical tools are essentials to obtain information about the operational conditions of the process. Some techniques as “neural networks”, “principal component analysis” (PCA), “prototypes method”, “hierarchical classification” and “auto-regressive models” (AR) are applicable. These algorithms have been broadly applied in chemical and petrochemical plants, including pipelines. The use of those resources have allowed the development of time series based fault detection systems which tend to present fast mathematical processing and high adaptive capacity. In this study, methodologies based on principal component analysis, hierarchical classification and prototype theory were developed to detect, identify and quantify the occurrence of operational anomalies on a pipeline network. Also developed numerical resources for leak detection based on auto-regressive, with external input, (ARX) predictors within a multi-input, multi-output (MIMO) context. System responses to two kinds of operational problems: “sensors faults” and “leakage occurrence”; were evaluated at different error severities. After appropriately tuned, the ARX-detector can be launched along process watching sessions, where it continuously generates predictions one instant ahead of the process. Supposing that no other kinds of abnormality may happen, leak detection can be ascertained when a statistically significant mismatch occurs between detector predictions and measured process outputs. After the leak was detected, we proceed to locate and quantify it, using ARX models too. On the other hand, also used PCA models to diagnose behavioral changes in the process. Models that describe processes variables typical behavior are useful to correlate the generated signal with the expected pattern. With a similar goal, hierarchical classification is applied over a time series data to separate the samples between “with” and “without” faults. The similarity metrics based on Euclidian distances, City Block and Minkowski of order 3 and 4 were used and better results were obtained from Minkowski order 4. PCA based tools showed superior detection capacity when compared with hierarchical classification. However, the classifiers turned to be key elements in determining the exact time of the abnormal event. Once classification is concluded, the samples classified as with or without anomalies, those with abnormal event were sub-classied according to the type of anomaly: leakage or sensors misfunction. With a success classification rate above 95%, this tool also estimates the severity of the abnormality found. Finally, to locate the problem’s origin, a hybrid approach was used, combining the hierarchical classification and the prototype method. Samples representing previously established abnormal operational conditions are known as prototypes and compared with the faulty data, allowing the identification of the leakage.

1. Introdução **14**

2. Revisão Bibliográfica **19**
2.1 Classificação dos Métodos de Detecção de Falhas 20
2.2 Métodos baseados em séries temporais 26
2.3 Métodos baseados em análise estatística 27
2.4 Modelos Auto Regressivos 28
2.5 Principal Component Analysys - PCA 32

3. Metodologias SPE, Q, Spca, Sdist, Classificadores Hierárquicos, Protótipos e SVI **45**
3.1 Metodologia de cálculo da análise por componentes principais (PCA) 46
3.2 Quantificação da aderência do modelo aos dados via SPE e Q 49
3.3 Quantificação da similaridade entre modelos PCA via Spca, Sdist e SF 51
3.4 Classificação hierárquica 56
3.5 Índice de validade dos sensores (SVI-D) 57
3.6 Método dos protótipos 59

4. Detecção, Localização e Quantificação de Vazamentos via Modelos ARX **60**

5. Rede de Escoamento **67**
5.1. Rede de Fornecimento de Nafta à Planta de Insumos Básicos BRASKEM-Camaçari 67
5.2. Redução da Rede de Escoamento do Sistema Nafta 75
5.3 Modelo de Simulação Dinâmica da Rede Reduzida 81

6. Resultados das Metodologias Baseadas em SPE, Q, Spca, Sdist, Classificadores Hierárquicos, Protótipos e SVI **91**
6.1 Detecção de falhas em sensores 91
6.2 Identificação do sensor em falha 113
6.3 Quantificação de falhas em sensores 116
6.4 Detecção de vazamento na rede de dutos 117
6.5 Identificação do tubo com vazamento 124
6.6 Quantificação de vazamentos 132
6.7 Diferenciação entre falhas em sensores e vazamentos 133

7. Resultados da Metodologia ARX para Detecção, Localização e Quantificação de Vazamentos **136**
7.1 Séries Temporais da Rede sob Ação de Três Vazamentos – Cenário R3L 136
7.2 Detecção, Localização e Quantificação de Vazamentos na Rede Reduzida 150

8 Conclusões **178**
9. Referências Bibliográficas **183**
ANEXO I – Implementação de preditor ARX em Linguagem LADDER **187**

LISTA DE FIGURAS

Figura 2.1. Seqüência de filtros

Figura 3.1: Visão geral das metodologias empregadas

Figura 3.2: Ilustração da técnica de PCA em R^2

Figura 3.3: Representação esquemática do posicionamento temporal

das fases A e B

Figura 4.1. Ciclo de predição composto da fase de treino e fase de utilização

Figura 4.2. Variáveis de entrada e de saída do modelo

Figura 5.1: Rede de Fornecimento de Nafta à Planta de Insumos Básicos da Braskem

Figura 5.2 : Curva Operacional do Canteiro de Bombas do Sistema Nafta

Figura 5.3: Perfis de Elevação (m) vs Vértices ao Longo de Caminhos de Escoamento

Figura 5.4 : Representação Planar em Escala da Rede Simplificada do Sistema Nafta

Figura 5.5 : Representação 3D em Escala da Rede Simplificada do Sistema Nafta

Figura 5.6 : Rede Inicialmente com *N* Vértices e *S* Arestas e após Ruptura e Vazamento

Figura 5.7 : Novas Dimensões de Arestas e Novas Especificações de Vértices após Ruptura e Vazamento em Rede Inicialmente com *N* Vértices e *S* Arestas

Figura 6.1: Sensores da rede de escoamentos NAFTA3

Figura 6.2: Representação em R^2 de uma situação de falha de sensor: (a) Séries Temporais com falha no instante 400 min; (b) Projeção dos dados em subespaço vetorial bidimensional

Figura 6.3: Representação em R^2 de uma situação de falha de sensor com severidade crescente: (a) Séries Temporais com falha de severidade variável; (b) Projeção dos dados em subespaço vetorial bidimensional.

Figura 6.4: Representação em R^2 de uma situação de falha de sensor com desvio positivo e negativo: (a) Séries Temporais com falha; (b) Projeção dos dados em subespaço vetorial bidimensional

Figura 6.5: Representação em R^2 de uma situação de falha de sensor anulada após 100 instantes de tempo: (a) Séries Temporais com falha; (b) Projeção dos dados em subespaço vetorial bidimensional

Figura 6.6: Representação esquemática do posicionamento temporal das fases ”A” e “B”

Figura 6.7: Grau de detecção via SPE em função do valor de lambda selecionado

Figura 6.8: Seqüência com falha a partir do instante 450

Figura 6.9: Seqüência com falha a partir do instante 700

Figura 6.10: Influência do instante de ocorrência da falha sobre o grau de detecção alcançado via SPE e Q.

Figura 6.11: Influência da severidade de falha e localização temporal no grau de detecção. Resultado baseado apenas no uso da ferramenta SPE

Figura 6.12: Influência da severidade na detecção de falhas via SPE e Q

Figura 6.13: Grau de detecção de falha por sensor

Figura 6.14: Resposta das métricas Spca, Sdist e SF a diferentes severidades de falha.

Figura 6.15: Posicionamento temporal das fases “A” e “B”

Figura 6.16: Influência do instante inicial da fase B na métrica Sdist

Figura 6.17: Seqüência de teste “A”, com falha a partir do instante temporal 300

Figura 6.18: Resposta de Sdist a diferentes severidades de falha para a seqüência “A”.

Figura 6.19: Seqüência de teste “B”, com falha a partir do instante temporal 450

Figura 6.20: Resposta de Sdist a diferentes severidades de falha para a seqüência “B

Figura 6.21: Seqüência de teste “C”, com falha a partir do instante temporal 550

Figura 6.22: Resposta de Sdist a diferentes severidades de falha para a seqüência “C”

Figura 6.23: Seqüência de teste “D”, com falha a partir do instante temporal 700.

Figura 6.24: Resposta de Sdist a diferentes severidades de falha para a seqüência “D”

Figura 6.25: Grau de acerto do classificador hierárquico para falhas de sensor

Figura 6.26: Relação entre a severidade da falha e a identificação de sua origem

Figura 6.27: Relação entre a distância dos agrupamentos e a severidade da falha

Figura 6.28: Relação entre a distância dos agrupamentos e a severidade da falha

Figura 6.29: Desempenho de detecção de vazamentos em função do valor de λ empregado

Figura 6.30: Influência do instante da falha na detecção de vazamentos

Figura 6.31: Influência do tubo com vazamento no sucesso da ferramenta detectora de vazamentos

Figura 6.32: Relação entre o grau de detecção via SPE e o diâmetro e comprimento do tubo com falha

Figura 6.33: Resposta em Sdist para Seqüência A

Figura 6.34: Resposta em Sdist para Seqüência B

Figura 6.35: Grau de acerto do classificador hierárquico para falhas de sensor

Figura 6.36: Vazamento no Tubo 1 provocado por um furo de diâmetro 5 centímetros

Figura 6.37: Vazamento no Tubo 2 provocado por um furo de diâmetro 5 centímetros

Figura 6.38: Vazamento no Tubo 4 provocado por um furo de diâmetro 5 centímetros

Figura 6.39: Vazamento no Tubo 5 provocado por um furo de diâmetro 5 centímetros

Figura 6.40: Vazamento no Tubo 6 provocado por um furo de diâmetro 5 centímetros

Figura 6.41: Vazamento no Tubo 7 provocado por um furo de diâmetro centímetros

Figura 6.42: Vazamento no Tubo 1 provocado por um furo de diâmetro 2 centímetros

Figura 6.43: Vazamento no Tubo 2 provocado por um furo de diâmetro 2 centímetros

Figura 6.44: Vazamento no Tubo 4 provocado por um furo de diâmetro 2 centímetros

Figura 6.45: Vazamento no Tubo 5 provocado por um furo de diâmetro 2 centímetros

Figura 6.46: Vazamento no Tubo 6 provocado por um furo de diâmetro 2 centímetros

Figura 6.47: Vazamento no Tubo 7 provocado por um furo de diâmetro 2 centímetros

Figura 6.48: Afastamentos entre clusters em função do diâmetro de furo presente

Figura 6.49: Resposta em Sdist e Spca para vazamentos e falhas em sensor

Figura 6.50: Representação gráfica dos dois componentes principais obtidos a partir do índice de validade dos sensores

Figura 7.1: Comportamento de Entradas Dinâmicas para R3L

Figura 7.2 : Comportamento Temporal de Variáveis para Cenário R3L

Figura 7.3 : Comportamento Temporal de Variáveis para Cenário R3L

Figura 7.4 : Comportamento Temporal de Variáveis para Cenário R3L

Figura 7.5 : Comportamento Temporal de Variáveis para Cenário R3L

Figura 7.6 : Desempenho de Detector ARX para R3LP7P12P15

Figura 7.7 : Desempenho de Detector ARX para R3LP7P15W1

Figura 7.8 : Desempenho de Detector ARX para R3LP7P15W1W5

Figura 7.9 : Séries Temporais de Entradas para RMEP7P15W1W5

Figura 7.10 : Séries Temporais de Sinais de Vazamentos em RMEP7P15W1W5

Figura 7.11 : Séries Temporais de Pressões em RMEP7P15W1W5

Figura 7.12 : Séries Temporais de Vazões de Aresta em RMEP7P15W1W5

Figura 7.13: Séries Temporais de Vazões Externas de Vértices em RMEP7P15W1W5

Figura 7.14: Séries Temporais de Vazões Externas de Vértices em RMEP7P15W1W5

Figura 7.15: Identificação do ARX(3,3,1)-Localizador em RMEP7P15W1W5 (P_7, P_{15})

Figura 7.16: Identificação do ARX(3,3,1)-Localizador em RMEP7P15W1W5 (W_1, W_5)

Figura 7.17: Aderência de ARX(3,3,1)-Localizador em RMEP7P15W1W5 Sem Receber Sinais de Vazamentos;

Figura 7.18: Aderência de ARX(3,3,1)-Localizador em RMEP7P15W1W5 Recebendo Sinais Corretos de Vazamentos;

Figura 7.19: Rede com Programação de Avaria no *Tubo* 9 Original Série Temporal R1LP7P15W1W5

Figura 7.20: Séries Temporais de Entradas Dinâmicas para R1LP7P15W1W5

Figura 7.21: Séries Temporais de Pressões para R1LP7P15W1W5

Figura 7.22: Séries Temporais de Vazões de Aresta e Externas de Vértices para R1LP7P15W1W5 (A Taxa de Vazamento é W_{17})

Figura 7.23: Predições do ARX-Detector vs Valores de Processo em R1LP7P15W1W5

Figura 7.24: Predições do ARX-Detector vs Valores de Processo em R1LP7P15W1W5

Figura 7.25: Predições do ARX-Detector vs Valores de Processo em R1LP7P15W1W5

Figura 7.26: Busca de Localização de Vazamento para *Tubo* 3

Figura 7.27: Busca de Localização de Vazamento para *Tubo* 6

Figura 7.28: Busca de Localização de Vazamento para *Tubo* 9

Figura 7.29: Busca de Localização de Vazamento para *Tubo* 11

LISTA DE TABELAS:

Tabela 5.1: Diâmetro Interno e Comprimento de *Tubos* do Sistema Nafta

Tabela 5.2: Elevações dos Vértices Principais do Sistema Nafta

Tabela 5.3: Sensores do Sistema Nafta ([#]: vértice)

Tabela 5.4: Diâmetro Interno e Comprimento de *Tubos* da Rede Reduzida

Tabela 5.5: Sensores da Rede Reduzida ([#]: vértice)

Tabela 6.1: Localização e tipo dos sensores considerados

Tabela 6.2: Grau de explicação acumulado por componente principal

Tabela 6.3: Percentagem de falhas detectadas em cada abordagem

Tabela 6.4: Resumo das seqüências empregadas para Sdist estático

Tabela 6.5: Valor médio da série temporal gerada por cada sensor sem falha

Tabela 6.6: Faixas temporais das fases B testadas

Tabela 6.7: Resultados de SVI para falha no sensor 1

Tabela 6.8: Similaridade entre os dados experimentais com falha e cada protótipo característico

Tabela 6.9: Percentagem de vazamentos detectados em cada abordagem

Tabela 6.10: Resultados de SVI para vazamento

Tabela 6.11: Resultados de SVI para vazamento

Tabela 6.12: Similaridade entre os dados experimentais com falha e cada protótipo característico

Tabela 6.13: Grau de acerto na identificação da origem do vazamento através do método dos protótipos

Tabela 6.14: Grau de sucesso obtido na identificação do tipo de anomalia via classificadores hierárquicos e método dos protótipos

Tabela 7.1: Programação de Vazamentos para o Cenário R3L

Tabela 7.2 : Resultados Finais de Localização de Vazamento para R1LP7P15W1W5

1. Introdução

Durante grande parte do século passado a indústria química ocupou uma posição invejável no mundo, sendo uma das principais atividades impulsionadoras do desenvolvimento da humanidade através dos benefícios advindos de seus produtos. Essa imagem positiva perante a sociedade passou a mudar com o crescimento da consciência social sobre as questões ambientais, especialmente após uma série de acidentes amplamente divulgados pela mídia nas décadas de 1970 e 1980. Acidentes como Sevezo na Itália em 1976, onde um vazamento de tetracloro-dibenzo-dioxina deixou mais de 200 mil feridos; Bhopal na Índia com 3800 mortos após vazamento de metil isocianato; Cidade do México com 490 mortos e 7000 feridos devido a explosão de GLP e Cubatão com 500 mortos devido a um incêndio provocado pelo vazamento de hidrocarbonetos. Destaca-se que estes três últimos acidentes ocorreram no mesmo ano: 1984.

Desenvolveu-se então maior percepção aos riscos inerentes as atividades da indústria química, exigindo desta melhorias em seus programas de ação, além de um maior controle por parte do governo. A legislação ambiental em todo o mundo e em especial no Brasil vem se tornando cada vez mais rigorosa, surgindo inclusive a figura da “teoria do risco assumido” e do “nexo causal”. O fundamento da teoria do risco assumido é que só se deve atuar numa área perigosa caso se tenha capacidade de assumir todos os riscos inerentes àquela atividade. Quem atua numa atividade perigosa deve ser responsável direto por qualquer dano causado: ou assume o risco ou não exerce a função **[GUSMÃO e DE MARTINI, 2009]**. Em paralelo tem-se o “nexo causal”: basta provar a existência de nexo causal entre a atividade produtora e o dano ambiental para configurar responsabilidade objetiva. Assim, a responsabilidade em relação ao dano ambiental existe independente da culpa ou prática de ato ilegal **[GUSMÃO e DE MARTINI, 2009]**. A Lei 9.605, de 12 de fevereiro de 1998 dispõe sobre as sanções penais e administrativas derivadas de condutas e atividades lesivas ao meio ambiente.

Desde seus primórdios a indústria química sempre se baseou no transporte de matérias-primas e produtos por longas distâncias através de dutovias. Estas encontram aplicação no deslocamento dos mais diversos produtos, tais como

petróleo, seus derivados e gases combustíveis. Dutovias são formas baratas, seguras, confiáveis e muito eficientes para o transporte de fluídos. Alternativas como o transporte ferroviário, naval ou rodoviário apresentam inúmeras desvantagens quando comparados ao uso de dutos. Embora altamente confiáveis, dutovias empregadas no transporte de produtos tóxicos e/ou inflamáveis apresentam um perigo operacional inerente, ocasionado pela conjunção da operação sob pressão e vazões normalmente elevadas com a sempre possível manifestação de fadiga e colapso de material das tubulações **[SILVA *et al.*, 1996]**. Tal conjunção levará a vazamentos que, dependendo de sua localização e magnitude, têm óbvia conotação desastrosa, gerando danos ambientais, danos materiais, perdas de inventário, incêndios, explosões e a possibilidade de perda de vidas humanas. Evitar tais falhas e, quando da sua ocorrência, minimizar a extensão de suas consequências, torna-se uma preocupação constante da indústria, agências reguladoras, e da sociedade como um todo. A legislação ambiental brasileira incorpora a exigência de implantação de sistemas de detecção de vazamentos (SDV) em dutovias. Segundo a Resolução CONAMA nº1, de 23 de Janeiro de 1986, **[CONAMA, 1986]** a instalação de oleodutos e gasodutos depende da elaboração de “Estudo de Impacto Ambiental” e “Relatório de Impacto Ambiental”. Já a Resolução nº 293 de 12 de Dezembro de 2001 **[CONAMA, 2001]** define a elaboração do “Plano de Emergência Individual” para acidentes que gerem vazamentos em instalações como dutos, terminais e plataformas.

Além de vazamentos, uma rede de escoamentos também pode sofrer com falhas em sensores, especialmente aquelas de maior extensão e número de sensores distribuídos. Em uma dutovia construída e operada de forma adequada, falhas em sensores por si só não são capazes de ocasionar acidentes, porém tendem a causar problemas operacionais que podem até mesmo resultar na interrupção da operação. Paralisar a operação acarreta prejuízos que alcançam facilmente a faixa dos milhões de dólares, seja por lucro cessante, seja por multas aplicadas por clientes não atendidos. É interessante salientar que, com a difusão do uso de sensores na indústria química, a possibilidade de ocorrer falha em algum deles também aumentou. Ou seja, as dutovias mais bem instrumentadas seriam exatamente aquelas com maior probabilidade de sofrer com esse tipo de problema.

Desse modo, detectar falhas do tipo “anomalias na instrumentação” ou a ocorrência de vazamentos são preocupações constantes na operação de uma dutovia.

Os métodos para detecção instantânea e localização de vazamentos (IDLL – “instantaneous detection and localization of leaks”) operam comparando medições “on-line” do processo com as predições correspondentes. É indício de possível vazamento a ocorrência de diferenças estatisticamente significativas entre o processo e o seu modelo operando em tempo real **[VAZ JUNIOR, ARAUJO e MEDEIROS, 2010].** Os métodos IDLL podem ser ordenados em três classes de acordo com a forma de geração da predição: (i) descrições fenomenológicas do processo, (ii) aproximações “caixa-preta” que não demandam descrição física do processo, e (iii) estratégias intermediárias, via linearizações, funções de transferência aproximadas, etc. Exemplos de IDLL da classe (ii) são redes neurais e métodos de predição baseados em aproximações de séries temporais, como o uso de preditores auto-regressivos (AR) e auto-regressivos e média móvel (ARMA). Os sistemas baseados em séries temporais para a detecção de vazamentos têm apresentado significativo sucesso em diversas aplicações.

Metodologias semelhantes podem ser empregadas na detecção de falhas em sensores. Tais falhas mudam a correlação entre o sinal do sensor defeituoso e os demais, possibilitando assim não apenas detectar a anomalia como também identificá-la. Diferentes autores empregam com sucesso ferramentas baseadas em análise de componentes principais (PCA) e teoria dos protótipos na detecção desse tipo de problema **[VAZ JUNIOR, ARAUJO e MEDEIROS, 2009]**. Na classificação acima, essas duas ferramentas também podem ser entendidas como aproximações “caixa-preta”.

O presente estudo aborda a elaboração, uso e avaliação de diversos sistemas de detecção de anomalias do tipo falha de sensor e vazamentos, em uma rede de dutos real operada pela BRASKEM, Camaçari. Como anomalias, entende-se a ocorrência de vazamentos e falhas da instrumentação instalada ao longo da dutovia.

No Capítulo 2, encontra-se uma detalhada revisão sobre os métodos de detecção de anomalias atualmente em uso. Especial atenção é dada aos métodos

baseados em séries temporais, embora aproximações puramente baseadas em hardware também sejam citadas. Dentre as técnicas que empregam séries temporais, maior destaque foi dado àquelas desenvolvidas a partir do método da Análise de Componentes Principais (PCA) e de modelos auto-regressivos.

O Capítulo 3 detalha os procedimentos matemáticos envolvidos na análise de Componentes Principais, metodologia que serve de base para o entendimento de diferentes métricas apresentadas na seqüência, tais como SPE (*square prediction error*), Q, Spca (*similarity* PCA) e Sdist (*distance similarity factor*). Abordagem dinâmica é então desenvolvida empregando as medidas de similaridade, permitindo superar inúmeras limitações inerentes ao método estático. Técnicas similares, tais como Métodos dos Protótipos, Classificadores Hierárquicos e Índice de Validade dos Sensores (SVI-D) também são descritos no Capítulo 3.

As abordagens baseadas em modelos auto-regressivos com componente externo (ARX) são tratadas no Capítulo 4, complementando as metodologias do capítulo anterior. Neste Capítulo, os modelos ARX são empregados não apenas na detecção de vazamentos, como também na identificação do tubo onde existe o furo e na estimativa do diâmetro do furo.

As ferramentas de detecção de falhas aqui desenvolvidas foram testadas em uma rede virtual que reproduz o comportamento de uma dutovia atualmente em operação no Brasil. O detalhamento estrutural dessa dutovia e o desenvolvimento de um simulador de redes de escoamento são assuntos tratados no Capítulo 5.

Os resultados encontrados são apresentados nos Capítulos 6 e 7. No Capítulo 6, descrevem-se os resultados para detecção de vazamentos e falhas em sensores empregando as metodologias baseadas em PCA e classificação hierárquica. Protótipos e SVI-D são usados na identificação da localização da falha, enquanto classificadores hierárquicos aparecem como opção na quantificação da

anomalia. O Capítulo 6 é encerrado com a busca pela diferenciação entre falhas de sensores e vazamentos.

No Capitulo 7, são introduzidos os modelos auto-regressivos, uma opção às abordagens anteriormente descritas. Resultados de detecção, localização e quantificação de vazamentos utilizando tal abordagem são detalhados. Os Capítulos 6 e 7 complementam-se, mostrando alternativas e possibilidades para a construção de ferramentas de detecção de anomalias baseadas em abordagens envolvendo o uso de mais de uma metodologia.

Finalmente, os Capítulos 8 e 9 apresentam respectivamente as conclusões alcançadas e as referências bibliográficas deste trabalho. O Anexo I apresenta a implementação da metodologia detectora ARX em liguagem LADDER, empregada em Controladores Lógico Programáveis (CLP).

2. Revisão Bibliográfica

Especialmente na última década do século passado o mercado de equipamentos eletrônicos voltados para indústria química e petroquímica passou por uma forte onda de crescimento. O barateamento da micro-eletrônica permitiu que esses equipamentos fossem aplicados em diferentes situações para as quais anteriormente seus custos eram proibitivos. Um grande número de sensores, tais como pressão, vazão e temperatura, passaram a ser empregados mesmo nas instalações industriais mais simples. No campo dos controladores, diferentes fabricantes passaram a comercializar CLPs (controlador lógico programável) de dimensões reduzidas, custos acessíveis, construção modular e operação cada vez mais simplificada.

O aumento do número total de sensores instalados teve como conseqüência direta um aumento na quantidade de dados acerca do comportamento do processo que, juntamente com o uso de *softwares* supervisórios e computadores cada vez mais desenvolvidos, ampliou enormemente a possibilidade de análises *on-line* e *off-line* visando à otimização da produção. O grande volume de dados permite o uso de ferramentas sofisticadas de análise, extraindo informação relevante a partir da base bruta. Mais do que possibilitar, o volume crescente de dados brutos passou a *exigir* a aplicação de ferramentas matemáticas de análise. Para **SINGHAL e SEBORG [2002]**, é correto afirmar que um banco de dados com informações detalhadas sobre a operação contém informações potencialmente valiosas. Porém, também é certo supor a dificuldade encontrada em se extrair informação útil a partir da enorme quantidade de dados brutos. Cria-se assim a figura do sistema “rico em dados, mas pobre em informações”.

A ampliação da instrumentação de plantas industriais permite aumentar o controle operacional, reduzindo perdas, evitando acidentes e maximizando o lucro. Por outro lado, a presença de um número maior de sensores eleva também a chance de ocorrência de falhas nesta malha de instrumentação. A explicação é simples: enquanto a probabilidade de um único sensor tornar-se defeituoso permanece baixa, a chance de algum sensor em um conjunto grande deles tornar-se defeituoso cresce em função do número de indivíduos do universo. Com o objetivo

de se garantir uma operação eficiente de qualquer processo é importante promover prontamente a detecção de equipamentos que apresentem mau funcionamento. Detectar e identificar falhas são os primeiros passos para uma rápida resolução do problema **[KANO *et al.* , 2000]**

Se por um lado a presença de um número elevado de sensores em uma instalação industrial eleva o número anual de falhas, essa grande quantidade de instrumentos acaba por prover valiosa redundância entre o sinal gerado por eles. É exatamente essa redundância que vem sendo fortemente explorada por diferentes metodologias de detecção de anormalidades **[DUNIA et al., 1996a]**. As séries temporais geradas pelos sensores apresentam-se em geral altamente correlacionadas quando obtidas sob condições normais de operação. Tal correlação advém dos fenômenos físicos e químicos que governam a operação do processo, tais como transferência de massa e energia. Consequentemente, a falha de um instrumento pode ser caracterizada pela perda de correlação com os demais sensores, permitindo a identificação da origem da falha **[DUNIA et al., 1996a]**.

A detecção de falhas a partir da perda de correlação entre os sinais gerados pelos sensores é apenas uma das diversas técnicas atualmente disponíveis para detecção de anomalias.

2.1 Classificação dos Métodos de Detecção de Falhas

Os sistemas de detecção de falhas podem ser classificados de diferentes modos. A divisão proposta por **ZHANG [1996]** é uma das mais simples e úteis, sendo até hoje empregada. As metodologias são separadas em três grandes grupos: técnicas biológicas (“método biológico”), técnicas baseadas em equipamentos físicos (“método de *hardware*”) e técnicas que utilizam programação computacional (“método de *software*”). Essa classificação foi inicialmente proposta por **ZHANG [1996]** para métodos de detecção de vazamentos, podendo facilmente ser generalizada para sistemas detectores de anomalias em geral. **ABHULIMEN e SUSU [2007]** descrevem as virtudes e limitações de cada abordagem.

2.1.1 Métodos biológicos

São as metodologias mais simples e baratas, podendo ser usadas independente da instrumentação disponível. Os métodos biológicos permitem identificar vazamentos através de inspeção visual, detecção de odor característico ou ruído sonoro **[ZHANG, 1996]**. Contudo, as abordagens baseadas em métodos biológicos não são possíveis de serem aplicadas na detecção de anomalias em equipamentos ou sensores, sendo este tipo de falha somente detectada por outras técnicas.

No contexto de dutovias, os métodos biológicos muitas vezes são baseados na experiência pessoal de operadores ou no treinamento de cães sensíveis ao odor das substâncias transportadas. Alterações na vegetação da superfície em áreas onde uma dutovia é subterrânea também é fator de alerta de vazamentos **[COLOMBO *et al*, 2009]**. Embora extremamente úteis em diversas situações, essas técnicas mostram-se altamente limitadas, sendo eficazes apenas quando aplicadas a pequenas redes de escoamento, ou nas extremidades finais de uma grande rede. Tais procedimentos também não são adequados para dutovias que transportam material inflamável ou altamente tóxico, as quais demandam abordagens mais sofisticadas.

2.1.2 Métodos de hardware

Muito mais desenvolvidos e precisos que os métodos biológicos são os chamados “métodos de hardware”. Estes utilizam diferentes dispositivos e instrumentos físicos de medição, tais como sensores de pressão, temperatura e vazão de fluido distribuídos ao longo da dutovia **[ZHANG, 1996].** Além destes, sensores de vibração ou acústicos também são empregados na detecção de vazamentos **[COLOMBO *et al.*, 2009].** Estes se baseiam na detecção das ondas sonoras ou de vibração geradas no instante do vazamento. O escoamento de gás através do orifício de vazamento induz ondas que se propagam tanto na superfície do duto na forma de vibração, quanto ondas sonoras. Através da propagação das

ondas é possível detectar e caracterizar a posição do vazamento. Esse método mostra-se especialmente eficaz para vazamentos em um raio de 250m do ponto de análise **[COLOMBO *et al.*, 2009]**. Detecções em pontos mais afastados ficam prejudicadas devido a perda de intensidade da onda. Curiosamente, este método não funciona bem para grandes vazamentos, pois estes não geram os ruídos de alta freqüência característicos **[COLOMBO *et al.*, 2009].** Embora atualmente disponha-se de metodologias mais sofisticadas, a detecção acústica ainda encontra emprego, especialmente nas vastas redes urbanas de abastecimento de água. Uma grande dificuldade na implementação desta metodologia está na necessidade da inclusão de diversos filtros eletrônicos para a remoção de ruídos provocados por agentes externos à tubulação, como o tráfego pesado de uma rodovia, o acionamento de bombas e válvulas, etc. Ressalta-se a importância de se remover corretamente as interferências externas de modo a não gerar falsos alarmes de vazamento, porém, a especificação inadequada de um filtro pode remover ruídos causados por um pequeno vazamento, havendo necessidade de um compromisso entre estes aspectos, comum em diversas metodologias de detecção de vazamentos: a sensibilidade a pequenos vazamentos aumenta significativamente o número de falsos alarmes. Estes, além de gerarem prejuízos financeiros devido à interrupção da operação da linha e ao início dos procedimentos de emergência, reduz a credibilidade da ferramenta. **[GAO *et al.*, 2005].** Para **ZHANG *et al* [2009]** detectar vazamentos em uma rede de escoamento não é mais o maior desafio, mas sim evitar os falsos alarmes.

Outro exemplo de métodologia de deteção de vazamentos baseada em hardware é detalhada por **BIMPAS *et. al.* [2010].** Os autores descrevem o uso de ondas de radar (GPR – *Ground Penetrating Radar*) na busca por vazamentos em redes subterrâneas de abastecimento de água. Tais radares já apresentam uso corrente em diferentes atividades de sondagem do solo, tais como na busca por minas terrestres ou sítios arqueológicos. Desse modo, a análise do perfil de reflexão da onda é capaz de relevar a ocorrência de uma anomalia na rede subterrânea.

De acordo com a volatilidade dos fluidos transportados pode-se adotar um sistema de monitoramento de gases para verificar a composição do ar atmosférico

nas proximidades da dutovia. Para esse fim, pode-se usar unidades fixas instaladas ao longo do duto ou unidades portáteis que possam ser transportadas, ou, ainda, empregar um pequeno tubo secundário instalado paralelamente à tubulação principal com a função de sugar o ar das proximidades da dutovia. A análise contínua do ar amostrado permite verificar a existência de vazamento. Essa ainda é considerada uma abordagem cara e de uso restrito **[ZHANG, 1996].**

ZHOU e ZENG [2006] descrevem o emprego de câmeras de vídeo infra-vermelhas na detecção de vazamentos de vapores orgânicos provenientes de equipamentos industriais. As câmeras de IR apontam mudanças no perfil de temperatura da região vigiada. Mudanças de temperatura são usuais em processos de vazamentos, de modo que a análise das imagens geradas pode ser uma alternativa economicamente favorável para a detecção de anomalias em plantas químicas e petroquímicas.

GAO *et al*, [2005] apresentam um sistema de detecção baseado na emissão/recepção de raio laser empregado em dutos para o transporte de gás natural. O laser é emitido de modo a “varrer” o ar próximo a tubulação, tendo sua intensidade medida por um detector. O sinal gerado pelo detector na presença de um vazamento é distinto daquele obtido durante o funcionamento normal da rede. Isso porque, em caso de vazamento, o laser irá cruzar a nuvem de gás natural formada, interferindo na passagem da luz.

Semelhante à detecção de vapores, encontra-se a detecção de líquidos que escapem dos dutos. Através da instalação de pequenas resistências metálicas ao longo da estrutura do duto, pode-se verificar mudanças na reflexão de pulsos elétricos como resultado da interação do metal com o líquido vazado. No lugar de resistências elétricas, também é possível usar fibra ótica com a mesma finalidade **[LÓPEZ *et al.*, 2002]**.

2.1.3 Métodos de software

Esses métodos compreendem técnicas baseadas em software que utilizam principalmente medições de pressão e vazão obtidas de sensores de campo. Os dados gerados são armazenados em uma base de dados do processo, sendo esta continuamente analisada pelo software especialista. Inúmeras tecnologias de detecção de anomalias atualmente em uso são baseadas na análise contínua de medições de pressão, temperatura, vazão e densidade **[ABHULIMEN e SUSU, 2004 e KAM, 2009].**

O rompimento da parede da tubulação provoca repentina queda de pressão no local do vazamento, que é rapidamente neutralizada, formando-se uma “onda de pressão”. A onda resultante percorre a tubulação em ambos os sentidos, podendo ser detectada nos sensores de pressão a montante e a jusante. **COLOMBO *et al.* [2009]** descrevem as propriedades transientes das onda de pressão. A resposta dos sensores às perturbações transientes geradas por um vazamento pode ser identificada através de métodos de tratamento de dados e assim empregada na detecção de falhas na rede. **COLOMBO *et al.* [2009]** ressaltam que o estudo das resposta hidráulica transiente revela uma quantidade muito maior de informações relevantes do que apenas a análise dos estados estacionários. É importante notar que os medidores de pressão e vazão distribuídos ao longo da rede geram dados que são então processados por softwares especialistas. Desse modo, a classificação do método de detecção como “por hardware” ou “por software” não é exata, sendo ambos usados em conjunto.

ZHANG *et al* [2009] demonstram o uso de um sistema de detecção de vazamentos baseado nos dados disponíveis no sistema SCADA do operador. O caso real analisado pelos autores baseia-se em uma rede extremamente bem instrumentada, contando com grande quantidade de informações sobre pressão, vazão (sensores ultra sônicos), e temperatura do fluido ao longo da dutovia. Os autores detalham a implementação de um sistema híbrido unindo metodologias diversas tais como “pressão negativa”, “mudanças de vazão e pressão”, “balanço de massa e volume”, “modelagem dinâmica”, entre outras.

Através do desenvolvimento de softwares específicos torna-se possível detectar não apenas vazamentos, como também diversas outras anomalias possíveis de ocorrer durante a operação de uma rede de escoamento. Os métodos baseados em software são especialmente úteis em dutovias de grande extensão, onde o uso dos outros métodos fica dificultado, como o caso do gasoduto que liga Niigata a Sendai no Japão, transportando gás natural por 250 quilômetros **[FUKUSHIMA *et al.*, 2000].** Para estruturas desse porte, o uso de métodos biológicos ou exclusivamente baseados em hardware se tornaria inviável.

Os métodos de software podem ser divididos em duas sub-classes: os baseados em equações fenomenológicas do processo, como balanço de massa e energia, e os métodos baseados em análise estatística dos dados. Nessa última subclasse encontram-se os modelos ARX, ARMAX e técnicas baseadas em análise de componentes principais e reconhecimento de padrões. Embora o presente trabalho seja focado nesta sub-classe, torna-se importante ressaltar a existência de inúmeras abordagens de detecção da vazamentos baseadas em modelos fenomenológicos, tal como aquela apresentada por **ABHULIMEN e SUSU [2007].** Recentemente **KAM [2010]** apresentou estudo sobre a modelagem de vazamentos em dutos submersos.

Antes de analisar mais detalhadamente métodos baseados em séries temporais, cabe ressaltar que a seleção da melhor metodologia de detecção depende de uma série de fatores que incluem: características da rede, do fluido, capacidade de instrumentação e comunicação de dados, e interesses econômicos **[ALASKA DEPARTMENT OF ENVIRONMENTAL CONSERVATION, 2006].** Fatores aparentemente simples, como a composição do duto, influenciam na escolha do método. A localização e extensão da dutovia pode tornar inviável, por exemplo, sua inspeção visual. Por fim, a toxidez e a inflamabilidade do material transportado também devem ser consideradas. Na busca por vazamentos na rede hidráulica de um edifício, o método visual supera os demais graças ao seu custo reduzido. O mesmo já não pode ser dito quando a detecção de vazamentos precisa ser efetuada em um duto de transporte de gás com quilômetros de extensão.

Apesar da diversidade de metodologias citadas aqui, qualquer sistema detector pode ser avaliado segundo critérios objetivos, tais como: “sensibilidade a pequenas anomalias”, “capacidade de localizar a origem do problema”, “adaptabilidade”, “disponibilidade total (7x24)”, “freqüência de falsos alarmes”, “dificuldade de manutenção”, “custo de instalação e operação” entre outros **[ZHANG, 1996].**

2.2 Métodos baseados em séries temporais

A ampliação da capacidade de memória e armazenamento dos computadores pessoais (PCs) permite o armazenamento de uma quantidade cada vez maior de dados sobre o processo. Assim, longas séries temporais com os dados operacionais adquiridos dos sensores instalados podem ser construídas. Diferentes métodos de software atualmente em uso para a formulação de sistemas de detecção de anomalias citados acima são baseados em séries temporais.

Antes de prosseguir cabe definir série temporal como sendo uma coleção de observações geradas seqüencialmente ao longo do tempo – a exemplo das bases de dados de processos químicos em sistemas supervisórios **[MORETTIN e TOLOI, 2004].** Os dados de uma série temporal são ordenados com respeito ao tempo e coletados com intervalos regulares e conhecidos. O intervalo entre duas observações pode variar da faixa de segundos e minutos até a faixa de meses e anos dependendo a aplicação a que se destinam. Enquanto sistemas de detecção de vazamentos (SDV) trabalham com espaçamento na faixa de segundos ou minutos, ferramentas de elaboração de previsões econômicas podem ser reguladas para intervalos maiores.

No desenvolvimento de softwares de detecção de falhas e/ou vazamentos, o objetivo principal costuma ser ajustar modelos que reproduzam adequadamente o comportamento do processo ao longo do tempo sob condições normais de operação. Esses modelos podem então ser usados na previsão do comportamento

futuro das variáveis de processo em situações de rotina. As previsões são então contrastadas com a realidade, permitindo a detecção de desvios operacionais. Modelos ARX e ARMAX de detecção de vazamentos operam deste modo **[VAZ JUNIOR *et al.,* 2010]**.

Em outra abordagem, os modelos são ajustados com dados referentes a janelas temporais distintas, onde sabidamente uma delas refere-se a operação normal da planta. A similaridade ou não entre as janelas temporais permite aferir se o processo permanece operando normalmente. Assim funcionam os sistemas que empregam PCA (Principal Component Analysis).

Modelos baseados em uma única variável são chamados de “modelos monovariáveis” e usam apenas uma única série histórica. Por outro lado, modelos que usam explicitamente mais de uma variável para descrever o comportamento do parâmetro desejado são chamados “modelos multivariáveis” ou “múltiplos” **[VANDAELE, 1983]**. Ressalta-se que a incorporação de outras variáveis que exerçam influência significativa sobre a variável prevista deve ser feita de forma criteriosa e com parcimônia. Sabe-se que a incorporação de uma variável relevante pode fortalecer a ferramenta de previsão, tornando-a mais robusta. Todavia a adição de um grande número de variáveis de forma desnecessária pode não apenas aumentar a carga de processamento computacional como inclusive dificultar a identificação de falhas.

2.3 Métodos baseados em análise estatística

Ao contrário das abordagens baseadas em modelos fenomenológicos, métodos baseados em análise estatística multivariável não necessitam de um sistema de equações que descreva as características físicas e químicas do processo. As abordagens estatísticas são capazes de manipular uma grande quantidade de dados de processo altamente correlacionados, tornando-se assim uma poderosa ferramenta na detecção de anormalidades. Muitos métodos baseados

em análise estatística empregam modelos desenvolvidos a partir de séries temporais na sua estrutura, conforme mencionado anteriormente.

Em diversas situações opta-se por abordagens baseadas em análise de dados, no lugar daquelas baseadas em equações fenomenológicas. Isso ocorre porque modelos fenomenológicos, quando existentes, tendem a apresentar elevada complexidade e demandam maior esforço computacional para sua simulação **[KANO *et al*., 2000].** Em algumas situações o modelo fenomenológico detalhado do processo não é nem mesmo conhecido, frente à complexidade operacional do sistema analisado. Por outro lado, modelos estatísticos precisam apenas das séries temporais. Outra vantagem da análise estatística é sua maior facilidade de adaptação de um processo para outro. Enquanto metodologias construídas a partir de modelos fenomenológicos dependem fortemente das características do processo estudado, os modelos estatísticos podem ser facilmente adaptados para diferentes aplicações industriais. Assim, as ferramentas de detecção de falhas desenvolvidas para dutovias podem facilmente ser adaptadas para a operação em caldeiras ou reatores.

Dentre as inúmeras ferramentas estatísticas aplicadas em análise de dados, duas são frequentemente empregadas em sistemas de detecção de falhas: modelos Auto Regressivos (ARIMA, ARX, ARMAX, entre outros) e análise dos componentes principais, ou PCA ("*Principal component analysis*"). Os itens seguintes apresentam uma rápida introdução a cada uma dessas abordagens, juntamente com uma visão geral das suas aplicações em diferentes áreas.

2.4 Modelos Auto Regressivos

Os modelos ARIMA resultam da combinação de três componentes denominados "filtros": o componente auto-regressivo (AR), o filtro de integração (I) e o componente de médias móveis (MA) **[BOX, JENKINS e REINSEL, 1994]**. Uma série temporal pode ser modelada pelos três filtros ou apenas um subconjunto deles **[WERNER e RIBEIRO, 2003]**. A Figura 2.1 resume, de modo esquemático, um exemplo de tratamento de dados usando os três filtros. Após aplicar o filtro de integração (I) aos dados observados (z_t), obtém-se a série filtrada (w_t). O filtro AR

produz então a série "e_t" e, finalmente, a série composta apenas por ruído aleatório, (a_t), é gerada após o uso do filtro MA **[VANDAELE, 1983]**. O objetivo da aplicação dessa seqüência de filtros é tornar a série original de dados em uma série de ruídos aleatórios. Encontrar os filtros necessários, juntamente com suas ordens, de modo a obter a série "a_t" equivale a achar estrutura, identificar a forma, e construir o modelo da série temporal **[VANDAELE, 1983]**.

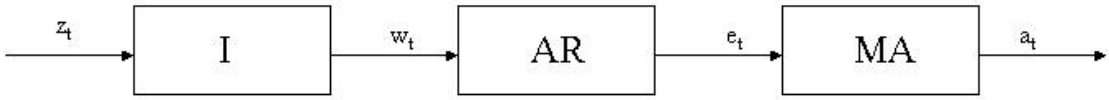


Figura 2.1. Seqüência de filtros

A metodologia mais tradicional de uso de modelos auto-regressivos em sistemas de detecção de falhas é descrita a seguir:

Identificação: é a etapa de definição dos filtros mais adequados para a série analisada. Utilizar os três filtros quando apenas o AR é necessário não promove melhores resultados **[VANDAELE, 1983].** Modelos ARIMA são de processamento matemático mais lento e, por terem mais parâmetros, demandam maior esforço na etapa de estimação.

Estimação: escolhidos os filtros, estimam-se as suas ordens. Novamente, é importante manter o modelo na sua forma mais simples, com a menor ordem em cada filtro. Ordens elevadas tornam o uso do modelo mais lento e elevam o risco de que funcione apenas com os dados para os quais foi treinado (sobre-parametrização). Um estudo completo sobre as etapas do processo de identificação e estimação foi feito por **VAZ JUNIOR [2006].**

Diagnóstico: determinar a acurácia do modelo estimado através da comparação dos dados previstos com os dados obtidos dos sensores de campo. Além de avaliar o comportamento do modelo em situações de

normalidade, a etapa de diagnóstico deve também testar a capacidade de detectar falhas conhecidas. São geradas falhas controladas em pontos estabelecidos da rede e verifica-se a capacidade do sistema de alertar sobre sua ocorrência.

Uso: após todas as fases anteriores terem sido efetuadas com sucesso, a ferramenta é colocada em uso.

Os sistemas de detecção de falhas via séries temporais baseiam-se na comparação dos valores preditos pelo modelo com os valores efetivamente medidos pelos sensores remotos. A capacidade preditora dos modelos auto-regressivos os tornam adequados para esta finalidade. O desenvolvimento de um bom sistema de previsão que, baseado nas informações disponíveis na base de dados, possa predizer valores futuros, é essencial para o sistema de detecção de falhas.

Os aplicativos de previsão baseados em modelos auto-regressivos são tradicionalmente usados na elaboração de previsões em diversos ramos do conhecimento, não estando limitados a aplicações de detecção de falhas em sensores ou vazamentos. Previsões econométricas baseadas em séries temporais podem ser abordadas usando modelos auto-regressivos, como desenvolvido por **VAZ JUNIOR e BARBABELA [2006, 2007],** no estudo do comportamento dos preços dos derivados do petróleo. Diversas tarefas administrativas e gerenciais baseiam-se em um sistema acurado de previsão de vendas semanais e/ou mensais, regulando a produção e a encomenda de matéria-prima em função das projeções de preço e demanda futura. Previsões neste campo são apresentadas por **WERNER e RIBEIRO [2003].** Órgãos governamentais utilizam previsões de demanda futura de energia elétrica para avaliar a necessidade da construção de novas unidades geradoras. Profissionais do mercado financeiro trabalham com estimativas de índices futuros de inflação, desemprego e outros parâmetros macroeconômicos. Indústrias químicas regulam suas previsões de lucro futuro também em função da previsão do valor de mercado de *commodities* como petróleo, aço, e insumos agrícolas **[SADORSKY, 2002]**.

WU et al. [2005] usam modelos tipo ARMAX (Autoregressive Moving Average Exogenous) para avaliar os parâmetros significativos na previsão de demanda de moeda por parte da população economicamente ativa do Taiwan. Os ARMAX são comparados a outras classes de modelos também baseadas em séries temporais, obtendo bom desempenho. Ainda na área econômica, modelos ARMAX são usados para prever a demanda mundial por viagens para a Turquia **[AKAL, 2004]**. Sabe-se que turismo internacional é uma atividade altamente sazonal, de modo que esse fator deve ser considerado no tratamento inicial dos dados. Estratégias de tratamento de séries temporais usando modelos da classe ARMAX foram consideradas as mais adequadas a esse tipo de previsão.

Também usando modelos auto-regressivos, **MADDISON [2005]** estima os efeitos de longo prazo da poluição do ar encontrada em grandes cidades em relação ao número de internações em hospitais. Dados relativos à cidade de Londres são usados. Efeitos sazonais devidos às diferentes estações do ano precisam ser removidos para diagnosticar a influência apenas da poluição do ar sobre a saúde da população. Séries de dados de gases tipicamente associados a poluição atmosférica, como SO_2, NO_2 e CO, são consideradas. Modelos ARMAX são, segundo o autor, especialmente indicados para diagnóstico de influências temporais de longo prazo.

FRAUSTO, PIETERS e DELTOUR [2003] aplicam modelos tipo ARX (Autoregressive Exogenous) e ARMAX para descrever a temperatura interna do ar em uma estufa. A previsão da temperatura em uma estufa é algo complexo, onde inúmeras variáveis apresentam alto nível de interação, em um sofisticado problema de troca térmica. Dados como temperatura externa, umidade relativa do ar, radiação solar e condições atmosféricas são usadas como variáveis de entrada. Os autores concluem que os modelos ARX conseguem performance superior aos ARMAX.

Em outra aplicação, **WARNES et al. [1996]** desenvolvem sensores virtuais através de modelos ARMAX. Esta aplicação de séries temporais auxilia a monitoração em uma situação em que a tecnologia disponível para sensores bioquímicos on-line não permitem que parâmetros significativos para o acompanhamento do processo possam ser medidos em tempo real. Em geral tais

parâmetros são medidos na forma off-line, em análises laboratoriais, causando atraso que impede ações corretivas em tempo real. Neste cenário, os sensores virtuais usam modelos matemáticos para estimar tais parâmetros a partir de medições auxiliares realizadas on-line.

Modelos ARMAX também são usados na identificação e construção de sistemas de controle adaptativos, especialmente para controladores auto-sintonizáveis **[GESSING e BLACHUTA, 1996]**. O trabalho de **RIVERA e GAIKWAD [1996]** usa modelos tipo ARX aplicados na sintonia de controladores digitais.

2.5 Principal Component Analysis - PCA

SMITH [2006] define PCA como sendo uma poderosa técnica estatística com aplicações em áreas como processamento e compressão de imagens e na localização de padrões em grandes sistemas de dados. O uso da análise de componentes principais não se mostra limitada ou restrita a aplicações de detecção de falhas, sendo extremamente amplo. **MARQUES E MARQUES [2005]** aplicam PCA para analisar uma base de dados envolvendo um conjunto de variáveis explicativas da economia brasileira. É importante ressaltar aqui que a possibilidade de aplicação em diferentes áreas não é exclusiva da técnica PCA, é na verdade uma característica das ferramentas estatísticas em geral, como observado para modelos auto-regressivos.

Quando aplicadas a processos industriais, ferramentas baseadas em PCA capturam a correlação existente entre os dados gerados pelos vários sensores disponíveis no processo. Graças à correlação natural existente entre o sinal gerado pelos sensores, e o fato que tal correlação ser alterada na presença de anomalias operacionais, esse tipo de análise permite que a técnica seja aplicada com sucesso em ferramentas de detecção de falhas.

Para **VENKATASUBRAMANNIAN *et al.* [2003],** a principal limitação do monitoramento baseado em PCA é que este modelo é invariante no tempo, enquanto processos reais tendem a apresentar alguma variação. Desse modo,

torna-se necessário implementar rotinas de re-validação do modelo após determinado intervalo temporal. A mesma limitação é observada quando modelos ARX e ARMAX são empregados, sendo necessária a adoção de janelas temporais de re-treinamento (ou re-sintonia) periodicamente. Tal limitação está presente também em metodologias baseadas em equações fenomenológicas, cujos parâmetros precisam ser reajustados para continuar a descrever o processo com precisão. O ajuste dos parâmetros de um modelo fenomenológico complexo pode ser uma tarefa tão ou mais complicada que a re-sintonia de um modelo PCA, ARX ou ARMAX.

KANO *et al.* [2000] descrevem uma abordagem para detectar a ocorrência de situações anormais de operação através do uso de modelos PCA. Esta aproximação inicia-se pelo ajuste de um modelo PCA aos dados. Ou seja, obtém-se um modelo PCA que represente corretamente os pontos experimentais durante uma condição de operação normal do sistema. Com o modelo validado para região da série temporal onde não existem anormalidades, inicia-se a fase de vigília, na qual a ferramenta detectora irá buscar a ocorrência de uma falha operacional ou anormalidade. A separação da janela temporal em fases, sendo o modelo ajustado em uma delas e validado ou testado na seguinte, é comum a diversas aplicações, sendo inclusive empregada em diversos trechos do presente estudo.

Na abordagem de **KANO *et al.* [2000],** a anormalidade será detectada quando os dados de processo não mais apresentarem aderência ao modelo previamente encontrado. Ou seja, durante a fase de vigília acredita-se que o modelo represente o comportamento correto e normal de operação do processo. Qualquer desvio significativo observado terá sido ocasionado por um comportamento anormal na planta, tratando-se, assim, de uma falha operacional. Desse modo, o método proposto por **KANO *et al.*, [2000]** se baseia em localizar instantes da série temporal onde os valores preditos pelo modelo e os dados do processo divirjam significativamente. **KANO *et al.*, [2000]** usam dados obtidos a partir do problema clássico de Tennessee Eastman para a avaliação da ferramenta de detecção. Ressalta-se que esse princípio é análogo ao empregado por **VAZ JUNIOR *et al.* [2009]**, quando do desenvolvimento de ferramenta detectora de vazamentos.

Graças às características dinâmicas dos processos químicos, **KANO *et al.*, [2000]** destacam a importância de determinar o tamanho da janela de vigília empregada até que um novo modelo PCA precise ser ajustado. Em um cenário de mudança das condições operacionais com o tempo, o sucesso de qualquer método baseado em equações que, de uma forma ou de outra descrevem o processo, está fortemente relacionado ao frequente re-treinamento do modelo.

Apesar de não implementado, **KANO *et al.*, [2000]** mencionam ainda que a metodologia de detecção apresentada poderia ser extrapolada para também identificar a origem da falha. **KANO *et al.*, [2000]** propõem a formação de uma base de dados composta por modelos gerados a partir de séries temporais representativas de diferentes falhas possíveis. Desse modo, os dados de uma situação de falha a ser identificada seriam comparados com cada um dos conjuntos representativos de falha. Através da similaridade entre os dados do processo e cada modelo de anomalia, seria possível identificar a origem da perturbação.

De modo semelhante a **KANO *et al.*, [2000],** o algoritmo de detecção de falhas implementado por **DUNIA *et al* [1996a]** também se baseia no erro de predição gerado por um modelo PCA previamente sintonizado. Do mesmo modo que **KANO *et al.*, [2000]**, a ocorrência de falha é detectada quando o erro rompe o intervalo de confiança pré-estabelecido. O principal destaque na abordagem proposta por **DUNIA *et al*. [1996a]** e **DUNIA *et al*. [1996b]** é a capacidade de, além de detectar, executar também a rápida identificação do sensor em falha. Em um processo industrial complexo, composto por dezenas de sensores, tão relevante quanto detectar uma anomalia é localizá-la. Um software capaz de identificar com precisão a origem do problema permite reduzir significativamente o tempo de manutenção da unidade, colaborando para o imediato retorno à operação normal. A metodologia proposta pelos autores reconstrói o sinal gerado por cada sensor baseando-se nos demais, usando para isso a correlação existente entre os sinais de dados. A partir da reconstrução dos dados e dos resíduos gerados nessa etapa, obtém-se o chamado “índice de validade do sensor” ou SVI. Este índice mede a participação de cada sensor no erro total de predição do modelo PCA. Assim, o índice indica o grau de confiabilidade no correto funcionamento de cada instrumento. Sensores com reduzida confiabilidade são os principais candidatos a estarem em falha. O índice de

validade do sensor foi usado neste estudo, e será descrito detalhadamente quando forem apresentadas as ferramentas empregadas.

A metodologia desenvolvida por **DUNIA *et al*. [1996a]** para identificação da origem da falha parte da premissa que apenas um único sensor pode falhar a cada momento. Premissa semelhante foi adotada por **VAZ JUNIOR *et al.* [2009]**, que consideram que ocorrências de múltiplas falhas simultâneas são consideradas extremamente raras.

Os trabalhos desenvolvidos por **KANO *et al.*, [2000]**, e **DUNIA *et al*. [1996a]** baseiam-se em princípios muito semelhantes, fazendo uso de métricas estatísticas tradicionais como T^2, Q e SPE na comparação entre os dados gerados pelo processo e o modelo PCA previamente ajustado. **KRZANOWSKI [1979]** e **SINGHAL e SEBORG [2002]**, por outro lado, empregam abordagens baseadas em princípios de similaridade entre modelos PCA no desenvolvimento de suas ferramentas detectoras de anomalias.

A abordagem proposta por **KRZANOWSKI [1979]** mede a similaridade entre dois conjuntos de dados através da comparação *dos modelos PCA* gerados para cada conjunto. Ou seja, não se comparam os dados diretamente, mas sim os modelos PCA obtidos a partir deles. Para quantificar a similaridade entre os modelos foi desenvolvido um índice denominado “Similarity PCA” ou “Spca”. Considerando dois conjuntos de dados que contém o mesmo grupo de n-variáveis, assume-se que o modelo PCA que descreve cada conjunto de dados é composto por k componentes principais, onde k é menor ou igual a n. A similaridade entre os dois grupos de dados é então quantificada através da comparação de seus componentes principais, via Spca. A métrica Spca compara os subespaços gerados através do cálculo do ângulo formado entre o “i-ésimo” componente principal do primeiro conjunto de dados e o “j-ésimo” componente principal do segundo conjunto.

Essa aproximação também emprega a separação da base de dados em duas janelas temporais, ou fases. Ajusta-se um modelo PCA para a fase de treino, na qual se garante a não ocorrência de anomalias. Após, um segundo modelo PCA é ajustado para a fase de vigília, na qual busca-se detectar a ocorrência de uma

anomalia. Caso o comportamento do processo seja idêntico nas duas fases, ou seja, o processo segue em operação normal também durante a fase de vigília, essa semelhança será refletida nos modelos ajustados. Nesse cenário, elevada similaridade entre os modelos será encontrada.

Por outro lado, caso o processo seja afetado pela presença de uma anormalidade ocorrida durante a fase de vigília, esta alteração será refletida no modelo PCA. Nessa situação, baixa similaridade entre os modelos será observada, ou seja os modelos serão dissimilares. Reduzida similaridade entre os modelos está portanto associada a presença de falhas durante o período de vigília.

De modo similar a **KRZANOWSKI [1979]**, **SINGHAL e SEBORG [2002]** propõem abordagem também baseada na similaridade entre modelos PCA. A métrica de similaridade adotada porém é distinta, sendo denominada “Distance Similarity Factor” ou “Sdist”. Enquanto o fator de similaridade Spca de **KRZANOWSKI [1979]** é influenciado pela orientação espacial do subespaço gerado pelos componentes principais, o fator “Sdist” de **SINGHAL e SEBORG [2002]** é usado para situações onde os conjuntos de dados tem orientação espacial similar mas estão localizados em posições fisicamente afastadas. **SINGHAL e SEBORG [2002]** define o fator Sdist como sendo a probabilidade que o centro do primeiro conjunto de dados esteja ao menos a uma distância predeterminada do centro dos dados do segundo conjunto.

O fator Sdist provê complemento natural às propriedades apresentadas por Spca, sendo assim natural o emprego de ambos de modo combinado, conforme introduzido por **SINGHAL e SEBORG [2002]**. A junção dos dois índices permite comparar os conjuntos de dados tanto em relação à orientação de seus subespaços quanto a distância entre eles. Ao combinar as duas métricas define-se uma nova, denominada “Similarity Factor”, ou “SF”. SF é definido como sendo a soma ponderada de ambos os índices, Sdist e Spca. A especificação do peso atribuído a cada fator é função do problema estudado. Nos testes realizados por **SINGHAL e SEBORG [2002]** a métrica Spca mostrou maior capacidade de capturar similaridade entre os conjuntos de dados do que o fator Sdist. Por esse motivo, maior peso foi

atribuído aquele fator. As métricas Spca, Sdist e SF foram amplamente testadas e avaliadas neste estudo.

Além de empregar Sdist, Spca e SF, **SINGHAL e SEBORG [2002]** incorporam uma metodologia bastante interessante para o monitoramento de processos: o uso de “janelas móveis”. As “janelas móveis” são usadas na busca por períodos de operação similares a um dado “período-base” previamente definido. O “período-base” pode tratar-se da resposta característica do processo a uma dada falha. Ao localizar trecho de dados no histórico do processo com padrão de comportamento similar ao do “período-base”, este trecho passa a ser considerado candidato a representar a ocorrência daquela falha. Essa metodologia permite não apenas detectar a ocorrência de uma falha como também identificá-la. Abordagem semelhante foi proposta, porém não implementada por **KANO [2000].**

A abordagem por janela móvel permite selecionar uma série de trechos da base de dados experimentais candidatos a estarem com falhas operacionais. Uma vez composta a lista de trechos, **SINGHAL e SEBORG [2002]** propõem que estes sejam então avaliados por um especialista. A avaliação humana irá confirmar ou não a identificação da falha em cada trecho candidato. A necessidade de contar com uma avaliação humana para confirmar ou não a ocorrência de falha é uma limitação deste método.

SINGHAL e SEBORG [2002] sugerem ainda duas métricas para quantificar a eficiência do método de detecção, denominadas “p” e “n”. A métrica “p” representa a razão entre o número de candidatos selecionados que realmente continham falha pelo número total de candidatos selecionados. Ou seja, caso método selecione 10 trechos como sendo similares ao comportamento do período-base, mas após verificação humana nota-se que apenas 3 destes trechos estavam corretamente identificado, tem-se que a métrica “p” obtida foi de 30%.

Por outro lado, a métrica “n” relaciona o número de candidatos selecionados que realmente continham falha, pelo número total de trechos de falhas que deveriam ter sido identificados. É importante ressaltar que este valor somente está disponível

durante o processo de teste da metodologia, quando as falhas foram propositalmente inseridas na base de dados.

SINGHAL e SEBORG [2002] observam que diferentes ajustes dos métodos de detecção podem privilegiar a métrica “n” ou “p”, dependendo da finalidade na qual ele é aplicado. Em processos nos quais nenhuma falha pode deixar de ser capturada, o método é ajustado para tornar-se bastante rigoroso na análise. A conseqüência negativa dessa regulagem é a geração de um numero elevado de candidatos que não apresentam falhas, elevando a carga de trabalho dos especialistas. Por outro lado, é possível configurar o método de modo a apresentar um número bastante reduzido de trechos candidatos, reduzindo o trabalho de análise posterior. O problema, neste caso, é deixar de detectar as anomalias de menor intensidade.

Apesar do sucesso obtido através do emprego de abordagens baseadas em PCA na detecção e identificação de falhas, é importante ressaltar a necessidade de implementar etapas preliminares de tratamento de dados com a finalidade de reduzir a ocorrência de falsos alarmes. Alarmes falsos ocorrem principalmente na presença de elevados ruídos nas medidas dos sensores. É importante lembrar porém que o uso de filtros também reduz a capacidade da ferramenta para identificar falhas de menor severidade. Ou seja, ao usar um filtro existe sempre a possibilidade de “filtrar” as falhas de menor severidade juntamente com o ruído.

Os parâmetros de avaliação propostos por **SINGHAL e SEBORG [2002]**, concentram-se na comparação entre os desempenhos na tarefa de classificar, não avaliando as exigências computacionais de cada metodologia. Como conseqüência, muitos classificadores que apresentam bom desempenho teórico não podem ser utilizados em aplicações reais porque sua exigência computacional é demasiadamente elevada. Nessa linha de raciocínio, **RIDDER et al., [2002]** buscam quantificar a complexidade computacional de vários classificadores estatísticos e mostrar a relação erro/complexidade apresentada.

VAZ JUNIOR *et al.* [2007, 2008] fazem uma revisão dos métodos apresentados **por KANO *et al.*, [2000]**, **DUNIA *et al* [1996a]**, **DUNIA *et al*. [1996b]**,

KRZANOWSKI [1979] e **SINGHAL e SEBORG [2002]**, adaptando-os ao contexto da detecção de falhas em redes de escoamentos. A comparação entre os métodos e a proposição de uma ferramenta integrada são as grandes contribuições destes trabalhos.

Outra metodologia de detecção de falhas baseada unicamente na análise de dados foi adotada por **DASH *et al.*, [2003]**. A ferramenta opera a partir da observação das séries temporais geradas pelos sensores, quanto a padrões de comportamento previamente definidos. Estes padrões são denominados “assinaturas”, e compõem o que os autores denominam de “alfabeto”. O objetivo é que, através de um alfabeto simples composto por poucas assinaturas, seja possível reconhecer qualquer estado de operação em que o processo se encontre. A identificação do estado de operação é realizada através do emprego de ferramentas de reconhecimento de padrões. A inspiração para tal metodologia encontra-se nos softwares de reconhecimento ótico de caracteres (Optical Character Recognition - OCR). O desenvolvimento de softwares de OCR começou em meados do século passado, muito antes da popularização dos sensores eletrônicos em plantas químicas. Essa mesma teoria vem sendo empregada atualmente no reconhecimento de placas de automóveis e de voz nos sistemas de auto-atendimento.

Os elementos fundamentais do alfabeto proposto são denominados “primitivos” ou “assinaturas”. Para **DASH *et al.*, [2003]**, uma idéia chave para o sucesso do método é manter os elementos fundamentais de descrição em sua quantidade mínima. Isto é, manter a linguagem em seu modo mais sucinto possível, permitindo todavia a descrição de qualquer situação, por mais complexa que seja. Os autores definem então sete primitivos: A(0,0), B(+,+), C(+,0), D(+,-), E(-,+), F(-,0), G(-,-), onde os sinais representam a derivada primeira e segunda, respectivamente.

Em função da escolha do “alfabeto” feita por **DASH *et al.*, [2003]**, a ferramenta proposta tornou-se altamente sensível a modificações na tendência do comportamento do processo. Os autores justificam tal escolha baseando-se no fato que a ocorrência de inúmeras falhas tem como conseqüência a geração de desvio de tendência. A natureza qualitativa do método de reconhecimento de padrões empregado por **DASH *et al.*, [2003]**, aliada a presença de ruídos nos sinais

experimentais, tornam imprecisas as classificações. Na visão dos autores, a busca por padrões altamente reprodutivos seria infrutífera, do mesmo modo que a busca pelo “caractere perfeito” em um software de OCR também resultaria em fracasso. Uma mesma falha pode apresentar comportamento ligeiramente diferente em ocasiões distintas devido a fatores altamente imprevisíveis, especialmente em complexas instalações industriais. Desse modo, ou se lida com a imprecisão da classificação ou busca-se gerar um alfabeto suficientemente extenso que cubra todas as eventualidades e suas combinações possíveis. Um alfabeto assim raramente é fisicamente realizável, e mesmo que o seja, seria complexo e extenso demais para ser gerenciado.

Para sobrepor essa limitação os autores propõem uma abordagem por lógica *fuzzy*, incorporando robustez a ferramenta classificadora e tornando possível lidar com as incertezas inerentes ao método. A lógica *fuzzy* permite lidar com as imprecisões de classificação e ao mesmo tempo manter a linguagem simples, com um número reduzido de assinaturas. De modo geral a lógica *fuzzy* age definindo um ajuste de classes com fronteiras imprecisas entre elas. Deste modo, geram-se as condições básicas para uma aproximação qualitativa da análise de sistemas complexos, nos quais variáveis lingüísticas são usadas na descrição de seu comportamento. Um dos atrativos da lógica *fuzzy* é a habilidade em converter dados numéricos em variáveis linguísticas. Um modelo *fuzzy* relaciona entradas e saídas usando uma combinação de regras e operadores lógicos. Regras de lógica são usadas para relacionar os dados temporais com estados de operação do processo em uma arquitetura do tipo “muitos para um”. Ou seja, informações advindas de vários sensores são usadas para caracterizar uma falha específica.

Outra limitação da abordagem proposta por **DASH *et al.*, [2003]** encontra-se em uma das premissas básicas assumidas: um mesmo tipo de falha deve apresentar sempre um mesmo tipo de resposta na forma de tendência, independente de sua severidade. A severidade da falha não deve assim distorcer, mas apenas ampliar, a resposta gerada na forma de tendência. Sob essa hipótese assume-se que cada tipo de falha seja representada por uma única assinatura na base de dados. Caso a falha seja simulada sobre três condições de severidade (fraca, média e forte), o armazenamento apenas da assinatura gerada por um dos experimentos é suficiente.

Apesar de fundamental, essa hipótese não é totalmente restritiva, já que seria possível incorporar na base de dados diferentes assinaturas para diferentes severidades de falha.

Operacionalmente a metodologia proposta por **DASH *et al.*, [2003]** divide-se em duas fases. Na primeira, denominada “identificação”, ocorre a comparação entre o comportamento atual do processo e as diversas assinaturas de falhas previamente estabelecidas. Em uma segunda fase, chamada “estimação”, a magnitude da falha detectada é então estimada.

Os autores utilizam um modelo de reator CSTR com controle de nível, temperatura e pressão, para testar o sistema de detecção de falhas. As falhas simuladas são em diferentes condições de operação e com severidades variando entre 5% e 40%. São compostos um total de 24 cenários de falhas. O método mostrou-se robusto diante da maior parte dos cenários testados, não sendo contudo eficiente para identificar falhas em situações de múltiplas ocorrências simultâneas.

Apesar de detectar corretamente a ocorrência da falha e estimar sua severidade, a metodologia proposta não se mostrou capaz de definir a sua origem. Ou seja, identificar dentre os sensores instalados qual encontra-se em falha. Essa dificuldade pode ser facilmente entendida, pois em diversas situações os sensores respondem de modo semelhante a diferentes falhas. Isto deixa as assinaturas de duas falhas demasiadamente parecidas, dificultando sua boa resolução. Uma vez que a técnica aplicada não utiliza informações fenomenológicas acerca do processo, seus resultados mostram-se altamente relacionados à qualidade da base de assinaturas disponível e ao alfabeto empregado.

Por melhor que sejam os resultados encontrados através de cada uma das metodologias descritas na acima, nenhuma delas terá eficiência absoluta em todos os cenários aos quais uma planta química está sujeita. Essa é a conclusão de **VENKATASUBRAMANNIAN *et al.*, [2003]**. Os autores promovem uma ampla revisão dos diferentes métodos de detecção de falhas baseados em séries temporais, comparando e avaliando as várias metodologias atualmente disponíveis a partir de uma série de características consideradas relevantes. Concluem que

nenhum método consegue, sozinho, satisfazer todas as exigências impostas a um sistema de diagnóstico de falhas.

Para **VENKATASUBRAMANNIAN *et al*., [2003]**, a utilização conjunta de duas ou mais metodologias pode tornar-se vantajosa na medida que une técnicas com características complementares. Essa abordagem, conhecida como “sistemas híbridos”, supera o desempenho alcançado através do uso de uma única técnica isoladamente. Torna-se assim recomendável considerar os diferentes métodos descritos na literatura como “mutuamente complementares”, ao invés de “mutuamente excludentes”.

Na mesma linha de raciocínio, **AMAND *et al.* [2001]** propõem o uso combinado de duas metodologias de modo a elevar a eficiência final obtida. Os autores empregam reconciliação de dados (“Data reconciliation”) e PCA no monitoramento de plantas químicas e na detecção de falhas. O objetivo é examinar como as duas técnicas podem ser empregadas em conjunto na tarefa de monitoração de processos.

A idéia por trás da reconciliação de dados é formular um modelo do processo baseado em diferentes restrições **[AMAND *et al.* 2001].** As restrições são deduzidas a partir de equações fenomenológicas de balanço de massa, energia, etc. Sob condições normais de operação o processo situa-se dentro das fronteiras impostas pelas restrições. Essas fronteiras são rompidas quando existe a presença de uma anormalidade. Durante a ocorrência de um vazamento, por exemplo, as restrições impostas pelas equações de balanço de massa não mais são obedecidas. Segundo os autores, o uso de reconciliação de dados de modo isolado é eficiente apenas na detecção de falhas grosseiras no processo. Porém, seu uso acoplado ao PCA permite elevar a eficiência da ferramenta detectora.

Além do uso em ferramentas de detecção de falhas em sensores ou vazamentos, sistemas de classificação também podem ser empregados para verificar o correto funcionamento de controles preditivos, sendo a operação dos controladores classificada como “normal” ou “anormal”. O método pode ainda indicar se uma perturbação não usual está presente na planta ou se suas condições

operacionais mudaram desde a sintonia dos controladores **[LOQUASTO e SEBORG, 2003]. LOQUASTO e SEBORG [2003]** propõem metodologias que permitam responder precisamente se o sistema de controle está funcionando corretamente e, caso negativo, se o problema encontra-se no modelo do processo ou numa anomalia operacional. Aferir o funcionamento do controle preditivo e verificar a causa da perda de eficiência evita que um complexo e custoso procedimento de re-sintonia seja deflagrado desnecessariamente. Os índices de similaridade Sdist, Spca e SF são empregados.

Na mesma linha de desenvolvimento, **CONNER e SEBORG [2005]** buscam diagnosticar quando o procedimento de re-identificação do modelo do processo torna-se obrigatório, de modo a obter uma sintonia mais adequada do controlador. Modelos PCA são usados nesse diagnóstico.

Ao longo do tempo uma planta química pode sofrer modificações que lhe afastam do comportamento previsto pelo modelo. Esse afastamento deve-se a inúmeros fatores, tais como, fadiga e desgaste operacional, modificação de condições operacionais, mudança de especificação de correntes de entrada, entre outros. Essa diferença de comportamento do processo pode tornar a sintonia original dos controladores pouco eficiente, prejudicando o desempenho e rendimento da planta.

Porém, a simples detecção de divergência entre o estimado pelo modelo e o comportamento observado não é sempre justificativa para uma completa re-identificação do modelo do processo. A ocorrência de pequenos desvios temporários ou a presença de ruídos é comum, mesmo após o desenvolvimento de um modelo bastante preciso, sendo muitas vezes causados por imprecisões da instrumentação ou fatores externos diversos.

Apesar da re-identificação poder sempre ser empregada, na prática é necessário que se tenha convicção de que tal procedimento é mesmo necessário. Ou seja, garantir antes de começar a re-sintonia que o modelo que virá a ser encontrado será estatisticamente diferente do então existente e irá prover incremento nas suas aplicações, tal como a sintonia de malhas de controle. Isso

porque o procedimento completo de identificação de um modelo que descreve o processo é, em geral, altamente custoso. Deste modo, a decisão por efetuar ou não a re-identificação completa é uma decisão importante. O objetivo de **CONNER e SEBORG [2005]** nesse trabalho é sistematizar uma metodologia quantitativa prática que permita decidir o momento em que uma re-identificação se faz necessária.

3. Metodologias SPE, Q, Spca, Sdist, Classificadores Hierárquicos, Protótipos e SVI

Este capítulo e o seguinte abordam as metodologias matemáticas empregadas ao longo do trabalho. Tais metodologias foram utilizadas para a implementação das diferentes abordagens de detecção, identificação e quantificação de anomalias. Os métodos descritos são aplicados de modo a permitir não apenas detectar a presença de uma falha de sensor ou vazamento na rede de dutos, como também identificar o tipo, origem e gravidade dessas anomalias. O desempenho obtido a partir da aplicação de cada metodologia é detalhado nos capítulos referentes aos resultados e conclusões.

A Figura 3.1 apresenta um panorama geral das técnicas empregadas. Com o intuito de detectar anomalias, utilizam-se as métricas ARX, SPE, Q, Spca, Sdist e classificadores hierárquicos. As metodologias SPE, Q, Spca e Sdist, baseadas em análise por componentes principais (PCA), e os classificadores hierárquicos, serão detalhados neste capítulo. A metodologia ARX é descrita no Capítulo seguinte.

Figura 3.1: Visão geral das metodologias empregadas

Para tornar possível a identificação da origem da falha, foram empregadas três metodologias distintas: SVI-D, método dos protótipos e modelos auto-regressivos. Os modelos ARX e os classificadores hierárquicos foram empregados na quantificação da severidade da falha ocorrida, seja estimando o grau de desvio do sinal do sensor, seja estimando o diâmetro de furo que gerou vazamento. A severidade neste caso não leva em consideração o tipo de fluido transportado, ignorando sua toxicidade e/ou inflamabilidade, por exemplo. Também não se considera as características dos trechos cortados pela dutovia, se existe população ou fontes de água potável. Para a diferenciação do tipo de anomalia ocorrida, ou "falha de sensor" ou "vazamento", foram testadas três aproximações: método dos protótipos, SVI-D e uma combinação de Sdist e Spca.

3.1 Metodologia de cálculo da análise por componentes principais (PCA)

A técnica PCA compõe o núcleo das mais diferentes metodologias aqui propostas, sendo desse modo fundamental para o entendimento destas. O método PCA pode ser entendido através do procedimento descrito por **DUNIA *et al* [1996a]**, apresentado a seguir. Inicialmente, dispõem-se os dados da série temporal na forma de matriz coluna. Ou seja, cada coluna representa os dados gerados por um sensor específico, e cada linha é uma observação distribuída ao longo do tempo. Após, os dados são normalizados, removendo-se a média e dividindo pelo desvio padrão. Prossegue-se então calculando-se os autovetores e autovalores da matriz de covariância dos dados ajustados.

A Figura 3.2 ilustra a disposição espacial dos autovetores em relação aos dados experimentais para uma situação particular envolvendo apenas duas variáveis, ou seja, ambiente R^2. Cada variável é representada em um dos eixos da figura. O autovetor com maior autovalor associado está representado pela linha vermelha, enquanto o segundo autovetor é representado pela linha azul.

Sabe-se que os autovetores, especialmente aqueles de maior autovalor associados, provêem importantes informações sobre o padrão de distribuição dos dados **[HART e DUDA, 2000]**. Qualitativamente, é possível observar que, enquanto

o autovetor representado em vermelho ajusta-se fortemente a distribuição dos dados, o segundo autovetor representa um padrão de distribuição que, apesar de existente, tem importância reduzida na descrição do comportamento dos pontos amostrais.

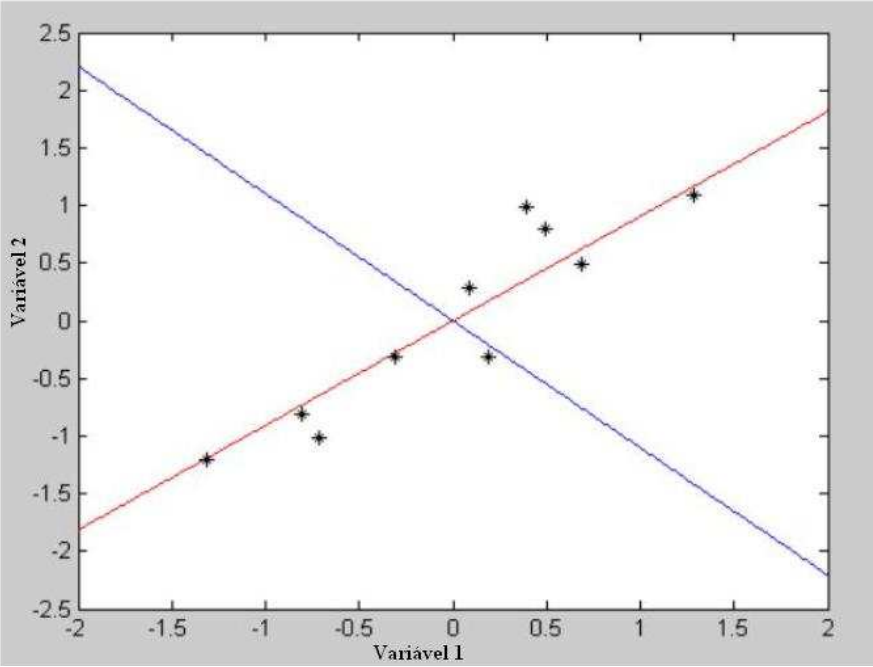


Figura 3.2: Ilustração da técnica de PCA em R^2

A quantificação dessa observação é realizada através dos autovalores, onde o autovetor que representa o padrão mais forte está relacionado ao autovalor mais elevado. Desse modo, ordenando os autovalores calculados é possível deduzir a significância do padrão apontado por cada autovetor. Denominando os autovetores de “componentes”, aqueles associados aos autovalores mais elevados são então chamados “componentes principais”. A análise via autovetores torna-se especialmente útil para os casos envolvendo R^N, onde é impossível qualquer observação gráfica do comportamento global dos dados.

Uma vez os autovetores estando ordenados através dos respectivos autovalores associados, procede-se a seleção dos n-componentes principais e a montagem da matriz de componentes, denominada “MC”. A seleção dos n-componentes principais tem como objetivo ignorar um ou mais autovetores pouco

significativos para a representação do comportamento do conjunto total de dados. Consegue-se reduzir a dimensão do problema sem causar perda de informação relevante.

É importante ressaltar que existem inúmeros procedimentos para definir o número de autovetores a serem selecionados para compor a matriz “MC”. Neste trabalho adotou-se a metodologia descrita por **SINGHAL e SEBORG [2002]**, onde correlaciona-se o valor do autovalor associado ao autovetor, a sua respectiva “taxa de explicação”. Define-se que o somatório total dos autovalores representa 100% da explicação sobre o comportamento dos dados. Desse modo, selecionam-se autovetores até que a taxa de explicação acumulada atinja o nível de explicação desejado. Por exemplo, pode-se optar por selecionar quantidade suficiente de autovetores de modo a representar 99% do comportamento do processo **[DUNIA et al. 1996a]**. Ressalta-se que não existe um critério único para selecionar o número de componentes principais a serem utilizados. Caso utilize-se um número muito pequeno de componentes, o resíduo gerado fica demasiadamente elevado, dificultando a identificação de pequenas falhas. Por outro lado, usar autovetores demais torna o processo lento, perdendo uma das mais relevantes vantagens do PCA, que é a redução na dimensão do problema **[DUNIA et al. 1996a].**

Os dados originais são então representados em um novo espaço vetorial, cuja base é composta pelos componentes principais selecionados. Obtém-se assim um novo conjunto de dados em um espaço vetorial de dimensão reduzida. Apesar da menor dimensão espacial, os dados guardam grande parte da informação relevante sobre o comportamento do conjunto original. O novo conjunto pode então ser comparado aos dados normalizados originais, de modo a determinar o erro do modelo PCA.

Na prática tornar-se mais fácil representar o modelo PCA através da matriz “C”, ao invés de se usar diretamente a matriz “MC”. A matriz “C” é obtida a partir da anterior, conforme apresentado na Equação 3.1. Tal matriz representa o modelo PCA, e pode ser aplicada diretamente aos dados experimentais ($\hat{x}\exp_k$) para se

obter os dados modelados ($\hat{x}\mathrm{mod}_k$), conforme Equação 3.2 **[DUNIA et al, 1996a]**.

$$C = MC * MC^T \qquad \text{eq. 3.1}$$

$$\hat{x}\mathrm{mod}_k = C\,\hat{x}\mathrm{exp}_k^T \qquad \text{eq. 3.2}$$

3.2 Quantificação da aderência do modelo aos dados via SPE e Q

Qualquer modelo matemático, por melhor ajustado que esteja, não reproduz perfeitamente os dados experimentais. A diferença entre o valor predito pelo modelo e o dado experimental é chamado de “erro”. Desse modo, o somatório quadrático dos erros pode ser usado para verificar a aderência do modelo aos dados. As métricas Q e SPE quantificam esse erro. Suas metodologias de cálculo são apresentadas abaixo.

3.2.1. Q

O cálculo de Q é executado a partir da Equação 3.3, e baseia-se no erro quadrático de predição, ou seja, na diferença entre o valor predito pelo modelo e o observado **[WISE e GALLAGHER, 1996]**:

$$Q = \hat{x}\mathrm{exp}_k (I - C)\hat{x}\mathrm{exp}_k^T \qquad \text{eq. 3.3}$$

onde a metodologia de calculo da matriz C foi apresentada anteriormente, $\hat{x}\mathrm{exp}_k$ é a matriz de dados normalizados, e “I” matriz identidade de dimensão apropriada.

3.2.2. SPE

O parâmetro SPE representa o erro quadrático da previsão ponderado com o tempo (*square prediction error*). Esta métrica possui “efeito memória”, ou seja, erros

passados têm maior ou menor influência sobre os valores mais recentes. O erro “e” é definido como sendo a diferença entre o valor experimental ($\hat{x}\exp_k$) e o obtido a partir do modelo ($\hat{x}\mathrm{mod}_k$) – Equação 3.4. O termo “esqr” é denominado erro quadrático (equação 3.5), enquanto a matriz “SPE” representa o valor final do SPE (equação 3.6) **[DUNIA et al, 1996a].**

$$e_k = (I - C)\hat{x}\exp_k \qquad \text{eq. 3.4}$$

$$esqr_k = e_k^2 \qquad \text{eq. 3.5}$$

$$SPE_k = (1-\lambda)SPE_{k-1} + \lambda esqr_k \qquad \text{eq. 3.6}$$

O efeito memória existente no cálculo de SPE é regulado através do parâmetro λ (lambda). Quanto mais próximo do valor unitário este parâmetro encontre-se, menor a influência do passado, sendo o SPE função principalmente dos valores mais recentes. Por outro lado, quanto mais próximo de zero o valor de λ, maior a influência do passado.

A detecção de anomalias ocorre quando o erro extrapola a faixa de operação normal. Para ambas as métricas, Q e SPE, o intervalo de confiança foi estabelecido como sendo 6 desvio-padrão calculados a partir de situações de operação normal da dutovia, ou seja, na ausência de anomalias. O uso de intervalos estreitos eleva o poder de detecção do método, pois o torna mais sensível a pequenas perdas de aderência entre o modelo e os dados experimentais, mas por outro lado incrementa o risco de falsos alarmes.

Para incorporar esta metodologia ao sistema de detecção de anomalias torna-se necessário que as séries temporais geradas pelo simulador de rede de escoamento sejam inicialmente separadas em duas fases, sendo a primeira denominada “Fase A” ou “fase de treino” e a seguinte chamada “Fase B” ou “fase de uso”, conforme Figura 3.3. Na primeira fase, ajustam-se os parâmetros do modelo PCA, promovendo a aderência deste aos dados experimentais. Torna-se assim fundamental garantir a ausência de qualquer anomalia operacional durante este período, ou seja, garantir que o treinamento seja efetuado em condições padrões de

operação. Após o término da fase de treino, inicia-se a “Fase B” ou “fase de uso”. Sobre os dados desta segunda fase aplica-se o mesmo modelo obtido na anterior, verificando-se então a sua aderência. Caso o modelo possua elevada aderência, isso indica com alto grau de certeza que o processo continua operando normalmente, pois o seu comportamento permanece similar ao observado na Fase “A”. Por outro lado, quando o modelo gerado na fase de treino não mais adere aos dados da fase de uso tem-se a indicação da ocorrência de uma anomalia durante a Fase “B”. O comportamento do processo agora diverge do observado durante a fase de treino. Os parâmetros SPE e Q são uma forma quantitativa de medir a aderência do modelo aos dados

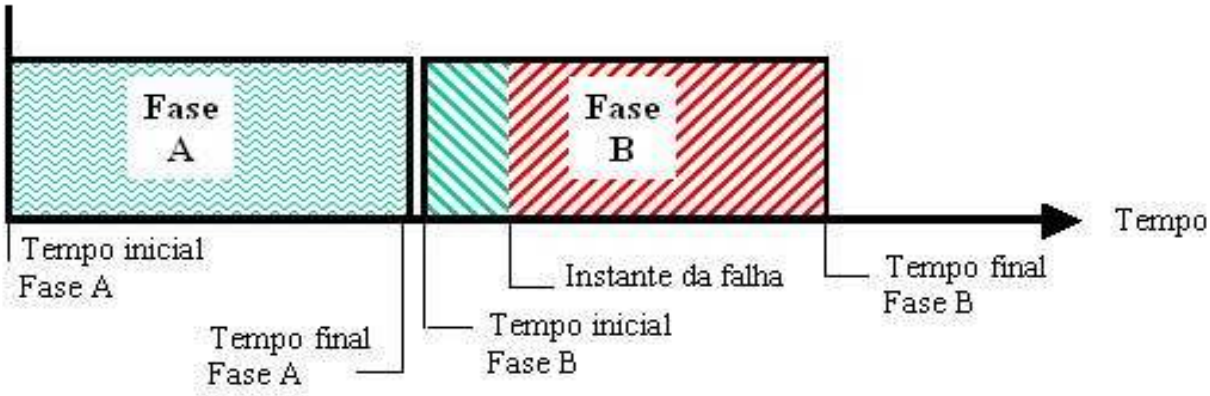


Figura 3.3: Representação esquemática do posicionamento temporal das fases A e B

Conforme mencionado, a Figura 3.3 representa a distribuição temporal das fases e do instante a partir do qual ocorre uma anormalidade. Na figura, a janela temporal sem falha é representada pelos reticulados verdes. O reticulado vermelho representa o trecho em que existe ocorrência de falha. Mais uma vez ressalta-se a importância da falha iniciar-se apenas durante a Fase “B”.

3.3 Quantificação da similaridade entre modelos PCA via Spca, Sdist e SF

Do mesmo modo que a quantificação da aderência do modelo via Q e SPE, a quantificação da similaridade entre dois modelos PCA via Spca, Sdist e SF também é baseada na comparação do comportamento do processo nas fases “A” e “B”

definidas anteriormente. Enquanto em Q e SPE existia apenas um modelo e o teste era feito através de sua aderência ou não aos dados da fase de uso, agora dois modelos PCA distintos são elaborados, um para cada fase. A detecção de anomalias se dá através da comparação, via índices de “simetria” ou “similaridade”, dos dois modelos obtidos. Uma reduzida similaridade entre os *modelos* das duas fases indica uma mudança no comportamento do processo, sendo indicação forte da ocorrência de falha na “fase B”.

Conforme mencionado anteriormente, os índices de similaridade Spca e Sdist serão usados ainda na tarefa de diferenciar o tipo de anomalia ocorrida. As três medidas de simetria empregadas neste trabalho são: Spca, Sdist e SF. Suas metodologias de cálculo estão apresentadas a seguir.

3.3.1 Spca

O método Spca (*similarity PCA*) quantifica a similaridade entre duas séries temporais através da comparação dos modelos PCA obtidos para cada conjunto de dados. Consideram-se dois conjuntos de dados que contém o mesmo número de variáveis (n). Assume-se que o modelo PCA que descreve cada conjunto é composto por “k” componentes principais, onde “k” é menor ou igual a “n”. A similaridade entre os dois grupos de dados é quantificada através da comparação de seus componentes principais. Mais precisamente, o Spca compara os subespaços gerados por cada modelo PCA através do cálculo do ângulo entre os componentes principais. O fator de similaridade é influenciado pela orientação espacial do subespaço gerado pelos componentes principais, sendo definido na Equação 3.7 **[SINGHAL e SEBORG, 2002]**.

$$Spca = \frac{1}{k}\sum_{i=1}^{k} \sum_{j=1}^{k} \cos^2 \theta_{ij} \qquad \text{eq. 3.7}$$

onde θ_{ij} é o ângulo formado entre o “i-ésimo” componente principal do primeiro conjunto de dados e o “j-ésimo” componente principal do segundo conjunto.

Na prática, tendo-se as matrizes de componentes MC_A e MC_B , compostas pelos “k” componentes principais que descrevem os conjuntos de dados das fases

“A” e “B” respectivamente, torna-se mais simples calcular Spca na sua forma matricial (Equação 3.8). Onde “trace” é a soma dos elementos da diagonal principal da matriz.

$$Spca = \frac{trace(MC_A^T MC_B MC_B^T MC_A)}{k} \qquad \text{eq. 3.8}$$

3.3.2 Sdist

Enquanto o fator de similaridade Spca é influenciado pela orientação espacial do subespaço gerado pelos componentes principais, o fator Sdist (ou “distance similarity factor”) é usado para situações onde os conjuntos de dados têm orientação espacial similar, mas estão localizados em posições distantes. O fator Sdist é a probabilidade que o centro do conjunto de dados A ($\bar{x}_A$) esteja ao menos a distância ϕ do centro dos dados B ($\bar{x}_B$). Calcula-se Sdist através da Equação 3.9 **[SINGHAL e SEBORG, 2002]**, a seguir**.**

$$Sdist = 2\left[1 - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\phi} e^{-\frac{z^2}{2}} dz\right] \qquad \text{eq. 3.9}$$

onde ϕ é a “distância de Mahalanobis”, calculada via equação 3.10:

$$\phi = \sqrt{(\bar{x}_A - \bar{x}_B)\sum_s^{*-1}(\bar{x}_A - \bar{x}_B)^T} \qquad \text{eq. 3.10}$$

Sendo $\sum$ a pseudo inversa da matriz de covariância dos dados S.

3.3.3 SF

Graças a sua definição, o fator Sdist provê complemento natural às propriedades apresentadas pelo Spca, sendo indicado seu uso combinado. Enquanto uma métrica é sensível a orientação espacial, outra reage mais fortemente ao afastamento espacial entre os modelos. A junção das duas métricas é denominada SF (*similarity factor*), e permite comparar os conjuntos de dados tanto em relação à orientação de seus subespaços quanto a distância entre eles. SF é

simplesmente a soma ponderada das duas métricas anteriores, conforme Equação 3.11 **[SINGHAL e SEBORG, 2002].**

$$SF = \alpha Spca + (1-\alpha)Sdist \qquad \text{eq. 3.11}$$

O parâmetro α representa a ponderação entre o fator Spca e Sdist.

A técnica de detecção de vazamentos ou falhas de sensores por similaridade, seja via Spca, Sdist ou SF, pode ser aplicada através de duas aproximações distintas: “estática” ou “dinâmica”. Estas estão detalhadas a seguir.

3.3.4 Abordagem estática

Do mesmo modo que aplicado na técnica de detecção via SPE e Q, a metodologia de detecção via similaridade baseia-se na comparação do comportamento da rede em duas janelas temporais distintas, denominadas “fase A” e “fase B”, ilustradas na Figura 3.3 junto com a disposição temporal das fases e de uma possível anomalia ocorrida.

Apesar de funcional, tal abordagem possui inúmeras limitações. A principal delas é a dificuldade de detectar anomalias ocorridas na parte final da fase de uso. Isto porque, quando a falha inicia-se muito próxima ao final da Fase “B”, sua influência sobre o modelo PCA desta fase torna-se muito pequena. O modelo irá refletir pouco a falha, sendo altamente similar ao encontrado para a Fase “A”, sem falhas. Logo, quanto mais próximo do final da Fase “B” a falha ocorrer, mais difícil será detectá-la através da abordagem estática. A abordagem dinâmica é uma tentativa de superar tal limitação.

3.3.5 Abordagem dinâmica (ou janela móvel)

Através desta abordagem a Fase “B” torna-se dinâmica, deslocando-se ao longo do eixo de tempo de modo a se afastar gradualmente do término da Fase “A”. A Figura 3.4 ilustra esse comportamento. Assim, supondo uma Fase “A” que se estenda do instante A até o instante B, ela será seguida por uma Fase “B” de mesma

extensão entre os instantes B+1 e C. Tornando a Fase “B” dinâmica, esta é deslocada de modo a afastar-se progressivamente da Fase “A”. Desse modo, a fase “B” seguinte estará localizada entre os instantes B+2 e C+1, supondo “passo unitário”. Mantendo a mesma lógica, as Fases “B” posteriores terão início nos instantes B+3, B+4, B+5, etc. A metodologia prossegue até a formação da última Fase “B”, que começará no instante D e se estenderá até o instante E.

Segundo **SINGHAL e SEBORG [2002],** um fator importante na qualidade e operacionalidade do método de janela móvel é a velocidade de deslocamento desta ao longo da série temporal. Uma velocidade de deslocamento lenta eleva o poder de identificação do método, mas também demanda maior esforço computacional. Assim, supõe-se que a Fase “B” move-se de um instante de observação para o imediatamente seguinte. Essa abordagem, apesar de apresentar a maior capacidade de identificação, muitas vezes torna-se inaplicável devido ao elevado número de comparações requeridas ao longo do estudo de uma extensa série temporal. Para reduzir a demanda computacional, a janela pode mover-se “w” observações a cada atualização. Ou seja, a cada atualização da posição da janela insere-se e remove-se “w” observações. O número “w” é denominado “taxa de movimento da janela” ou “passo”. **SINGHAL e SEBORG [2002]** sugerem adotar valor de “w” entre 1/10 e 1/5 da extensão da janela. Neste estudo o “passo” adotado foi o menor possível, ou seja, unitário. Tal ajuste privilegiou o desempenho de detecção da ferramenta, aceitando a maior carga computacional daí decorrente.

A Figura 3.4 representa ainda a ocorrência de uma falha no instante F. Observa-se que a medida que a Fase “B” desloca-se, a falha vai preenchendo uma parcela crescente da sua janela temporal. Esse fenômeno, não observado na abordagem estática, melhora o grau de detecção da ferramenta independente do instante de ocorrência da falha **[SINGHAL e SEBORG, 2002]**.

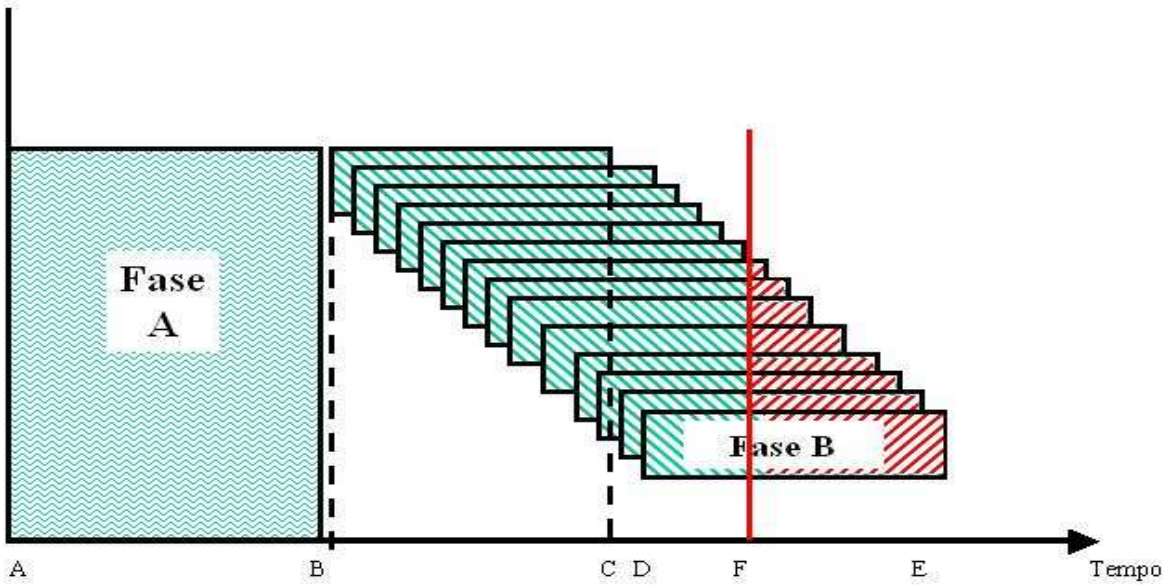


Figura 3.4: Representação da abordagem dinâmica

3.4 Classificação hierárquica

Os classificadores hierárquicos são usados nas mais diferentes áreas do conhecimento com o objetivo de separar e/ou classificar de forma útil observações contidas em um vasto conjunto de dados. A separação é promovida através da formação de agrupamentos, ou *clusters*, cujos membros apresentam elevada similaridade entre si. Objetos similares são alocados em um mesmo grupo, enquanto objetos "não similares" são posicionados em agrupamentos distintos **[VAZ JUNIOR *et al.*, 2009]**. No presente trabalho, aplica-se essa abordagem para separar, de um conjunto total de amostras temporais onde existe a ocorrência de falha, dois grandes conjuntos: "com" e "sem" presença de anomalia. Ao promover a classificação dos pontos amostrais torna-se possível definir com precisão o momento da ocorrência de falha operacional.

A similaridade entre dois pontos pode ser entendida como sendo inversamente proporcional à distância espacial entre esses pontos. Em um espaço vetorial R^2 tradicionalmente aplica-se a métrica Euclidiana para quantificar a distância entre o ponto "i" e o ponto "j" no plano XY. Esta métrica pode ser ampliada para um espaço vetorial genérico de dimensão "n" (R^n) através da Equação 3.12.

$$d_{ij} = \left[\sum_{k=1}^{n} \left(x_{i,k} - x_{jk} \right)^2 \right]^{1/2} \qquad \text{eq 3.12}$$

Apesar de amplamente utilizada, a distância Euclidiana não é a única definição possível para a distância entre dois pontos. As métricas conhecidas como City Block (Equação 3.13) e Minkowski (Equação 3.14), descritas por **DUDA E HART (2000),** muitas vezes apresentam maior eficácia na caracterização de similaridade entre amostras. Nota-se que as métricas City Block e Euclidiana são casos particulares de Minkowski assumindo “r” igual a 1 e 2 na Equação 3.14

$$d_{ij} = \sum_{k=1}^{n} \left| x_{ik} - x_{jk} \right| \qquad \text{eq 3.13}$$

$$d_{ij} = \left(\sum_{k=1}^{n} \left| x_{ik} - x_{jk} \right|^r \right)^{1/r} \qquad \text{eq 3.14}$$

Uma vez definida a métrica, o procedimento de classificação em si é bastante simples. Elabora-se uma matriz de distâncias entres todas as observações disponíveis e busca-se o par de pontos com menor afastamento entre si. Ambas as amostras são então unidas, dando origem a um novo ponto. Uma nova matriz de distâncias é então elaborada, selecionando-se novamente o par de menor afastamento. O procedimento segue até que a população esteja totalmente alocada em dois grandes agrupamentos. Idealmente as amostras pertencentes a um mesmo agrupamento devem apresentar elevada similaridade entre si, e forte dissimilaridade quando comparadas a amostras pertencentes ao outro *cluster*. Por fim, a identificação do *cluster* “com falha” e “sem falha” é promovida classificando uma observação de comportamento previamente conhecido. Além de permitir a detecção de falhas, a classificação hierárquica também possibilita a quantificação da severidade da anomalia ocorrida. Isso porque o afastamento entre os dois clusters formados guarda relação direta com a gravidade da falha.

3.5 Índice de validade dos sensores (SVI-D)

As plantas químicas modernas estão atualmente equipadas com grande quantidade de sensores. Enquanto parte desta instrumentação é empregada na

composição de malhas de controle, outra parcela é aplicada apenas no monitoramento e supervisão do processo. A disponibilidade de um vasto número de sensores promove então redundância entre seus valores, criando a possibilidade de detectar e identificar ocorrências de falhas em um sensor quando este perde a correlação esperada em relação aos demais. Tal correlação é um fenômeno natural, e deve-se ao fato do processo como um todo ser governado por princípios básicos da química e da física, tais como a conservação de massa e energia.

Aproveitando essa correlação natural entre os sinais gerados pela instrumentação, **DUNIA *et al* [1996a]** e **DUNIA *et al* [1996b]** propõem um índice de validade dos sensores (SVI). Este índice seria capaz de quantificar a confiança existente na precisão do sinal gerado por cada sensor. A metodologia proposta pelos autores sugere a reconstrução do sinal de cada sensor baseando-se em modelo PCA previamente ajustado. A comparação entre os dados reconstruídos e os experimentais gera um resíduo, a partir do qual encontra-se o SVI.

O SVI é útil nas etapas de identificação do tipo de anomalia e na sua localização. Na primeira, tem-se como objetivo identificar se a anomalia encontrada trata-se de um vazamento ou de uma falha de sensores. Vazamentos tendem a degradar o sinal de inúmeros sensores em sua vizinhança, enquanto uma falha de sensor afeta apenas o sinal do instrumento defeituoso. A segunda etapa visa estimar o posicionamento da falha na dutovia.

O procedimento de cálculo do índice é apresentado nas Equações 3.15 a 3.17 **[DUNIA *et al* 1996a].** A matriz C foi definida na equação 3.1, durante a apresentação do método PCA. O valor de SVID varia entre 0 e 1, sendo 1 indicativo de máxima confiança no correto funcionamento do sensor. Os sensores que apresentam menores SVID são os principais suspeitos de estarem apresentando falha. As ocorrências de vazamentos tendem a afetar o índice de validade de uma série de sensores em sua proximidade. Conhecendo os sensores cujo SVI-D foram mais afetados, torna-se possível estimar a localização do furo.

$$SVID = 1 - (1 - C_{ii})\frac{D_i}{A} \qquad \text{eq. 3.15}$$

onde:

$$D_i = \frac{(\hat{x}\,\mathrm{mod}_i - \hat{x}\,\mathrm{exp}_i)^2}{(1 - C_{ii})^2} \qquad \text{eq. 3.16}$$

$$A = \sum_{h=1}^{m} D_h (1 - C_{hh})^2 \qquad \text{eq. 3.17}$$

3.6 Método dos protótipos

Outra abordagem para a identificação e localização da anomalia é o método dos protótipos **[VAZ JUNIOR *et al.*, 2008]**. Para implementar tal abordagem são previamente selecionados todos os casos representativos das principais anomalias possíveis de ocorrer na rede em estudo. Ou seja, as anomalias do tipo “falha em sensor” – uma por sensor, e do tipo “vazamento” – uma por tubo. Esses casos representativos são então denominados “protótipos”. Apesar de tratarem de uma situação específica, espera-se que o protótipo seja adequado para identificar situações nas quais o instante temporal da falha ou a sua severidade sejam distintas. Tal hipótese é semelhante aquela adotada por **DASH *et al.*, [2003]** em seu trabalho, conforme detalhado anteriormente.

Uma vez criados os protótipos e uma anomalia operacional tendo sido detectada, as observações experimentais relativas a essa falha são então comparadas com cada um dos padrões, na busca pela maior similaridade. Através do protótipo de maior similaridade com os dados experimentais torna-se possível não apenas estimar o tipo de anomalia ocorrida como também a sua origem.

4. Detecção, Localização e Quantificação de Vazamentos via Modelos ARX

Os modelos ARX, também chamadas de “preditores” são totalmente baseados em séries temporais, utilizando esses dados de processo para as etapas de treinamento e operação **[VANDAELE, 1983]**. Pode-se dizer que essas duas fases compõem um “ciclo de predição”, conforme Figura 4.1. Ressalta-se a semelhança entre a figura a seguir e aquelas apresentadas no Capítulo 3.

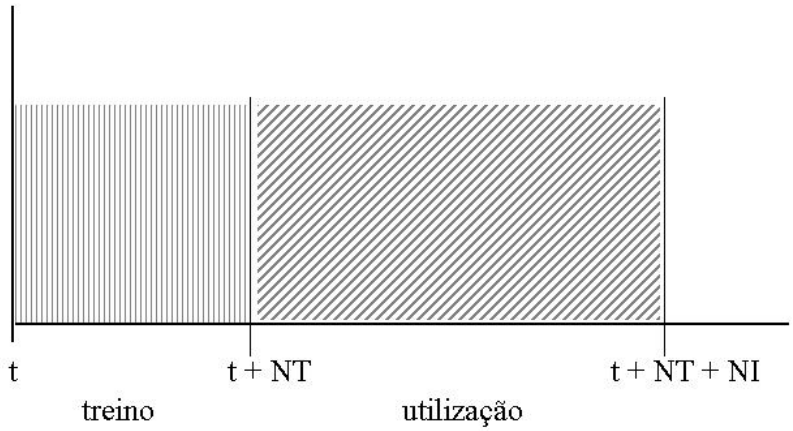


Figura 4.1. Ciclo de predição composto da fase de treino e fase de utilização

Os sensores disponíveis na dutovia são divididos em dois grandes grupos: entrada (SU) e saída (SY) (Figura 4.2).

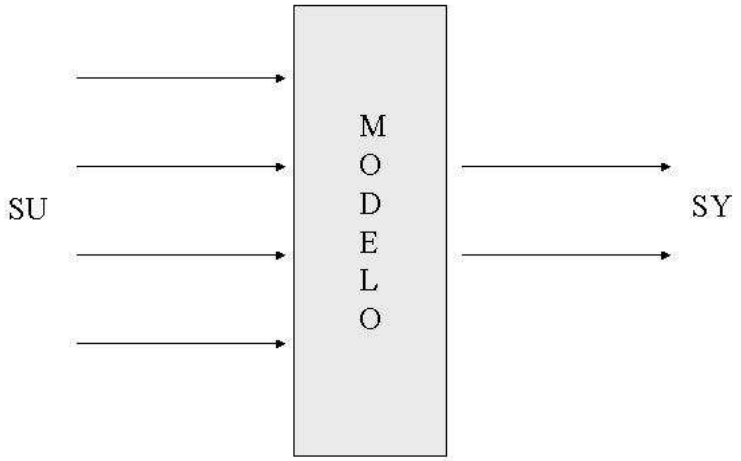


Figura 4.2. Variáveis de entrada e de saída do modelo

Os preditores ARX-MIMO têm a forma:

$$\underline{\hat{Y}}(t+1) = \underline{\underline{B}}(q)\,\underline{U}(t+1) + q\left[1 - \underline{\underline{A}}(q)\right]\underline{Y}(t) \qquad \text{eq 4.1}$$

onde $\underline{\hat{Y}}(t+1)$ é o vetor de respostas preditas para o instante *t+1*, *q* representa o operador de deslocamento para frente (*forward shift operator*), e *A(q)* e *B(q)* são filtros definidos por:

$$\underline{\underline{A}}(q) = 1 + \sum_{k=1}^{n} \underline{\underline{A}}_k q^{-k}(.) \qquad \text{eq. 4.2}$$

$$\underline{\underline{B}}(q) = \sum_{k=1}^{m} \underline{\underline{B}}_k q^{-k}(.) \qquad \text{eq. 4.3}$$

onde *n* e *m* são a ordens dos filtros (usualmente 3, 4 ou 5) **[VAZ JUNIOR, 2006]** , $\underline{\underline{A}}_k$ e $\underline{\underline{B}}_k$ são matrizes constantes com tamanhos (*ny* por *ny*) e (*ny* por *nu*), respectivamente. A etapa de identificação do modelo ARX **[VAZ JUNIOR, 2006]** consiste em estimar as matrizes $\underline{\underline{A}}_k$ e $\underline{\underline{B}}_k$ a partir dos valores gerados de $\underline{Y}$ e $\underline{U}$ durante a janela de treino de *NT* instantes de tempo.

O preditor é usado nos *NI* instantes de tempo seguintes (Figura 4.1) para monitorar as respostas do processo. Essas duas etapas são repetidas seguidamente, consistindo-se do “ciclo de predição”. A qualidade do treinamento e o grau de degradação do preditor ao longo de sua fase de utilização, podem ser medidas através da largura do intervalo de confiança de suas predições. Os principais parâmetros envolvidos na janela de treinamento são estudados abaixo. Os parâmetros ARX $(\underline{\hat{\theta}})$ são estimados via máxima verossimilhança, correspondendo a minimização da soma dos quadrados dos resíduos entre valor observado e valor predito, Equação (4.4). A solução dessa equação e predições na janela de treino são geradas pelas equações subseqüentes, Equações (4.5) a (4.10). A matriz $\underline{\underline{\Xi}}$

acumula um “conhecimento básico” sobre recentes relações entre entradas e saídas do processo **[VAZ JUNIOR *et al.*, 2010]**

$$\underset{\{\underline{\underline{\hat{\theta}}}\}}{Min} \quad (1/2)Trace((\underline{\underline{\hat{Y}}} - \underline{\underline{\Psi}})^T \underline{\underline{G}}(\underline{\underline{\hat{Y}}} - \underline{\underline{\Psi}})) \qquad \text{eq. 4.4}$$

$$\underline{\underline{U}} = [\underline{U}(t-(m-1)) \ldots \underline{U}(t) \ldots \underline{U}(t+NT-1)]^T \qquad \text{eq. 4.5}$$

$$\underline{\underline{Y}} = [\underline{Y}(t-(n-1)) \ldots \underline{Y}(t) \ldots \underline{Y}(t+NT-1)]^T \qquad \text{eq. 4.6}$$

$$\underline{\underline{\Psi}} = [\underline{Y}(t+1) \ldots \underline{Y}(t+NT)]^T \qquad \text{eq. 4.7}$$

$$\underline{\underline{\hat{Y}}} = [\underline{\hat{Y}}(t+1) \ldots \underline{\hat{Y}}(t+NT)]^T \qquad \text{eq. 4.8}$$

$$\underline{\underline{\Xi}} = [-\underline{\underline{S}}(1,n)\underline{\underline{Y}} \ \ldots \ -\underline{\underline{S}}(n,n)\underline{\underline{Y}} \ \ \underline{\underline{S}}(1,m)\underline{\underline{U}} \ \ldots \ \underline{\underline{S}}(m,m)\underline{\underline{U}}] \qquad \text{eq. 4.9}$$

$$\underline{\underline{\hat{\theta}}} = [\underline{\underline{A}}_1 \ \ldots \ \underline{\underline{A}}_n \ \underline{\underline{B}}_1 \ \ldots \ \underline{\underline{B}}_m]^T \quad , \quad \underline{\underline{\hat{Y}}} = \underline{\underline{\Xi}}\,\underline{\underline{\hat{\theta}}} \quad , \quad \underline{\underline{\hat{\theta}}} = [\underline{\underline{\Xi}}^T \underline{\underline{G}}\,\underline{\underline{\Xi}}]^{-1}[\underline{\underline{\Xi}}^T \underline{\underline{G}}\,\underline{\underline{\Psi}}] \qquad \text{eq. 4.10}$$

$$\underline{S}_R = diag((\underline{\underline{\hat{Y}}} - \underline{\underline{\Psi}})^T \underline{\underline{G}}(\underline{\underline{\hat{Y}}} - \underline{\underline{\Psi}}))/(NT - n*ny - m*nu) \qquad \text{eq. 4.11}$$

$$\underline{\underline{Y}}_U = [\underline{Y}_U(t+1) \ldots \underline{Y}_U(t+NT)]^T \quad , \quad \underline{\underline{Y}}_L = [\underline{Y}_L(t+1) \ldots \underline{Y}_L(t+NT)]^T \qquad \text{eq. 4.12}$$

$$\underline{\Lambda} = diag(\underline{\underline{\Xi}}[\underline{\underline{\Xi}}^T \underline{\underline{G}}\,\underline{\underline{\Xi}}]^{-1}\underline{\underline{\Xi}}^T) \qquad \text{eq. 4.13}$$

$$\underline{\underline{Y}}_U = \underline{\underline{\hat{Y}}} + T_{1-\alpha/2}\sqrt{\underline{\Lambda}\,\underline{S}_R^T} \qquad \text{eq. 4.14}$$

$$\underline{\underline{Y}}_L = \underline{\underline{\hat{Y}}} - T_{1-\alpha/2}\sqrt{\underline{\Lambda}\,\underline{S}_R^T} \qquad \text{eq. 4.15}$$

Nas equações acima, as seguintes observações se aplicam: (i) símbolo ^ refere-se a valores preditos ou estimados, (ii) matriz $\underline{\underline{S}}(k,p)$ é um seletor esparso para extração dos elementos do pós-multiplicador (com *NT+p-1* linhas) correspondente aos instantes de tempo *t+1-k* $\rightarrow$ *t+NT-k*, (iii) vetor $\underline{S}_R$ contém o somatório dos quadrados dos desvios entre valores preditos e observados *Y*

divididos pelos graus de liberdade, (iv) $\underline{\underline{G}}$ é uma matriz de escala, (v) o operador *diag* extrai a diagonal principal da matriz quadrada, (vi) $\pm T_{1-\alpha/2}$ são valores de *t-Student* (para *NT-n*ny-m*nu* graus de liberdade) definindo um intervalo com probabilidade de *(1-α)*100 %* com, por exemplo, *α=0.01*, (vii) $\underline{\underline{Y}}_U$, $\underline{\underline{Y}}_L$ são os limites de confiança superior e inferior, com *(1-α)*100 %* de probabilidade, para o valor correto da matriz de predição $\hat{\underline{\underline{Y}}}$ dentro da janela de treinamento.

Nas janelas de predições seguintes (ou seja: *t+NT+1* $\leq$ *k+1* $\leq$ *t+NT+Nl*) as seguintes relações são aplicadas para predição no instante *(k+1)*, $\hat{\underline{Y}}(k+1)$, e com fronteiras de confiança *(1-α)*100 %* $\underline{Y}_U(k+1)$ e $\underline{Y}_L(k+1)$:

$$\hat{\underline{Y}}(k+1) = \hat{\underline{\underline{\theta}}}^T \underline{\underline{\Xi}}(k) \qquad \text{eq. 4.16}$$

$$\underline{\underline{\Xi}}(k)^T = [-\underline{Y}(k)^T \ \ldots \ -\underline{Y}(k-n+1)^T \ \underline{U}(k)^T \ \ldots \ \underline{U}(k-m+1)^T] \qquad \text{eq. 4.17}$$

$$\underline{Y}_U(k+1) = \hat{\underline{Y}}(k+1) + T_{1-\alpha/2}\sqrt{\underline{S}_R \, \underline{\underline{\Xi}}(k)^T [\underline{\underline{\Xi}}^T \underline{\underline{G}}\underline{\underline{\Xi}}]^{-1} \underline{\underline{\Xi}}(k)} \qquad \text{eq. 4.18}$$

$$\underline{Y}_L(k+1) = \hat{\underline{Y}}(k+1) - T_{1-\alpha/2}\sqrt{\underline{S}_R \, \underline{\underline{\Xi}}(k)^T [\underline{\underline{\Xi}}^T \underline{\underline{G}}\underline{\underline{\Xi}}]^{-1} \underline{\underline{\Xi}}(k)} \qquad \text{eq. 4.19}$$

Um vazamento pode ser detectado, na janela de utilização, quando observada uma forte discrepância instantânea e estatisticamente significante entre a predição e a resposta medida. Por “estatisticamente significante” assumem-se valores maiores que 3 vezes a metade da largura do intervalo de confiança. A sensibilidade do detector deve depender:

(a) do conjunto de respostas monitoradas ($\underline{Y}$) e da sua proximidade física com o local do vazamento,

(b) das ordens e tamanhos (*n, m, ny, nu, NT, Nl*) que caracterizam o detector,

(c) a intensidade do ruído agindo sobre o valor medido, e

(d) a intensidade do vazamento.

Graças à natureza intrínseca do modelo de detector de vazamentos, torna-se especialmente sujeito a falhas quando, durante a janela de treinamento, ocorrer um vazamento. Dessa forma, durante a fase de treino deve-se garantir a não ocorrência de vazamentos. Essa mesma condição já havia sido levantada anteriormente quando outras metodologias baseadas em séries temporais foram propostas. Durante a fase de utilização, caso ocorra vazamento, após sua detecção será necessário repetir a etapa de treinamento para garantir novamente a eficiência do detector. Ou seja, após a detecção do vazamento, o detector torna-se incapaz de detectar qualquer novo vazamento.

A metodologia de detecção de vazamentos consiste em construir um preditor ARX-MIMO (*multiple input multiple output*) – o detector – que é colocado para "assistir" um conjunto de respostas do processo. Esse conjunto de respostas que serão assistidas consistem de um sub-conjunto das variáveis de processo disponíveis. Na ocorrência de um vazamento na fase de uso do preditor, haverá um expressivo e rápido descolamento entre valor predito pelo ARX e valor medido, rompendo o intervalo de confiança.

As fases da metodologia de detecção são **[VAZ JUNIOR *et al*, 2010]**:

(i) O detector é construído com dados de processo (*SU*) e (*SY*) para instantes $t+1,...,t+NT$ (fase de treino);

(ii) Os limites de confiança e a predição são gerados sucessivamente durante a fase de uso do preditor, $t+NT+1, ..., t+NT+NI$;

(iii) A detecção do vazamento é associada com a discrepância entre a predição e o valor medido;

(iv) Depois de $t+N+NI$ (ou depois de um vazamento ter sido detectado), o detector precisa ser "rejuvenescido" através de uma nova sessão de treinamento.

As fases (i) e (ii) descritas acima foram programadas em linguagem LADDER, tradicionalmente utilizada nos Controladores Lógico Programáveis (CLP) disponíveis no mercado. Uma descrição mais detalhada sobre a linguagem LADDER, seus recursos e limitações foge ao escopo deste trabalho, porém, o código LADDER do preditor ARX implementado é apresentado no Anexo I. A possibilidade de programar tal abordagem de detecção de vazamentos diretamente em ambiente CLP representa avanço importante em sua incorporação ao ambiente real de controle de dutovias.

Controladores Lógico Programáveis processam a linguagem Ladder de forma rápida e dedicada, sendo altamente estáveis e seguros. Porém, a linguagem LADDER é extremamente limitada em termos matemáticos, não dispondo de cálculos matriciais ou quaisquer operações matemáticas avançadas. Desse modo, em muitos casos a metodologia de detecção de falhas não é passível de ser implementada diretamente nesta linguagem.

Nestes casos torna-se necessário usar recursos auxiliares, tais como o desenvolvimento de *scripts* construídos em ambiente Windows para rodarem acoplados ao software supervisório, tradicional ponto de instabilidade no sistema. Outra opção é programar o método de detecção em um software de processamento matemático, tal como o MATLAB. Estabelecendo então comunicação em tempo real entre este software e a base de dados do supervisório. Esta comunicação em ambiente Windows pode ser feita seguindo o protocolo OLE for Process Control (OPC), conforme descrito por **ZHANG *et al.* [2009]**.

Para a ferramenta de localização de vazamentos é usado outro preditor ARX-MIMO para $\underline{Y}$ com um filtro extra *C* (ordem *p*) para um segundo conjunto de entradas ($\underline{H}$, tamanho *2h*) correspondendo às alterações conhecidas do processo durante um vazamento **[VAZ JUNIOR *et al*, 2010]**. Cada tubo *k (k=1,...h, h≤E)* com diâmetro D_k e comprimento L_k, candidato a desenvolver vazamentos, envia um vetor de sinal de vazamento ($\underline{H}_k$) com dois números adimensionais. Esses dois números caracterizam a intensidade e a posição do vazamento: (i) o sinal *UD* dado

por $(D_F / D_k)^{2.5}$, (ii) o sinal UL dado por $(D_F / D_k)(1 + L_F / L_k)$, onde D_F é o diâmetro do orifício e L_F sua posição. Assim, $\underline{H}^T = [\underline{H}_1^T \ldots \underline{H}_h^T]$. O localizador é escrito como uma expansão da Equação (4.20), de acordo com:

$$\hat{\underline{Y}}(t+1) = \underline{\underline{B}}(q)\underline{U}(t+1) + \underline{\underline{C}}(q)\underline{H}(t+1) + q\left[1 - \underline{\underline{A}}(q)\right]\underline{Y}(t) \qquad \text{eq. 4.20}$$

$$\underline{\underline{C}}(q) = \sum_{k=1}^{p} \underline{\underline{C}}_k q^{-k}(.) \qquad \text{eq. 4.21}$$

Assume-se que o tubo candidato *K* não tem vazamento, $\underline{H}_k = \underline{0}$. O localizador é designado para aderir à resposta do processo quando esta é informada com o correto vetor de sinal de vazamento. Seus parâmetros precisam ser identificados previamente para qualquer evento, através do uso de série temporal simulada contendo: (i) entradas típicas da rede (<u>*U*</u>), (ii) sinais de vazamento <u>*H*</u> correspondentes a um conhecido programa de vazamento que afete os tubos candidatos, e (iii), respostas <u>*Y*</u> geradas por simulação.

O procedimento para localização de vazamentos começa após a detecção da ocorrência de um vazamento no instante T_D, de acordo com os seguintes passos:

(i) Preparação de uma ocorrência de vazamento

(ii) Para cada tubo candidato *K* sujeito a desenvolver vazamento, o vetor $\underline{H}_k$ é estimado para maximizar a aderência do localizador para a ocorrência. Essa maximização da aderência é feita via minimização da função objetivo (Φ_K) dado pela soma quadrática dos resíduos preditos. Isso configura uma busca bidimensional no domínio das variáveis D_F e L_F. Ou seja: minimiza-se o resíduo para otimizar as <u>indicação do diâmetro de furo e localização do vazamento</u>

(iii) A ocorrência do vazamento é localizada no tubo *K*, que possui função objetivo mínima (Φ_K), tendo parâmetros D_F e L_F, calculados da estimativa H_k.

5. Rede de Escoamento

Devido à impossibilidade técnica de treinar um sistema de detecção de anomalias em redes industriais reais, torna-se necessário o desenvolvimento de uma ferramenta para geração de séries temporais, com e sem falhas. Essa ferramenta comporta-se como uma versão virtual da rede de escoamento real. Com este recurso pode-se avaliar o comportamento da rede em variadas condições operacionais, além de testar os procedimentos de detecção de vazamentos e falhas em sensores propostos. Os treinos e testes de detectores – bem como todo o arcabouço de sistemas detectores aqui desenvolvidos – são conduzidos usando-se séries temporais geradas pelo recurso de simulação da rede de escoamento. Não se trabalhou com dados reais de processo durante o desenvolvimento deste trabalho.

5.1. Rede de Fornecimento de Nafta à Planta de Insumos Básicos BRASKEM-Camaçari

A rede de escoamento aqui estudada é esquematizada em digrafo na Figura 5.1. Os vértices deste digrafo correspondem fisicamente a locações especiais do sistema, como junções, pontos de instrumentação, pontos de ramificação, pontos com válvulas de emergência acionadas remotamente, tancagens intermediárias, pontos de carga e de descarga. Os vértices são identificados por números entre colchetes "[ ]".

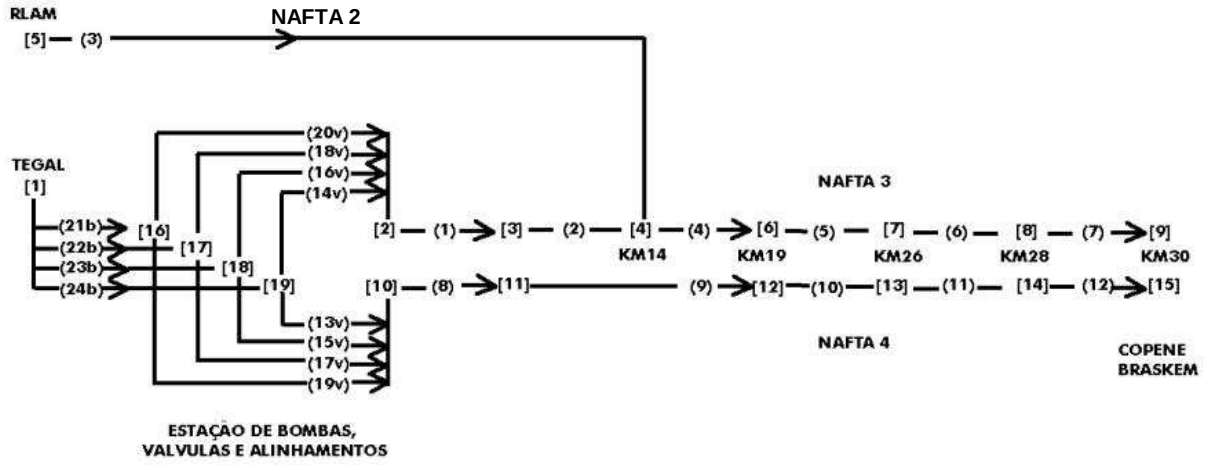


Figura 5.1: Rede de Fornecimento de Nafta à Planta de Insumos Básicos da Braskem

As arestas orientadas do digrafo correspondem a trechos de escoamento conectando vértices. Fisicamente as arestas correspondem a tubulações com seção constante, bombas, e tubulações curtas valvuladas. Note-se que estas válvulas são utilizadas para efetuar os alinhamentos necessários no Canteiro de Bombas que caracterizam uma campanha do sistema. Na Figura 5.1 as arestas são identificadas por números entre parênteses "( )" que podem ser acompanhados das letras "b" ou "v" nos casos de arestas de bombas e arestas valvuladas, respectivamente. Deve-se atentar para a seguinte convenção utilizada neste Capítulo: as arestas tubulares, as arestas de bombas e as arestas valvuladas serão referenciadas (em itálico) simplesmente como "*tubos*", "*bombas*", e "*válvulas*" respectivamente.

Resumidamente a rede abrange dois pontos principais de alimentação, dois pontos (próximos) de destino e duas longas dutovias dorsais, paralelas, de grande diâmetro, para o transporte de nafta líquida, em temperatura ambiente, à Planta de Insumos Básicos da BRASKEM S.A. em Camaçari, Bahia. Os pontos de alimentação da rede são o Terminal TEGAL no Porto de Aratu e a Refinaria RLAM da PETROBRAS S.A. em Mataripe, Bahia. Os dois pontos de alimentação podem ser usados alternada ou simultaneamente, de acordo com a conveniência da Companhia em função de oportunidades associadas a oscilações de preços dos fornecedores de nafta, das flutuações dos custos portuários e do cumprimento de contratos de longa duração com a Refinaria. As dutovias dorsais são denominadas NAFTA3 e NAFTA4. NAFTA2 é a denominação da dutovia auxiliar conectando a refinaria a um vértice do NAFTA3. Todo o sistema é denominado Sistema Nafta.

O regime operacional principal corresponde à alimentação do sistema a partir de navios em TEGAL. Visando à otimização de investimentos e custos de arrendamento, não há tancagem em TEGAL. Por este motivo, e também para minimizar custos de atracação quando alimentadas a partir de um ou mais navios, as dutovias NAFTA3 e NAFTA4 foram projetadas para operação intermitente, bombeando nafta diretamente à tancagem em Camaçari com elevadas vazões.

As quatro bombas de alta potência empregadas estão instaladas no Canteiro de Bombas em um sítio BRASKEM à pequena distância do Terminal (aproximadamente 1km). Normalmente há um rodízio na operação do Canteiro de Bombas: três bombas operam em paralelo e a restante é mantida desligada. A Figura 5.2 apresenta as curvas de operação do Canteiro de Bombas no caso em que uma única bomba está em serviço e no caso em que três delas operam em paralelo. A vazão máxima do sistema atinge, portanto, o expressivo valor de 1.8 m^3/s (aproximadamente 1300 kg/s). O *Head* máximo atinge 631m de fluido (aproximadamente 45 bar).

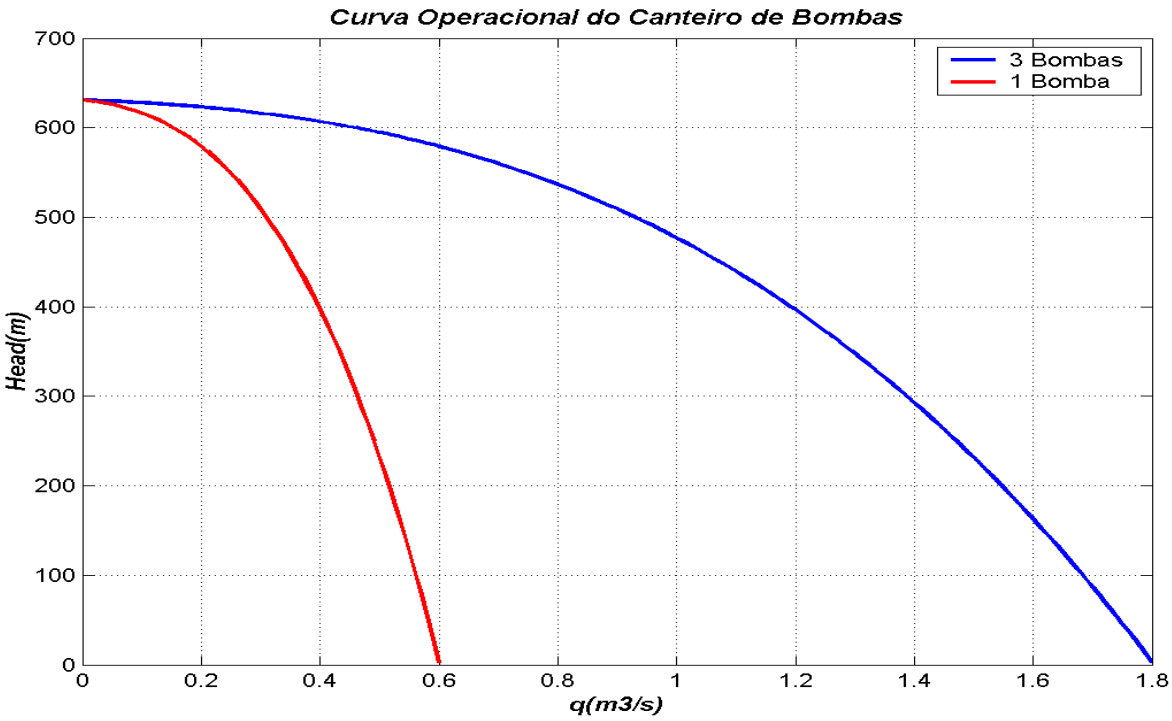


Figura 5.2 : Curva Operacional do Canteiro de Bombas do Sistema Nafta

O envio de nafta através de um duto de pequeno comprimento do(s) navio(s) ao Canteiro de Bombas deve ser sustentado pelas bombas da(s) própria(s) embarcação(ões). A partir do Canteiro a nafta poderá seguir ao longo de mais de 30km até Camaçari através do NAFTA 3, do NAFTA 4, ou de ambos simultaneamente. Esta seleção é implementada através do conjunto de 8 *válvulas* localizadas no Canteiro de Bombas (Figura 5.1).

O regime operacional secundário, ou complementar, corresponde à alimentação do sistema a partir da RELAM. No caso de apenas esta fonte operar, a transferência ocorre apenas através do NAFTA2 e posteriormente pelo NAFTA 3. Como já mencionado, entretanto, o sistema poderá operar com todas as fontes de alimentação e com ambas as dutovias dorsais NAFTA3 e NAFTA4 em serviço.

Nos documentos do Sistema obtiveram-se as curvas de bombas e os parâmetros:

- Comprimento (*L*) e diâmetro (*D*) de *tubos* (Tabela 5.1);
- Rugosidade de parede de tubos dada por $\varepsilon = 2.54.10^{-5}\,m$;
- Viscosidade do fluido no valor constante de $\mu = 0.56.10^{-3}\,Pa.s\ (0.56cP)$;
- Densidade do fluido no valor constante de $\rho = 720kg/m^3$;
- Elevações (*z*) dos vértices principais relativas ao nível do mar (Tabela 5.2).

Outras variáveis e parâmetros da modelagem da rede seguem abaixo (unidades S.I.; pressões (estáticas) absolutas; aceleração gravitacional $g = 9.81$ m/s^2):

- N : Número de vértices da rede;
- E : Número de arestas da rede;
- t : Instante de tempo;
- nt : Número de instantes de tempo em série temporal;
- P_k : Pressão (*bar*) no vértice *k*;
- $\underline{P}$: Vetor (*N x 1*) de pressões da rede;
- q_j : Vazão mássica *(kg/s)* na aresta *j* (positiva se a favor da aresta);
- $\underline{q}$: Vetor (*E x 1*) de vazões de arestas da rede;
- W_k : Vazão mássica *(kg/s)* externa no vértice *k* (positiva se alimentação);
- $\underline{W}$: Vetor (*N x 1*) de vazões mássicas externas de vértices da rede;
- $\underline{X}$: Vetor (*2N+E x 1)* de variáveis da rede ($\underline{X}^T = [\,\underline{P}^T\ \underline{q}^T\ \underline{W}^T\,]$)
- f_j, Re_j : Fator de atrito de Darcy e Número de Reynolds na aresta *j*;

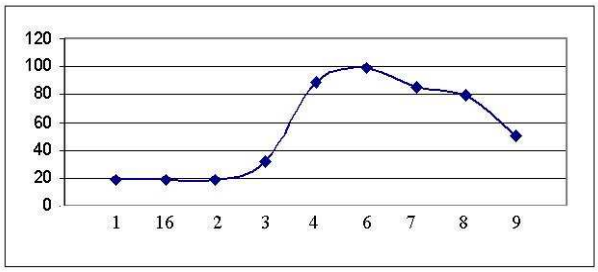
- $\underline{f}$: Vetor (*E x 1*) de fatores de atrito das arestas da rede;
- L_j : Comprimento (*m*) da aresta *j*;
- $\underline{L}$: Vetor (*E x 1*) de comprimentos;
- L_F : Posição (*m*) de uma ruptura ("furo") em uma aresta a favor de sua orientação;
- D_j : Diâmetro (*m*) da aresta *j*;
- $\underline{D}$: Vetor (*E x 1*) de diâmetros da rede;
- D_F : Diâmetro (*m*) de uma ruptura ("furo") em uma certa aresta;
- $\underline{K}$: Vetor (*E x 1*) de coeficientes hidráulicos das arestas da rede $(kg.m)^{-1}$;
- z_k : Elevação (*m*) do vértice *k*;
- $\underline{z}$: Vetor (*N x 1*) de elevações da rede;
- $\underline{\underline{G}}$: Matriz de seleção (*N x 2N+E*) de especificações da rede;
- G_{kn} : Termo *kn* de $\underline{\underline{G}}$ ($G_{kn}=1 \Rightarrow X_n$ especificado no vértice *k*; senão $G_{kn}=0$);
- $\underline{\phi}(t)$: Vetor (*N x 1*) de especificações de vértices da rede;
- $\underline{U}(t)$: Vetor (*nu x 1*) de *entradas* dinâmicas da rede ($\underline{U}(t) \subset \underline{\phi}(t)$);
- $\underline{Y}(t)$: Vetor (*ny x 1*) de *saídas* dinâmicas da rede ($\underline{Y}(t) \subset \underline{X}(t)$)
- $\underline{\underline{M}}$: Matriz de Incidência (*N x E*) da rede;
- M_{kj} : Termo *kj* da Matriz de Incidência da rede;
  $M_{kj}=1$ se aresta *j* chega ao vértice *k*;
  $M_{kj}=-1$ se aresta *j* sai do vértice *k*;
  $M_{kj}=0$ se aresta *j* não toca o vértice *k*.

Os sensores atualmente instalados no Sistema Nafta e suas respectivas localizações estão apresentados na Tabela 5.3. A localização de um sensor de pressão é dada pelo vértice em que ele se encontra. A localização de um sensor de vazão refere-se ao vértice no qual está instalado, sendo medida: (i) a vazão externa no caso de um vértice de extremidade; ou (ii) a vazão da aresta principal pelo vértice (por exemplo, a aresta de saída em uma junção ou a aresta de entrada em uma bifurcação). Ressalta-se que não há informações acerca de instrumentação no ramal RLAM (NAFTA2). Claramente esta instrumentação deve existir mas não houve acesso a estes dados por estar a gerência do duto a cargo de outra companhia. O

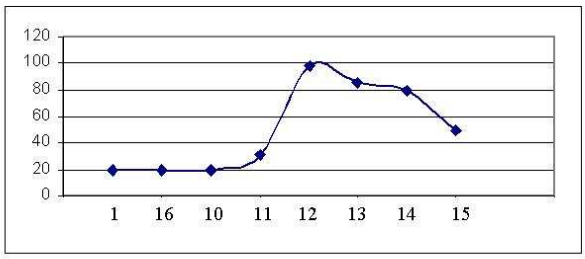
perfil das elevações ao longo de alguns caminhos de escoamento nas dutovias “NAFTA4”, “NAFTA3 via TEGAL” e “NAFTA3 via RLAM-NAFTA2” estão representados na Figura 5.3.

Tabela 5.1: Diâmetro Interno e Comprimento de *Tubos* do Sistema Nafta

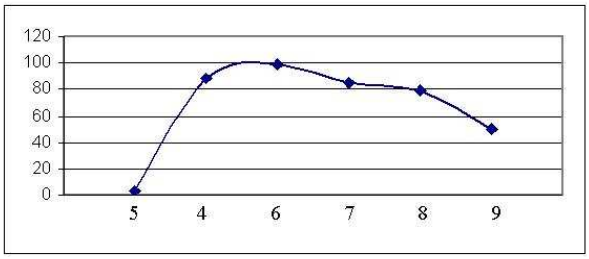
Tubo	Diâmetro (m)	Comprimento (m)
1	0,733	3876
2	0,485	4713
3	0,338	13860
4	0,338	13140
5	0,338	2000
6	0,338	1000
7	0,338	6800
8	0,880	3876
9	0,485	17853
10	0,485	2000
11	0,485	1000
12	0,485	6800



TEGAL – Nafta3 – Camaçari



TEGAL – Nafta4 – Camaçari



RLAM – Nafta3 – Camaçari

Figura 5.3: Perfis de Elevação (m) vs Vértices ao Longo de Caminhos de Escoamento

Tabela 5.2: Elevações dos Vértices Principais do Sistema Nafta
(Ref. Nível do Mar)

Vértices	Elevação (m)
1	19,1
2	19,1
3	31,4
4	89
5	4
6	98,5
7	85,5
8	79
9	50
10	19,1
11	31,4
12	98,5
13	85,5
14	79,5
15	50

Tabela 5.3: Sensores do Sistema Nafta ([#]: vértice)

Código	Tipo	Localização
PI-20 2 PI29103	Pressão	[2]
PI-20 3 PI29104	Pressão	[3]
PI-20 4 PI29421	Pressão	[4]
PI-20 10 PI29101	Pressão	[10]
PI-20 11 PI29102	Pressão	[11]
PI-20 16 PI0931A	Pressão	[16]
PI-20 17 PI0931B	Pressão	[17]
PI-20 18 PI0931C	Pressão	[18]
PI-20 19 PI0931D	Pressão	[19]
FI-10 1 XA29115	Vazão	[2]
FI-10 2 XA29124	Vazão	[3]
FI-10 4 XA29436	Vazão	[4]
FI-10 8 XA29101	Vazão	[10]
FI-10 9 XA29109	Vazão	[11]
FI-20 9	Vazão	[9]
FI-20 15	Vazão	[15]

5.2. Redução da Rede de Escoamento do Sistema Nafta

Foram implementadas simplificações estruturais na rede real visando a desenvolver uma plataforma simples e prática de simulação do sistema, ao mesmo tempo em que se conservam seus aspectos mais importantes, como comprimentos, diâmetros, elevações e capacidades. Um objetivo secundário consistiu em gerar-se uma rede com topologia similar à da original, oferecendo representação gráfica em escala de forma simples.

Como não se dispõe de coordenadas espaciais precisas dos vértices da rede, apenas das distâncias entre vértices e das suas elevações, definiu-se com base na Figura 5.1 um eixo *x* imaginário, solidário ao principal eixo de escoamento da rede, a

saber a direção e sentido associados às dutovias NAFTA3 e NAFTA4 (que talvez não sejam retilíneas na realidade). O eixo *y* é associado à extensão lateral imaginária do sistema. O eixo *z* permanece alinhado verticalmente, para cima, a partir do nível do mar.

Adicionalmente definiu-se um novo Vértice 1 gerado pela condensação do Vértice 1 original (TEGAL) com o Canteiro de Bombas. Com isto omite-se a complexidade desnecessária das conexões e *by-passes* do Canteiro de Bombas e evita-se lidar com a modelagem de *bombas* e *válvulas* nas equações do modelo do Sistema. Secundariamente, distende-se um pouco a rede ao longo dos eixos *y* e *x* para permitir distanciamento entre vértices apropriado à visualização gráfica em escala. As cotas verticais de vértices são mantidas nos valores originais devido à óbvia importância destes parâmetros no escoamento.

As arestas *bombas* e *válvulas* são portanto removidas. As modificações no Canteiro de Bombas acarretaram:

- Remoção dos vértices de distribuição 16 a 19
- Remoção de arestas de *Bombas* 21 a 24
- Remoção de arestas de *Válvulas* 15 a 20
- Arestas 13 e 14 sem válvulas, com suas tomadas diretamente no vértice 1
- Escoamento induzido por especificação de pressão (ou vazão externa) no vértice 1

Com respeito às modificações estruturais, descritas a seguir, os efeitos globais foram:

- Distensão axial da Rede em aproximadamente 7000m (eixo *x*)
- Distanciamento lateral das dutovias dorsais da Rede em 3000m (eixo *y*)
- Distensão lateral da "aba" RELAM em 5000m (eixo *y*)

As modificações em NAFTA3 e no Ramal RELAM resultaram em:

- Deslocamento de vértices 2,3,6 em +4000m no eixo *x* e +1500m no eixo *y*

- Deslocamento do vértice 4 em +4000m no eixo *x* e +6500m no eixo *y*
- Deslocamento de vértices 7,8,9 em +7000m no eixo *x* e +1500m no eixo *y*
- Uniformização aproximada de distancias entre vértices 6,7,8,9
- Deslocamento do vértice 5 (RELAM) para manter constante sua distância ao vértice 4

As modificações em NAFTA4 resultaram em:

- Deslocamento de vértices 10,11,12 em +4000m no eixo *x* e -1500m no eixo *y*
- Deslocamento de vértices 13,14,15 em +7000m no eixo *x* e -1500m no eixo *y*
- Uniformização aproximada de distancias entre vértices 12,13,14,15

A rede simplificada é representada em vista planar na Figura 5.4. A rede reduzida terá *N=15* vértices e *E=14* arestas. A mesma rede com visualização 3D é apresentada na Figura 5.5. A Tabela 5.4 apresenta os comprimentos e diâmetros dos *tubos* da nova rede. As elevações de vértices são idênticas às da Tabela 5.2.

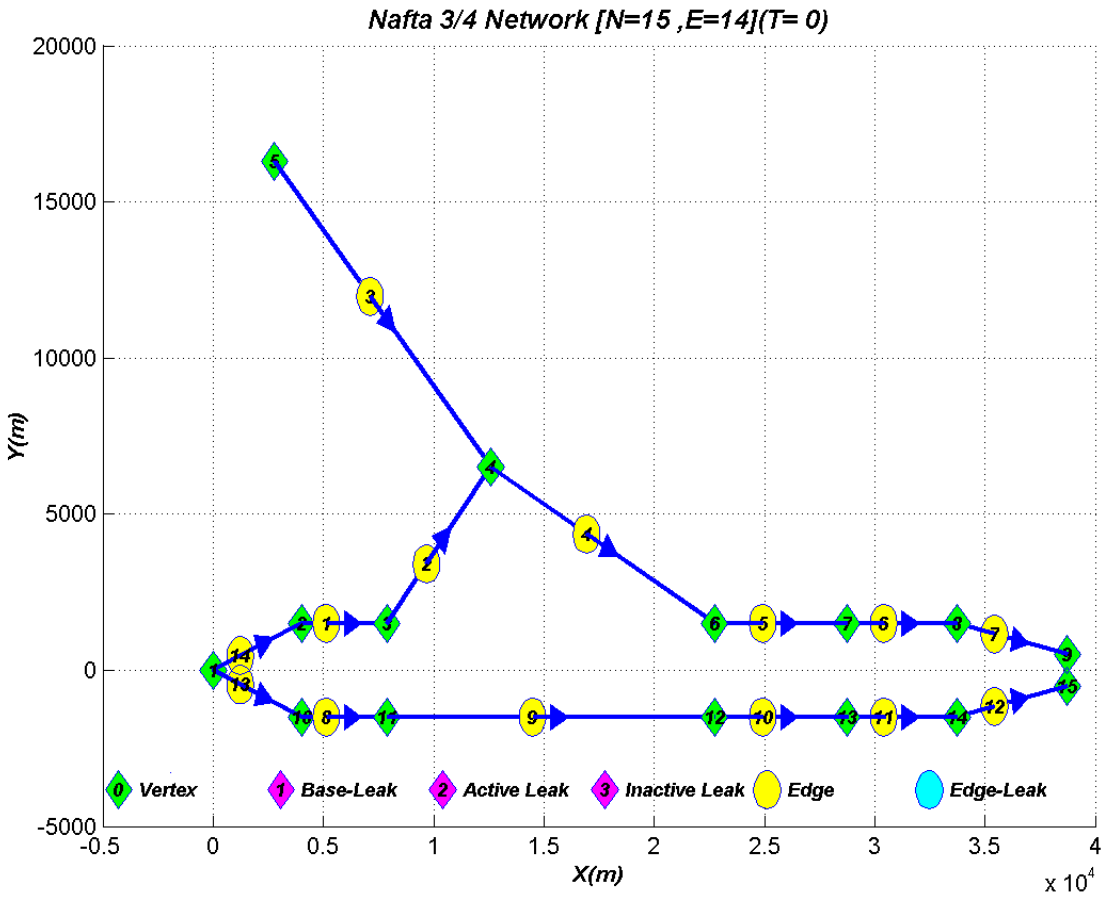


Figura 5.4 : Representação Planar (vista superior) em Escala da Rede Simplificada do Sistema Nafta

Na nova representação, a pressurização e tráfego no interior do sistema passam a ser definidos via especificação de pressão (*P*) e de vazão mássica externa (*W*) diretamente nos vértices de despacho (o novo vértice 1 TEGAL e o vértice 5 RELAM) e nos vértices de destino (os vértices 9 e 15 em Camaçari). Desta forma admite-se que *todos* os vértices de despacho e de recebimento atuam simultaneamente e que todas as arestas estão disponíveis para escoamento do fluido (isto é, dispensou-se a noção de válvula fechada).

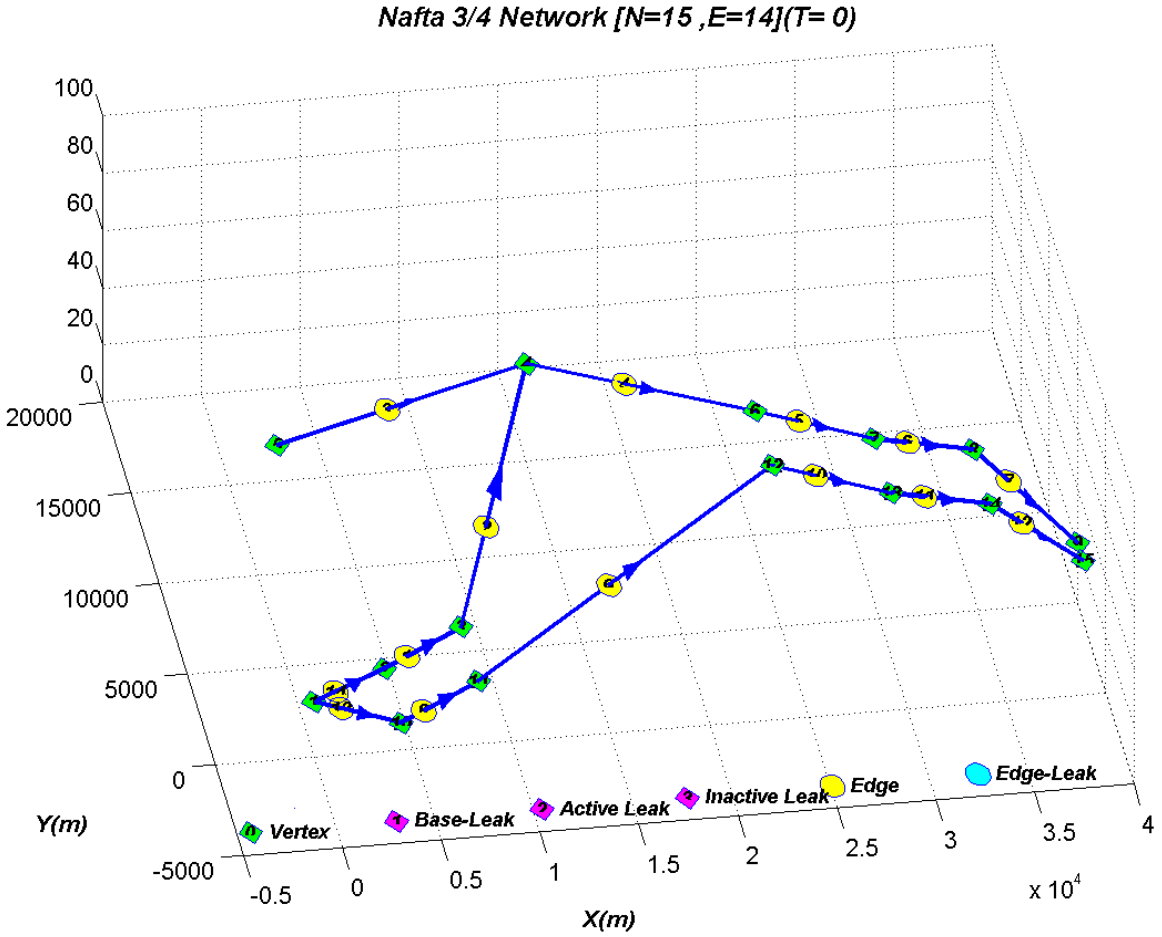


Figura 5.5 : Representação 3D em Escala da Rede Simplificada do Sistema Nafta

Uma vez que a definição do escoamento na rede exige *uma especificação por vértice*, a ativação (ou desativação) de caminhos de escoamento específicos será estabelecida via especificações apropriadas de pressão (*P*), de vazão externa nos

vértices de despacho (*W>0*) e vazão externa nos vértices de recebimento (*W<0*), sendo os demais vértices intermediários especificados com vazões externas nulas (*W=0*). Por exemplo, a especificação $P_1 > 0$, $W_5 = 0$, $W_{15} = 0$, $P_9 < P_1$, $W_k = 0\ (k = 2..4,6..8,10..14)$ estabelecerá (a menos da contribuição gravitacional) o caminho de escoamento TEGAL – NAFTA3 – Camaçari, com NAFTA2 e NAFTA4 pressurizados porém com vazões nulas. Analogamente, a especificação $P_5 > 0$, $W_1 = 0$, $W_{15} = 0$, $P_9 < P_5$, estabelecerá (a menos da contribuição gravitacional) o caminho de escoamento RELAM – NAFTA2 – NAFTA3 – Camaçari, com o trecho inicial de NAFTA3 e todo NAFTA4 pressurizados porém com vazões nulas. Por fim, a especificação $P_1 > 0, P_5 > P_1$, $W_{15} < 0$, $P_9 < P_1$, estabelecerá (a menos da verificação da contribuição gravitacional) simultaneamente todos os caminhos de escoamento do sistema, a saber:

- RELAM – NAFTA2 – NAFTA3 – Camaçari;
- TEGAL – NAFTA3 – Camaçari; e
- TEGAL – NAFTA4 – Camaçari (este com vazão pré-definida).

Tabela 5.4: Diâmetro Interno e Comprimento de *Tubos* da Rede Reduzida

Tubo	**Diâmetro (m)**	**Comprimento (m)**
1	0.7334504	3876
2	0.4857496	6871,4
3	0.3381248	13860
4	0.3381248	11305,7
5	0.3381248	6000
6	0.3381248	5000
7	0.3381248	5099
8	0.88014048	3876
9	0.4857496	14853,2
10	0.4857496	6000
11	0.4857496	5000
12	0.4857496	5099
13	0.88014048	4272
14	0.7334504	4272

5.2.1 Sensores da rede reduzida

Os sensores da rede real foram re-designados para as condições da rede reduzida. Com isto pretende-se adaptar o quadro de sensores de modo a garantir densidade similar de instrumentação entre as redes original e reduzida. Todavia, novos sensores foram incluídos visando a comparar desempenhos de diversos esquemas SDVs.

As mudanças conduzidas podem ser exemplificadas com o caso dos 4 sensores de pressão nos antigos vértices 16, 17, 18 e 19, que serão condensados em apenas um sensor de pressão no novo vértice 1. Para compensar esta redução, novos sensores de pressão são incluídos. Outra modificação estabelece sensores de vazão nas correntes externas (*W*) das 4 extremidades da rede (vértices 1,5,9,15), embora alguns destes já estivessem presentes no Sistema Nafta original. A Tabela abaixo apresenta a nova distribuição de sensores.

Tabela 5.5: Sensores da Rede Reduzida ([#]: vértice)

Código	Tipo	Localização
PI-1	Pressão	[1]
PI-2	Pressão	[2]
PI-3	Pressão	[3]
PI-4	Pressão	[4]
PI-5	Pressão	[5]
PI-7	Pressão	[7]
PI-9	Pressão	[9]
PI-10	Pressão	[10]
PI-11	Pressão	[11]
PI-12	Pressão	[12]
PI-15	Pressão	[15]
FI-1	Vazão	[1]
FI-2	Vazão	[2]
FI-4	Vazão	[4]
FI-5	Vazão	[5]
FI-9	Vazão	[9]
FI-10	Vazão	[10]
FI-12	Vazão	[12]
FI-15	Vazão	[15]

5.3 Modelo de Simulação Dinâmica da Rede Reduzida

A rede será simulada no tempo admitindo-se equações de escoamento em regime estacionário submetidas a um quadro temporalmente variável de parte das suas especificações. Cada vértice da rede exige a definição de uma especificação. Normalmente boa parte destas especificações são trivialmente associadas a valores constantes. Este é o caso das vazões externas nulas dos vértices de junção ou de ramificação da rede que acarretam a especificação $W=0$. Este também é o caso de vértices com *head* hidráulico fixo, como nos casos de reservatórios de nível

constante ou aberturas da rede para a atmosfera (i.e. o caso de vértices de vazamento) que acarretam a especificação $P=P_0$. Todavia, há uma parte das especificações do sistema que naturalmente variam no tempo em função de ruídos ou variáveis não-determinísticas que as influenciam. Este pode ser o caso de pressões e/ou vazões externas de extremidades da rede relacionadas aos pontos de bombeamento ou de consumo.

A estratégia de gerar-se séries temporais de respostas do sistema utilizando-se equações de estado estacionário respondendo *instantaneamente* a especificações transientes – as entradas dinâmicas do sistema, i.e. seus "inputs" – é obviamente uma simplificação, sendo conhecida como simulação em *modo pseudo-estacionário* **[SOUZA, 2004; COSTA *et al.*, 2002]**. Tal tipo de abordagem tende a ser aceitável em sistemas de dinâmica rápida com pequenos valores de tempo morto.

Embora sempre existam ponderáveis valores de tempo morto em redes de escoamento, admitiu-se aqui que a simulação em *modo pseudo-estacionário* é aceitável para descrever a dinâmica de resposta de redes de escoamento incompressível de porte médio ou pequeno, onde os dutos estão pressurizados e não existem folgas, bolsões de fase vapor ou vazios (i.e. não ocorre "slack-flow"). Este é aproximadamente o caso do escoamento no Sistema Nafta (e na versão reduzida desta rede) *após atingida* a condição de pressurização em todo o sistema.

5.3.1 Equações em Regime Estacionário para Redes Pressurizadas de Escoamento Incompressível

Têm-se **[COSTA *et al.,* 2002]** as seguintes equações para redes pressurizadas de escoamento incompressível em regime estacionário onde bombas não estão presentes e acidentes podem ser desprezados (ou convertidos em comprimentos equivalentes de tubulação):

Balanço Material em Vértices (*N* equações*)*:

$$\underline{\underline{M}}\,\underline{q} + \underline{W} = \underline{0} \qquad \text{eq. 5.1}$$

Balanço de Energia Mecânica em Arestas (*E* equações):

$$\underline{\underline{M}}^T \underline{P} + \underline{f} \bullet \underline{K} \bullet \left|\underline{q}\right| \bullet \underline{q} + \rho g \underline{\underline{M}}^T \underline{z} = \underline{0} \qquad \text{eq. 5.2}$$

Equações de Especificação de Vértices (*N* equações):

$$\underline{\underline{G}}\underline{X} - \underline{\phi}(t) = \underline{0} \qquad \text{eq. 5.3}$$

Estas *2N+E* equações são estabelecidas para as *2N+E* variáveis da rede $\underline{X}^T = [\,\underline{P}^T\ \underline{q}^T\ \underline{W}^T\,]$. Nas equações acima as multiplicações são interpretadas no sentido matricial convencional, excetuando-se aquelas (bem como as divisões) associadas ao operador "$\bullet$" que devem ser interpretadas no contexto de operação componente a componente entre dois vetores. O vetor de coeficientes hidráulicos (*(kg.m)*$^{-1}$) das arestas da rede é definido a partir da densidade do fluido e dos comprimentos e diâmetros de arestas pela equação 5.4 a seguir:

$$\underline{K} = (8/\rho\pi^2)\underline{L} \bullet / \underline{D}^5 \qquad \text{eq. 5.4}$$

Fatores de atrito de arestas são estimados pela Eq. de Churchill (**COSTA et al., 2002**) a partir das rugosidades relativas e do vetor de Números de Reynolds (<u>*Re*</u>) de arestas:

$$\underline{f} = \underline{f}(\varepsilon \bullet / \underline{D}, \underline{Re}), \quad \underline{Re} = 4\underline{q} \bullet / \pi \underline{D} \qquad \text{eq. 5.5}$$

As equações 5.1, 5.2 e 5.3 podem ser simbolizadas sob a forma funcional de um sistema de equações ressaltando-se o papel das variáveis $\underline{X}(t)$ e das especificações temporais $\underline{\phi}(t)$, segundo:

$$\underline{F}(\underline{X}(t),\underline{\phi}(t)) = \underline{0} \qquad \text{eq. 5.6}$$

A partir de série temporal de especificações da rede $\{\underline{\phi}(t_1),\underline{\phi}(t_2),...,\underline{\phi}(t_n)\}$, equação 5.6 pode ser numericamente resolvida para obter-se série temporal de variáveis de estado da rede $\{\underline{X}(t_1),\underline{X}(t_2),...,\underline{X}(t_n)\}$.

A resolução numérica da equação 5.6 foi implementada pelo Método Newton-Raphson com matriz jacobiana analítica e algoritmo de amortecimento. Para inicialização da busca Newton-Raphson de $\underline{X}(t_k)$, utiliza-se o valor imediatamente anterior existente na série temporal ($\underline{X}(t_{k-1})$). Quando este valor não se encontra disponível, a inicialização é executada resolvendo-se a versão linear da equação 5.6 correspondente à condição de escoamento laminar (**COSTA *et al.,* 2002**).

Definem-se as *Entradas* ($\underline{U}(t)$) e as *Saídas ou respostas* ($\underline{Y}(t)$) da rede como vetores relacionados a $\underline{\phi}(t)$ e $\underline{X}(t)$, tais que:

- $\underline{U}(t)$ é vetor (*nu x 1*) contendo as componentes de $\underline{\phi}(t)$ tipicamente transientes e variando de forma não-determinística no tempo ($nu \leq N$);
- $\underline{Y}(t)$ é vetor (*ny x 1*) contendo componentes observáveis (ou funções observáveis) de $\underline{X}(t)$, i.e. variáveis de estado do sistema (ou funções destas) para as quais há sensores habilitados. No presente caso, $\underline{Y}(t)$ sempre corresponderá a uma pequena parte das variáveis de estado da rede ($ny << 2N + E$).

A partir das séries temporais $\{\underline{\phi}(t_1),\underline{\phi}(t_2),...,\underline{\phi}(t_n)\}$ e $\{\underline{X}(t_1),\underline{X}(t_2),...,\underline{X}(t_n)\}$, podem ser extraídas as séries temporais de *Entradas* $\{\underline{U}(t_1),\underline{U}(t_2),...,\underline{U}(t_n)\}$ e de *Saídas* $\{\underline{Y}(t_1),\underline{Y}(t_2),...,\underline{Y}(t_n)\}$ do sistema.

5.3.2 Geração de Séries Temporais de Entradas para Redes Pressurizadas de Escoamento Incompressível

Séries temporais de entradas do sistema, $\{\underline{U}(t_1),\underline{U}(t_2),...,\underline{U}(t_n)\}$, podem ser criadas via modelos simples de séries temporais não-determinísticas existentes na literatura (**MEDEIROS *et al.*, 2001**). Existem diversas abordagens para este tipo de representação **(SOUZA, 2004**). A que foi admitida aqui é uma das mais simples existentes, caracterizada por ruído seguindo distribuição normal, com variância constante e esperança nula, atuando sobre o valor de referência da entrada, também suposto constante no tempo. Adicionalmente introduziu-se dependência não-determinística do valor atual em termos do passado da variável de entrada através da utilização de probabilidade de repetição de valor passado. Admitiu-se também que as séries temporais das diversas entradas não são correlacionadas.

Neste trabalho, as séries dependem de parâmetros arbitrários característicos da entrada em questão, definidos abaixo juntamente com outros símbolos:

- U_k^{Ref} :Valor de referência da entrada U_k;
- Γ_{U_k} : Probabilidade de repetição de U_k em instantes sucessivos;
- λ_{U_k} : Variância de U_k em torno de U_k^{Ref}
- $N(\mu,\lambda)$: Distribuição Normal com esperança μ e variância λ
- $Rnd(0,1)$: Distribuição Uniforme no intervalo [*0, 1*]

O modelo aqui utilizado corresponde à aplicação do seguinte algoritmo para a geração da série temporal para a *entrada* $U_k(t)$ $(k = 1 \ldots nu, t = 0 \ldots nt)$:

- $t = 0$
    - $U_k(t) = U_k^{Ref}$
    - Fim
- $t > 0$
    - $Gerar\ w \rightarrow Rnd(0,1)$
    - $w \leq \Gamma_{U_k}$
        - $U_k(t) = U_k(t-1)$
        - Fim
    - $w > \Gamma_{U_k}$
        - $Gerar\ \zeta \rightarrow N(0, \lambda_{U_k})$
        - $U_k(t) = U_k^{Ref} + \zeta$
        - Fim

Valores utilizados de parâmetros de geração de séries temporais de entradas são os seguintes:

- $\Gamma_{U_k} = 70\%$ (para todas as entradas U_k, em todos os exemplos);
- $\sqrt{\lambda_{U_k}} = 10\%\ de\ U_k^{Ref}$ (para todas as entradas U_k, em todos os exemplos)

5.3.3 Modificação do Modelo da Rede de Escoamento devido à Ação de Vazamentos

As equações apresentadas acima devem ser modificadas a partir da manifestação de um (ou mais) vazamento(s) em algum ponto do sistema. Novamente existem diversas maneiras de abordar a representação matemática da ocorrência deste evento, tanto do ponto de vista da dinâmica associada quanto da estimativa da vazão estacionária de vazamento (**SOUZA, 2004**). Neste trabalho adotou-se uma descrição bastante simples para enquadrar a ocorrência.

Na manifestação do vazamento considera-se que a rede de escoamento, originalmente com *N* vértices e *S* arestas, é alterada pela criação de 2 novos vértices e 2 novas arestas. A Figura 5.6 descreve a transformação associada ao surgimento

de um vazamento na aresta *k* ligando, originalmente, os vértices *n* e *m*. O vazamento surge na posição L_F correspondendo à posição fracionária $x = L_F / L_k$ (*0< x <1*). No ponto do vazamento, é instalado um vértice denominado de "raiz" do vazamento (*N+1*). A nova aresta *k* é associada ao trecho de duto do vértice *n* até *N+1*, mantendo o mesmo diâmetro e orientação anterior. A continuação da aresta danificada até o vértice *m* é associada a uma nova aresta *S+1*, de igual diâmetro e orientação com respeito à antiga *k*. Partindo do vértice "raiz" emerge uma pequena aresta *S+2* conectando *N+1* a um novo vértice *N+2* aberto para o exterior. A aresta *S+2* é orientada no sentido provável do escoamento de vazamento, tendo diâmetro dado pelo diâmetro da ruptura ($D_F = d$) e sendo seu comprimento simbolicamente escolhido em um valor muito pequeno (por exemplo, o próprio valor *d*).

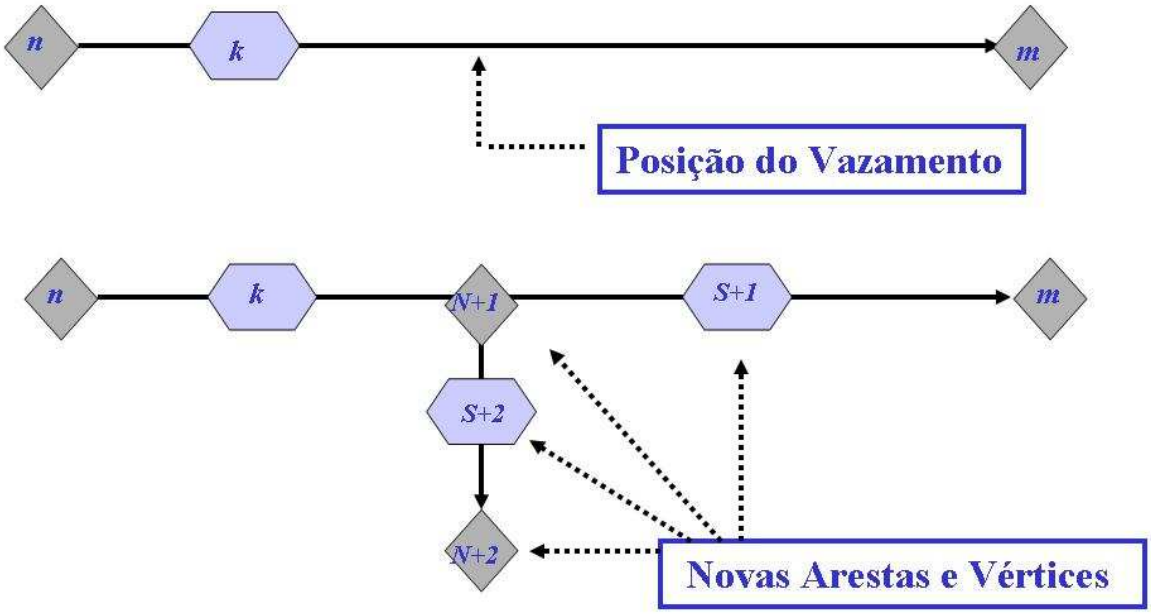


Figura 5.6 : Rede Inicialmente com *N* Vértices e *S* Arestas e após Ruptura e Vazamento

Estas transformações devem ser codificadas na matriz de incidência da rede afetada, sendo necessário atualizar-se as listas de arestas e vértices, bem como os comprimentos e diâmetros das arestas envolvidas (*k*, *S+1*, *S+2*), e as elevações e especificações dos vértices envolvidos (*N+1*, *N+2*). Estas inclusões são esquematizadas na Figura 5.7. As especificações dos novos vértices *N+1* e *N+2* correspondem a fixar vazão externa nula para o primeiro ($W_{N+1} = 0$) e o valor da

pressão externa para o segundo ($P_{N+2} = P_{ext}$). Por fim, devem ser atualizadas também as listas de variáveis da rede ($\underline{P}, \underline{q}, \underline{W}$), onde cada vetor deverá ganhar duas novas posições. A atualização do sistema de equações da rede afetada, equação 5.6, (devendo o mesmo ser feito para a respectiva matriz jacobiana) compreenderá:

- Equações anteriores de balanço de massa e de especificação para todos os vértices exceto *m;*
- Equações anteriores de balanço de energia mecânica para todas as arestas exceto *k*;
- Novas equações de balanço de massa e de especificação para vértices *m, N+1, N+2*;
- Novas equações de balanço de energia mecânica para arestas *k*, *S+1*, *S+2*;

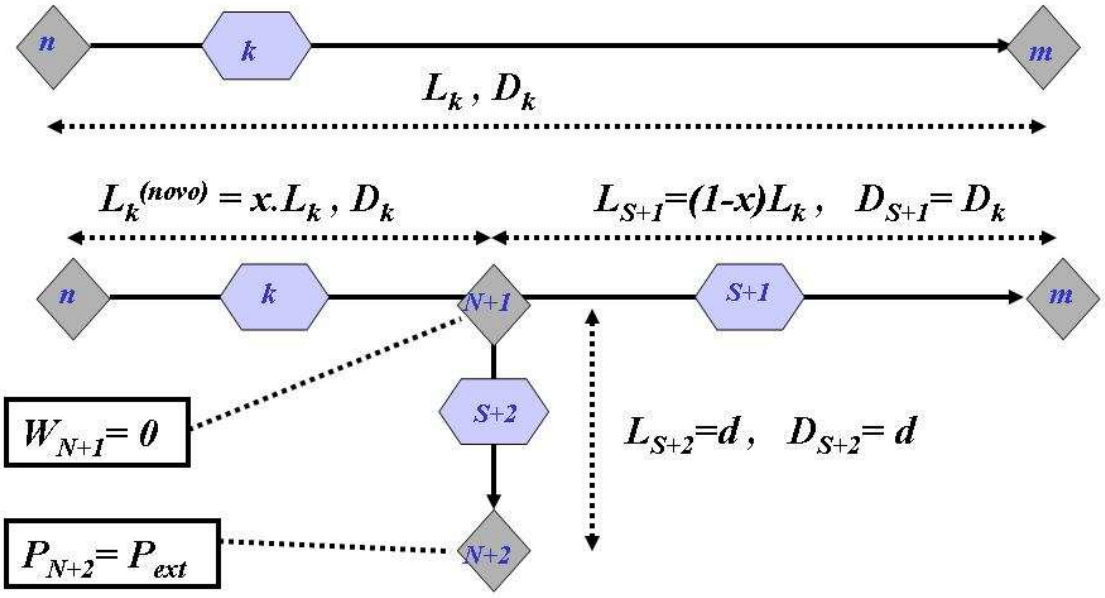


Figura 5.7 : Novas Dimensões de Arestas e Novas Especificações de Vértices após Ruptura e Vazamento em Rede Inicialmente com *N* Vértices e *S* Arestas

Após instalado o vazamento na representação matemática do sistema, as equações da rede danificada (equação 5.6) passarão a ser utilizadas na continuação

da série temporal em que o vazamento se manifestou, de modo a serem previstos os reflexos pertinentes. A vazão do vazamento recém-instalado corresponde aos valores das novas variáveis q_{S+2} e W_{N+2} (esta última sob a forma de um número negativo), as quais, evidentemente, não terão medidor associado.

Cabe também notar que a intervenção da ruptura no sistema promove uma transformação abrupta no mesmo, mas não interfere com o quadro de *entradas* dinâmicas da rede, pois os novos vértices criados apenas demandam especificações constantes que não são consideradas entradas. Isto é, o quadro de entradas dinâmicas continua agindo sobre o sistema no mesmo padrão não-determinístico anterior. A detecção do vazamento, dependerá, portanto, de variações perceptíveis nas respostas geradas pela rede danificada a estas entradas.

Sob simulação em *modo pseudo-estacionário*, a ocorrência do vazamento gerará um impacto abrupto nas variáveis medidas do sistema em função do diâmetro da ruptura e da proximidade de sensores ativos ao vazamento. É esta transição no quadro de medições da rede que será utilizada pelo sistema de detecção para: (i) primeiramente detectar o vazamento; e (ii) posteriormente determinar a aresta afetada (k), e estimar a localização fracionária da ruptura ($x = L_F / L_k$) e o respectivo diâmetro ($D_F = d$).

Na metodologia aqui empregada, considera-se que o vazamento possa ser "desligado" após algum tempo de ação. Caracteriza-se, em geral, um vazamento pelos seguintes números:

- k : Número da aresta afetada;
- $x = L_F / L_k$: Posição fracionária de localização na aresta afetada;
- D_F : Diâmetro da ruptura;
- t_E : Instante de tempo de "Entrada" (i.e. manifestação) do vazamento;
- t_D : Instante de tempo de "Desligamento" (i.e. cessação) do vazamento.

Note-se que o "desligamento" do vazamento é obtido com baixíssimo esforço de programação comparativamente às alterações necessárias para criá-lo. A topologia da rede não demanda modificações. Bastará reduzir-se o diâmetro da ruptura (D_F) a um valor pequeno o bastante (porém não nulo, evitando-se perda de posto na jacobiana da rede) capaz de "vazar" desprezivelmente e não perturbar a marcha do algoritmo de resolução (por exemplo, usou-se $D_F = 0.25mm$ *ou* $0.5\ mm$ para "desligar" vazamentos). Nos gráficos de séries temporais de variáveis medidas, os instantes de "Entrada" e de "Desligamento" do *k*-ésimo vazamento na campanha, são indicados com flechas perpendiculares ao eixo do tempo com as inscrições "E_k" e "D_k".

6. Resultados das Metodologias Baseadas em SPE, Q, Spca, Sdist, Classificadores Hierárquicos, Protótipos e SVI

Neste capítulo e no seguinte, são apresentados os resultados obtidos para cada uma das metodologias de detecção, identificação e quantificação de anomalias propostas, empregadas conforme descrito na Figura 3.1. Ferramentas que utilizam as metodologias SPE, Q, Spca, Sdist, classificadores hierárquicos, protótipos e SVI terão seus resultados apresentados neste capítulo **[VAZ JUNIOR *et al.*, 2009]**, enquanto o desempenho das abordagens baseadas em modelos auto-regressivos será descrito no Capítulo 7 **[VAZ JUNIOR *et al.*, 2010]**.

Para melhor se avaliar cada uma das metodologias, diferentes cenários são propostos, variando-se tanto a localização quanto a severidade da falha do sensor ou vazamento. De modo a permitir que o desempenho de cada abordagem seja avaliado frente a uma vasta gama de situações, opta-se por utilizar aqui apenas um trecho da dutovia previamente apresentada. Deste modo, os resultados deste Capítulo são baseados no trecho da via denominado “NAFTA 3”, sendo este responsável pela ligação do Vértice 1 (TEGAL) ao Vértice 9 (Braskem), conforme previamente descrito.

Nestes testes, são assumidos que todos os vértices e tubos da dutovia contêm sensores de pressão e vazão instalados. A lista completa de sensores do NAFTA 3 aqui utilizados é apresentada na Tabela 6.1, e sua disposição espacial dos sensores ao longo da via encontra-se na Figura 6.1. Devido a fatores econômicos, a rede real descrita no Capítulo 5 contém um número menor de sensores instalados (Tabela 5.3), fato que contribui negativamente para a eficiência dos métodos de detecção.

6.1 Detecção de falhas em sensores

Os resultados aqui apresentados derivam da aplicação de duas abordagens distintas e complementares com a finalidade de detectar anormalidades operacionais em uma dutovia. A primeira metodologia baseia-se em técnica de estatística multivariável denominada “análise dos componentes principais” (PCA). Propriedades

físicas inerentes ao processo de escoamento promovem elevada correlação entre o comportamento dos dados de pressão e vazão gerados pelos sensores posicionados ao longo da rede. Tal correlação é modificada diante da ocorrência de uma anomalia, sendo este o princípio fundamental da detecção de falhas via PCA. Complementarmente a esta aproximação, uma segunda abordagem, baseada em classificadores hierárquicos, é empregada. Classificadores hierárquicos buscam separar diferentes amostras em agrupamentos, usando para isso algum critério de similaridade previamente definido.

Enquanto a análise por componentes principais mostra-se eficaz na identificação de modificações no padrão de comportamento do sistema e, portanto, na detecção de anomalias, esta abordagem por si só não é capaz de indicar com precisão o momento de início da falha operacional. Para que essa identificação possa ser feita com exatidão, as observações temporais anteriores e posteriores à ocorrência da anomalia são então submetidas a um classificador, sendo desse modo alocadas em dois grandes agrupamentos, ou “clusters”. Enquanto pontos sem anomalia são alocados em um agrupamento, amostras representantes de instantes com falha operacional são posicionadas no outro. A técnica permite ainda estimar a severidade da falha a partir do afastamento espacial dos clusters formados. A abordagem híbrida busca assim unir as vantagens provenientes das duas técnicas.

Tabela 6.1: Localização e tipo dos sensores considerados

Número do Sensor	Código do Sensor	Tipo	Localização
1	FI 20-1	Vazão (m^3/s)	[1]
2	FI 20-9	Vazão (m^3/s)	[9]
3	PI 20-1	Pressão (bar)	[1]
4	PI 20-2	Pressão (bar)	[2]
5	PI 20-3	Pressão (bar)	[3]
6	PI 20-4	Pressão (bar)	[4]
7	PI 20-6	Pressão (bar)	[6]
8	PI 20-7	Pressão (bar)	[7]
9	PI 20-8	Pressão (bar)	[8]
10	PI 20-9	Pressão (bar)	[9]
11	FI 10-1	Vazão (m^3/s)	(1)
12	FI 10-2	Vazão (m^3/s)	(2)
13	FI 10-4	Vazão (m^3/s)	(4)
14	FI 10-5	Vazão (m^3/s)	(5)
15	FI 10-6	Vazão (m^3/s)	(6)
16	FI 10-7	Vazão (m^3/s)	(7)

Nota: [ ] representa vértice, ( ) representa tubo

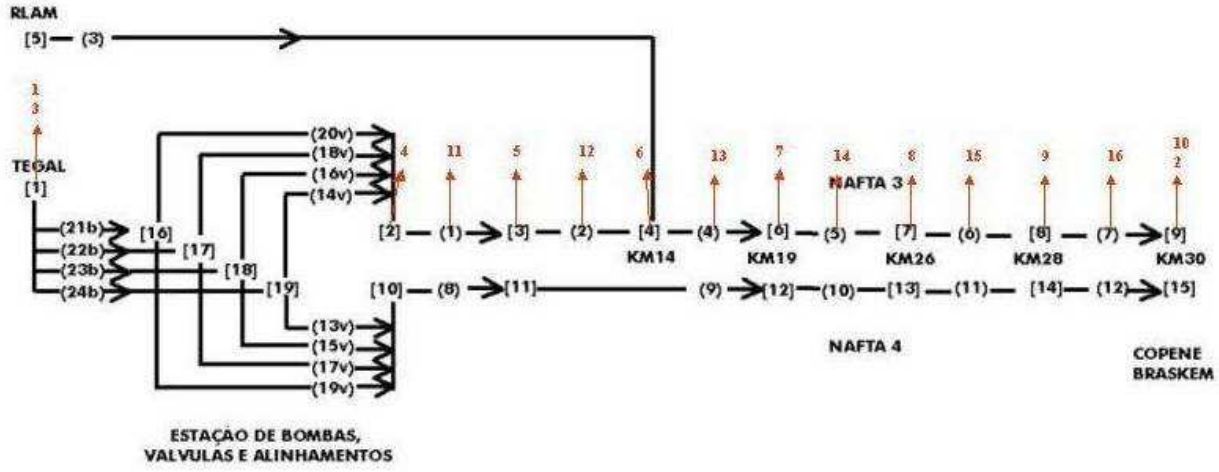
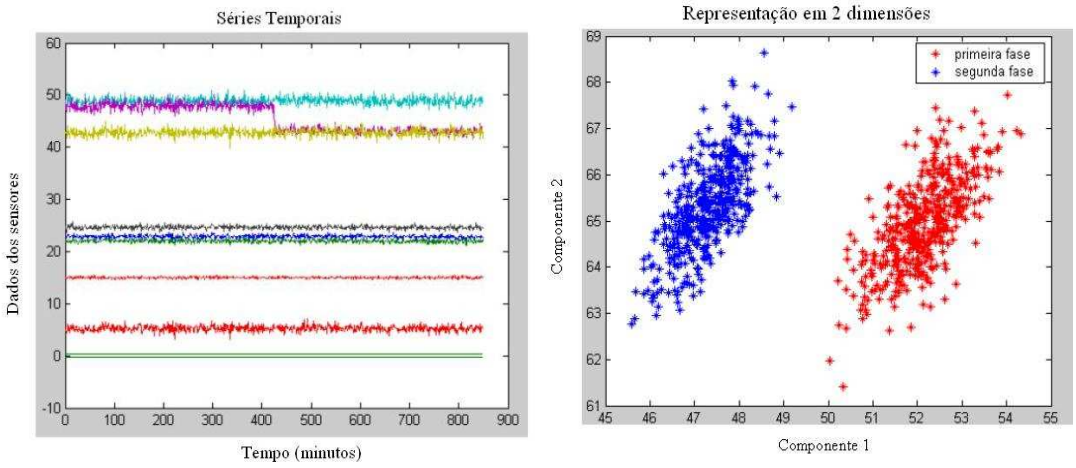


Figura 6.1: Sensores da rede de escoamentos NAFTA3

6.1.1 Detecção por análise dos componentes principais (PCA)

Aplicando-se inicialmente a técnica de PCA como instrumento de visualização dos dados, observa-se que esta permite resolver corretamente os instantes temporais anteriores e posteriores a ocorrência de uma anomalia. A Figura 6.2(a) representa a ocorrência de falha em um sensor no instante temporal de 400 minutos, contido em uma série temporal de 850 instantes de tempo. A projeção dos dados em um subespaço vetorial de dimensão dois, cuja base vetorial é composta pelos dois componentes principais de maior relevância promove a separação das amostras em dois grandes agrupamentos (Figura 6.2(b)). No primeiro encontram-se situadas as observações temporais representantes de situação sem falha, enquanto o outro contém as amostras relativas a janela com falha.

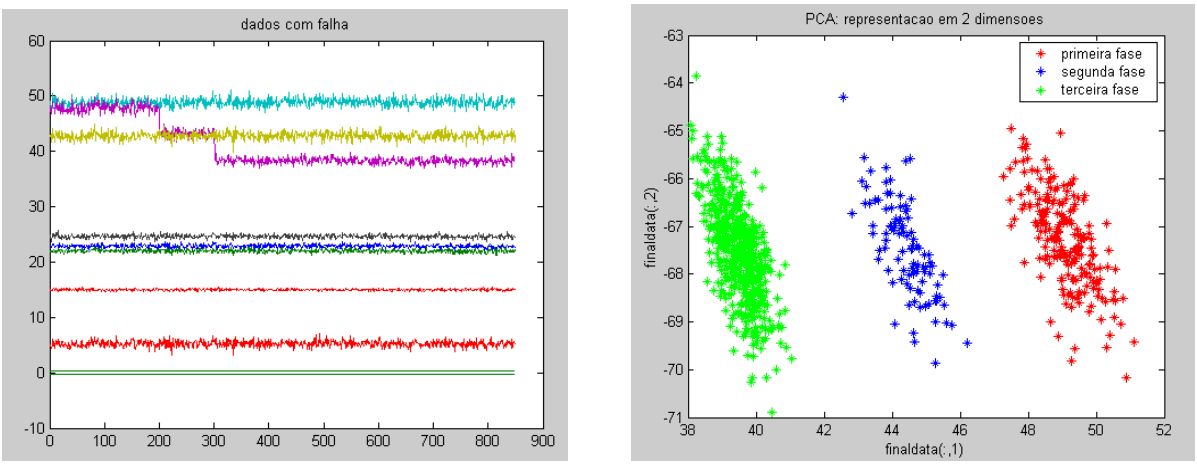
A boa resolução visual apresentada permite adiantar uma provável eficiência da técnica PCA na detecção de falhas deste tipo. Interessante notar ainda que diante da falha aqui exemplificada, uma projeção em R^1 através do 1º componente principal já seria suficiente para obter distinção entre os dois agrupamentos formados.



(a) (b)

Figura 6.2: Representação em R^2 de uma situação de falha de sensor: (a) Séries Temporais com falha no instante 400 min; (b) Projeção dos dados em subespaço vetorial bidimensional.

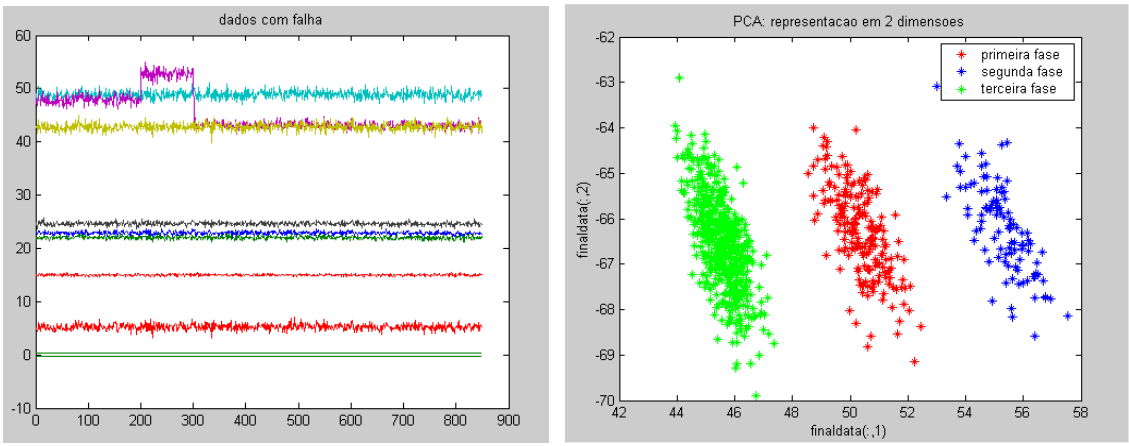
O afastamento relativo entre os *clusters* mostra-se função da severidade da anomalia estudada. Deste modo, aplicando-se inicialmente uma falha de sensor de severidade 5% no instante 200, e na seqüência, um desvio de 10% em relação ao valor correto a partir do instante 300, promove-se a separação das amostras em três grandes agrupamentos. A Figura 6.3(a) apresenta a série temporal gerada pelos sensores, sendo possível observar a ocorrência das duas anomalias descritas. Os três *clusters* formados utilizando PCA como instrumento de visualização estão representados na Figura 6.3(b). Verifica-se que, quanto maior a severidade da falha, maior o afastamento entre os agrupamentos formados.



a **b**

Figura 6.3: Representação em R^2 de uma situação de falha de sensor com severidade crescente: (a) Séries Temporais com falha de severidade variável; (b) Projeção dos dados em subespaço vetorial bidimensional.

As falhas de sensores analisadas até este ponto foram desvios negativos das medidas corretas. A técnica é igualmente válida para variações positivas no sinal dos sensores. A Figura 6.4(a) ilustra tal situação, onde desvios positivos e negativos são aplicados sucessivamente nos instantes 200 e 300. A separação resultante da visualização PCA encontra-se na Figura 6.4(b).

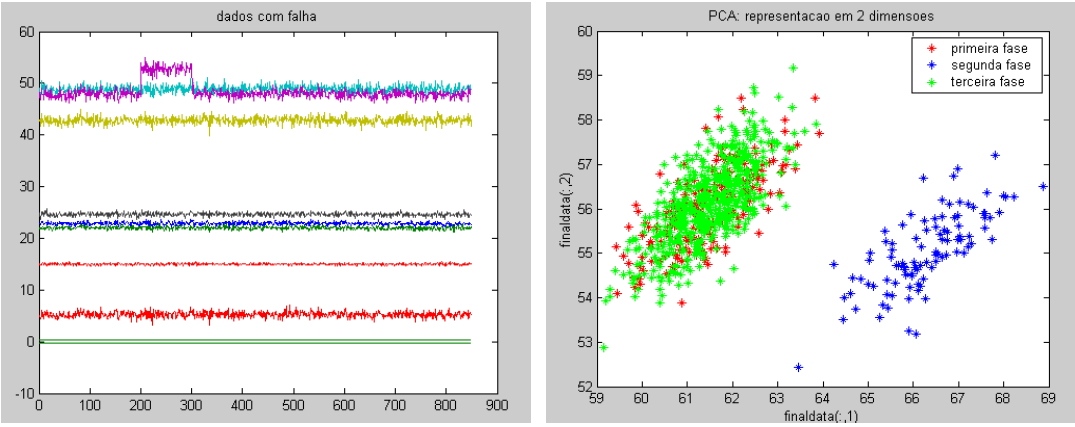


a **b**

Figura 6.4: Representação em R^2 de uma situação de falha de sensor com desvio positivo e negativo: (a) Séries Temporais com falha; (b) Projeção dos dados em subespaço vetorial bidimensional

Por fim, a Figura 6.5(a) ilustra a situação na qual a falha é corrigida instantes após ter se iniciado. Mais especificamente, o sinal de um dos sensores da dutovia

sofre desvio positivo de 10% no instante 200, porém tal desvio é rapidamente corrigido e o instrumento volta a operar corretamente. Na visualização via PCA, (Figura 6.5(b)) as duas fases de operação normal, antes e após a janela com falha, situam-se um único agrupamento. As amostras relativas ao período com falha localizam-se em um *cluster* afastado.



a **b**

Figura 6.5: Representação em R^2 de uma situação de falha de sensor anulada após 100 instantes de tempo: (a) Séries Temporais com falha; (b) Projeção dos dados em subespaço vetorial bidimensional

Apesar de eficiente nas situações ilustradas, a visualização PCA apresentada baseia-se apenas nos dois componentes principais de maior relevância, sendo incapaz de resolver corretamente anomalias de baixa severidade. Para tais casos se faz necessário o uso de um maior número de componentes principais, tornando inviável a avaliação visual do resultado. Metodologias avançadas de análise, como as descritas anteriormente, precisam ser empregadas. Os resultados para detecção de falhas em sensores são detalhados a seguir.

6.1.1 a) Via Q e SPE

Nesta abordagem, assume-se a fase “A” sem a presença de falhas, com duração de 250 instantes de tempo, estendendo-se do instante inicial 0 ao instante final 250. A fase “B” inicia-se no instante 251, prosseguindo até o instante 800. Falha em um dos sensores foi implementada a partir do instante 301, ou seja, 50 instantes

de tempo após o início da fase “B”. A Figura 3.3, apresentada no Capítulo 3, ilustra a disposição das fases e a ocorrência de falha ao longo do eixo temporal.

Para a rede de escoamento empregada, a matriz de dados referentes à fase A possui dimensão 250x16, ou seja, 250 instantes de tempo e 16 sensores. Ordenando-se os autovetores no sentido do maior autovalor para o menor, foram selecionados os 12 primeiros vetores, denominados “componentes principais”. Com esse número de autovetores atinge-se um grau de explicação relativo à variabilidade dos dados superior a 90%. A Tabela 6.2 apresenta o grau de explicação por componente acumulado. Na tabela estão apresentados os resultados obtidos para dois experimentos independentes.

Tabela 6.2: Grau de explicação acumulado por componente principal

Componente	**% acumulada exp 1**	**% acumulada exp 2**	**Componente**	**% acumulada exp 1**	**% acumulada exp 2**
1	47,85	46,11	9	83,13	83,43
2	54,31	52,89	10	86,28	86,78
3	59,68	58,48	11	89,19	89,77
4	64,13	63,18	**12**	**92,03**	**92,61**
5	68,39	67,71	13	94,68	95,37
6	72,51	71,98	14	97,20	97,77
7	76,37	76,16	15	99,58	99,55
8	79,89	79,95	16	100,00	100,00

Conforme previamente descrito, a aplicação desta metodologia envolve o ajuste de um único modelo, que deve descrever com precisão os dados relativos à fase de treino, ou fase “A”. Este modelo PCA é então aplicado aos dados da fase de uso (fase “B”). Sendo a fase de uso apenas uma continuidade da operação normal observada na fase anterior, o modelo ajustará corretamente também os dados da fase “B”. Porém, no caso da fase “B” apresentar anormalidades que desviem o processo da normalidade, o modelo PCA não mais se ajustará corretamente aos dados desta fase, sendo este um indicativo forte da ocorrência de problema operacional. Falhas em sensores e vazamentos são exemplos de situações anormais que podem ser detectadas através da aplicação desta metodologia.

A qualidade do ajuste do modelo aos dados da fase de uso pode ser medida através das métricas Q e SPE. Valores elevados destes indicam a não aderência do modelo aos dados e, neste caso, a presença de falha operacional. Um dos parâmetros matemáticos de ajuste do método SPE é a variável λ (efeito memória), tornando-se assim interessante verificar a influência deste sobre o desempenho final do sistema de detecção. Para cada um dos dezesseis sensores presentes na rede de escoamento, foram implementadas falhas de severidades 5%, 10% e 20%. E para cada combinação "sensor/severidade" verificou-se a eficiência da ferramenta de detecção quanto aos valores de λ de 0.25, 0.50 e 0.75. Deste modo, a detecção de falha de severidade 5% no sensor 1, por exemplo, foi testada para ajustes de λ de 0.25, 0.50 e 0.75. O mesmo procedimento foi adotado para as demais combinações.

A partir dos testes efetuados conclui-se que λ's de valor reduzido contribuem para um maior nível de detecção de falhas. Ou seja, elevar a influência do passado sobre o valor presente de SPE incrementa sua capacidade detectora. A Figura 6.6 apresenta os resultados alcançados para cada escolha de λ. Optou-se neste trabalho por fixar o valor de λ em 0,25.

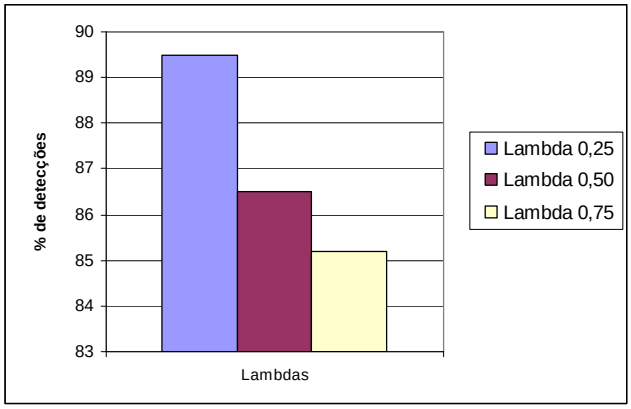


Figura 6.6: Grau de detecção via SPE em função do valor de lambda selecionado

Os resultados encontrados para detecção de falhas em sensores via SPE e Q estão reunidos na Tabela 6.3 e, entre outros fatores, demonstram que a distância entre o final da fase "A" e o instante inicial de ocorrência da falha parece exercer influência sobre o grau de detecção alcançado. Essa influência será investigada a seguir. Na Tabela 6.3, o índice sobrescrito, utilizado para a abordagem Sdist estática, corresponde ao código de seqüências da Tabela 6.4.

Ao se comparar o desempenho alcançado para o cenário de falha a partir do instante 301, com o grau de detecção alcançado em situações onde a falha inicia-se após o instante 451, observa-se que ao afastar o instante da falha do final da fase “A” promove-se redução no índice de detecção. Essa redução é especialmente forte para falhas de severidade baixa, supostamente as de mais difícil detecção. A situação na qual a falha inicia-se no instante 700, apresentou o pior desempenho, comprovando mais uma vez a influência do instante de ocorrência da anomalia sobre a capacidade detectora do método. As Figuras 6.7 e 6.8 exibem a disposição das fases “A” e “B”, além da janela de falha, ao longo do eixo temporal. Ilustram, assim, os diferentes afastamentos entre a fase “A” e o início da janela de falha.

Tabela 6.3: Percentagem de falhas detectadas em cada abordagem

Severidade (%)	**Instante da falha (min)**	**Q**	**SPE (lambda 0.25)**	**Sdist (estático)**	**Sdist (dinâmico)**
2	301			100 [A]	100
2	451			0 [B]	100
2	550			100 [C]	100
2	700			31,25 [D]	18,75
5	301	75,00	81,25	100 [A]	100
5	451	37,75	68,75	62,5 [B]	100
5	550			100 [C]	100
5	700	31,25	37,50	93,75 [D]	93,75
10	301	87,50	87,50	100 [A]	100
10	451	75,00	93,75	93,75 [B]	100
10	550			100 [C]	100
10	700	62,50	93,75	100 [D]	93,75
20	301	100	100	100 [A]	100
20	451	100	100	100 [B]	100
20	550			100 [C]	100
20	700	100	100	100 [D]	100

Tabela 6.4: Resumo das seqüências empregadas para Sdist estático

	Fase A (instante inicial / final) (minutos)	**Fase B (instante inicial / final) (minutos)**	**Falha (instante da ocorrência) (minutos)**
Seqüência A	0 / 250	251 / 501	301
Seqüência B	0 / 250	251 / 501	451
Seqüência C	0 / 250	500 / 750	550
Seqüência D	0 / 250	500 / 750	700

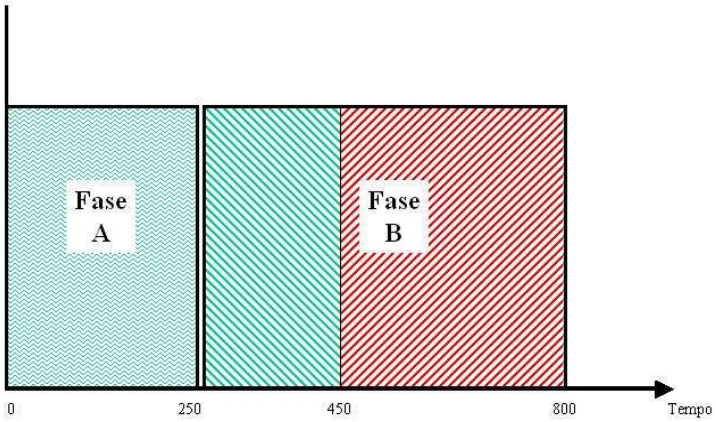


Figura 6.7: Seqüência com falha a partir do instante 450

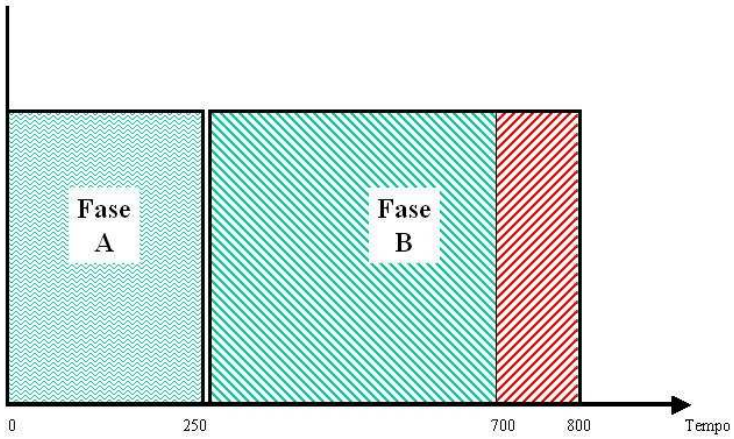


Figura 6.8: Seqüência com falha a partir do instante 700

A Figura 6.9 resume a influência do instante de ocorrência de falha sobre o grau de detecção da ferramenta: anomalias localizadas na parcela final da fase “B” apresentam mais difícil detecção. É possível entender essa diferença de desempenho sabendo que o peso dos pontos com falha no cálculo do erro total, via SPE ou Q, fica reduzido nas situações que a janela de falha é pequena. Ou seja,

nestes casos a janela temporal com falha apresenta influência reduzida sobre o valor total final do erro.

A eficiência dos métodos baseados em SPE e Q não é afetada pelo afastamento temporal entre o final da fase “A” e o início da anomalia, mas sim pela razão entre o instante de falha e o instante final da fase “B”. Quando tal razão é baixa, ou seja, a falha ocorre na parcela inicial da fase “B”, a eficiência de detecção torna-se elevada. Porém para razões elevadas, próximas do valor unitário, o desempenho do método fica bastante prejudicado.

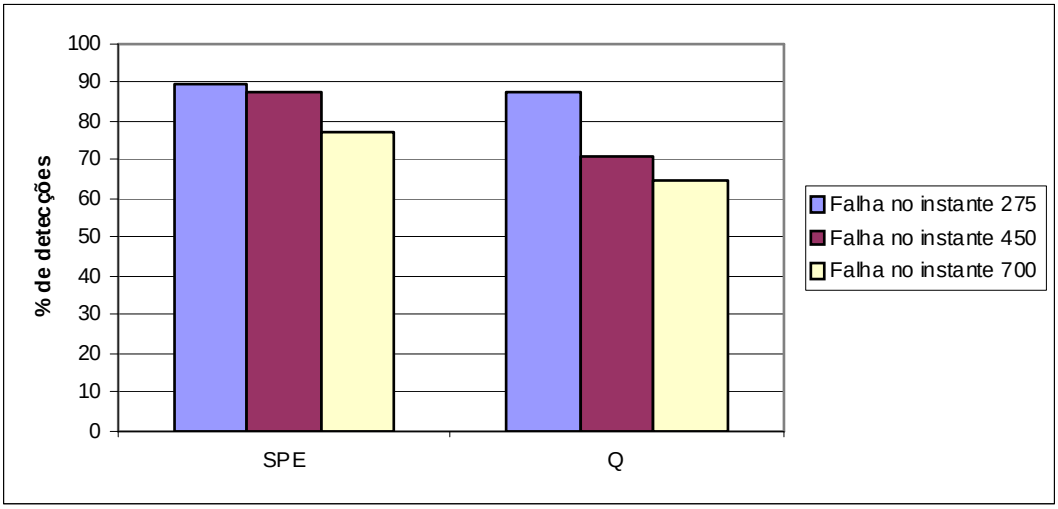


Figura 6.9: Influência do instante de ocorrência da falha sobre o grau de detecção alcançado via SPE e Q.

Quando se combina a análise de severidade da falha com o afastamento em relação ao final da fase “A”, torna-se possível concluir que o afastamento prejudica especialmente a detecção das falhas de pequena significância, conforme ilustrado na Figura 6.10.

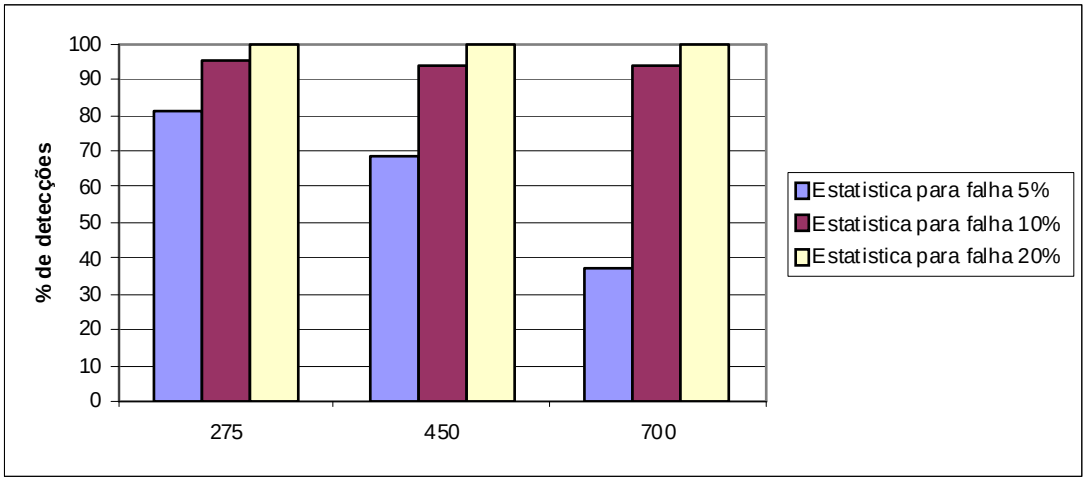


Figura 6.10: Influência da severidade de falha e localização temporal no grau de detecção. Resultado baseado apenas no uso da ferramenta SPE.

A Tabela 6.3 demonstra que falhas de maior severidade são mais facilmente detectadas. Esse comportamento, além de esperado, é comum a todas as ferramentas de detecção de falhas relatadas na literatura. Falhas de grande severidade, que apresentam desvio de 20% ou superior, foram detectadas usando SPE ou Q independente da disposição temporal. Para anomalias de baixa severidade, com 5% de desvio no valor medido, o grau de detecção variou entre 30% e 80%, dependendo do instante de ocorrência da falha (Tabela 6.3). A Figura 6.11 ilustra a relação entre severidade de falha e o grau de detecção alcançado (assumindo falhas começando no instante 301).

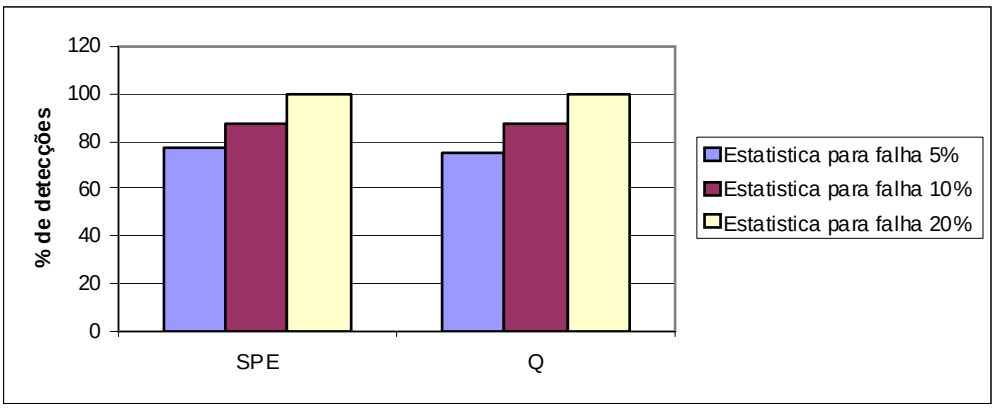


Figura 6.11: Influência da severidade na detecção de falhas via SPE e Q

Além da severidade da falha e do instante temporal no qual esta ocorre, é possível que fatores como tipo de sensor (pressão ou vazão) e o seu

posicionamento ao longo da dutovia também alterem a capacidade detectora. Os resultados de detecção por sensor estão representados na Figura 6.12. A tarefa mostrou-se especialmente difícil nos sensores de 3 e 10, ambos do tipo “sensor de pressão” e localizados nos extremos da rede de dutos (Figura 6.1 e Tabela 6.1). As técnicas aqui avaliadas, SPE e Q, sofreram influência análoga em decorrência do tipo e localização dos sensores.

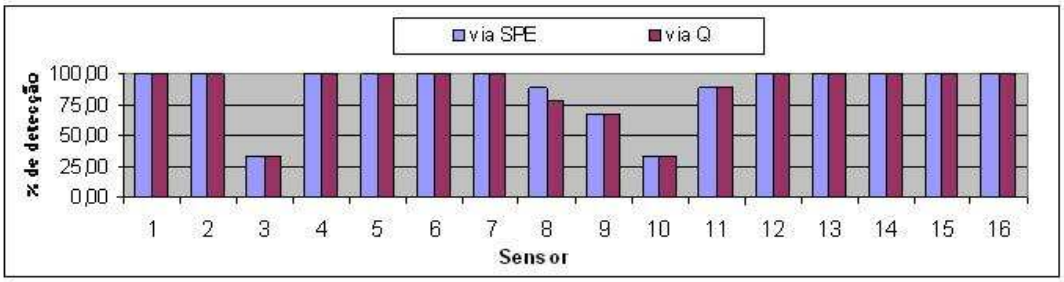


Figura 6.12: Grau de detecção de falha por sensor

Possível justificativa para os diferentes graus de acerto observados por sensor seria a ordem de grandeza das medidas geradas. Ou seja, sensores cujas variáveis medidas têm valores elevados teriam suas falhas mais facilmente detectadas do que aqueles cujas medidas apresentassem valores reduzidos. A Tabela 6.5 abaixo apresenta o valor médio gerado por sensor. Ao comparar o grau de acerto do detector e os valores médios das medidas dos sensores, não existe correlação aparente entre as duas informações. Sendo assim possível atribuir à localização do sensor na rede a maior ou menor facilidade na detecção de sua falha.

Tabela 6.5: Valor médio da série temporal gerada por cada sensor sem falha

Sensor	Média do valor medido	Posição no vetor	Média do valor medido
1	0,2808	9	21,9915
2	0,2806	10	14,9924
3	51,7380	11	0,2807
4	48,8086	12	0,2809
5	47,8110	13	0,2807
6	42,7310	14	0,2810
7	24,5808	15	0,2810
8	22,8652	16	0,2805

6.1.1 b) Via Sdist e Spca

Conforme descrito anteriormente, **SINGHAL e SEBORG [2002]** propõem o uso de três métricas para estabelecer similaridade entre modelos PCA: Sdist, Spca e

SF. Através dos resultados aqui obtidos observou-se que a métrica Spca não se mostra sensível às modalidades de anomalia implementadas. O parâmetro Sdist, por outro lado, responde fortemente aos desvios operacionais, conforme ilustrado na Figura 6.13. Uma vez que a métrica Spca não respondeu adequadamente, torna-se também dispensável o uso de SF, por este tratar-se apenas da soma ponderada de Sdist e Spca.

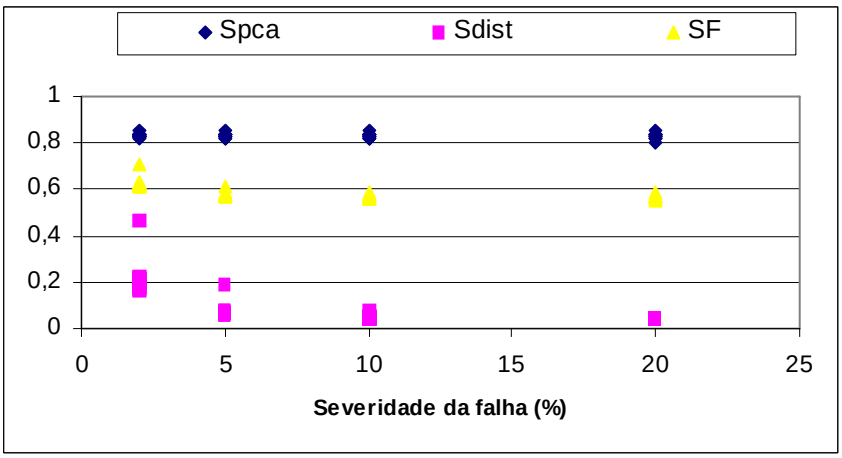


Figura 6.13: Resposta das métricas Spca, Sdist e SF a diferentes severidades de falha.

Ao contrário da abordagem por Q e SPE, as ferramentas desenvolvidas a partir das metodologias Sdist, Spca e SF baseiam-se no ajuste de dois modelos PCA distintos, um para cada fase. Promove-se então a comparação destes modelos. Caso sejam similares tem-se que o comportamento das duas fases também é semelhante. Como a fase “A” representa a operação normal do sistema, sem falha, uma fase “B” similar indica baixa probabilidade de falha. Por outro lado, modelos altamente distintos revelam comportamento das duas séries temporais também distintos, sendo indicativo de operação anormal no decorrer da fase “B”.

Deste modo, na ausência de anomalias em ambas as fases, os dois modelos PCA deveriam ser perfeitamente similares, gerando valor unitário para as métricas Sdist, Spca e SF. Essa idealidade porém não ocorre, pois os dados nos quais os modelos são baseados apresentam diversos ruídos e pequenos desvios normais a operação industrial. Por essa razão, inicialmente foram efetuados diferentes experimentos para medir o nível de similaridade alcançado na ausência de falhas. Para esse estudo, algumas disposições temporais da fase “B” foram testadas. A

Figura 6.14 ilustra o procedimento. A Tabela 6.6 resume os instantes iniciais e finais das fases “B” testadas.

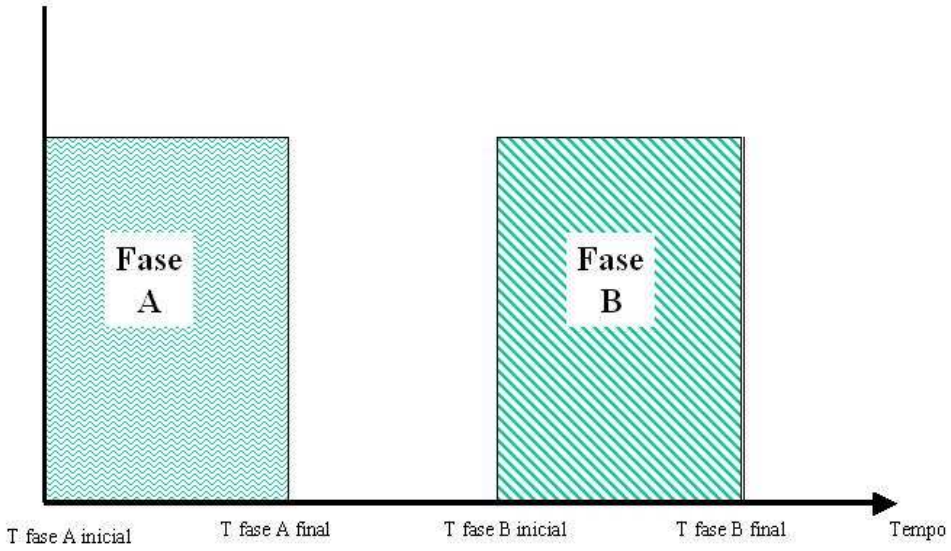


Figura 6.14: Posicionamento temporal das fases “A” e “B”

Tabela 6.6: Faixas temporais das fases B testadas

Início da fase B	Final da fase B
251	501
275	525
300	550
(...)	(...)
660	910
675	925
690	940

Por questões de similaridade, ambas as fases apresentam a mesma extensão temporal em todos os cenários, variando apenas o afastamento entre elas. Valores de Sdist elevados, variando entre 0.8 e 0.7 foram encontrados para diferentes instantes iniciais da fase B, conforme Figura 6.15. Observa-se porém que em situações nas quais o afastamento entre as fases supera 250 instantes de tempo, ou seja: a fase B inicia-se após o instante 500, a similaridade decresce fortemente mesmo na ausência de anomalia operacional.

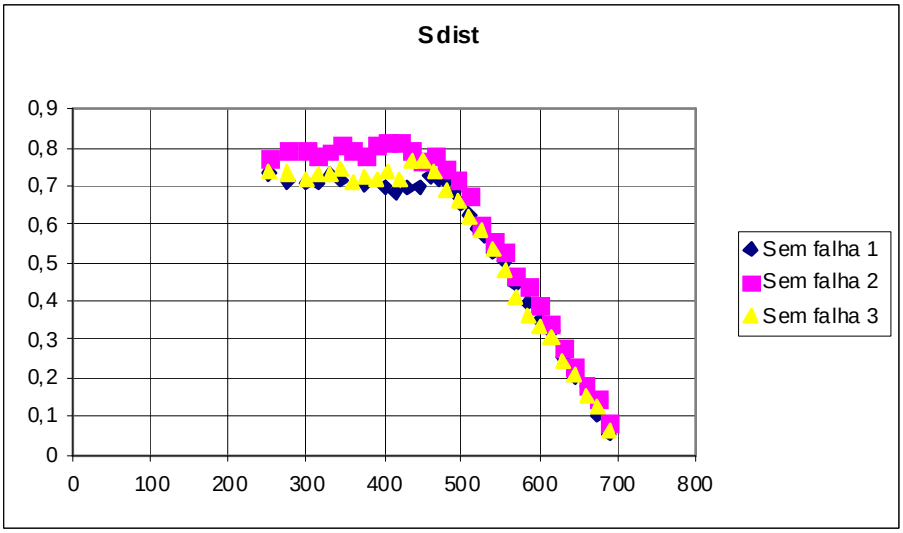


Figura 6.15: Influência do instante inicial da fase B na métrica Sdist

A partir do comportamento exibido pela métrica Sdist, adotam-se os seguintes critérios: a) a fase "B" pode iniciar-se entre os instantes 250 e 500; b) valores de Sdist inferiores a 0.6 revelam elevado indício de falha.

Uma vez estabelecida faixa confiável para valores de similaridade, aplica-se a "abordagem estática" descrita anteriormente. Quatro seqüências de experimentos (Tabela 6.4) envolvendo falhas de sensores com severidades de 5%, 10% e 20% na Fase "B" são implementadas. A duração de cada uma das fases e o instante de ocorrência da falha estão apresentados na Tabela 6.4, enquanto a severidade da falha é descrita na Tabela 6.3. Ressalta-se que em virtude do bom desempenho apresentado pela métrica Sdist, optou-se por incorporar também falhas da ordem de 2%.

Para a "seqüência A" obtem-se êxito em todos os testes implementados (Tabela 6.3). A Figura 6.16 ilustra a disposição temporal das fases de treino e uso, assinalando também a janela de falha de sensor. Mesmo falhas de baixa severidade são detectadas por essa abordagem, ou seja, resultam em valores reduzidos de similaridade via Sdist (Figura 6.17). A aproximação por Sdist mostra-se superior a abordagem por SPE e Q.

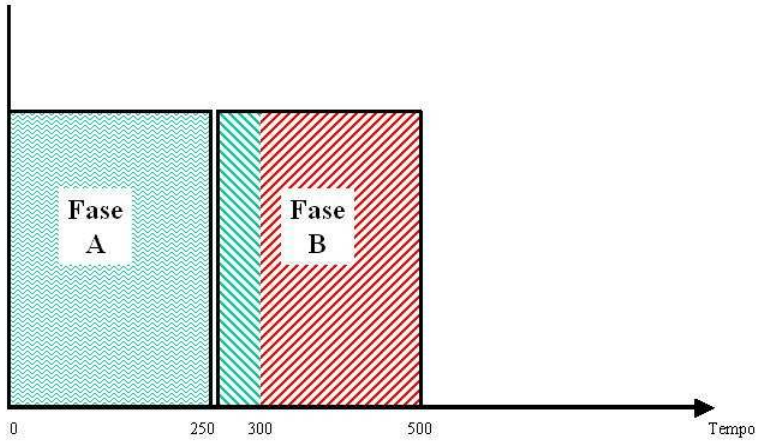


Figura 6.16: Seqüência de teste “A”, com falha a partir do instante temporal 300.

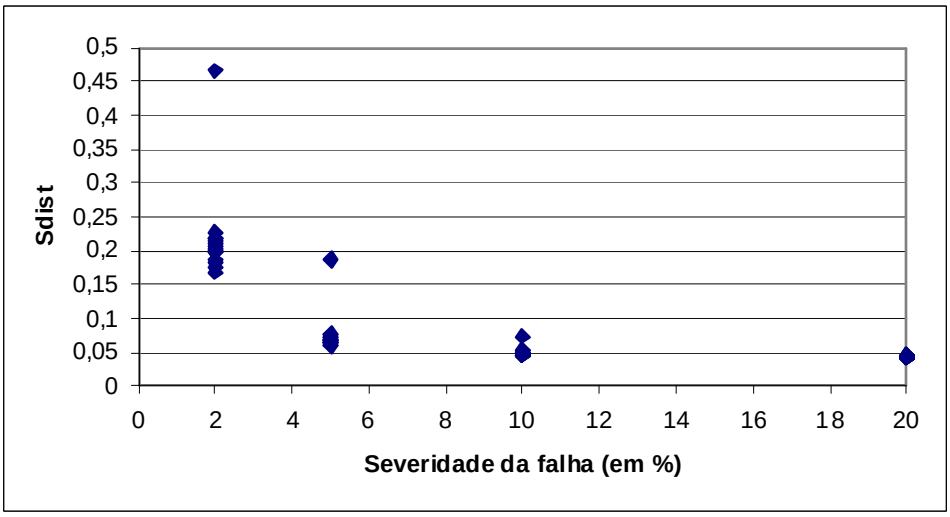


Figura 6.17: Resposta de Sdist a diferentes severidades de falha para a seqüência “A”.

Com a finalidade de aferir a influência do instante de ocorrência da falha no desempenho da ferramenta de detecção, aplica-se a “seqüência B”. Nesta utiliza-se a mesma disposição temporal para as fases “A” e “B”, modificando apenas o instante de ocorrência da anomalia. Enquanto na “seqüência A” a falha originava-se a partir do instante 301, para a seqüência “B” a anomalia inicia-se apenas no instante 451 (Figura 6.18). Ao contrário do caso anterior, nesta nova condição algumas das falhas testadas não puderam ser detectadas, tendo produzido similaridades superiores a 0.6 (Figura 6.19). Ou seja, ao comparar os testes efetuados com falhas nos instantes 301 (seqüência A) e 451 (seqüência B) constata-se menor índice de detecção no segundo cenário. Duas explicações tornam-se possíveis: (i) a posição relativa da

falha dentro da fase “B” afeta o desempenho do método, ou (ii) a distância entre o término da fase “A” e o instante que ocorre a falha é responsável pela mudança na capacidade detectora. Ambas as hipóteses são então investigadas.

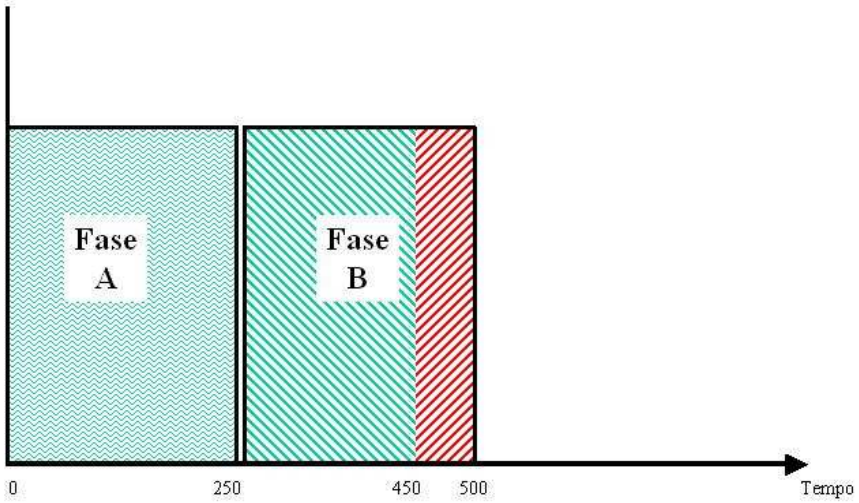


Figura 6.18: Seqüência de teste “B”, com falha a partir do instante temporal 450.

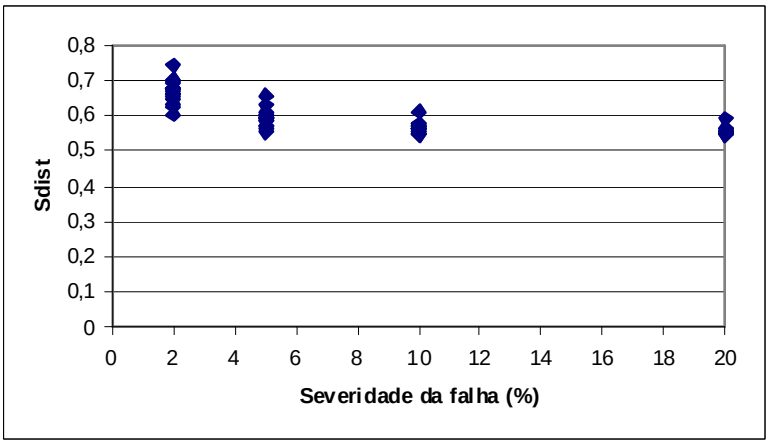


Figura 6.19: Resposta de Sdist a diferentes severidades de falha para a seqüência “B”

Enquanto nas seqüências “A” e “B” o posicionamento das fases de treino e de uso era idêntico, apenas variando o momento de início da falha, na seqüência “C” a fase de uso somente se inicia a partir do instante 500 (Tabela 6.4). Ou seja, existe um afastamento entre o final da fase de treino e o início da fase “B” (Figura 6.20). A falha de sensor foi configurada de modo a ser iniciada a partir do instante 550. Esse

posicionamento temporal da falha implica no distanciamento entre a sua ocorrência e o final da fase “A”, condição similar a imposta pela seqüência “B”. Porém, graças ao deslocamento da fase de uso, a anomalia passa a ocorrer no início desta, em uma condição neste aspecto similar a testada na seqüência “A”. Todas as falhas testadas na seqüência “C” são detectadas (Figura 6.21). Assim, é possível supor que a posição relativa da falha dentro da fase “B” exerce influência mais importante sobre a capacidade detectora do método do que a distância absoluta entre o final da fase “A” e a falha. A seqüência “D” é usada para comprovar essa hipótese.

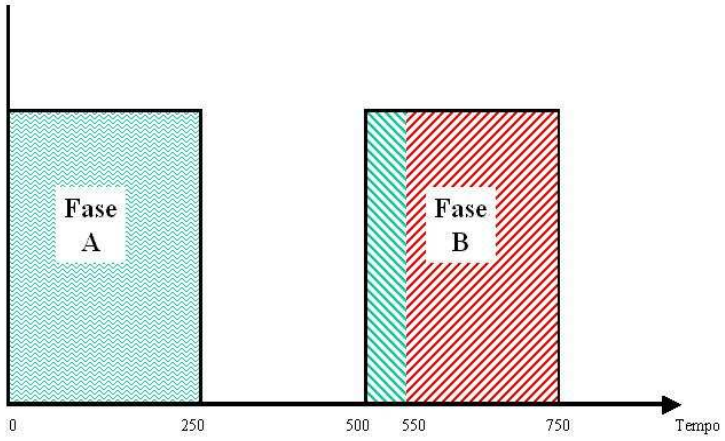


Figura 6.20: Seqüência de teste “C”, com falha a partir do instante temporal 550.

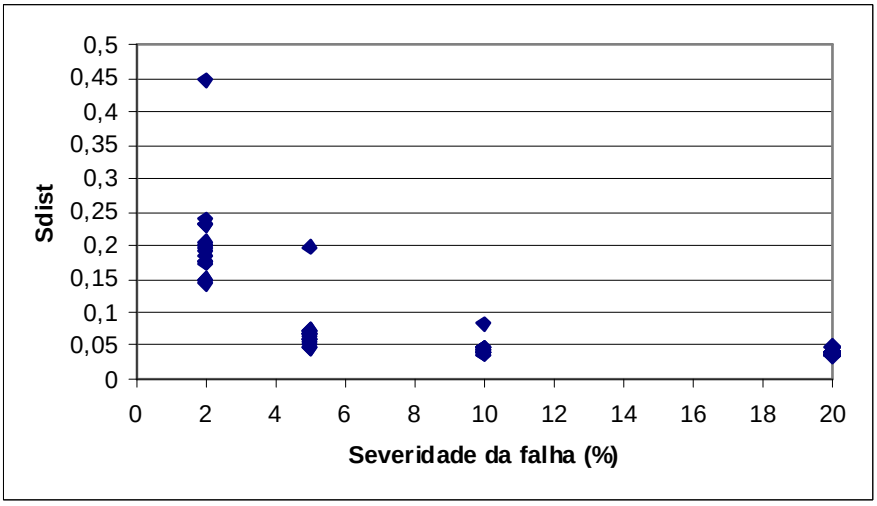


Figura 6.21: Resposta de Sdist a diferentes severidades de falha para a seqüência “C”

Na “seqüência D” utiliza-se o mesmo posicionamento das fases “A” e “B”, porém modificando o início da janela de falha para o instante 700 (Figura 6.22).

Deste modo a anomalia passa a situar-se na parte final da fase de uso. Um pior desempenho de detecção foi assim obtido, conforme Figura 6.23. Tais resultados mostraram-se semelhantes aos encontrados para a “seqüência B”. Mais uma vez os resultados sugerem que o desempenho de detecção parece estar mais fortemente relacionado à posição relativa da falha na fase “B” do que ao seu afastamento em relação ao final da fase “A”. Este resultado é análogo ao obtido para a ferramenta detectora baseada em SPE e Q.

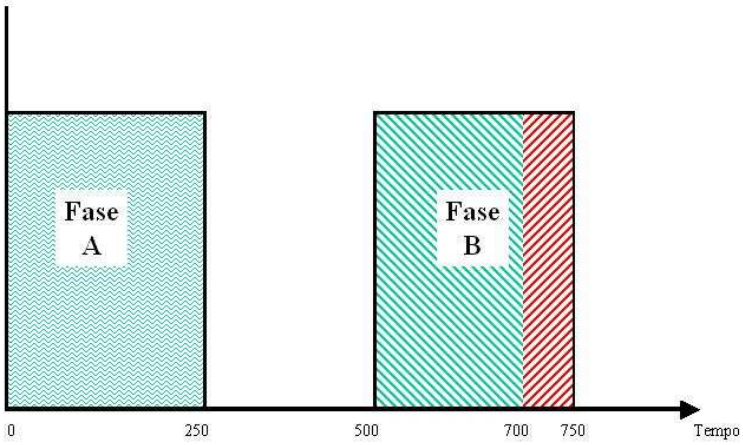


Figura 6.22: Seqüência de teste “D”, com falha a partir do instante temporal 700.

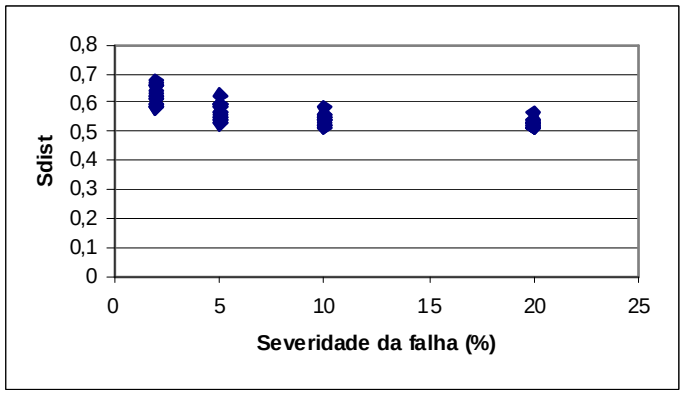


Figura 6.23: Resposta de Sdist a diferentes severidades de falha para a seqüência “D”

Uma vez avaliado o desempenho da abordagem estática, parte-se para a aproximação “dinâmica”, ou seja, com fase “B” móvel. Algumas condições particulares são necessárias ao correto funcionamento desta metodologia. Por basear-se no índice de similaridade Sdist, ambas as fases a serem comparadas

(fase “A” e “B”) devem, preferencialmente, apresentar mesma extensão temporal. Nos cenários aqui testados as fases foram reguladas para apresentarem uma extensão de 250 instantes de tempo cada. Outra condição imposta é o espaçamento máximo entre o término da fase “A” e o início de uma fase “B”, tendo este também sido regulado em 250 instantes de tempo. Ou seja, para cada 250 instantes de treino, a ferramenta de detecção pode ser utilizada por até 500 instantes de uso. Assumiu-se que valores de Sdist inferiores a 0.6 são considerados indicadores de falha. Tais condições são análogas aquelas empregadas na abordagem estática. Nesse trabalho adotou-se “passo unitário” para a movimentação da janela de uso, ou seja, entre uma fase “B” e a seguinte, a janela temporal é deslocada de um instante de tempo.

Com a finalidade de comparar a abordagem dinâmica com a estática, repetiram-se as posições temporais das anomalias testadas nas seqüências “A”, “B”, “C” e “D” (Tabela 6.4). Em relação às seqüências “A”, “B” e “C” a abordagem dinâmica se mostrou totalmente satisfatória, com um índice de detecção de 100% para todas as severidades testadas (Tabela 6.3), sendo deste modo superior as técnicas via Q, SPE e Sdist estático. Porém, quando a falha é implementada apenas no instante 700, esta abordagem foi incapaz de detectar inúmeras anomalias de menor severidade

Explica-se a maior capacidade detectora exibida pela abordagem dinâmica através da “habilidade” que tal metodologia possui de posicionar as falhas ocorridas nos instantes 301, 451 e 550 no início de alguma das diversas fases de uso. Essa capacidade eleva de forma significativa o grau de detecção da ferramenta, pois modifica a razão entre o instante de falha e o instante final da fase “B”, tornando-a mais adequada. À medida que a fase móvel desloca-se, a razão diminui, favorecendo a detecção. Esta característica torna a abordagem por “Sdist dinâmica” a mais eficaz dentre as aqui testadas.

6.1.2 Detecção por classificação hierárquica

A abordagem de detecção de falhas via análise de componentes principais consegue atingir elevado grau de eficiência, especialmente através da metodologia

Sdist dinâmica. Essa metodologia porém é incapaz de apontar com precisão o instante inicial da falha. Com a finalidade de superar tal limitação optou-se por empregar ferramenta de classificação hierárquica. O classificador separa a série de observações temporais em dois grandes conjuntos: “com” e “sem” falha.

Baseando-se nas condições da “seqüência B” da Tabela 6.4, porém sem a divisão em fases, utilizou-se conjunto de 500 observações, das quais 450 representam situação de operação normal, e as 50 restantes retratam anormalidade. Conforme Figura 6.24, para falhas com severidades de 10% ou 20%, mais de 90% das observações foram corretamente classificadas através da métrica “Minkowski r4”.

Na busca pela métrica mais adequada, além de verificar a eficiência das distâncias “City Block” e “Euclidiana”, foram testados ainda dois valores de “r” para a métrica “Minkowski”. Valores de “r” igual a 3 e 4 foram então avaliados. Constatou-se desempenho superior ou equivalente da métrica Minkowski com r igual a 4 para todas as severidades de falha implementadas.

A correta classificação hierárquica das observações permite localizar com elevado grau de precisão o instante inicial de falha. Porém, o grau de acerto dessa técnica é baixo quando lidando com falhas de menor severidade. Por essa razão, seu uso como ferramenta de detecção de anomalias somente se justifica quando aliado a outros métodos, como Sdist dinâmico. Nesse caso, o desenvolvimento de uma ferramenta híbrida permite obter melhor resultado final.

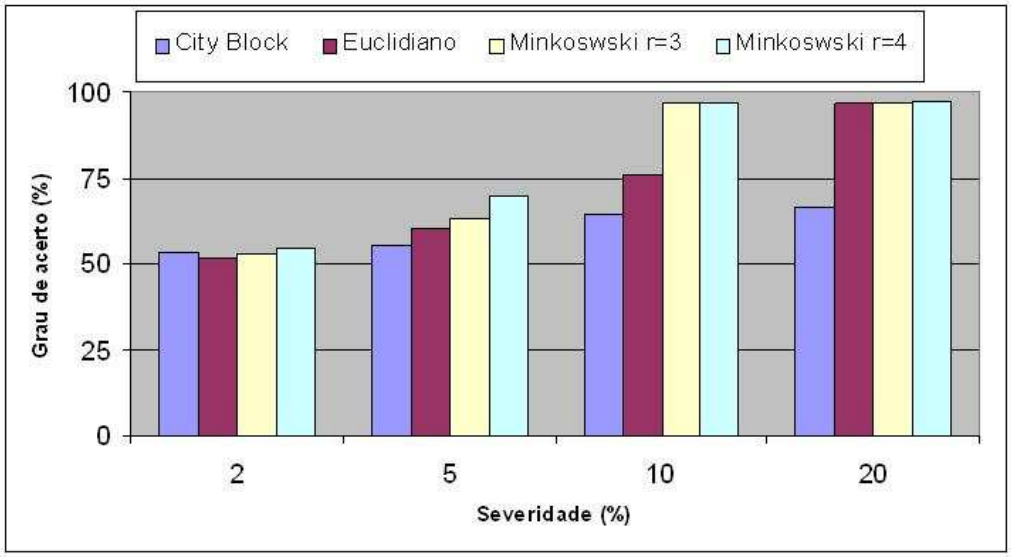


Figura 6.24: Grau de acerto do classificador hierárquico para falhas de sensor

6.2 Identificação do sensor em falha

Uma vez detectada a ocorrência de uma falha de sensor iniciam-se os procedimentos para a sua correta localização. Tão importante quanto detectar é localizar com alguma precisão a origem desse problema. A identificação contribui para minimizar o tempo de reparo, especialmente quando se lida com uma dutovia extensa, com trechos de difícil acesso físico.

Com essa finalidade foram abordadas duas metodologias distintas: “índice de validade do sensor” (SVI) e “método dos protótipos”. Os resultados destas aproximações são analisados neste Capítulo.

6.2.1 Identificação do sensor em falha via SVI-D

A metodologia aplicada na identificação do sensor em falha utiliza o “índice de validade dos sensores” – SVI-D, descrito por **DUNIA *et al* [1996a].** O seu cálculo é baseado na quantificação da participação de cada sensor no erro total de predição. Quanto maior a participação de um sensor no erro, menor a confiabilidade deste instrumento. Para uma falha no “sensor 1” (FI20-1) por exemplo,obtém-se os valores de SVI-D apresentados na Tabela 6.7. O menor valor de SVI-D foi obtido para o “sensor 1”, o que é coerente com o fato deste estar em falha.

Aplicando-se esta metodologia para falhas ocorridas no instante 300 minutos, tem-se que, para 51% dos casos analisados, a origem da falha foi detectada imediatamente através do primeiro candidato. Ou seja, o sensor apontado como de menor confiabilidade realmente estava em falha. Considerando-se os dois sensores de menor confiabilidade o grau de acerto sobe para 81%. Resultados semelhantes são obtidos para falhas implementadas nos instantes 450 e 700 minutos.

Conforme esperado, a severidade influencia diretamente a precisão na identificação da origem do desvio. Anomalias que provocam desvios pequenos são mais difíceis de terem sua origem determinada, como apresentado na Figura 6.25.

Tabela 6.7: Resultados de SVI para falha no sensor 1

Código do Sensor	Número do sensor	SVI-D
FI 20-1	**1**	**0,75258**
FI 20-9	2	0,98808
PI 20-1	3	0,99611
PI 20-2	4	0,98195
PI 20-3	5	0,97706
PI 20-4	6	0,78474
PI 20-6	7	0,99885
PI 20-7	8	0,92613
PI 20-8	9	0,82938
PI 20-9	10	0,99338
FI 10-1	11	0,99718
FI 10-2	12	0,99242
FI 10-4	13	0,87699
FI 10-5	14	0,96037
FI 10-6	15	0,96015
FI 10-7	16	0,98463

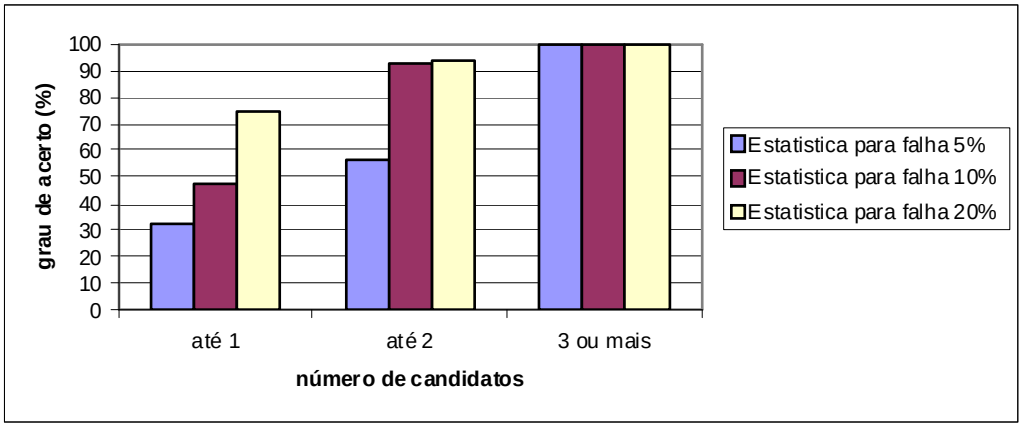


Figura 6.25: Relação entre a severidade da falha e a identificação de sua origem

6.2.2 Identificação do sensor em falha via protótipos

O método dos protótipos é outra abordagem possível para a localização do sensor em falha. Para implementar tal abordagem foram selecionados 22 casos representativos de anomalias, sendo dezesseis anomalias do tipo “falha em sensor” – uma por sensor; e seis situações de vazamento – uma por tubo. Esses 22 casos são aqui denominados “protótipos”. Nos protótipos a anomalia ocorre a partir do instante 300, e tem severidade de 5% no caso de falhas em sensores e diâmetro de furo de 0.03 metros para os protótipos representativos de situações de vazamento. Uma vez dispondo dos protótipos e uma falha operacional tendo sido detectada, esta falha é comparada com cada um dos vinte e dois padrões, buscando-se a similaridade mais elevada. A comparação é realizada usando a métrica Spca.

Como exemplo ilustrativo, considera-se a situação na qual tem-se a ocorrência de uma falha de 2% no sinal do sensor 16. A Tabela 6.8 exibe as medidas de similaridade entre os dados experimentais com falha e cada um dos protótipos. Os dados apresentam maior similaridade com o protótipo “sen300S16F05”, sendo este o protótipo característico de falhas no sensor 16. Deste modo, para a situação ilustrada, a abordagem foi eficaz na detecção da origem da falha.

Tabela 6.8: Similaridade entre os dados experimentais com falha e cada protótipo característico

Protótipo	Spca	Protótipo	Spca
vaz300T1B03	0.83065	sen300S6F05	0.83543
vaz300T2B03	0.8134	sen300S7F05	0.82248
vaz300T4B03	0.77818	sen300S8F05	0.85526
vaz300T5B03	0.74073	sen300S9F05	0.83108
vaz300T6B03	0.74975	sen300S10F05	0.82097
vaz300T7B03	0.82859	sen300S11F05	0.81994
sen300S1F05	0.84492	sen300S12F05	0.82357
sen300S2F05	0.82996	sen300S13F05	0.80933
sen300S3F05	0.80965	sen300S14F05	0.83324
sen300S4F05	0.82417	sen300S15F05	0.83881
sen300S5F05	0.8462	**sen300S16F05**	**0.85323**

Porém, repetindo o procedimento acima para falhas de 1% e 2% de severidade no sinal dos demais sensores, a eficiência do método foi muito baixa. Essa metodologia foi capaz de indicar com sucesso a origem de uma falha de sensor em menos de 7% das situações analisadas, desempenho significativamente inferior ao obtido através da abordagem SVI-D.

6.3 Quantificação de falhas em sensores

Além de poderem ser usados com sucesso em uma abordagem de detecção de falhas, os classificadores hierárquicos permitem ainda estimar a severidade da falha a partir do afastamento existente entre os centros dos dois agrupamentos formados. Ou seja, a dissimilaridade observada entre os *clusters* “com” e “sem” falha pode ser correlacionada com a severidade da anomalia ocorrida. A Figura 6.26 retrata a relação entre amplitude de falha e afastamento entre grupos. Enquanto para severidades baixas, tais como 2% e 5%, a distância média entre os agrupamentos é menor que 5, para falhas severas o afastamento atingiu valores até quatro vezes maiores. Devido ao maior êxito observado na etapa de detecção através do uso da métrica Minkowski r4, apenas essa distância foi empregada neste procedimento de quantificação.

Além de mostrar-se função da severidade da falha, a dissimilaridade aferida entre os dois *clusters* mostrou-se função também do sensor em falha. Na Figura 6.26 é retratado o comportamento da distância para cada sensor em função da severidade.

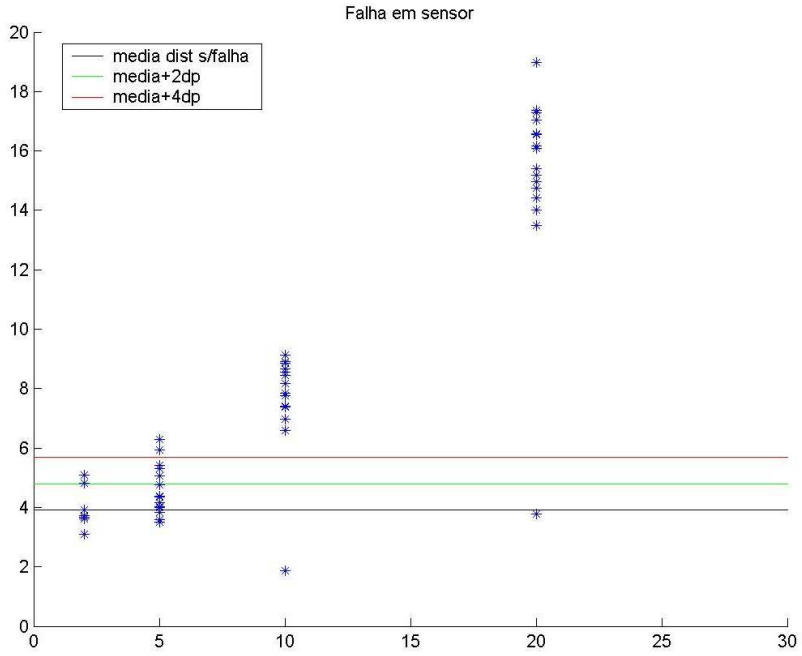


Figura 6.26: Relação entre a distância dos agrupamentos e a severidade da falha

O conhecimento prévio de qual dos sensores apresenta-se em falha contribui para elevar a precisão da estimativa de sua severidade, sendo possível encontrar relações lineares entre a distância entre *clusters* e a severidade da falha. Cada sensor apresenta sua própria relação característica.

6.4 Detecção de vazamento na rede de dutos

Esta seção descreve o desempenho das ferramentas quanto à detecção de vazamentos. Os resultados apresentados neste Capítulo referem-se a duas metodologias: baseada em PCA e classificadores hierárquicos **[VAZ JUNIOR *et al.*, 2009]**. Modelos auto-regressivos também são empregados com o objetivo de detectar vazamentos, porém seu desempenho é avaliado no Capítulo 7 **[VAZ JUNIOR *et al.*, 2010]**.

6.4.1 Detecção por análise dos componentes principais (PCA)

De modo similar ao observado para falhas em sensores, diferentes abordagens baseadas em PCA também podem ser empregadas na detecção de vazamentos. Nos cenários de maior severidade, a representação com apenas dois

componentes principais é suficiente para promover separação visual adequada, resultado então em dois grandes agrupamentos: “com vazamento” e “sem vazamento”. O espaçamento entre os *clusters* também se mostra função da severidade do problema.

6.4.1 a) Via Q e SPE

A série temporal gerada a partir do simulador de redes de dutos é dividida em duas fases, denominadas “A” e “B” – Figura 6.6. Assume-se a “fase A”, ou “fase de treino”, como tendo duração de 250 instantes de tempo, estendendo-se do instante inicial ao instante 250. Por outro lado, a “fase B”, ou “fase de uso”, inicia-se no instante 251, prosseguindo até o instante 800. Diferentes cenários de vazamento foram implementados, variando o instante inicial (301, 450, 550 e 700), a severidade (0.01, 0.02 e 0.05 metros) e a localização. A Tabela 6.9 apresenta uma descrição detalhada.

Tabela 6.9: Percentagem de vazamentos detectados em cada abordagem

Diâmetro de furo (m)	Instante do vazamento (min)	Q	SPE (lambda 0,25)	Sdist (estático)	Sdist (dinâmico)
0,01	301	16,66%	66,66%	100% [A]	100%
0,01	451	0%	16,66%	66,67% [B]	100%
0,01	550	0%	16,66%	100% [C]	100%
0,01	700	0%	0%	50,00% [D]	16,66%
0,02	301	100%	100%	100% [A]	100%
0,02	451	50%	100%	100% [B]	100%
0,02	550	66,66%	83,33%	100% [C]	100%
0,02	700	83,33%	83,33%	83,33% [D]	83,33%
0,05	301	100%	100%	100% [A]	100%
0,05	451	100%	100%	100% [B]	100%
0,05	550	100%	100%	100% [C]	100%
0,05	700	100%	100%	100% [D]	100%

Nota: o índice sobrescrito corresponde ao código de seqüências da Tabela 6.4.

Do mesmo modo que no desenvolvimento de ferramenta detectora de falhas em sensores, diferentes valores de λ são testados, de modo a aferir a influência do peso dado a instantes passados na detecção de vazamentos. Novamente conclui-se que λ’s de valor reduzido contribuem para um maior nível de detecção de anomalias, conforme apresentado na Figura 6.27. Ou seja, elevar a influência do passado sobre

o valor presente de SPE incrementa a capacidade de detecção tanto de falhas em sensores quanto de vazamentos. O valor de trabalho para o parâmetro λ foi fixado em 0.25.

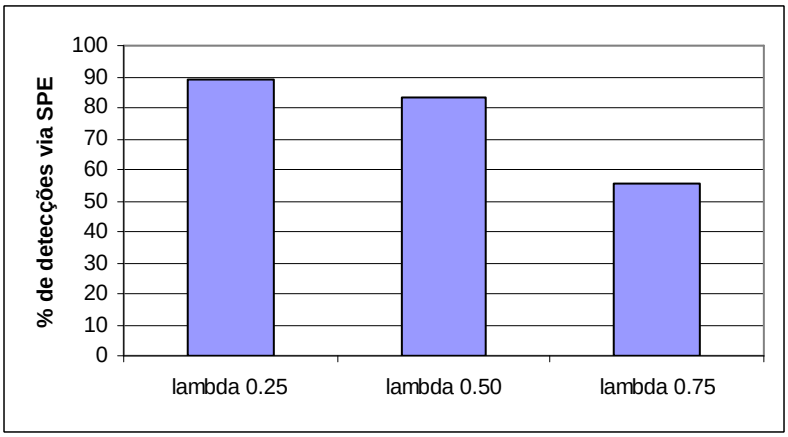


Figura 6.27: Desempenho de detecção de vazamentos em função do valor de λ empregado

Ressalta-se que o desempenho da ferramenta de detecção de vazamentos baseada na métrica SPE com λ em 0,25, superou aquele alcançado empregando Q (Tabela 6.9).

Os resultados encontrados confirmam ainda que o posicionamento temporal do vazamento na fase “B” exerce forte influência sobre o grau de detecção (Tabela 6.9). Para um diâmetro de furo de 2 centímetros, o método foi capaz de detectar vazamentos em todos os dutos da rede quando o instante de falha utilizado foi 301 ou 451, ou seja, situações nas quais a anomalia ocorre na fração inicial da fase “B”. Por outro lado, quando o vazamento ocorreu apenas a partir do instante 550 ou 700, para o mesmo diâmetro de vazamento, o desempenho da ferramenta foi significativamente inferior. A Figura 6.28 exibe o comportamento encontrado. É possível concluir que, de forma análoga as falhas em sensores, vazamentos com maior diâmetro de furo e na parcela inicial da fase de uso são mais facilmente detectados.

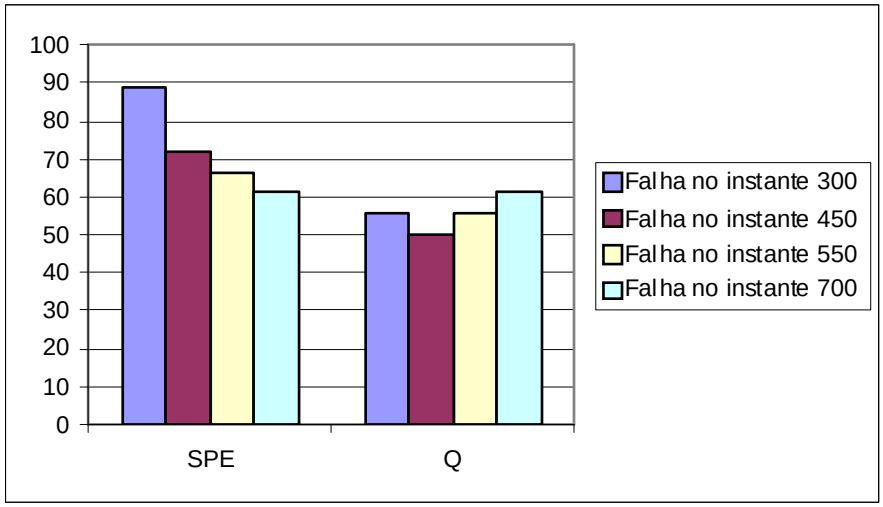


Figura 6.28: Influência do instante da falha na detecção de vazamentos

Adicionalmente, investigou-se o efeito do posicionamento espacial do furo na possibilidade deste ser detectado. Ou seja, vazamentos em determinados tubos podem ser mais ou menos fáceis de serem percebidos. Como havia sido observado durante o estudo de falhas em sensores, anomalias ocorridas na parte central da dutovia são de mais fácil detecção, quando comparadas àquelas ocorridas em pontos extremos da rede. A Figura 6.29 confirma esse desempenho superior do método de detecção para tubos centrais. A mesma figura também confirma a superioridade da abordagem por SPE em relação a aproximação via Q.

Investiga-se também uma possível influência do diâmetro e/ou comprimento do tubo no desempenho da ferramenta. Nenhuma relação direta foi encontrada (Figura 6.30).

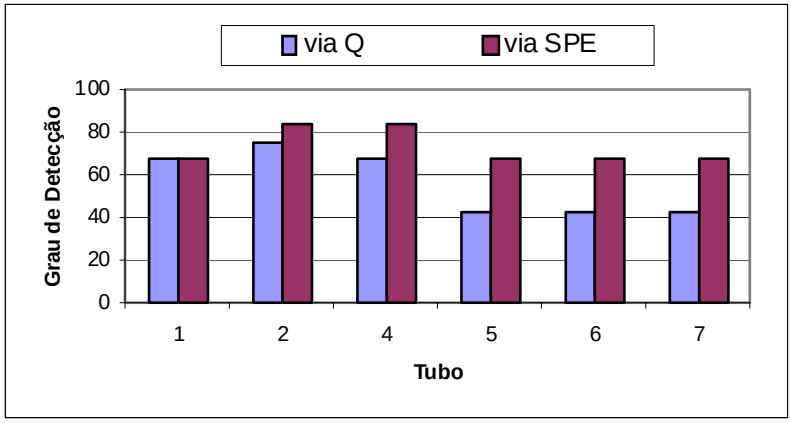


Figura 6.29: Influência do tubo com vazamento no sucesso da ferramenta detectora de vazamentos

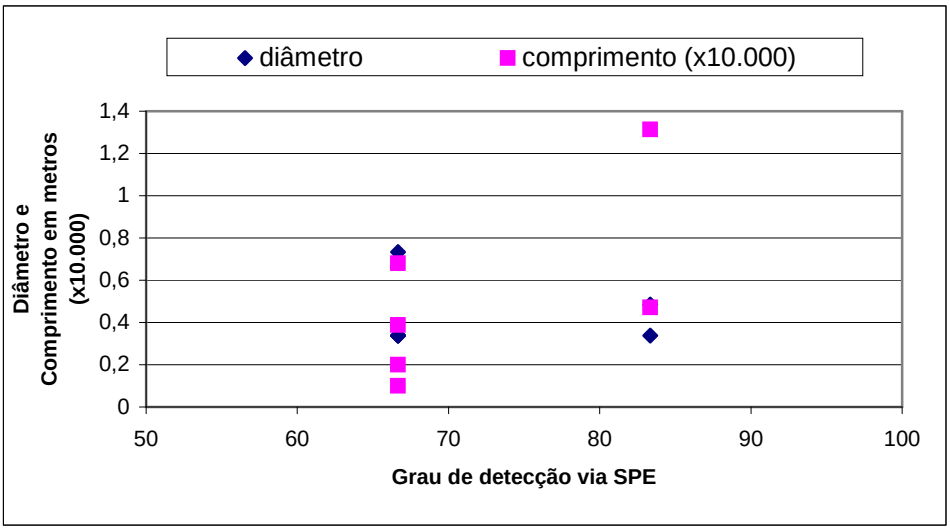


Figura 6.30: Relação entre o grau de detecção via SPE e o diâmetro e comprimento do tubo com falha

6.4.1 b) Via Spca e Sdist

Para os cenários de vazamentos testados a métrica Spca foi considerada inadequada para a sua detecção. Opta-se assim pelo emprego da métrica Sdist.

A aplicação do parâmetro Sdist se deu através das duas abordagens já conhecidas, “estática” e “dinâmica”. Na abordagem estática foram propostas quatro disposições temporais para as fases de uso e de vazamento. Na seqüência “A” posicionou-se a fase de uso entre os instantes de tempo 251 e 501, com vazamento ocorrendo a partir do instante 301. Todos os vazamentos analisados para essa seqüência produziram valores de Sdist bastante inferiores à faixa de normalidade, promovendo eficiência total na detecção de vazamentos independente do diâmetro de furo empregado – Figura 6.31 e Tabela 6.9. A “faixa de normalidade” aqui considerada é idêntica a aplicada para detecção de falhas em sensores, de modo que valores de Sdist abaixo de 0.6 são considerados indicadores de vazamento.

O mesmo resultado foi encontrado para vazamentos implementados de acordo com a Seqüência “C”, com 100% de sucesso de detecção (Tabela 6.9).

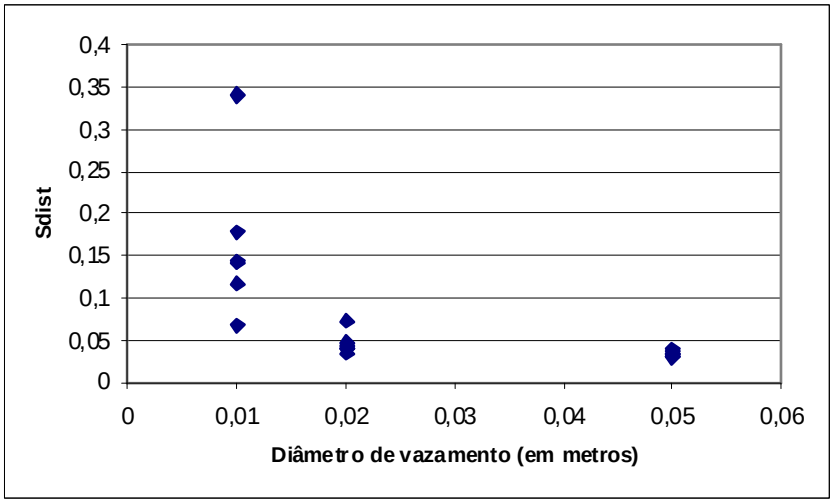


Figura 6.31: Resposta em Sdist para Seqüência A

Por outro lado, vazamentos baseados nas seqüências “B” e “D” foram de mais difícil detecção via Sdist estático. Algumas ocorrências não puderam ser identificadas, tendo o método mostrado especial dificuldade na detecção dos vazamentos de menor severidade (Figura 6.32). Para diâmetro de vazamento 0.01 metros inúmeras amostras estão situadas acima de 0.6, ou seja, não podem ser consideradas como indicação da presença de anomalia na rede. Por outro lado, para a seqüência “A” (Figura 6.31) todas as amostras apresentavam valores significativamente abaixo de 0.6.

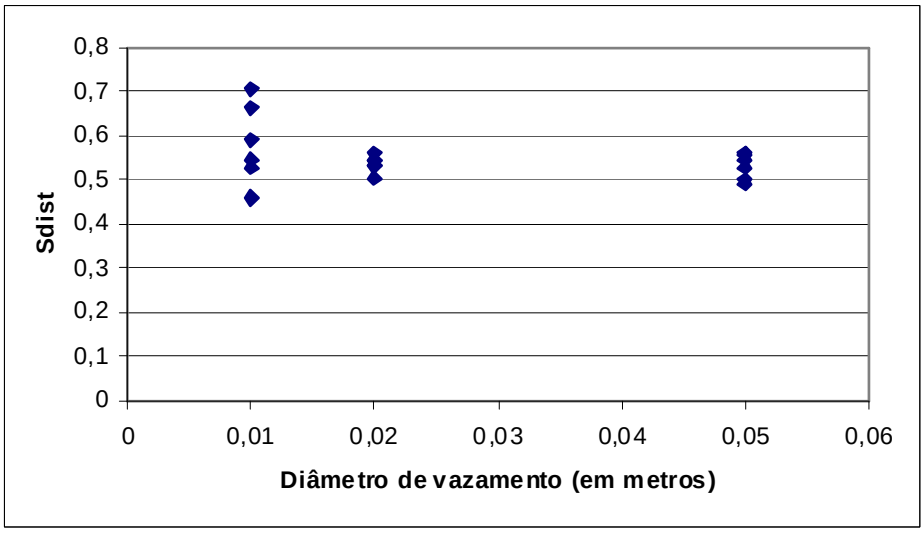


Figura 6.32: Resposta em Sdist para Seqüência B

A abordagem dinâmica para Sdist é aplicada com sucesso para as seqüências “A”, “B” e “C”, onde o grau de detecção foi de 100% (Tabela 6.9). Para a

disposição temporal da seqüência “D”, a detecção mostrou-se eficaz apenas em vazamentos severos, promovidos por furos de largo diâmetro. A mesma limitação já havia sido observada no estudo de falhas de sensores.

6.4.2 Detecção por classificação hierárquica

Para estabelecer a métrica que melhor resolve os *clusters* formados com observações temporais “com” e “sem” vazamento foram testadas quatro definições de distância: “Euclidiana”, “City Block”, “Minkoswki r3” e “Minkoswki r4”.

Baseando-se na “seqüência B” utilizou-se conjunto de 500 observações, das quais 450 representam situação de operação normal e as 50 restantes retratam anormalidade. A Figura 6.33 expõe os resultados encontrados para diâmetros de furo de 1, 2 e 3 centímetros. Nenhuma das métricas apresentou desempenho adequado para vazamentos ocasionados por furos de 1 centímetro. Ou seja, o método de detecção via classificador hierárquico não foi capaz de detectar vazamentos de pequeno volume, tendo atingido desempenho extremamente inferior ao alcançado usando Sdist dinâmico. Para diâmetros de vazamento superiores o desempenho foi mais satisfatório, especialmente quando aplicada métrica “Minkoswki r4”.

O desempenho da ferramenta de classificação hierárquica quando comparado ao atingido com o uso de Sdist, permite concluir que sua aplicação somente se justifica em uma abordagem híbrida. Nela, o classificador permite precisar o instante em que a falha teve início, enquanto a detecção em si é promovida pela abordagem Sdist dinâmica.

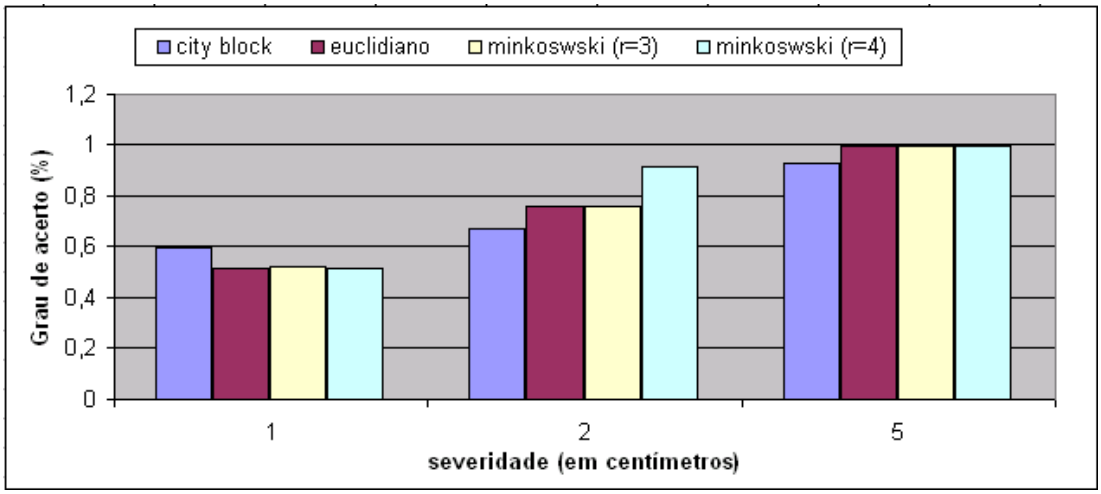


Figura 6.33: Grau de acerto do classificador hierárquico para falhas de sensor

6.5 Identificação do tubo com vazamento

A possibilidade de localizar rapidamente um vazamento mostra-se de extrema importância. Além de permitir o retorno da operação da dutovia no menor tempo possível, agilizando a manutenção, saber a correta localização do vazamento também permite definir claramente as ações que serão tomadas. Muitas vezes o impacto ambiental é função do trecho no qual a dutovia se rompeu, ou seja, da área que será afetada. A localização permite elaborar uma estimativa inicial sobre a existência de vítimas ou não no local. Todas essas informações tornam-se rapidamente disponíveis para as equipes de emergência quando um sistema eficiente de identificação da origem da falha está disponível.

Assim, uma vez detectada a ocorrência de um vazamento iniciam-se os procedimentos para a sua correta localização. O “índice de validade do sensor” (SVI) foi testado com tal finalidade, porém sem sucesso. Uma abordagem através do “método dos protótipos”, semelhante à empregada para falhas em sensores, também foi empregada para localizar vazamentos. No Capítulo 7, modelos ARX também são empregados com tal finalidade.

6.5.1 Identificação do tubo com vazamento via SVI-D

De forma análoga àquela empregada na localização de falhas em sensores, a metodologia aplicada na identificação do tubo com vazamento utiliza o “índice de validade dos sensores” – SVI-D, descrito por **DUNIA et al [1996a]**. Como definido, o índice mede a validade do sinal de cada sensor, não sendo desse modo um “índice de validade do tubo” diretamente. A idéia é que o vazamento em um dado tubo reflita-se mais fortemente nos sensores localizados em suas proximidades, afetando o SVI-D destes sensores.

Para um vazamento severo, com diâmetro de furo de 5 centímetros, a Tabela 6.10 apresenta os três sensores de menor validade para cada cenário de vazamento. Desse modo, para um vazamento ocorrido no Tubo 1, os sensores com menor SVI são: PI 20.2 (4), FI 20.1 (1) e FI 10.2 (12). Já para o caso de um vazamento no Tubo 2, os sensores de menor validade são: PI 20.8 (9), FI 10.2 (12) e FI 10.6 (15).

Tabela 6.10: Resultados de SVI para vazamento

tubo com vazamento	**1**	**2**	**4**	**5**	**6**	**7**
sensores candidatos	PI 20.2 (4)	PI 20.8 (9)	PI 20.9 (10)	FI 10.6 (15)	FI 20.9 (2)	FI 20.9 (2)
	FI 20.1 (1)	FI 10.2 (12)	FI 10.5 (14)	FI 10.5 (14)	FI 10.1 (11)	PI 20.8 (9)
	FI 10.2 (12)	FI 10.6 (15)	PI 20.2 (4)	FI 10.2 (12)	FI 10.4 (13)	PI 20.4 (6)

Nota: o número entre parênteses corresponde ao número que cada sensor recebeu neste Capítulo.

As Figuras 6.34 a 6.39 ilustram o posicionamento dos sensores candidatos para cada um dos cenários de vazamento apresentados na Tabela 6.10.

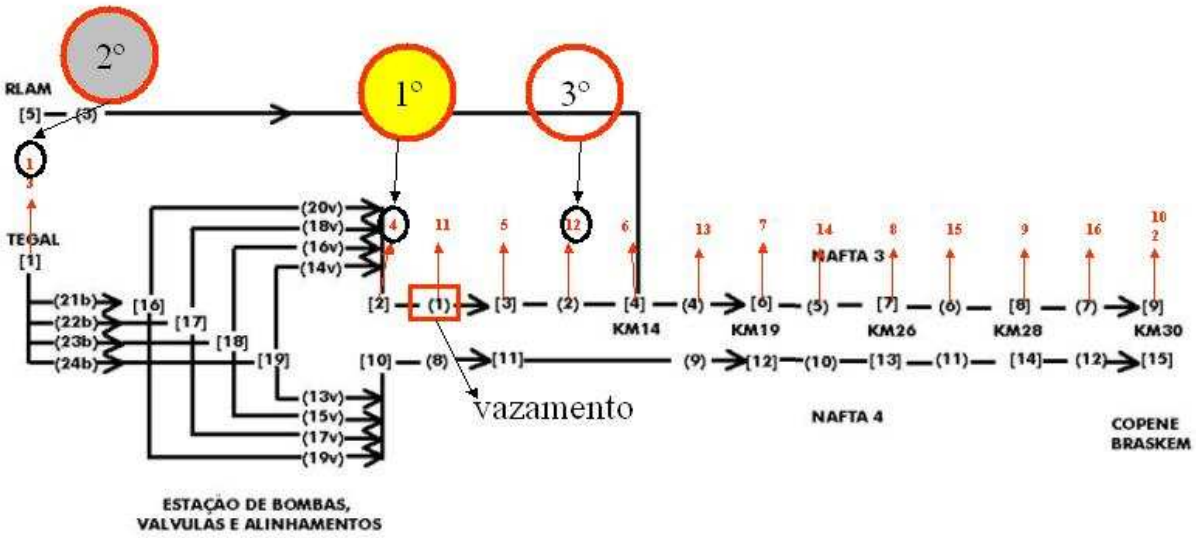


Figura 6.34: Vazamento no Tubo 1 provocado por um furo de diâmetro 5 centímetros

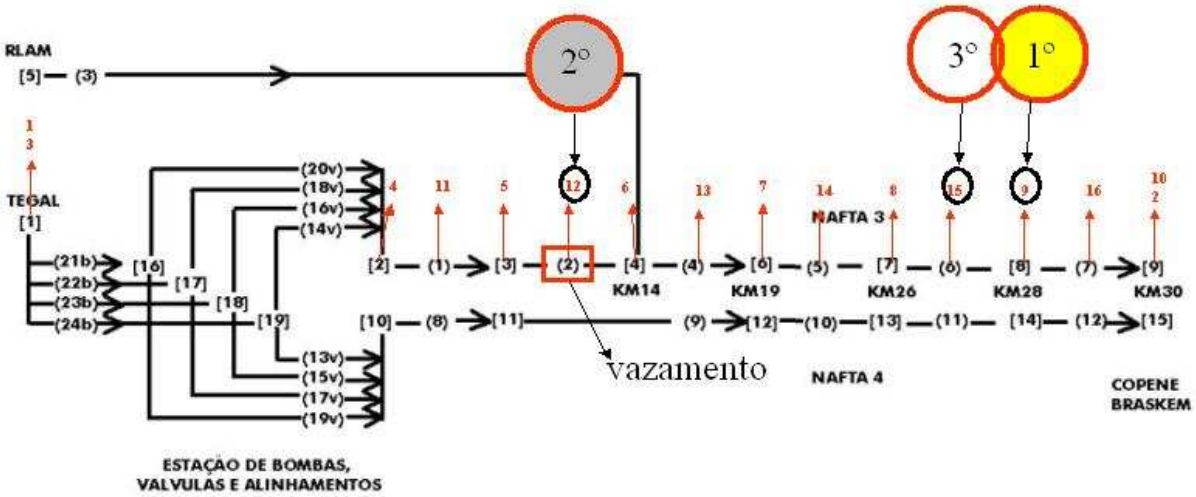


Figura 6.35: Vazamento no Tubo 2 provocado por um furo de diâmetro 5 centímetros

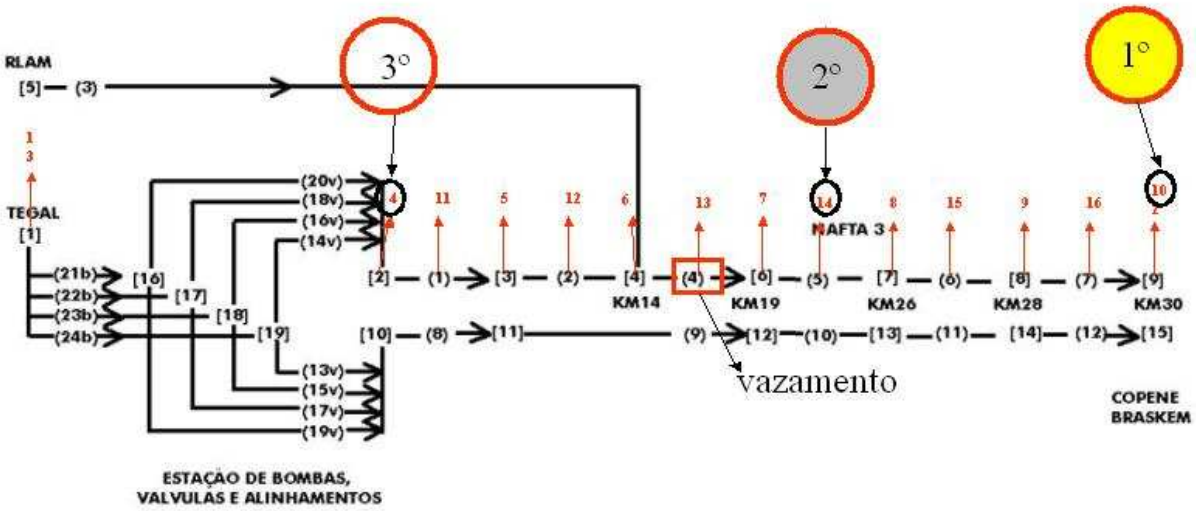


Figura 6.36: Vazamento no Tubo 4 provocado por um furo de diâmetro 5 centímetros

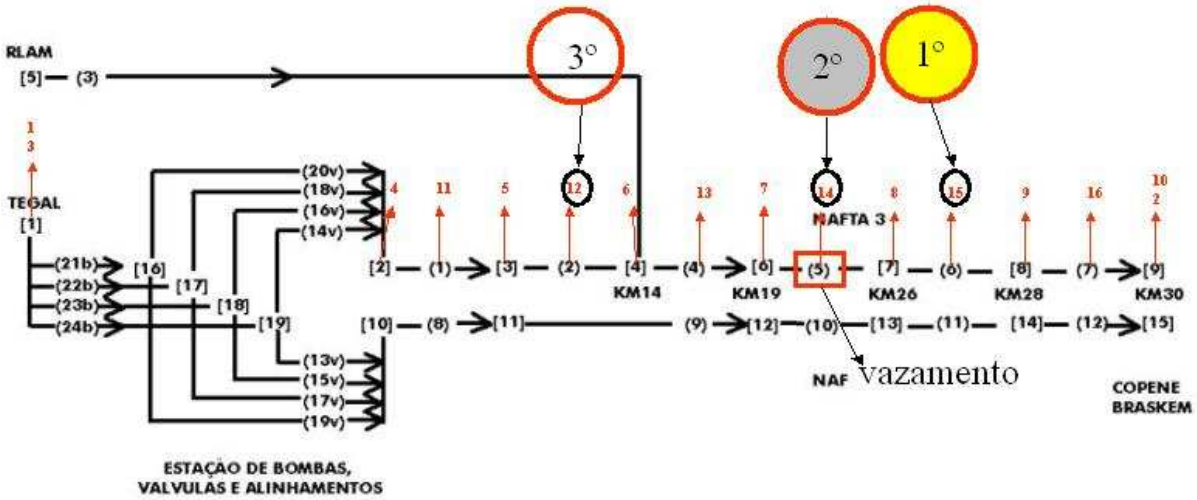


Figura 6.37: Vazamento no Tubo 5 provocado por um furo de diâmetro 5 centímetros

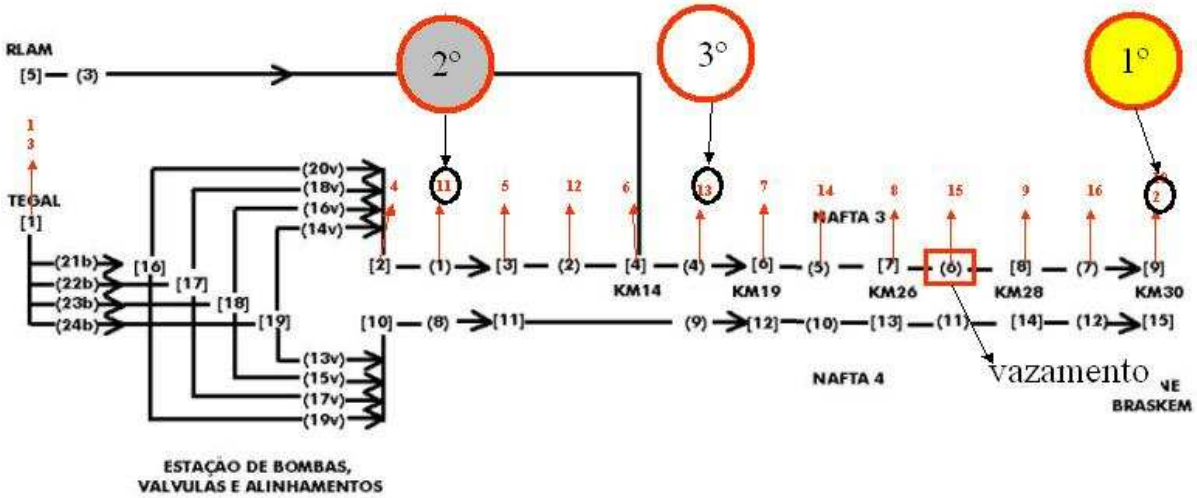


Figura 6.38: Vazamento no Tubo 6 provocado por um furo de diâmetro 5 centímetros

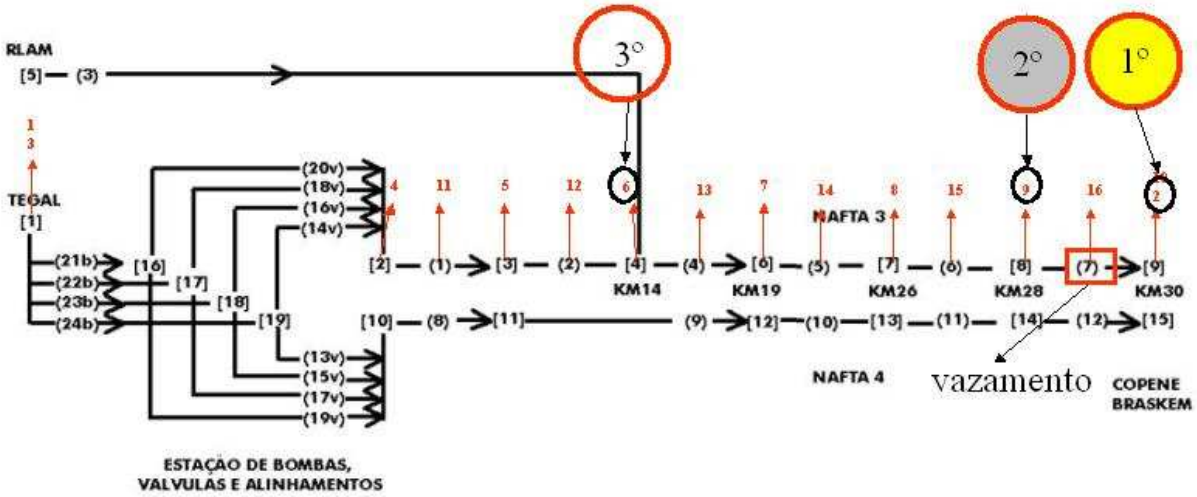


Figura 6.39: Vazamento no Tubo 7 provocado por um furo de diâmetro 5 centímetros

Ao reduzir o diâmetro de furo para 2 centímetros, a listagem de sensores cujo sinal apresenta baixa confiabilidade modificou-se fortemente, conforme a Tabela 6.11 e as Figuras 6.40 a 6.45.

Tabela 6.11: Resultados de SVI para vazamento

tubo com vazamento	1	2	4	5	6	7
candidatos	FI 10,7 (16) FI 10,1 (11) PI 20,4 (6)	FI 20,1 (1) PI 20,9 (10) PI 20,6 (7)	FI 10,1 (11) PI 20,2 (4) FI 20,9 (2)	FI 20,1 (1) PI 20,8 (9) PI 20,4 (6)	FI 10,6 (15) FI 10,7 (16) PI 20,2 (4)	FI 10,6 (15) FI 20,1 (1) FI 20,9 (9)

Nota: o número entre parênteses corresponde ao número que cada sensor recebeu neste Capítulo.

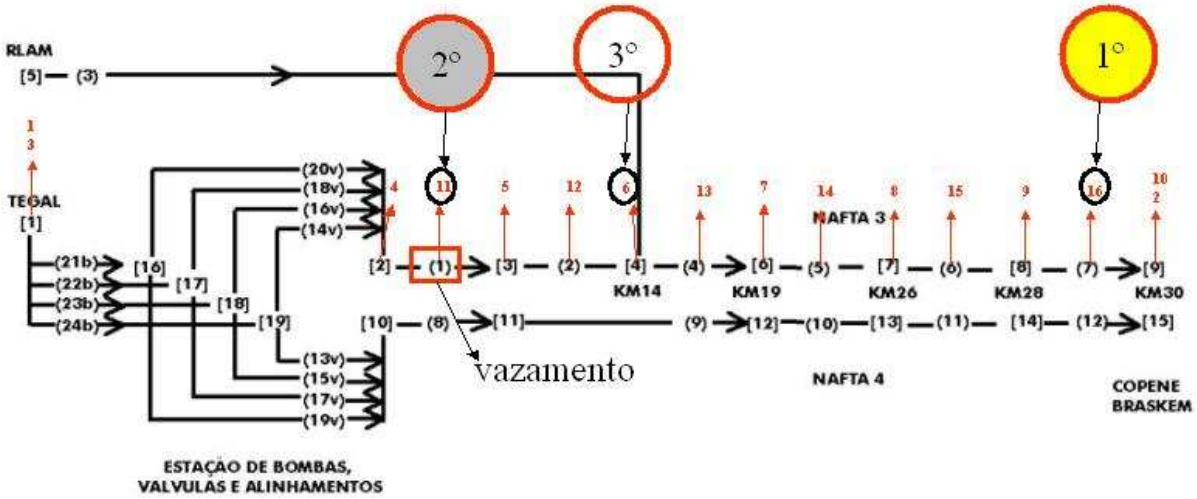


Figura 6.40: Vazamento no Tubo 1 provocado por um furo de diâmetro 2 centímetros

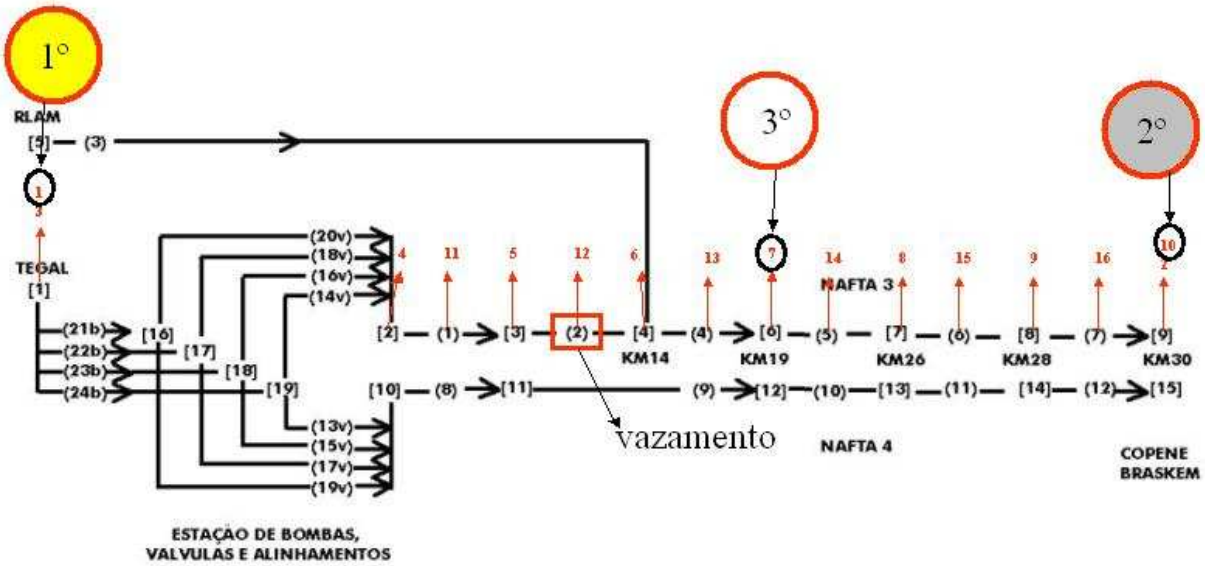


Figura 6.41: Vazamento no Tubo 2 provocado por um furo de diâmetro 2 centímetros

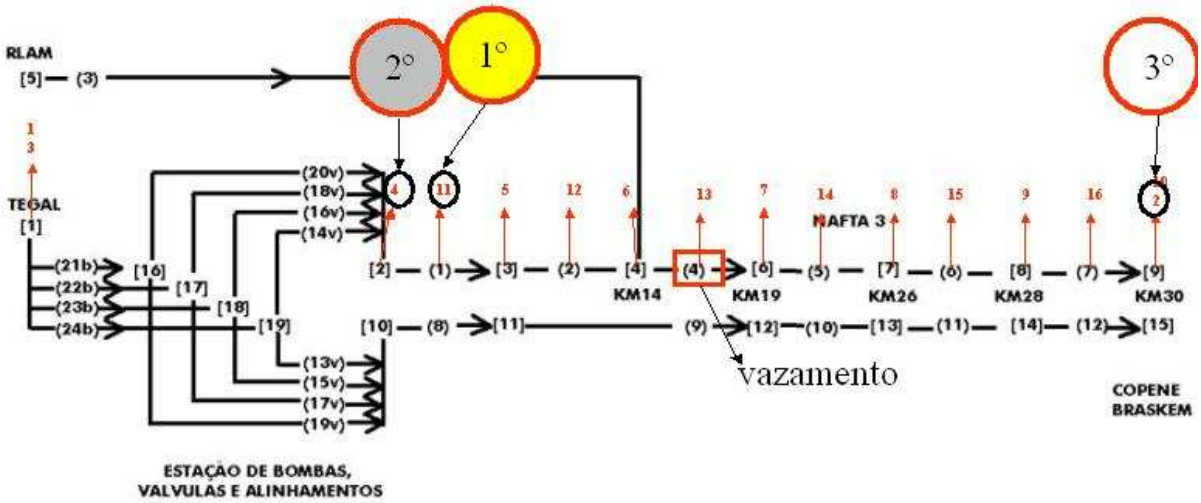


Figura 6.42: Vazamento no Tubo 4 provocado por um furo de diâmetro 2 centímetros

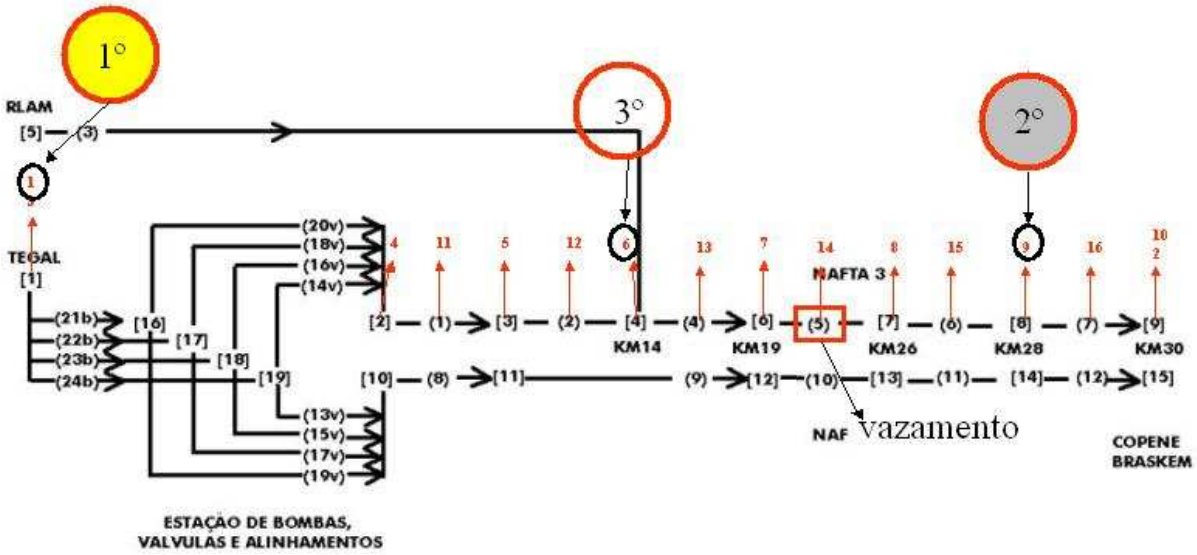


Figura 6.43: Vazamento no Tubo 5 provocado por um furo de diâmetro 2 centímetros

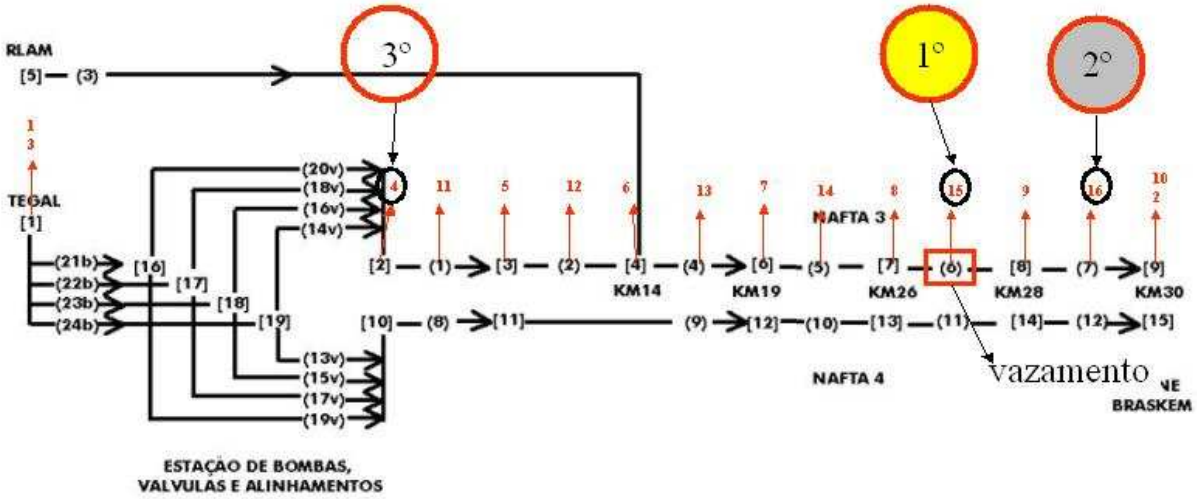


Figura 6.44: Vazamento no Tubo 6 provocado por um furo de diâmetro 2 centímetros

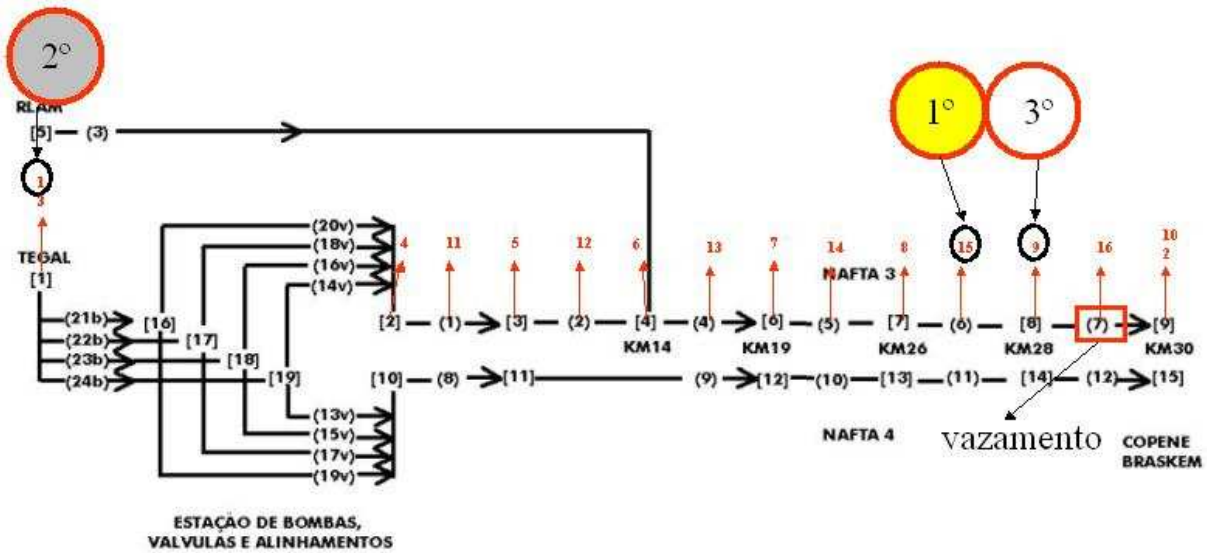


Figura 6.45: Vazamento no Tubo 7 provocado por um furo de diâmetro 2 centímetros

Através da disposição espacial dos sensores candidatos e a forte diferença entre os resultados apresentados em ambas as tabelas, modificando apenas a severidade da anomalia, conclui-se que a tentativa de relacionar o índice de confiabilidade do sinal dos sensores e o local de vazamento foi insatisfatória.

6.5.2 Identificação do tubo com vazamento via protótipos

Nesta abordagem foram utilizados os mesmos 22 casos representativos de anomalias que foram empregadas na identificação de falhas em sensores. Uma vez dispondo dos protótipos e um vazamento tendo sido detectado, o sinal gerado é comparado com cada um dos padrões em uma busca pela maior similaridade, permitindo a identificação do tubo com falha.

Ilustrativamente, considera-se a situação na qual se tem a ocorrência de um vazamento gerado por um furo com 2 centímetros de diâmetro no Tubo 5. A Tabela 6.12 exibe as medidas de similaridade entre os dados experimentais com vazamento e cada um dos protótipos. Conforme apresentado na tabela, os dados guardam maior similaridade com o protótipo “vaz300T5B03”, sendo este característico de vazamentos no tubo 5. Para a situação ilustrada, a abordagem foi eficaz na detecção da origem da anomalia.

Tabela 6.12: Similaridade entre os dados experimentais com falha e cada protótipo característico

Protótipo	Spca	Protótipo	Spca
vaz300T1B03	0.76291	sen300S6F05	0.80012
vaz300T2B03	0.80003	sen300S7F05	0.80719
vaz300T4B03	0.79987	sen300S8F05	0.78641
vaz300T5B03	**0.83792**	sen300S9F05	0.78321
vaz300T6B03	0.82393	sen300S10F05	0.79694
vaz300T7B03	0.7584	sen300S11F05	0.80338
sen300S1F05	0.79104	sen300S12F05	0.79092
sen300S2F05	0.77618	sen300S13F05	0.79437
sen300S3F05	0.79561	sen300S14F05	0.79647
sen300S4F05	0.81124	sen300S15F05	0.78969
sen300S5F05	0.79214	sen300S16F05	0.79597

Repetindo o procedimento acima para vazamentos gerados a partir de furos com diâmetros de 1, 2 e 4 centímetros, a eficiência do método foi elevada. Para diâmetros de furo de 2 centímetros a metodologia localizou corretamente a origem do vazamento em 67% dos casos (Tabela 6.13).

A localização da origem do vazamento via método dos protótipos mostrou eficiência superior a abordagem via SVI, sendo esse resultado oposto ao encontrado para falhas em sensores. Ressalta-se que esta é uma das poucas diferenças significativas na abordagem de ambas as anomalias.

Tabela 6.13: Grau de acerto na identificação da origem do vazamento através do método dos protótipos

Diâmetro de furo (cm)	Acertos (tubos)	Erros (tubos)	Grau de acerto (%)
1	1	5	16,67
2	4	2	66,67
4	4	2	66,67

6.6 Quantificação de vazamentos

Além dos modelos ARX (Capítulo 7), classificadores hierárquicos também foram empregados para estimar a severidade da anomalia a partir do afastamento existente entre os centros dos dois agrupamentos formados. Ou seja, a dissimilaridade observada entre os *clusters* “com” e “sem” falha pode ser

correlacionada com a severidade da anomalia ocorrida. Emprega-se o mesmo princípio na estimação do diâmetro de furo, ou seja, na quantificação do vazamento.

A Figura 6.46 retrata a relação entre amplitude de falha e afastamento entre grupos. Enquanto em severidades baixas, tais como 1 e 2 centímetros, a distância média entre os agrupamentos é menor que 5, para falhas severas o afastamento atingiu valores até seis vezes maiores. Devido ao maior êxito observado na etapa de detecção através do uso da métrica Minkowski r4, apenas essa distância foi empregada neste procedimento de quantificação.

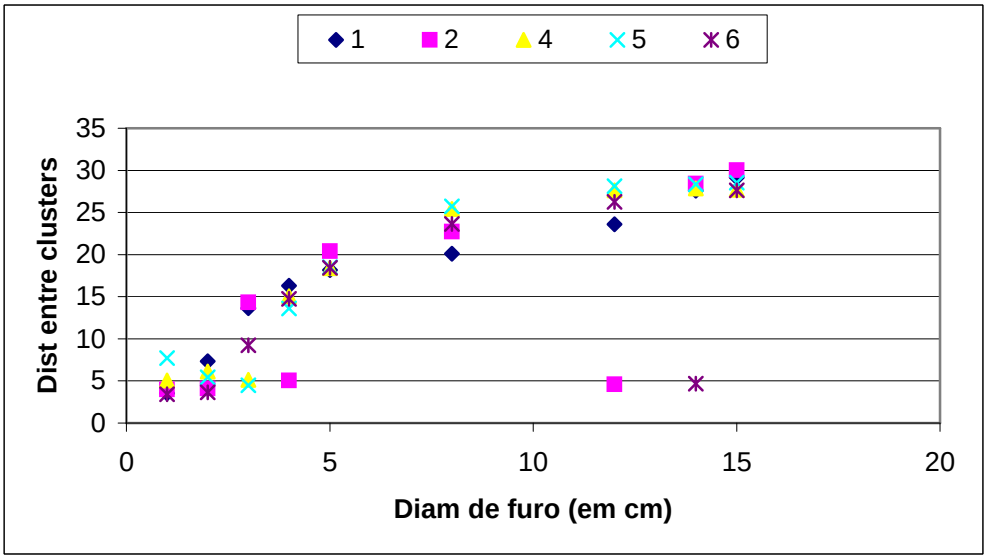


Figura 6.46: Afastamentos entre clusters em função do diâmetro de furo presente

6.7 Diferenciação entre falhas em sensores e vazamentos

Três abordagens baseadas PCA e classificadores hierárquicos foram testados com o objetivo de diferenciar entre ocorrências de falhas em sensores e vazamentos. A diferenciação entre essas duas anormalidades é importante pois retratam situações de gravidade operacionais distintas. Enquanto falhas em sensores causam transtorno operacional e, na maior parte dos casos, não significam risco imediato, vazamentos são situações de risco extremo.

Na primeira abordagem supõe-se que diferentes tipos de anomalias promovem respostas distintas nas métricas Sdist e Spca **[SINGHAL e SEBORG,**

2002]. Buscou-se então diferenciar as duas espécies de falhas através de uma análise visual dos *clusters* formados pela combinação de Spca e Sist. O método porém não foi eficaz, pois as falhas em sensores e vazamentos promoveram respostas similares, como exibido na Figura 6.47.

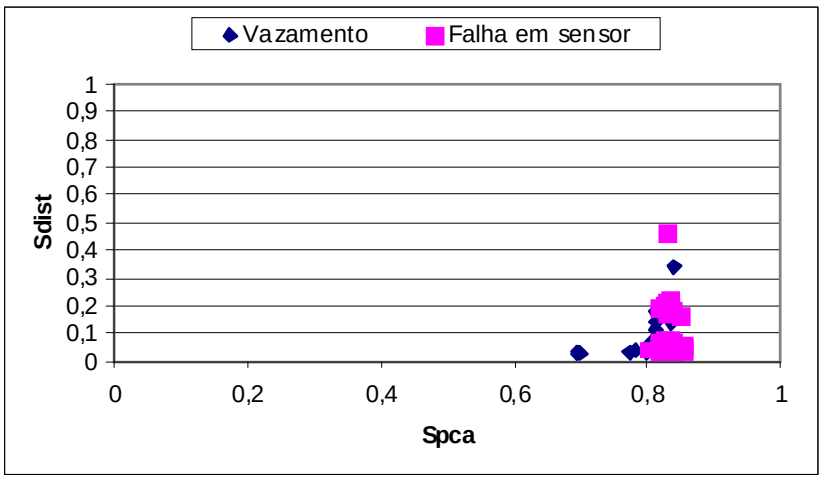


Figura 6.47: Resposta em Sdist e Spca para vazamentos e falhas em sensor

A segunda abordagem baseou-se no efeito da anomalia sobre o parâmetro SVI. Nos cenários de falha da instrumentação tem-se a ocorrência de um único sensor em falha (falhas simultâneas não foram consideradas). Por outro lado, devido às propriedades físicas inerentes a ocorrência de vazamento, este provoca reflexos em diversos sensores localizados nas proximidades do ponto de rompimento. Assim, supõe-se que cenários de vazamentos tendam a se refletir no índice SVI de dois ou mais sensores. Essa diferença de resposta no SVI poderia então ser aplicada na metodologia de diferenciação.

Baseando-se neste princípio buscou-se elaborar metodologia de diferenciação via resposta em SVI. Foi organizada base de dados com 22 situações, sendo 16 delas falhas de elevada severidade em sensores e as restantes vazamentos severos. Cada observação é composta por 16 variáveis – um valor de SVI para cada sensor. Tal sistema, de dimensão 22x16, foi então processado novamente através da técnica de PCA na busca de agrupamentos que pudessem ser identificados. Porém, representando graficamente os dois componentes principais, não foi possível resolver corretamente as amostras em dois agrupamentos (Figura 6.48).

Por fim, uma abordagem híbrida baseada em classificação hierárquica e teoria dos protótipos foi aplicada. Os vinte e dois protótipos representando falhas em sensores e vazamentos são utilizados aqui para possibilitar a identificação do tipo de anomalia encontrada. Os dados contendo a anomalia em estudo, juntamente com os 22 protótipos, são submetidos a procedimento classificatório. O grau de similaridade entre as observações temporais e cada um dos protótipos é capaz de identificar o tipo de anomalia em questão. Essa técnica obteve êxito na diferenciação de falhas em sensores e vazamentos. A Tabela 6.14 apresenta os resultados obtidos. Para falhas em sensores, o grau de acerto no estabelecimento do tipo de falha usando esta abordagem foi absoluto, tanto para desvios de 1% ou 2%. Em cenários de vazamento a eficiência foi superior a 80%, sendo a única exceção as situações geradas por diâmetro de furo de 1 centímetro.

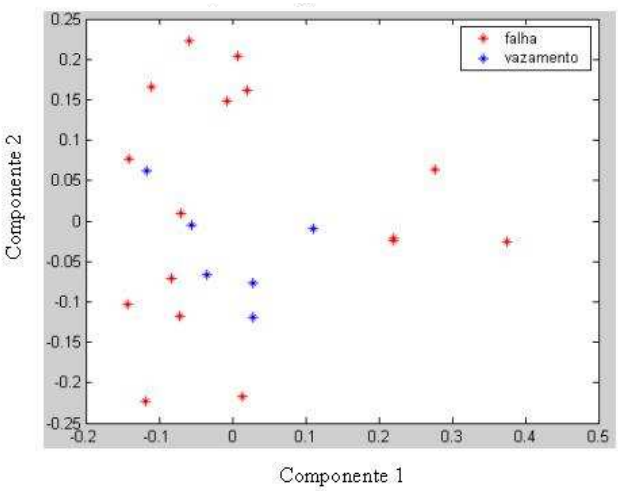


Figura 6.48: Representação gráfica dos dois componentes principais obtidos a partir do índice de validade dos sensores

Tabela 6.14: Grau de sucesso obtido na identificação do tipo de anomalia via classificadores hierárquicos e método dos protótipos

Diâmetro de furo (cm)	Acertos	Erros	Grau de acerto (%)
1	2	4	33,33
2	5	1	83,33
4	6	0	100,00
5	5	1	83,33
8	5	1	83,33
12	6	0	100,00
14	6	0	100,00
Severidade da falha em sensor (%)			
1	16	0	100,00
2	16	0	100,00

6.8 Escolha das Melhores Abordagens

A partir dos resultados encontrados neste capítulo é possível propor a metodologia mais adequada para cada uma das tarefas do sistema detector de anomalias:

Detecção:
falha em sensor: Sdist dinâmico
vazamento: Sdist dinâmico

Identificação da Origem:
falha em sensor: SVI-D
vazamento: protótipos

Quantificação da Severidade:
falha em sensor: classificação hierárquica
vazamento: classificação hierárquica

Diferenciação:
Diferenciação de vazamento e falha em sensor: protótipos

7. Resultados da Metodologia ARX para Detecção, Localização e Quantificação de Vazamentos

Este Capítulo apresenta os resultados obtidos a partir do uso de modelos auto-regressivos nas tarefas de detectar, localizar e quantificar vazamentos. Complementam-se assim os resultados apresentados no Capítulo 6. É importante salientar mais uma vez que as diferentes abordagens estudadas não devem ser encaradas como mutuamente excludentes, mas sim como ferramentas complementares.

7.1 Séries Temporais da Rede sob Ação de Três Vazamentos - Cenário R3L

Com a finalidade de gerar-se um caso de referência para estudos de Detecção, Localização e Quantificação de vazamentos usando modelos ARX, procedeu-se a escolha de um cenário base de simulação. Neste cenário – identificado como R3L – três vazamentos serão designados para acontecer em uma série temporal da rede com duração de 1000 instantes de tempo (nt=1000). A rede empregada neste Capítulo foi aquela denominada “reduzida”, descrita no Capítulo 5, Item 2, com os respectivos sensores (Tabela 5.5).

Em R3L os dois pontos de fornecimento e os dois pontos de recebimento da rede estarão em serviço, juntamente com todas as dutovias NAFTA2, NAFTA3 e NAFTA4. As especificações de vértices intermediários corresponderão a vazões externas nulas (*W=0*). Os valores de referência das especificações dos vértices de extremidade foram escolhidos de acordo com: $P_1^{Ref} = 35bar;\quad P_5^{Ref} = 40bar;$ $P_9^{Ref} = 4bar;\quad W_{15}^{Ref} = -10kg/s$. Em R3L P_9 é constante no tempo no seu valor de referência P_9^{Ref}. As demais especificações de extremidades oscilam no tempo. R3L responderá, portanto, a três entradas dinâmicas:

$$\underline{U}(t) = \begin{bmatrix} P_1(t) \\ P_5(t) \\ W_{15}(t) \end{bmatrix} \qquad \text{eq. 7.1}$$

Três vazamentos são programados para ocorrência de acordo com os valores da Tabela 7.1. Estes vazamentos afetarão, respectivamente, as dutovias NAFTA2, NAFTA3 e NAFTA4. Os diâmetros de furo empregados nestes testes são semelhantes àqueles utilizados para testar as ferramentas baseadas em PCA.

Os comportamentos não-determinísticos das entradas são apresentados na Figura 7.1. Os instantes de manifestação de rupturas são assinalados com flechas e com as inscrições E_1, E_2, E_3. Os respectivos instantes de cessação são assinalados com flechas e com as inscrições D_1, D_2, D_3.

Tabela 7.1: Programação de Vazamentos para o Cenário R3L

	Vazamento 1	**Vazamento 2**	**Vazamento 3**
Aresta Danificada	3	9	6
Diâmetro da Ruptura (D_F)	15mm (1,5 cm)	20mm (2,0 cm)	10mm (1,0 cm)
Posição Frac. da Ruptura ($x = L_F / L_k$)	30%	40%	70%
Instante de Ocorrência (t_E)	250	500	800
Instante de Cessação (t_D)	930	955	985

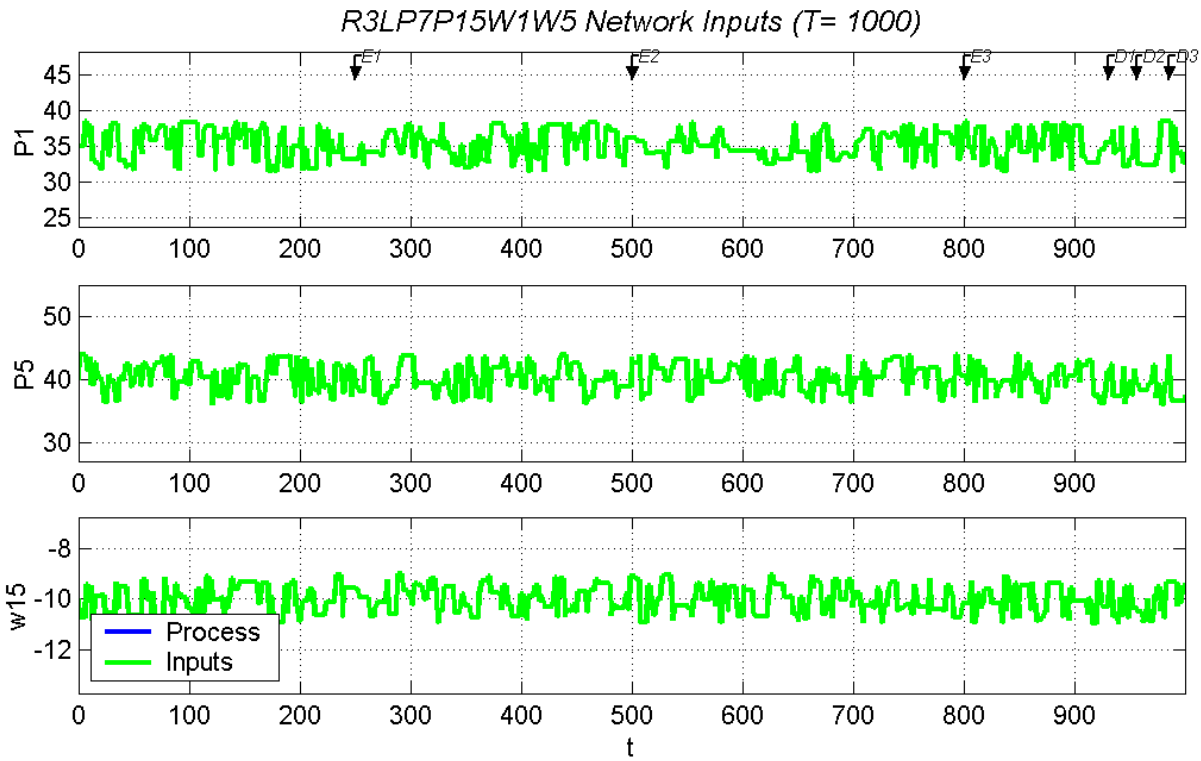


Figura 7.1: Comportamento de Entradas Dinâmicas para R3L

O comportamento de diversas variáveis do processo são mostrados nas Figuras 7.2 a 7.5. O impacto dos vazamentos em pressões de vértices, vazões de arestas e vazões externas de vértices podem ser apreciados. As vazões propriamente de vazamentos, correspondem às variáveis W_{17}, W_{19} e W_{21} (todas negativas por serem saídas de fluido da rede). Nestas figuras certas variáveis (associadas aos vértices e arestas novas) são criadas ao longo da campanha, apresentando, em função disto, séries temporais caracterizadas por uma linha base nula seguida de um salto associado à aparição de um vazamento (por exemplo, o caso de P_{16} e de P_{18}).

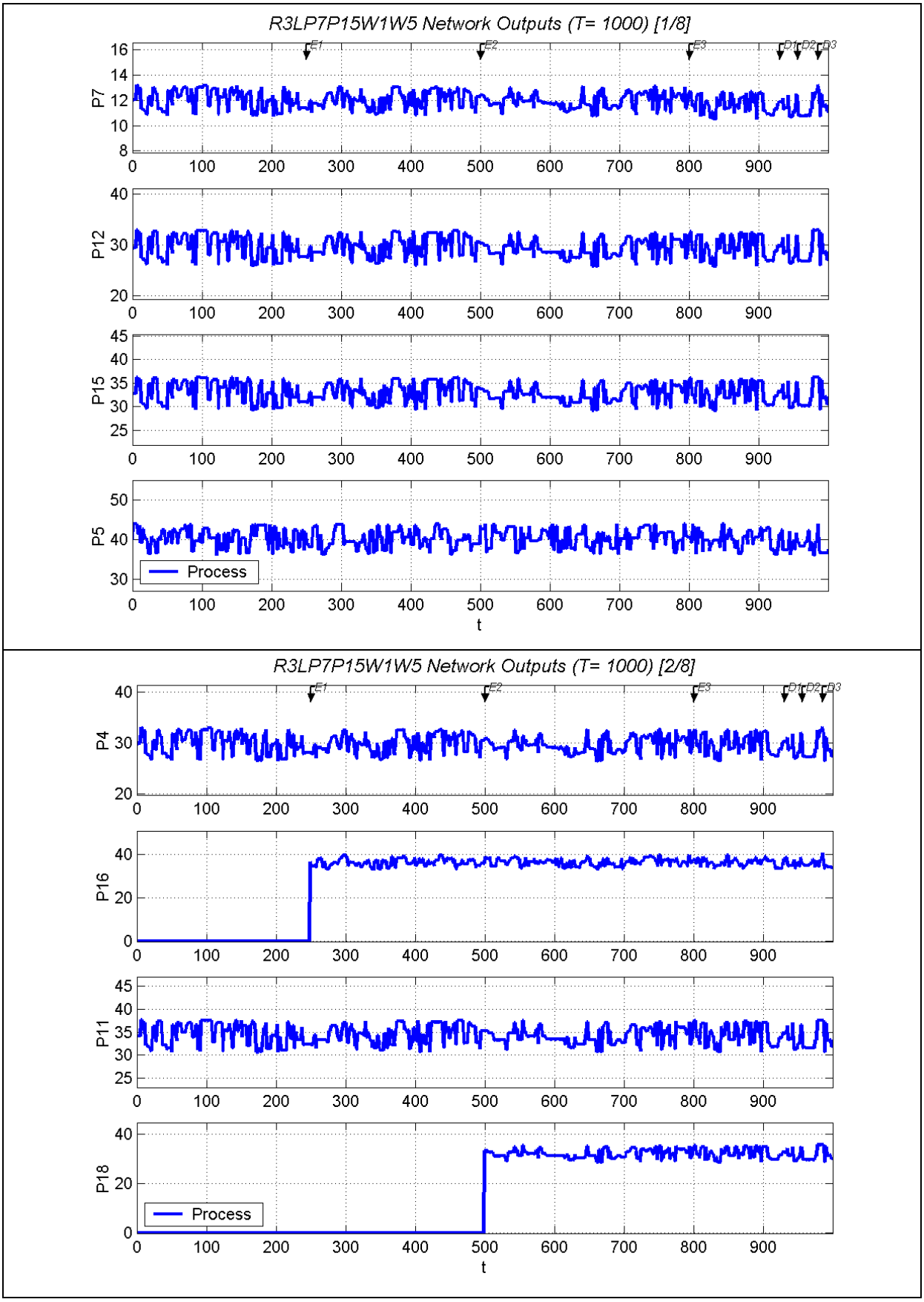


Figura 7.2 : Comportamento Temporal de Variáveis para Cenário R3L

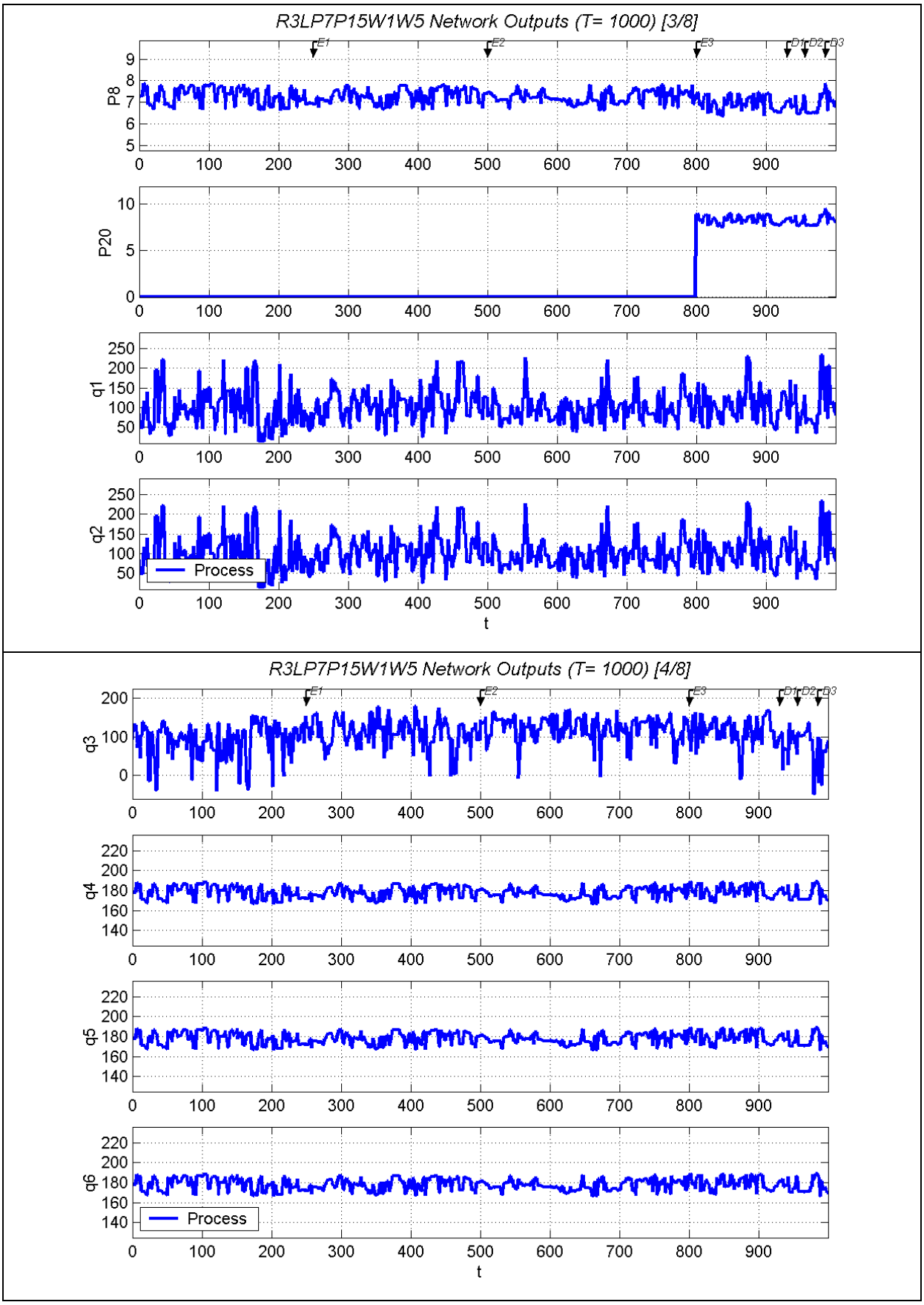


Figura 7.3 : Comportamento Temporal de Variáveis para Cenário R3L

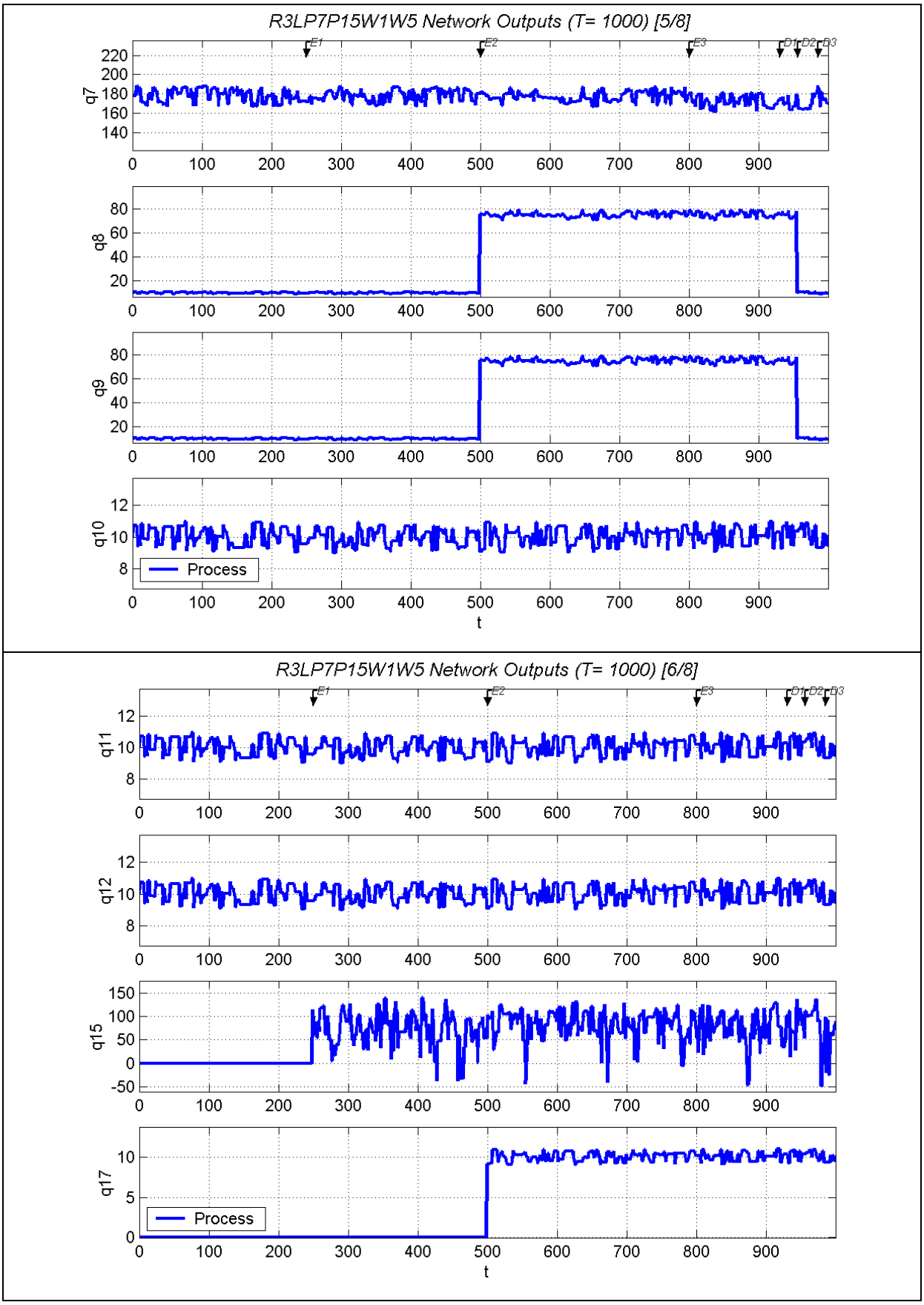


Figura 7.4 : Comportamento Temporal de Variáveis para Cenário R3L

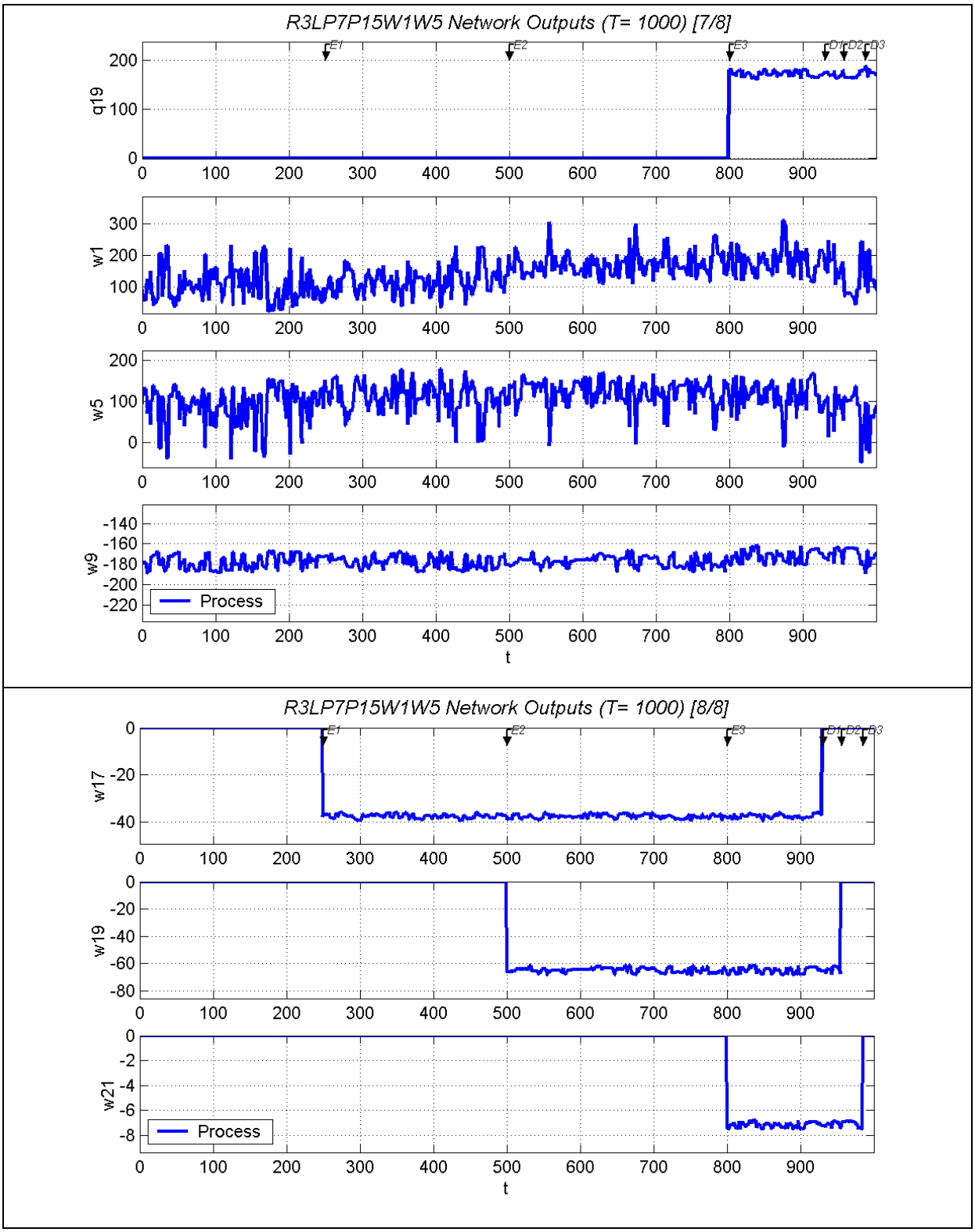


Figura 7.5 : Comportamento Temporal de Variáveis para Cenário R3L

Para construção de estratégias de detecção no Cenário R3L, conjuntos de variáveis de *saída* ($\underline{Y}$) devem ser escolhidos. Três conjuntos de medições foram considerados como propostas de variáveis de *saída* do processo (são indicados variáveis e respectivos sensores):

- P7 (PI-7), P12 (PI-12) e P15 (PI-15)
- P7 (PI-7), P15 (PI-15) e W1 (FI-1)
- P7 (PI-7), P15 (PI-15), W1 (FI-1) e W5 (FI-5)

O primeiro conjunto utiliza apenas medidas de pressão como *saídas*. As corridas de simulação associadas a este conjunto são identificadas por R3LP7P12P15. O segundo conjunto utiliza duas das pressões anteriores e a vazão externa do vértice 1. As corridas associadas a este conjunto são identificadas por R3LP7P15W1. O terceiro conjunto emprega as duas pressões do conjunto anterior e as vazões externas de vértices 1 e 5. As corridas associadas a este conjunto são identificadas por R3LP7P15W1W5.

A metodologia de detecção de vazamentos a partir de preditores ARX-MIMO foi apresentada anteriormente. O preditor ARX-MIMO (Equação 7.2) aqui empregado opera com ordens *m=3* e *n=3* nos filtros matriciais de entrada ($\underline{\underline{B}}(q)$) e auto-regressivo ($\underline{\underline{A}}(q)$). Estudo detalhado sobre a definição da ordem dos filtros foi desenvolvido por **VAZ JUNIOR [2006].**

$$\hat{\underline{Y}}(t+1) = \underline{\underline{B}}(q)\underline{U}(t+1) + q\left[1 - \underline{\underline{A}}(q)\right]\underline{Y}(t) \qquad \text{eq. 7.2}$$

Em todos os casos estudados o número de *entradas* é 3 (*nu=3*). Nos casos R3LP7P12P15 e R3LP7P15W1 o número de *saídas* é 3 (*ny=3*), enquanto o caso R3LP7P15W1W5 utiliza 4 *saídas* (*ny=4*). Nos dois primeiros casos os filtros $\underline{\underline{B}}(q)$ e $\underline{\underline{A}}(q)$ têm ambos tamanhos *3 x 3*, ao passo que no último caso têm tamanhos *4 x 3* e *4 x 4*, respectivamente.

Nos três casos os ARX(3,3)-Detectores operam com sessões contínuas de treinamento e vigilância. As sessões de treinamento duram 100 instantes (*NT=100*),

enquanto as sessões de vigilância estendem-se subseqüentemente por 240 instantes (*NI=240*). Ao longo das séries de 1000 instantes, será possível ocorrerem 3 ciclos de treino e vigilância com 340 instantes cada (o último período de vigilância não é completado). Nos diagramas de detecção flechas assinalam os instantes iniciais das sessões de treino com a inscrição T_k para o *k*-ésimo treino.

O Cenário R3L foi simulado em 1000 instantes consecutivos para teste das estratégias de detecção R3LP7P12P15, R3LP7P15W1 e R3LP7P15W1W5. Como a geração de *entradas* é não-determinística, as três séries temporais não são exatamente idênticas, porém os padrões probabilísticos que atuam nos três casos são os mesmos.

De modo análogo aquele empregado para fazer detecção pelos métodos PCA, o mecanismo de detecção usando ARX emprega a identificação de uma discrepância indubitável entre as predições e os valores de processo. As predições ARX ($\underline{\hat{Y}}(t)$) avançam um instante adiante do processo, isto é, são geradas um instante antes da chegada da nova amostragem ($\underline{Y}(t)$). Entende-se como "discrepância indubitável" em *t*, um desvio superior a 4 desvios padrões da variável vigiada (para, pelo menos, uma das variáveis vigiadas) conforme descrito na Equação 7.3:

$$\left|\hat{Y}_k(t) - Y_k(t)\right| > 4\sigma_{Yk} \qquad \text{eq. 7.3}$$

onde as estatísticas *desvio padrão* e *expectância* da variável vigiada correspondem a valores históricos do processo em condições normais de serviço, sendo calculados por séries anteriores à série em uso, de acordo com:

$$\sigma_{Yk} = \sqrt{\frac{\sum_{t=1}^{nt}(Y_k(t) - <Y_k>)}{nt-1}} \quad , \quad <Y_k> = \frac{\sum_{t=1}^{nt} Y_k(t)}{nt} \qquad \text{eq. 7.4}$$

Espera-se, em geral, que o as predições do detector apresentem aderência ao processo na janela de treino. Esta aderência deverá perdurar de forma

sustentável na próxima fase de vigilância. Naturalmente esta fase não poderá estender-se indefinidamente, de modo que um novo treino será necessário após algum tempo. As condições que definem as proporções das fases de treino e vigilância não serão discutidas neste trabalho.

A aderência do detector ao processo durante a fase de vigilância, deverá ser destruída ou significativamente deteriorada após a manifestação de um evento capaz de criar uma avaria não desprezível na rede. Esta deterioração costuma afetar as predições de todas as saídas do ARX-detector e costuma perdurar, ou mesmo se agravar, até que se atinja a próxima fase de treino. Em condições adequadas de utilização a perda de aderência do ARX-Detector após o evento de ruptura é tão marcante que as predições chegam a "fugir" da escala dos gráficos correspondentes. Naturalmente, o evento de ruptura não poderá ocorrer durante a fase de treino do detector, pois, neste caso, ele será assimilado como comportamento a "ser imitado" pelo mecanismo de predição.

As Figuras 7.6, 7.7 e 7.8 apresentam os resultados de detecção para as *saídas* de R3LP7P12P15, R3LP7P15W1 e R3LP7P15W1W5, respectivamente. Os valores de processo estão em azul, as predições do ARX-Detector em treino estão em vermelho, enquanto as predições do ARX-Detector em vigilância estão em lilás. Neste sentido, verifica-se que R3LP7P12P15 não é capaz de cumprir nem mesmo a primeira fase de treino do detector. A Figura 7.6 mostra que alguma aderência só se estabelece após o terceiro treinamento, e mesmo assim, esta aderência não é capaz de ser perdida na ocorrência do terceiro evento. Antes, as predições apresentam-se completamente descoladas do processo real, localizando-se fora da escala do gráfico.

A Figura 7.7 revela que substituição da *saída* P_{12} por W_1, traz expressiva melhora no desempenho do ARX-Detector. Neste sentido R3LP7P15W1 é capaz de cumprir suas três fases de treino, sendo capaz de detectar E_2 e E_3. No entanto, mostra-se incapaz de detectar o vazamento E_1 que atinge o *tubo* 3 pertencente ao ramal RELAM (NAFTA2). Parece claro que deve-se incluir na estrutura de detecção pelo menos uma medida amostrada no trecho NAFTA2.

A inclusão da vazão externa do vértice RELAM (W_5) dá origem ao ARX-Detector R3LP7P15W1W5 que se mostra razoavelmente efetivo conforme mostrado na Figura 7.8. Neste caso, todos os treinos são cumpridos adequadamente e os três eventos (E_1, E_2 e E_3) são precisamente detectados. Registre-se, com respeito à ocorrência E_3, que três saídas deste detector respondem fortemente ao evento, caracterizando inequivocamente a ruptura em questão. Todavia, suas predições $\hat{P}_{15}$ têm certa dificuldade para "descolar" das amostragens P_{15} (embora apresentem um salto perceptível no preciso instante do evento E_3). Isto parece razoável porque este evento age sobre o *tubo* 6, ao final da linha NAFTA3, enquanto que P_{15} é amostrado no ponto terminal de NAFTA4. Uma conclusão óbvia que resulta destas comparações, corresponde à noção de que estruturas de detecção de vazamentos apresentam desempenho superior quando alimentadas com medições confiáveis de vazões em lugar de pressões.

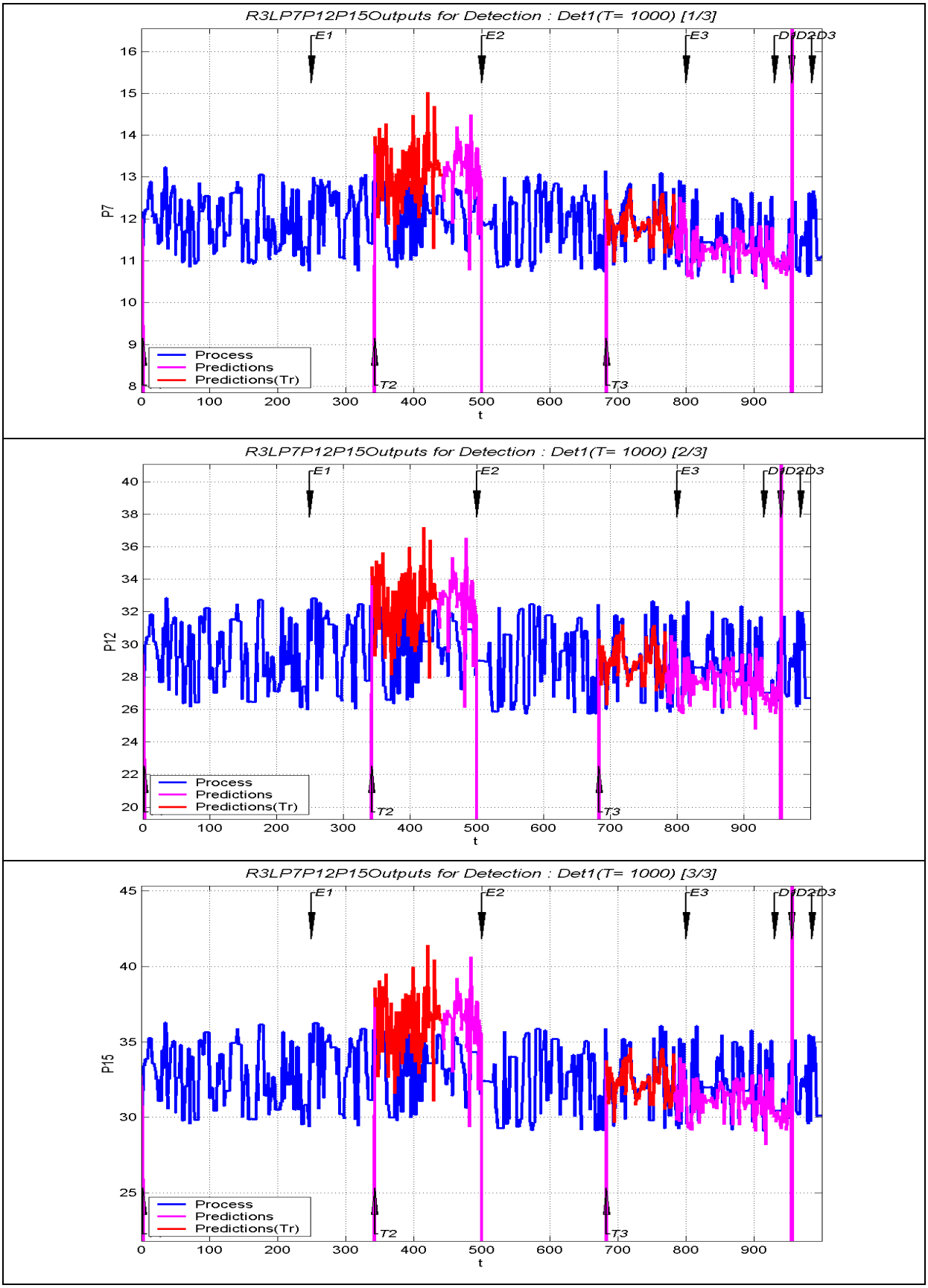


Figura 7.6 : Desempenho de Detector ARX para R3LP7P12P15
Azul: Processo; Vermelho: Predições em Treino; Lilás: Predições em Vigilância

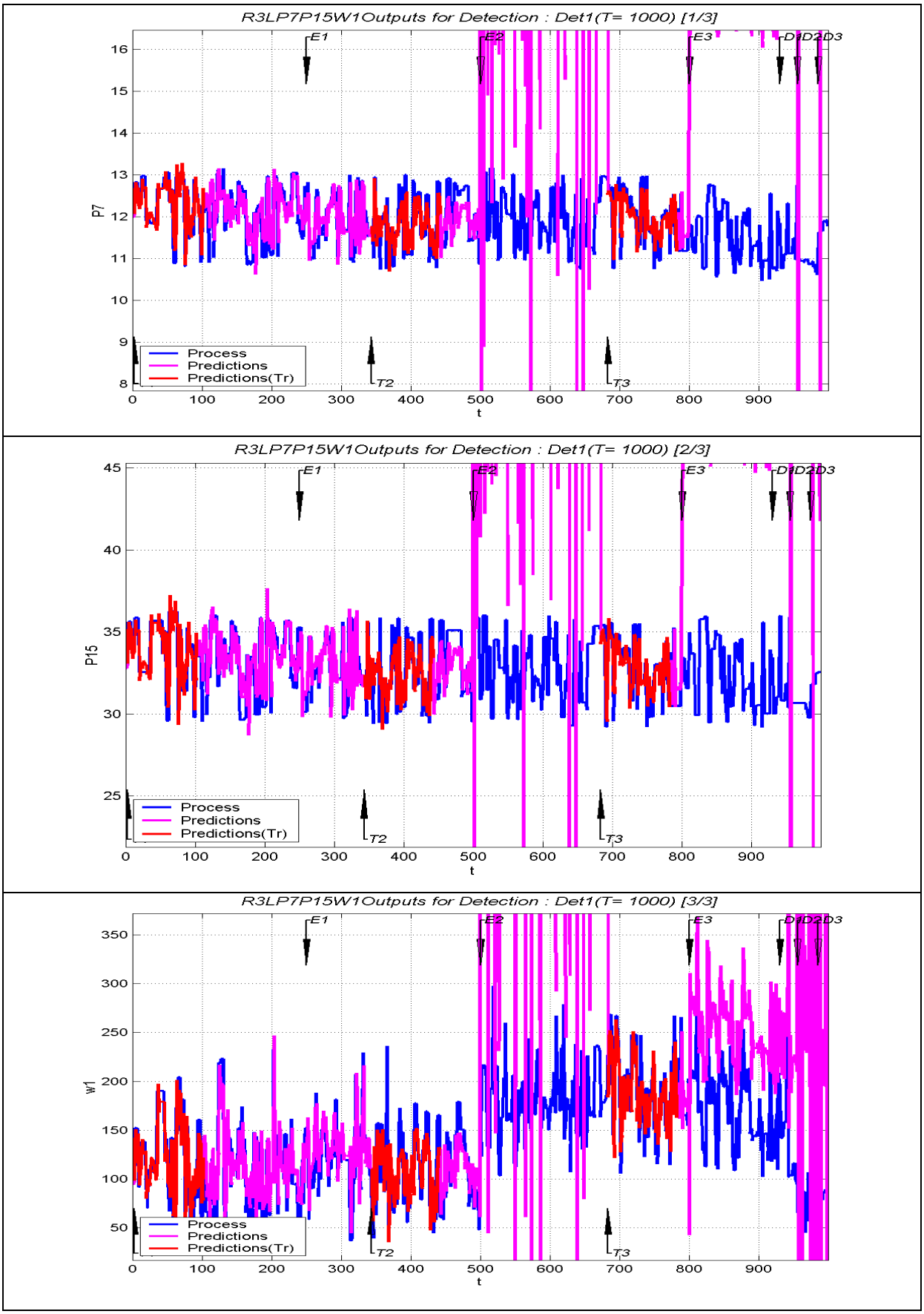


Figura 7.7 : Desempenho de Detector ARX para R3LP7P15W1
Azul: Processo; Vermelho: Predições em Treino; Lilás: Predições em Vigilância

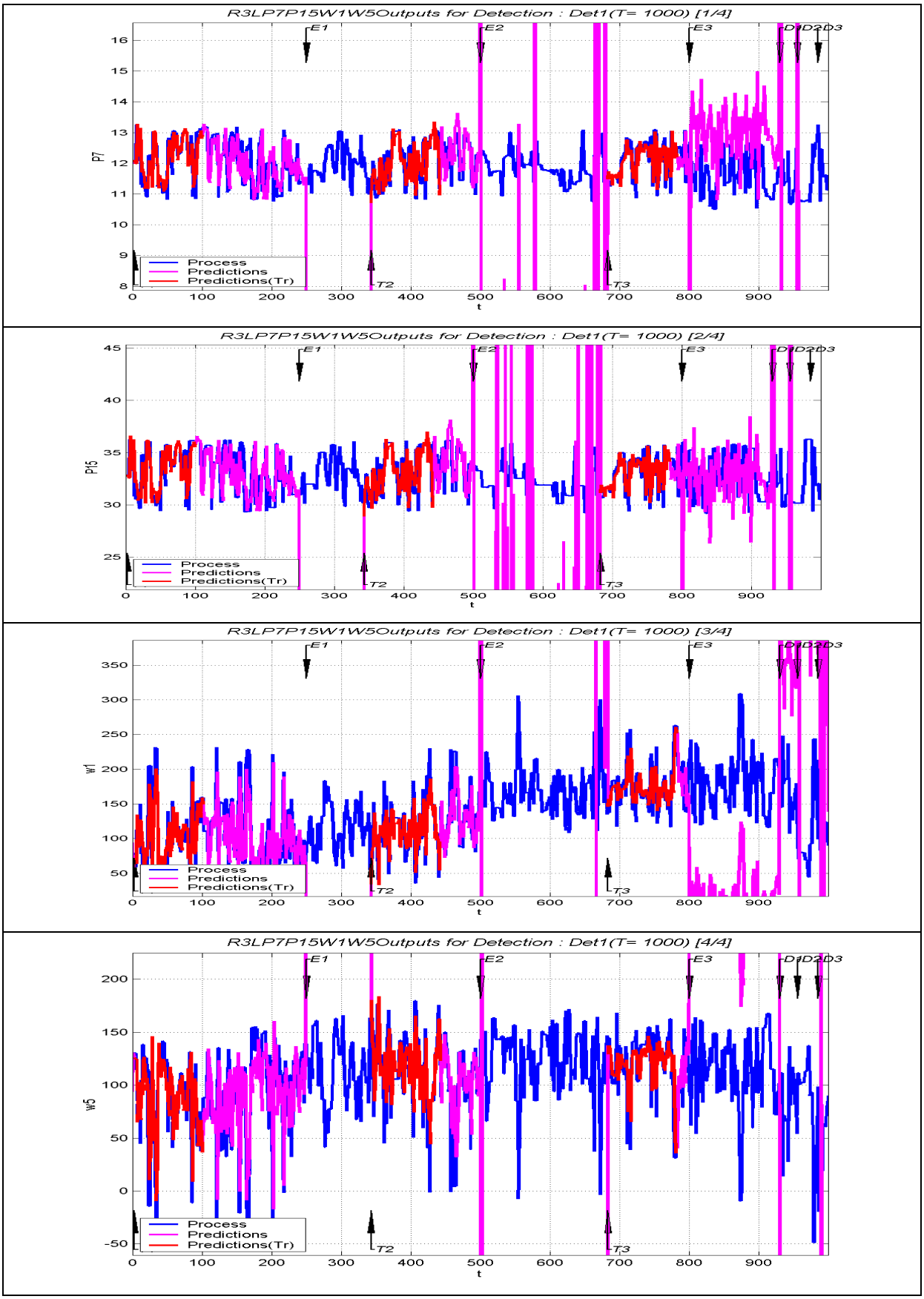


Figura 7.8 : Desempenho de Detector ARX para R3LP7P15W1W5
Azul: Processo; Vermelho: Predições em Treino; Lilás: Predições em Vigilância

7.2 Detecção, Localização e Quantificação de Vazamentos na Rede Reduzida

Nesta seção serão discutidas as ferramentas de Localização (identificação do *tubo k* afetado e posição fracionária do vazamento, $x = L_F / L_k$) e Quantificação (determinação do diâmetro de ruptura, D_F) baseadas em modelos ARX.

Admitindo-se que a detecção identificou corretamente o instante de manifestação de uma ruptura, deve-se lançar o procedimento de Localização e Quantificação. O preditor ARX-MIMO para esta missão é chamado de Localizador. Sua diferença básica em relação ao ARX-MIMO Detector consiste em possuir um segundo filtro *X* (ordem *p*) para processamento de um segundo "feixe" de entradas associadas aos "sinais" de vazamento *UD* e *UL* "emitidos" pelos *tubos* avariados. A estrutura do preditor ARX-MIMO DUAL utilizado nos Localizadores é mostrada na equação 7.5 abaixo:

$$\underline{\hat{Y}}(t+1) = \underline{\underline{B}}(q)\underline{U}(t+1) + \underline{\underline{C}}(q)\underline{H}(t+1) + q\left[1 - \underline{\underline{A}}(q)\right]\underline{Y}(t) \qquad \text{eq. 7.5}$$

Onde $\underline{U}(t)$, $\underline{Y}(t)$ são, respectivamente, os mesmos vetores de *entradas* e de *saídas* do processo usados no ARX-Detector. A equação 7.5 mostra que o Localizador é alimentado com um segundo vetor de *entradas* $\underline{H}(t)$, tamanho *nh x 1*, o qual é processado pelo filtro matricial $\underline{\underline{C}}(q)$ com tamanho *ny x nh*.

O vetor *dual* de *entradas*, $\underline{H}(t)$, corresponde às alterações conhecidas do processo com respeito a um conjunto com *h tubos* pré-selecionados como mais propensos a manifestar vazamentos graves. Nesta lista, cada *tubo k (k=1,...h, h ≤ E)*, com diâmetro D_k e comprimento L_k, considerado candidato a desenvolver vazamentos, emite um vetor de sinal de vazamento ($\underline{H}_k$) com dois números adimensionais: os sinais $UD_k(t)$, $UL_k(t)$. Estas informações são apresentadas abaixo:

$$\underline{H}^T = [\,\underline{H}_1^T \ldots \underline{H}_h^T\,] \quad , \quad \underline{H}_k(t) = \begin{bmatrix} UD_k(t) \\ UL_k(t) \end{bmatrix} \qquad \text{eq. 7.6}$$

O sinal de vazamento $\underline{H}_k$, "emitido" pelo *tubo k*, é identicamente nulo enquanto não houver vazamento. A partir do instante t_E em que um vazamento (D_F, L_F) lá se instalar, $\underline{H}_k$ sofrerá alteração de valor conforme descrito na equação 7.7 (admitimos que o vazamento cessará para $t \geq t_D$):

$$\underline{H}_k(t) = \begin{bmatrix} UD_k(t) \\ UL_k(t) \end{bmatrix} = \begin{cases} \begin{bmatrix} 0 \\ 0 \end{bmatrix} & t < t_E,\ t \geq t_D \\ \\ \begin{bmatrix} (D_F / D_k)^{2.5} \\ (D_F / D_k)(1 + L_F / L_k) \end{bmatrix} & t_D > t \geq t_E \end{cases}$$

eq. 7.7

onde $D_F(t)$ e $L_F(t)$ são, respectivamente, o diâmetro da ruptura (*m*) e sua posição (*m*) no *tubo k* tomada a favor da orientação do mesmo.

Foram testadas várias fórmulas empíricas para *UD* e *UL*. Aparentemente, o formato utilizado no lado direito da equação acima apresentou melhores características de robustez com respeito à identificação do ARX-Localizador e com respeito à sua capacidade de resolver o problema de Localização e Quantificação da ruptura. Note que resolver o problema consiste em estimar o valor correto de $\underline{H}(t)$ para um cenário de Detecção bem sucedida. Estima-se assim tanto a localização comparando a severidade do vazamento. Note também, que durante a "vigência" do vazamento no *tubo k* o valor de $\underline{H}_k$ é não nulo e constante no tempo.

Ao longo da busca a ferramenta tentará “sintonizar” seu $\underline{H}_k$ de modo a buscar aderência à série temporal da rede danificada. Isto será conduzido através de problemas seqüenciais de otimização para cada *tubo* candidato a romper. Isto é, admitindo-se vazamento no *tubo k,* busca-se o valor ótimo de $\underline{H}_k$ de modo a minimizar uma medida da falta de aderência de predições ($\hat{\underline{Y}}$) da ferramenta com respeito aos registros ($\underline{Y}$) da série temporal da rede danificada. O *tubo* para o qual melhor grau de aderência foi atingido é o escolhido como provável portador de

vazamento. O valor correspondente de $\underline{H}_k$ contém a solução estimada do problema de Localização/Quantificação.

7.2.1 Identificação do Localizador em Série Temporal Multi-Eventos na Rede Reduzida

O Localizador/Quantificador (ou apenas "Localizador") deve ter seus parâmetros dos filtros $\underline{\underline{A}}(q)$, $\underline{\underline{B}}(q)$, $\underline{\underline{C}}(q)$ estimados em um árduo treinamento que envolve gerar uma série temporal do processo admitindo-se diversas incidências de vazamentos (variando-se tubo, diâmetro de furo e posição da ruptura) ao mesmo tempo em que são aplicadas *entradas* típicas de serviço normal e registradas valores das *saídas* habituais do sistema. Dada a impossibilidade de varrer-se exaustivamente todas as condições de vazamento, os parâmetros de ruptura submetidos serão sorteados a partir de faixas de valores razoáveis para D_F e L_F .

Consideramos que na rede apenas quatro tubos (*h=4*) são sujeitos a apresentarem rompimentos. A partir disto, a ferramenta será treinada com os dados de Série Temporal com Programação Multi-Eventos caracterizada por:

- Identificação da Série Temporal : RME P7P15W1W5
- Duração da Série : $nt = 2150$ instantes
- Especificações nas extremidades da rede: $P_1^{Ref} = 35bar;$ $P_5^{Ref} = 40bar;$ $P_9^{Ref} = 4bar;$ $W_{15}^{Ref} = -10kg/s$, significando utilização de todos os caminhos de escoamento na rede, com os dois vértices fontes e os dois vértices de recebimento em serviço;
- *Entradas* dinâmicas: $P_1(t)$, $P_5(t)$, $W_{15}(t)$, P_9 Fixo
- *Saídas* do processo : *P7, P15, W1, W5*,
- *Tubos* a receberem rupturas : 3, 6, 9 e 11 (*h=4*)
- Os primeiros 80 e os últimos 80 instantes da Série temporal não apresentam vazamentos ativos;

- Cada tubo recebe 4 rupturas ativas em faixas de tempo não superpostas, com parâmetros D_F, L_F sorteados uniformemente nos domínios: $D_F \in [5mm, 25mm]$, $L_F / L_k \in [10\%, 90\%]$;
- A ciclagem de sorteio de rupturas corresponde a percorrer a lista de *tubos* 3, 6, 9, 11 quatro vezes até completar 4 rupturas por tubo;
- Serão geradas 4 X 4 = 16 rupturas ao longo da série, cada uma delas com aproximadamente 120 instantes de vigência;
- Não haverá rupturas concomitantes; duas rupturas vizinhas em tempo estarão separadas por pelo menos 7 instantes de tempo;

As entradas dinâmicas $\underline{U}(t), \underline{H}(t)$ aplicadas no Localizador, ao longo da Série Temporal são mostradas nas Figuras 7.9 e 7.10. A Figura 7.9 apresenta a distribuição temporal de entradas do processo em RMEP7P15W1W5. A Figura 7.10 apresenta a distribuição temporal de Sinais de Vazamento (UD, UL) ao longo de RMEP7P15W1W5. As Figuras 7.11, 7.12, 7.13 e 7.14 apresentam as séries temporais de diversas variáveis do processo ao longo de RMEP7P15W1W5.

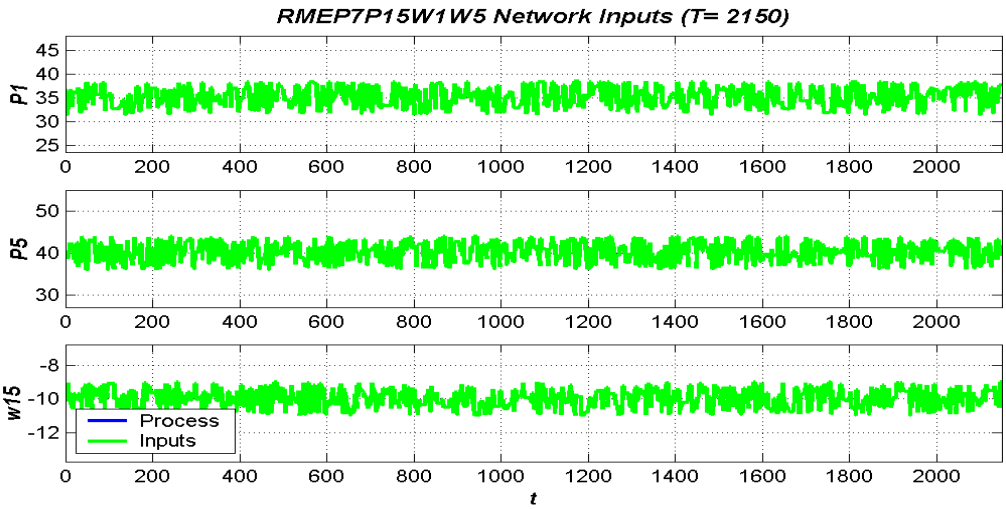


Figura 7.9 : Séries Temporais de Entradas para RMEP7P15W1W5

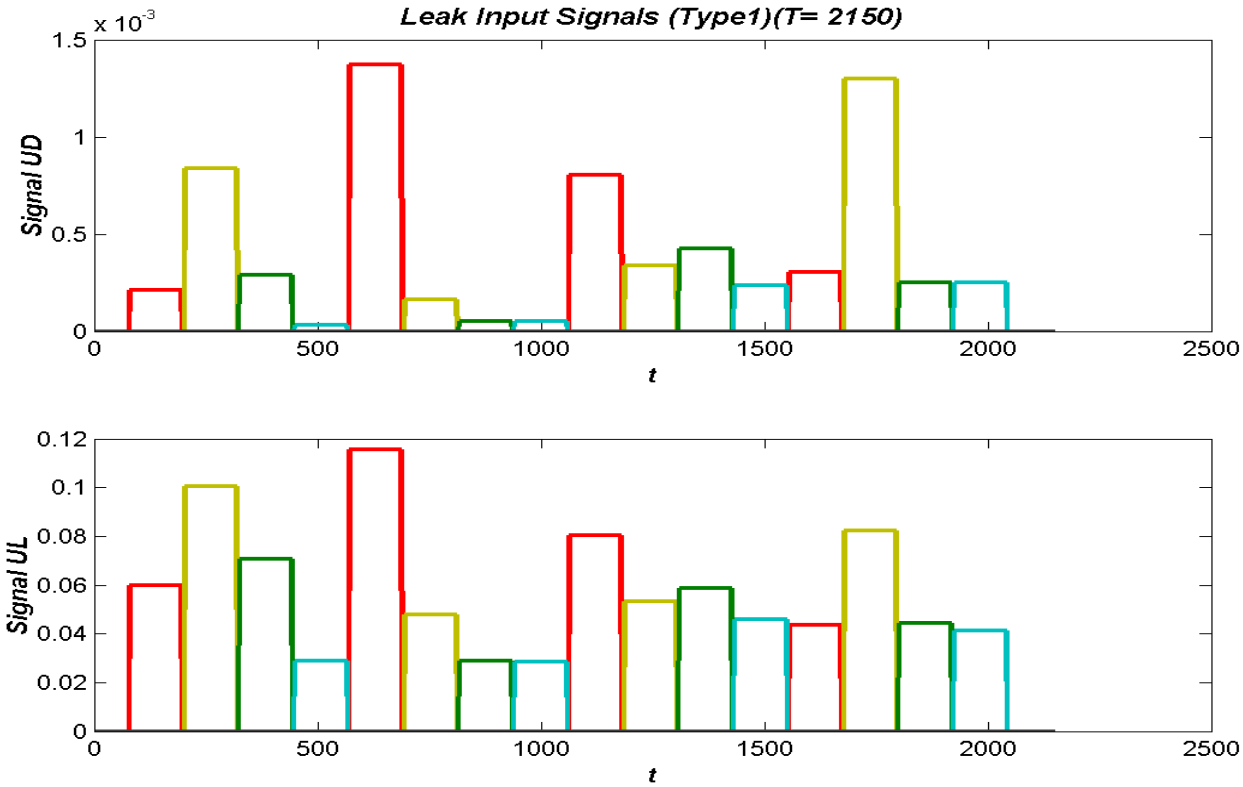


Figura 7.10 : Séries Temporais de Sinais de Vazamentos em RMEP7P15W1W5 (*Tubo* 3: Vermelho; *Tubo* 6: Amarelo; *Tubo* 9: Verde; *Tubo* 11: Azul Claro)

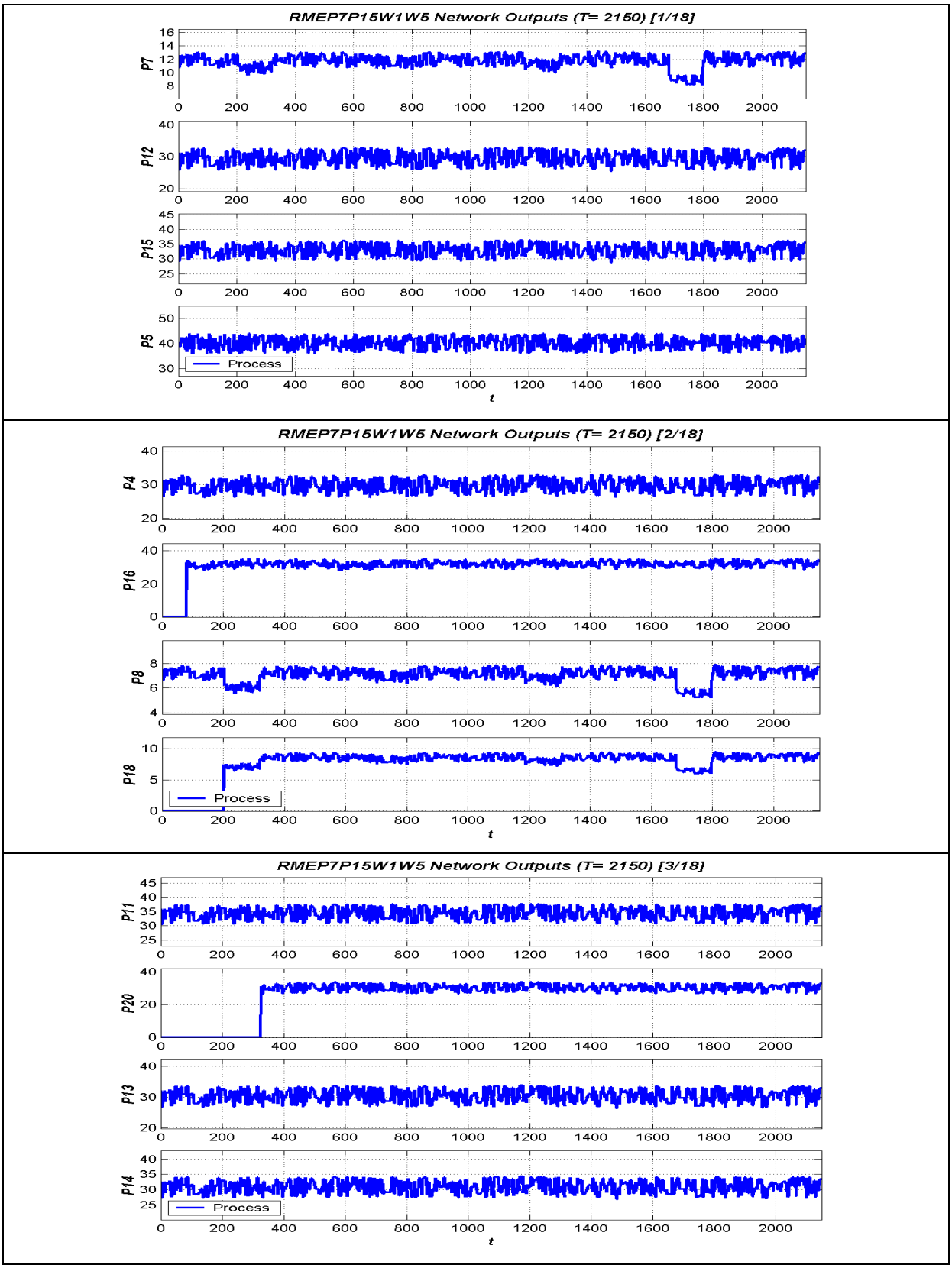


Figura 7.11 : Séries Temporais de Pressões em RMEP7P15W1W5

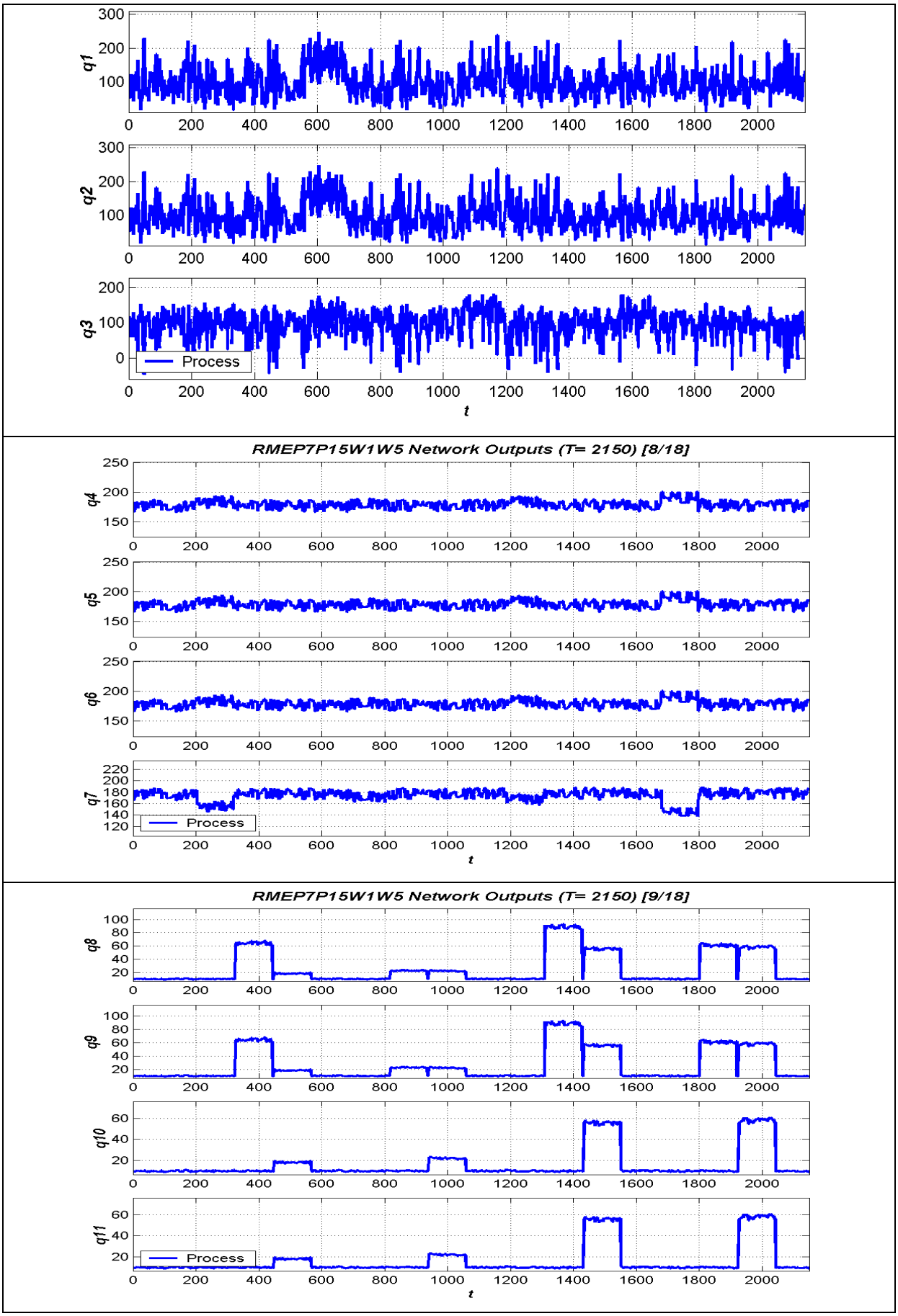


Figura 7.12 : Séries Temporais de Vazões de Aresta em RMEP7P15W1W5

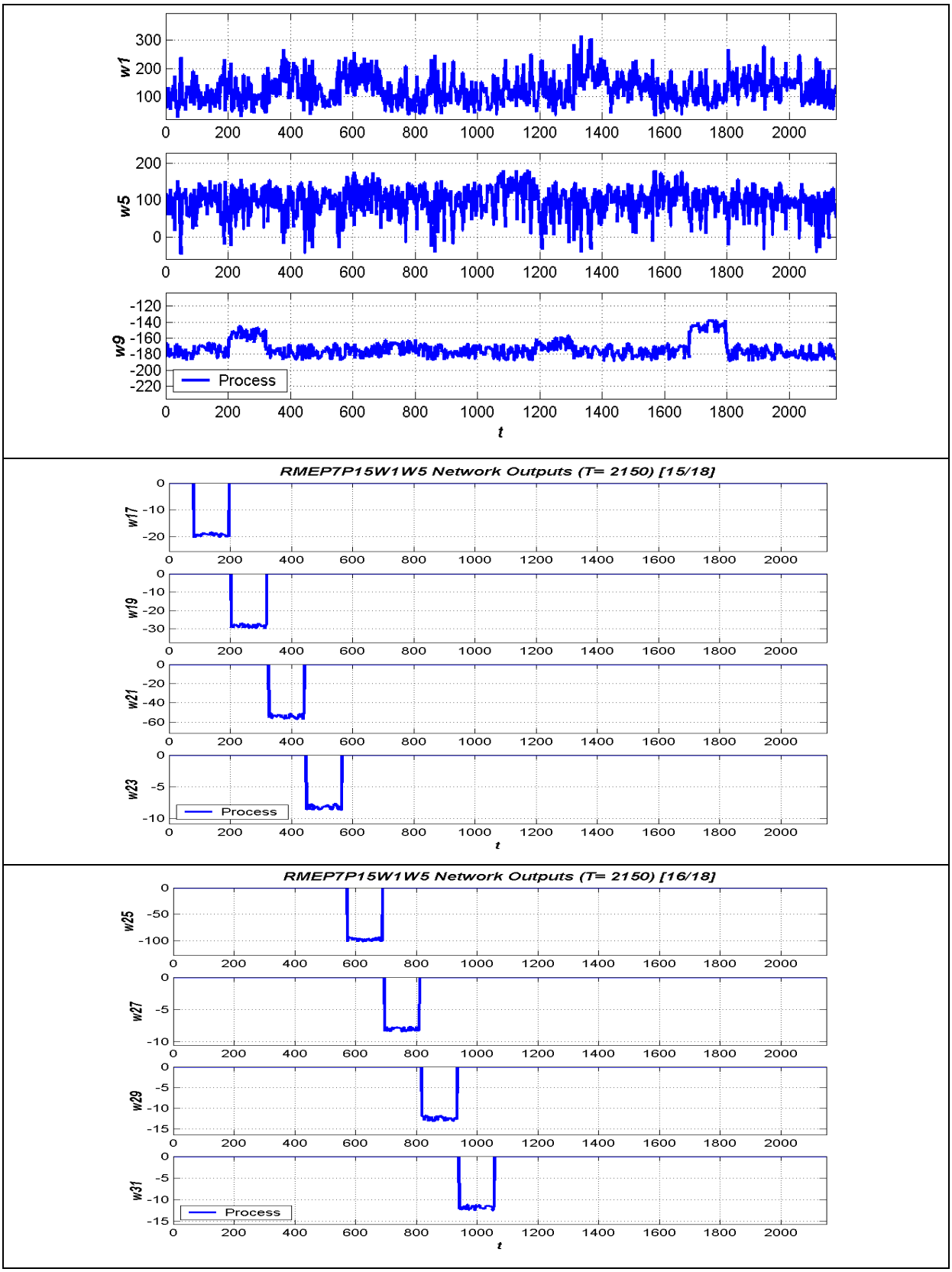


Figura 7.13: Séries Temporais de Vazões Externas de Vértices em RMEP7P15W1W5

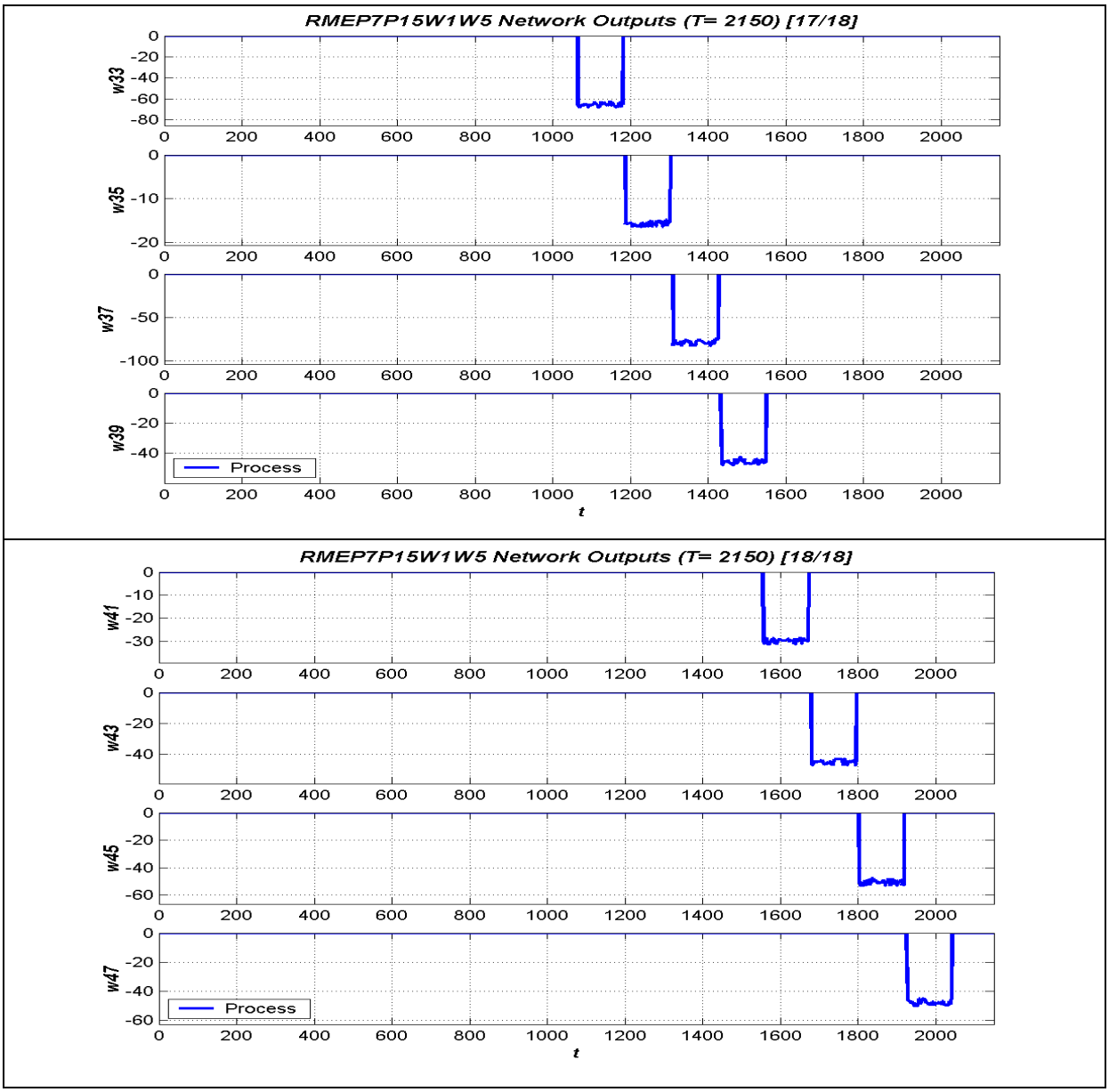


Figura 7.14: Séries Temporais de Vazões Externas de Vértices em RMEP7P15W1W5

A série RMEP7P15W1W5 é utilizada para identificar parâmetros do Localizador pelo procedimento usual de Identificação ARX onde minimiza-se a discrepância entre a série de amostragens do processo $\{\underline{Y}(t_1),\cdots,\underline{Y}(t_{nt})\}$ e a série correspondente de predições $\{\hat{\underline{Y}}(t_1),\cdots,\hat{\underline{Y}}(t_{nt})\}$. As ordens dos filtros $\underline{\underline{A}}(q)$, $\underline{\underline{B}}(q)$, $\underline{\underline{C}}(q)$ do Localizador foram escolhidas, respectivamente, como $n=3,\ m=3,\ p=1$. Os tamanhos dos vetores de *entradas dinâmicas* ($\underline{U}(t)$), *entradas de vazamentos* ($\underline{H}(t)$) e *saídas* ($\underline{Y}(t)$) são, respectivamente, *nu=3, nh=8, ny*=4. A ordem 1 do filtro dual justifica-se pelo fato dos sinais de vazamento serem

constantes no tempo uma vez estabelecidos. Os tamanhos dos filtros $\underline{\underline{A}}(q)$, $\underline{\underline{B}}(q)$, $\underline{\underline{C}}(q)$ são respectivamente *4 x 4, 4 x 3* e *4 x 8*. Levando-se em conta as ordens (3,3,1) acima, a identificação do Localizador consiste em um problema de estimação (linear) de 16*3+12*3+32 = 116 parâmetros. A construção do Localizador utilizará faixa de treino de 2000 instantes (*NT=2000)* e uma pequena faixa de validação (*NI=150*). Desta forma, os primeiros 2000 instantes da série RMEP7P15W1W5 (total de 2150 instantes) serão usados nesta identificação, correspondendo a 2000 X 4 = 8000 predições sob ajuste.

Os resultados da identificação do ARX(3,3,1)-Localizador são apresentados nas Figuras 7.15.e 7.16 em termos de predições em treino (vermelho) versus valores amostrados (azul) para as 4 saídas de processo *P1, P15, W1* e *W5.* Na pequena região de validação as predições do Localizador estão em lilás.

É visível que a Identificação do Localizador foi de boa qualidade. Sua aderência é perfeita ao longo dos 2000 instantes de treino e na região de validação.

A qualidade do Localizador produzido pode ser testada através do seguinte artifício. Uma vez concluído seu processo de identificação, sua aderência ao processo RMEP7P15W1W5 somente ocorre em função do Localizador ser corretamente informado acerca de vazamentos através do feixe dual de entradas ($\underline{H}(t)$). Portanto, se as predições do Localizador forem testadas contra RMEP7P15W1W5 com fornecimento sistemático de $\underline{H}(t)=\underline{0}\ \ (\forall t)$ (isto é, sinais informam que *não há* vazamentos) uma significativa perda de aderência deverá ocorrer comparativamente ao mesmo teste em que os sinais de vazamento são alimentados corretamente. Obviamente, quanto mais bem sintonizado estiver o Localizador maior a perda de aderência observada neste teste.

A Figura 7.17 expõe razoável perda de aderência nas predições para as saídas *P1, P15, W1* e *W5* quando o Localizador é testado contra a mesma RMEP7P15W1W5 utilizada para identificá-lo, sem, contudo ser informado a respeito dos 16 vazamentos sucessivos ($\underline{H}(t)=\underline{0}\ \ (\forall t)$). Esta situação deve ser comparada com o quadro que se obtém quando o Localizador é informado corretamente acerca

do fluxo de vazamentos através do seu filtro dual, conforme mostrado na Figura 7.18.

No primeiro caso (sem sinais de vazamentos) o histograma de resíduos de predição contabiliza 5200 resíduos de pequena estatura, sendo a maior amplitude de resíduo da ordem de 180 com várias instâncias deste porte. Estando o Localizador corretamente informado acerca do fluxo de vazamentos, o histograma de resíduos de predição contabiliza quase 6000 resíduos de pequena estatura, com pouquíssimas instâncias de amplitude de resíduo superior a 180 e com inferior dispersão na distribuição destas amplitudes.

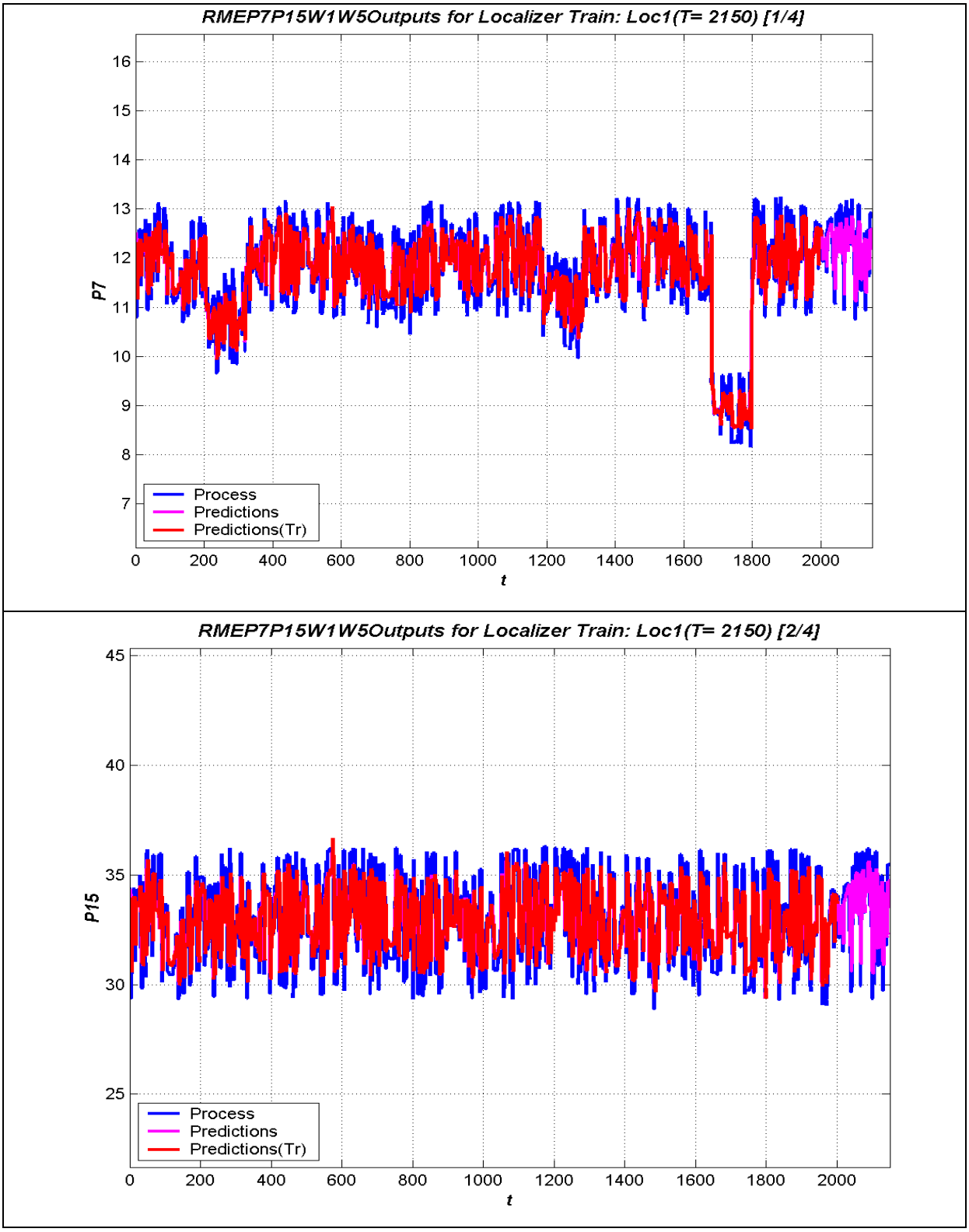


Figura 7.15: Identificação do ARX(3,3,1)-Localizador em RMEP7P15W1W5 (P_7, P_{15})
Azul: Processo; Vermelho: Predições em Treino; Lilás: Predições em Validação

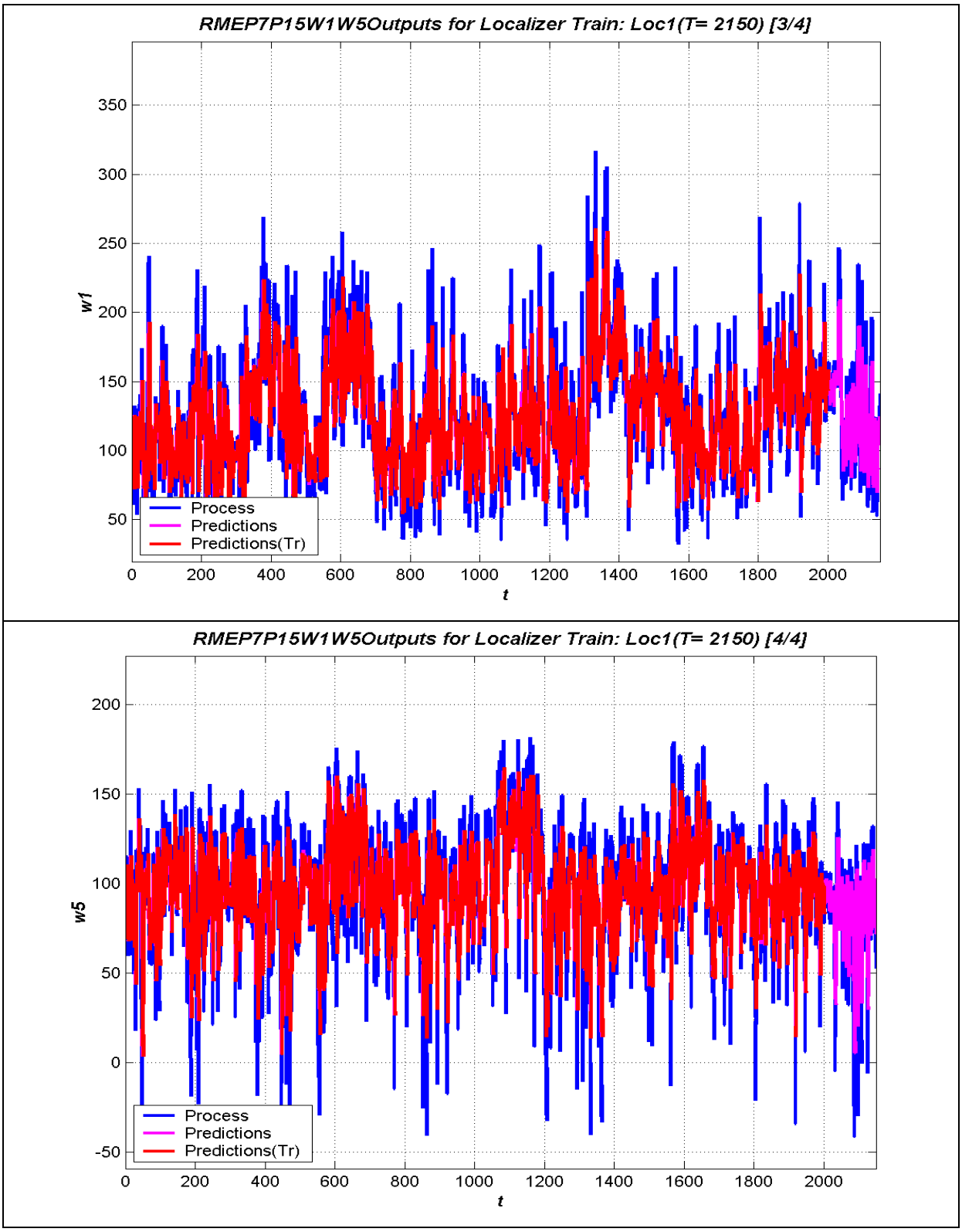


Figura 7.16: Identificação do ARX(3,3,1)-Localizador em RMEP7P15W1W5 *(W_1, W_5)*

Azul: Processo; Vermelho: Predições em Treino; Lilás: Predições em Validação

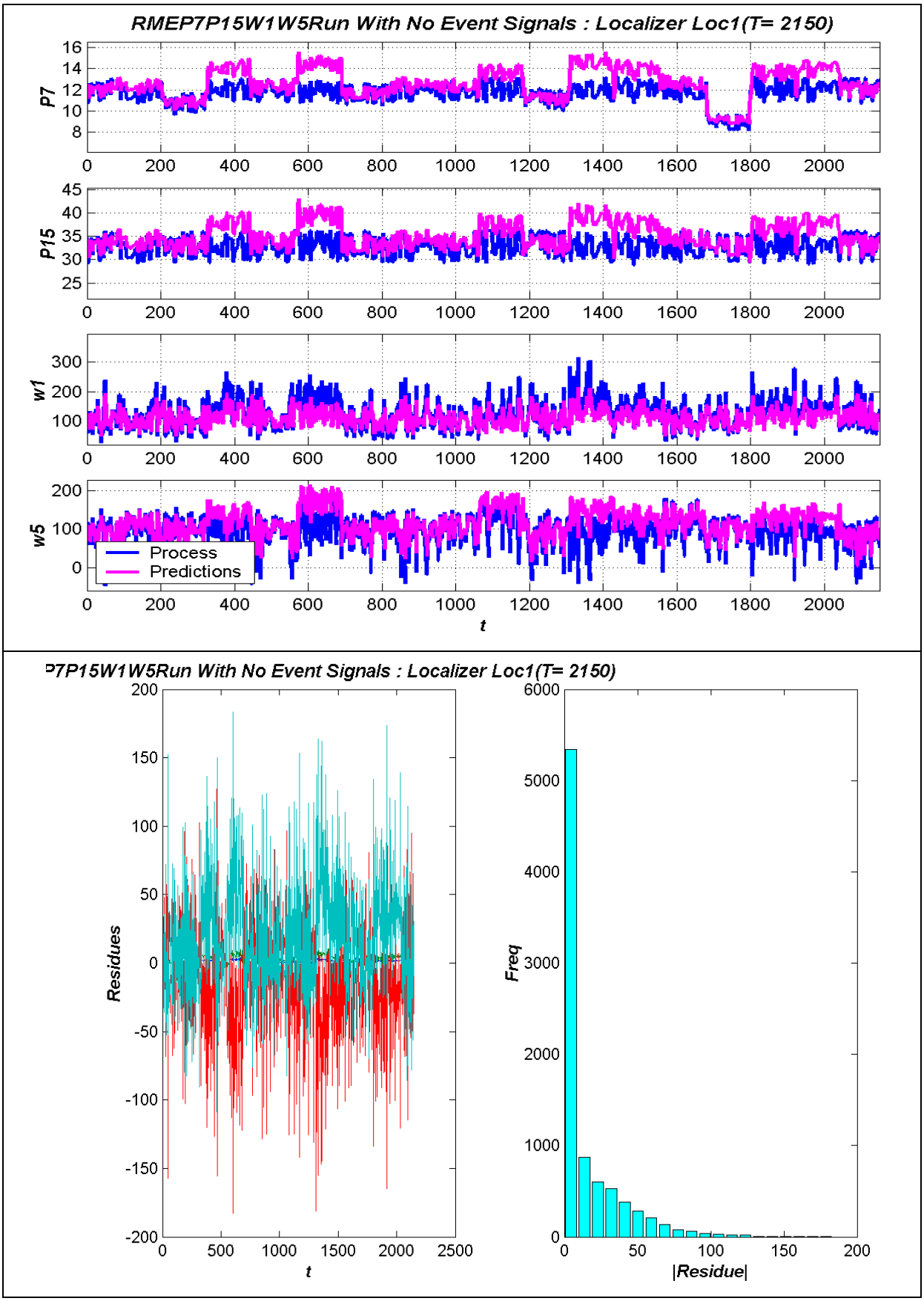


Figura 7.17: Aderência de ARX(3,3,1)-Localizador em RMEP7P15W1W5
Sem Receber Sinais de Vazamentos; Azul: Processo; Lilás: Predições em Validação

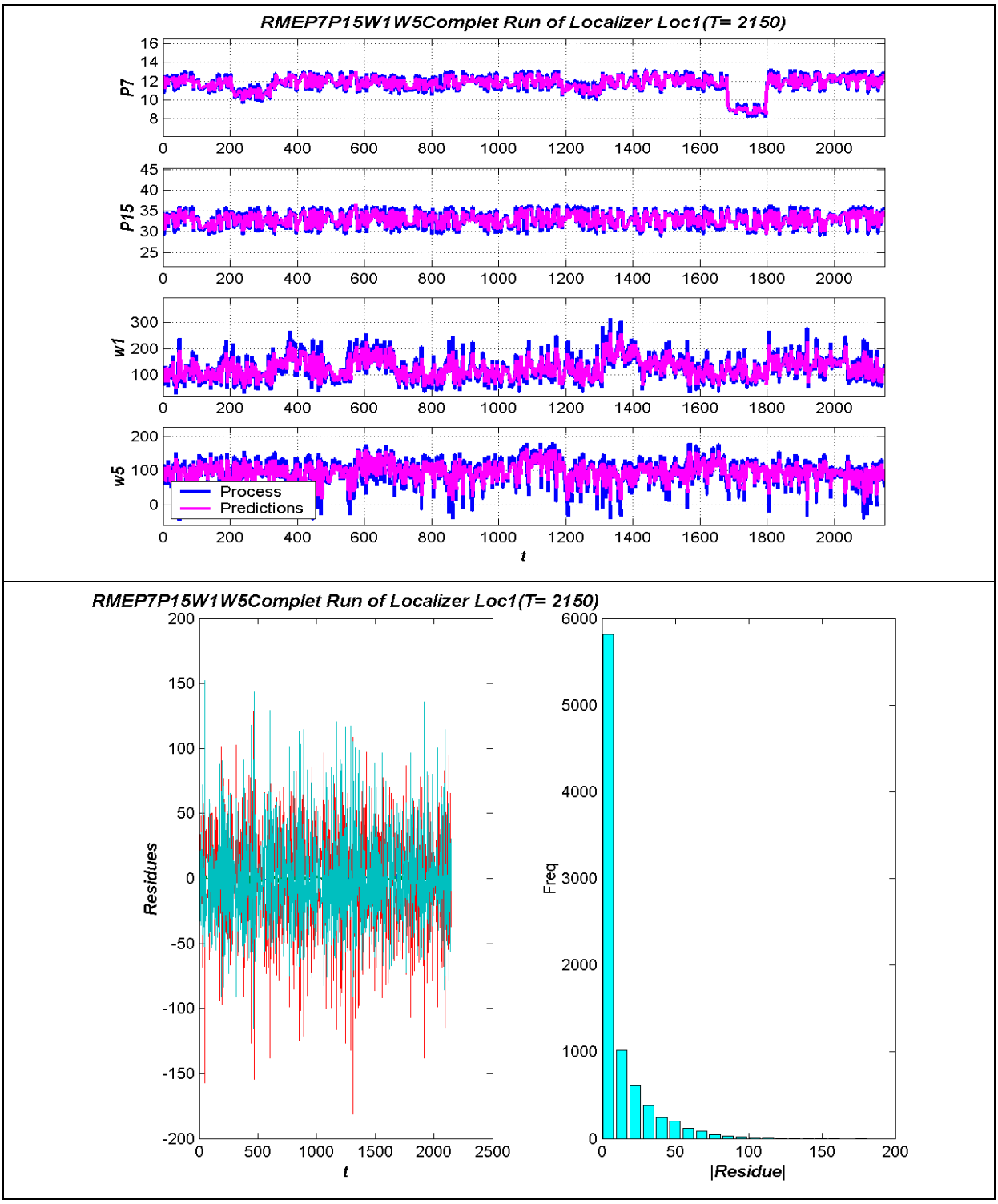


Figura 7.18: Aderência de ARX(3,3,1)-Localizador em RMEP7P15W1W5 Recebendo Sinais Corretos de Vazamentos; Azul: Processo; Lilás: Predições Validação

7.2.2 Teste dos Procedimentos ARX para Detecção, Localização e Quantificação

O procedimento completo de Detecção, Localização e Quantificação de vazamentos usando ferramenta ARX é testado aqui para a rede reduzida sob a ação de apenas um vazamento.

Utiliza-se série temporal descrita no item anterior, com 1000 instantes de tempo. Os Vazamentos 1 e 3 (ver Tabela 7.1) terão os seus instantes de ocorrência re-programados para datas *t>1000* de modo que não atuarão na série. Apenas o Vazamento 2 permanecerá com seus dados originais de programação: *Tubo* 9 danificado, na posição 40% com D_F de ruptura de 20mm em t_E=500. Adicionalmente, por uma questão de simetria, move-se o instante de cessação do vazamento para t_D=*800.* O Cenário deste teste é identificado como R1LP7P15W1W5 já que as séries de trabalho estarão operando sobre as saídas *P7, P15, W1* e *W5*. A representação da rede avariada é feita na Figura 7.19.

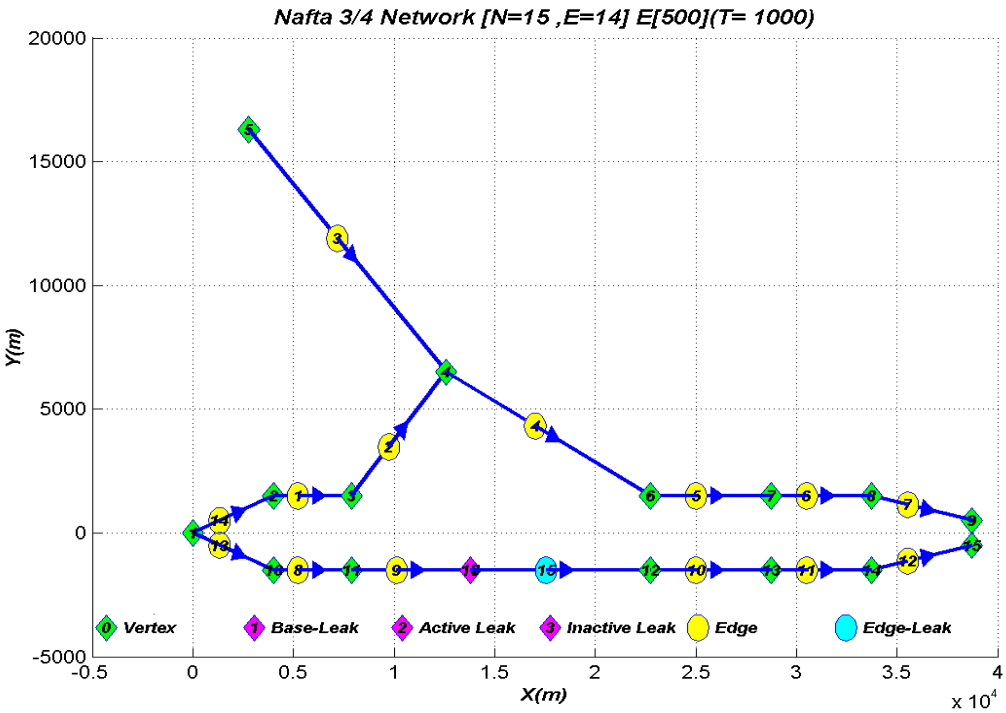


Figura 7.19: Rede com Programação de Avaria no *Tubo* 9 Original Série Temporal R1LP7P15W1W5

A Figura 7.20 apresenta as séries temporais de entradas dinâmicas de R1LP7P15W1W5. Séries de variáveis de processo são exemplificadas nas Figuras 7.21 e 7.22. Os instantes de "entrada" e de "desligamento" do vazamento são assinalados por flechas. Na maioria dos diagramas é difícil perceber-se os efeitos do vazamento por pura inspeção dos perfis temporais.

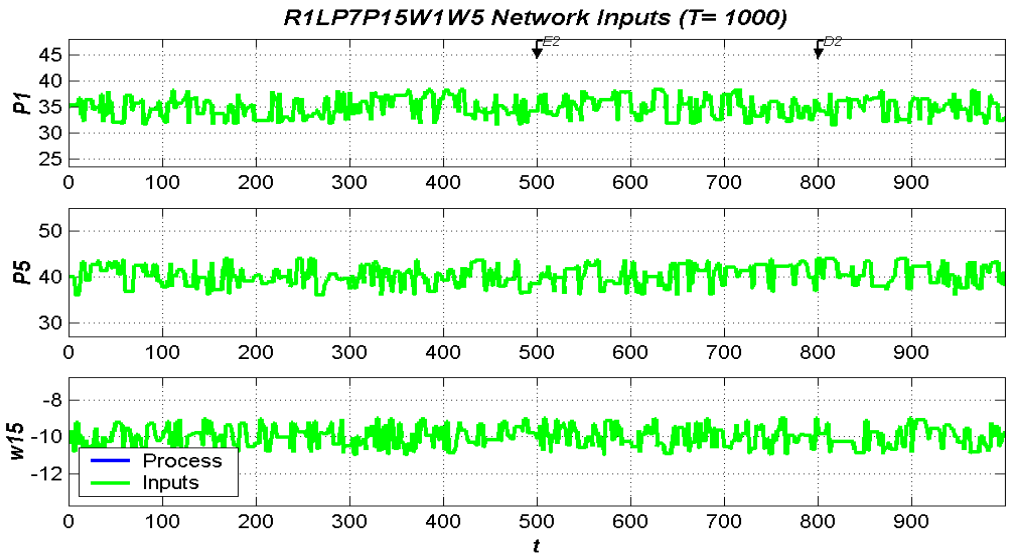


Figura 7.20: Séries Temporais de Entradas Dinâmicas para R1LP7P15W1W5

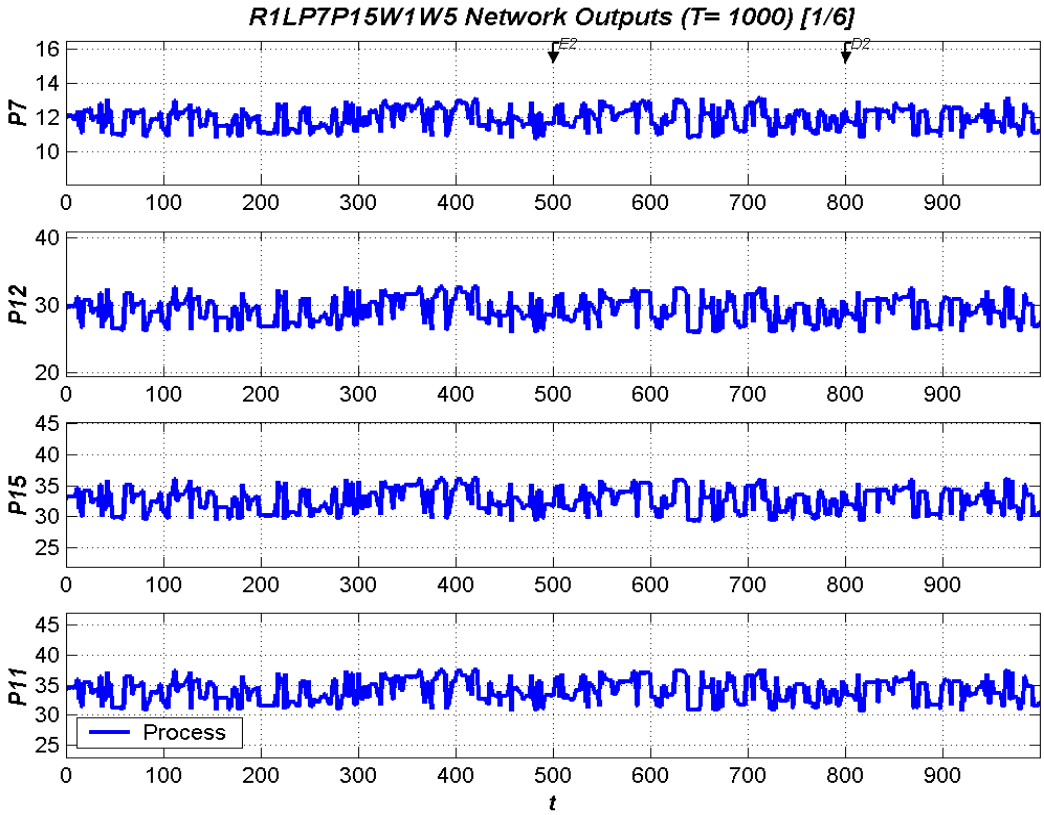


Figura 7.21: Séries Temporais de Pressões para R1LP7P15W1W5

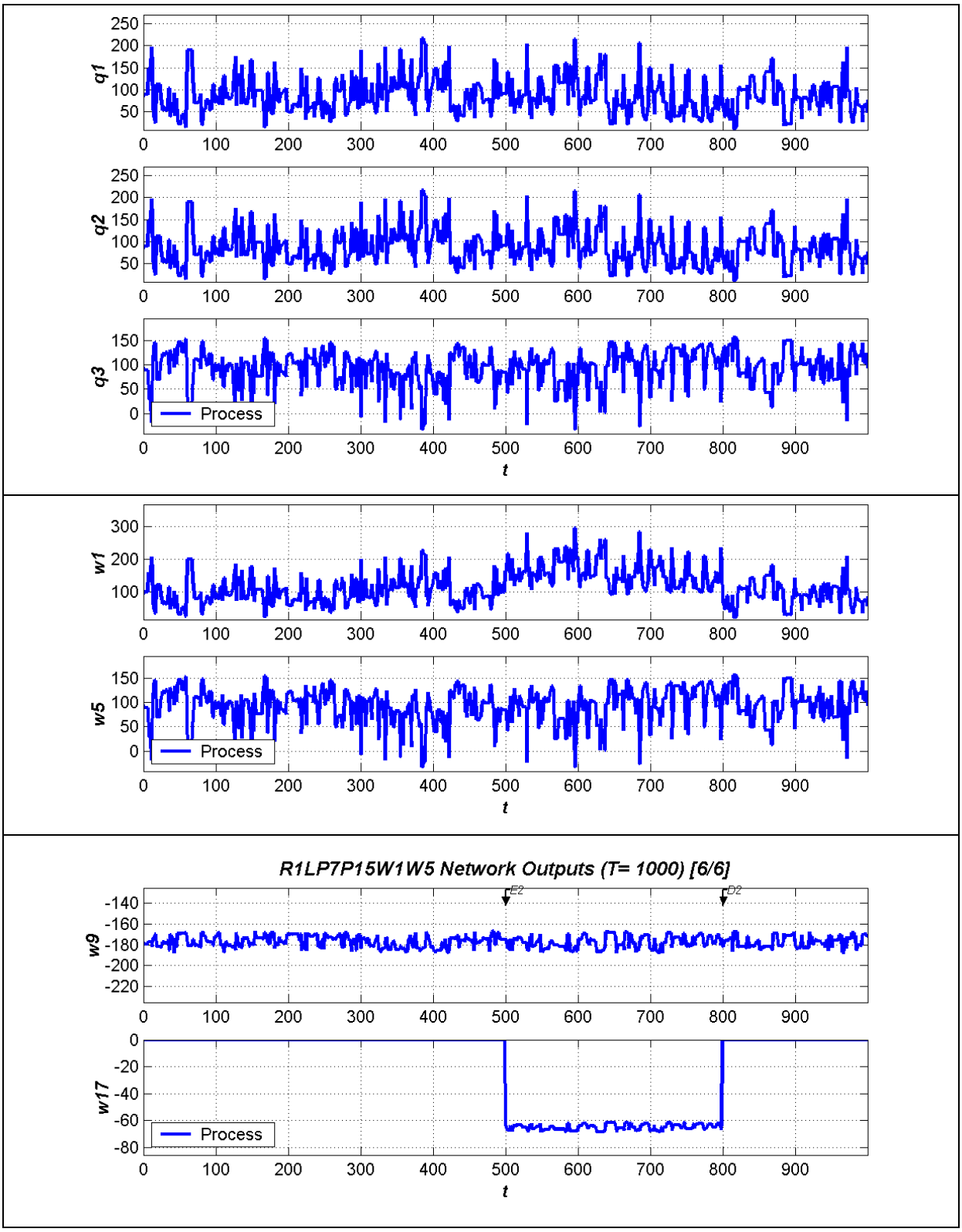


Figura 7.22: Séries Temporais de Vazões de Aresta e Externas de Vértices para R1LP7P15W1W5 (A Taxa de Vazamento é W_{17})

i. Fase de Detecção do Vazamento

O ARX(3,3)-Detector será programado para treino de 250 instantes (*NT=250*) e vigilância de 800 instantes (*NI=800*). Desta forma só haverá uma fase de treino que se encerrará a 250 instantes antes do evento de ruptura. Os resultados de detecção são mostrados nas Figuras 7.23 e 7.24 para as 4 saídas *P7, P15, W1* e *W5*.

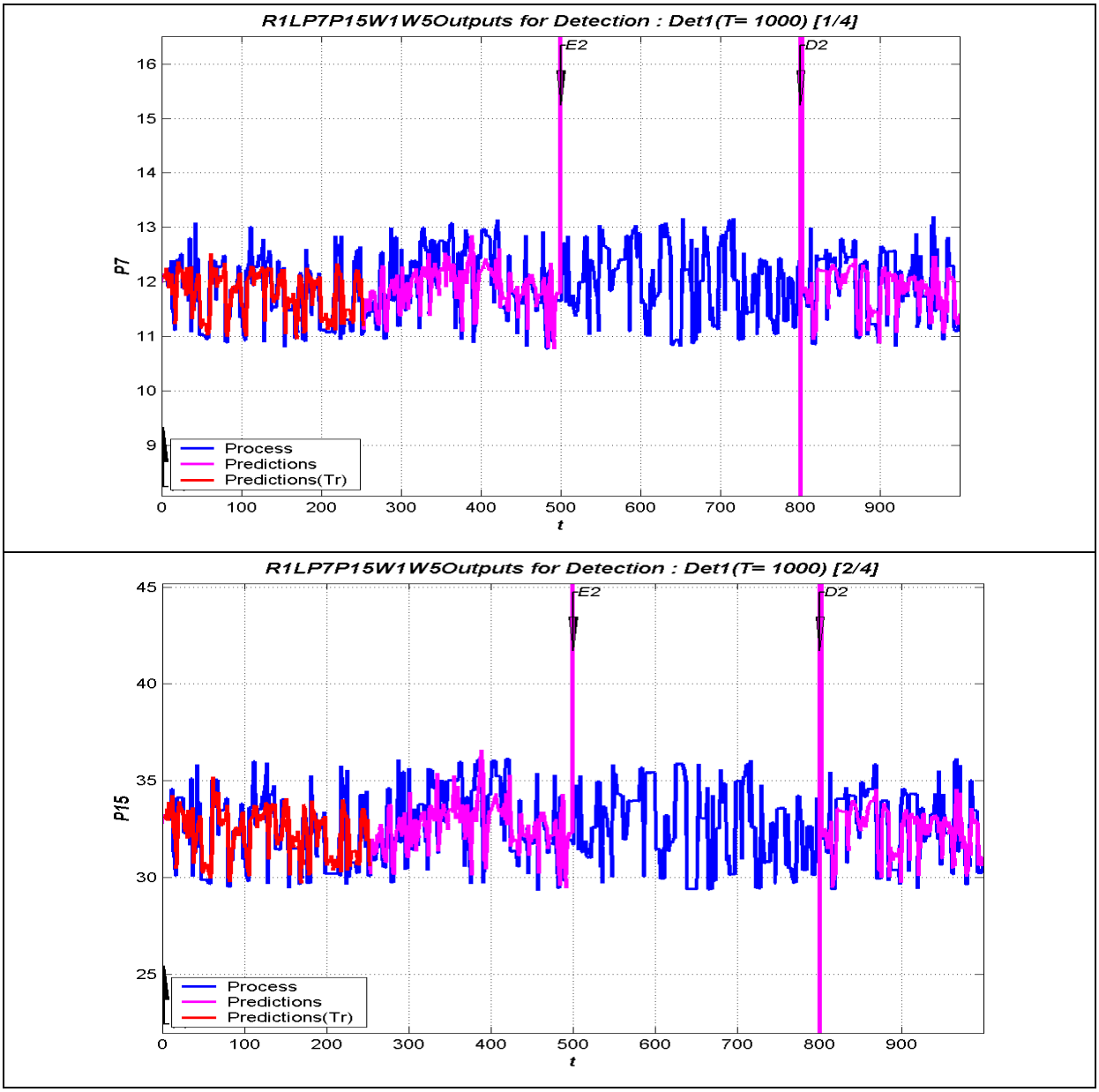


Figura 7.23: Predições do ARX-Detector vs Valores de Processo em R1LP7P15W1W5
Azul: Processo; Vermelho: Predições em Treino; Lilás: Predições em Vigilância

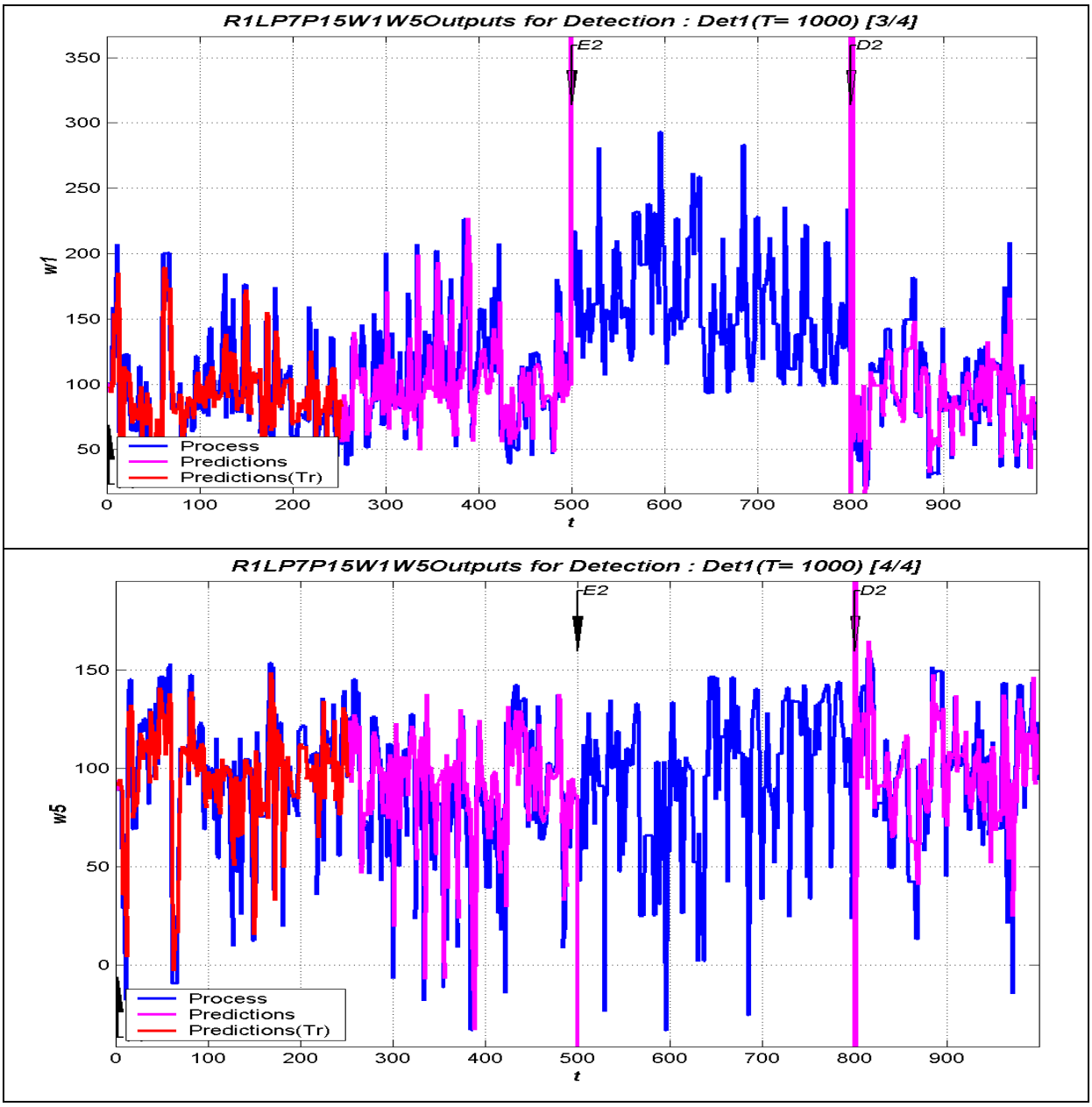


Figura 7.24: Predições do ARX-Detector vs Valores de Processo em R1LP7P15W1W5
Azul: Processo; Vermelho: Predições em Treino; Lilás: Predições em Vigilância

A Figura 7.25 resume a informação das Figuras 7.23 e 7.24, adicionando os limites de confiança 95% para predições corretas (em verde). Esta figura tem o objetivo de mostrar que na região de vigência do vazamento as predições do Detector ultrapassam as janelas de confiança 95%, descolando-se fortemente das séries de valores do processo. A detecção é convincente, precisa (em t=500) e totalmente bem-sucedida com respeito às quatro saídas. Um fato interessante: ao cessar a vigência do vazamento, as predições do ARX voltam a aderir perfeitamente às séries amostradas, indicando que se trata de um preditor bem ajustado à

essência do processo. Isto é, a forte perda de aderência está associada apenas à ação do fator agressor da rede.

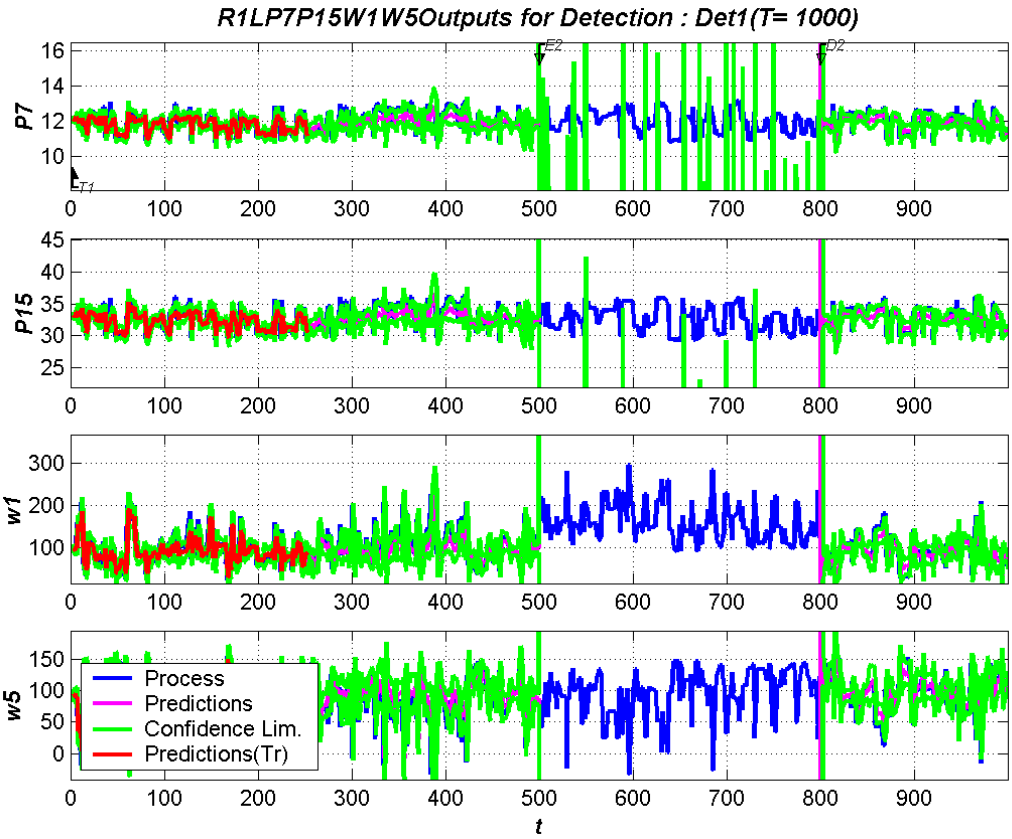


Figura 7.25: Predições do ARX-Detector vs Valores de Processo em R1LP7P15W1W5
Azul: Processo; Vermelho: Predições Treino; Lilás: Predições Vigilância; Verde: Intervalos de Confiança 95% para Predições Corretas

ii. Fase de Localização e Quantificação do Vazamento

O ARX(3,3)-Localizador/Quantificador calibrado será posto em ação para localizar e estimar parâmetros do vazamento capturado pelo detector. A validade de sua utilização está condicionada à manutenção da integridade física da rede, à manutenção do padrão probabilístico de *entradas* dinâmicas e à ocorrência de vazamento de características cobertas pela faixa de treino do localizador/quantificador (o que inclui o fato de que só haverá possibilidade de reconhecer vazamentos sobre os *tubos* 3, 6, 9 e 11).

Além disto, como este localizador foi construído com informações de vazamentos seqüenciais, não concomitantes, é razoável supor que ele somente tenha capacidade de lidar com o caso de apenas *um vazamento* atuante na rede; isto é, a tentativa de localizar uma segunda avaria na rede seria, provavelmente, mal-sucedida.

A data de incidência do evento foi determinada corretamente pelo detector como $t_E=500$. Este valor é transferido à fase de localização. São criados recortes das 4 séries temporais de amostragens (isto é, do processo) das saídas *P7, P15, W1* e *W5*. ao redor do instante estimado da ocorrência ($t_E=500$). No presente caso tais recortes foram definidos de $t=450$ até $t=700$. Note-se que é interessante que o recorte inclua trechos da série temporal *antes* da ocorrência e avance mais prolongadamente na faixa de tempo em que a rede está danificada.

Em seguida são executados problemas de otimização seqüenciais para cada *tubo* candidato a romper (no caso 3, 6, 9 e 11). Em cada um destes processos, busca-se o valor ótimo de $\underline{H}_k(t)$ ($k=1..h$) de forma a maximizar a aderência das predições do Localizador ao recorte extraído acima. O processo com máxima aderência define a solução do problema de Localização.

Sendo $t_E - a = 450,\ t_E + b = 700$ as extremidades em tempo do recorte de localização, o problema de estimação de $\underline{H}_k(t)$ corresponde à formulação:.

$$\underset{\{\underline{H}_k\}}{Min}\ \Phi_k = \left(\frac{1}{2}\right)\sum_{t=t_E-a}^{t_E+b} (\hat{\underline{Y}}(t,\underline{U},\underline{Y},\underline{H}_k) - \underline{Y}(t))^T\, \underline{\underline{\Omega}}(t)(\hat{\underline{Y}}(t,\underline{U},\underline{Y},\underline{H}_k) - \underline{Y}(t)) \qquad \text{eq. 7.8}$$

Nesta expressão, $\hat{\underline{Y}}(t,\underline{U},\underline{Y},\underline{H}_k)$ representa a predição de *saídas* feita pelo Localizador no instante *t* a partir da série de *entradas dinâmicas* ($\underline{U}$), da série de amostragens do processo ($\underline{Y}$), e do valor de parâmetros do vazamento supostamente sobre o *tubo k*. A matriz $\underline{\underline{\Omega}}(t)$ é um ponderador de tamanho *ny x ny*

arbitrado para reduzir o problema de escalas muito diferentes nos valores de variáveis de saída ao longo da série (eventualmente, usa-se $\underline{\underline{\Omega}}(t) = \underline{\underline{I}}$).

Na presente aplicação o problema foi resolvido para cada um dos *tubos* candidatos em seqüência. Em cada minimização foi empregado o Método Simplex Flexível para otimização não-linear em 2 variáveis (UD_k, UL_k) com transformação elementar de variáveis para impor limites inferior e superior em cada variável. Tais limites de busca correspondem aproximadamente aos limites de D_F e L_F usados na construção do localizador ($5mm \leq D_F \leq 30mm$; $10\% \leq L_F / L_k \leq 90\%$). O número máximo de cálculos do objetivo foi estipulado em 200 para todos os casos. O Método Simplex Flexível é disponibilizado através da rotina FMINSEARCH do Matlab R12.

Resultados acerca do trajeto de otimização são mostrados nas Figuras 7.26, 7.27, 7.28 e 7.29, para as buscas de vazamento nos *tubos* 3, 6, 9 e 11, respectivamente. A Tabela seguinte sumariza os resultados obtidos.

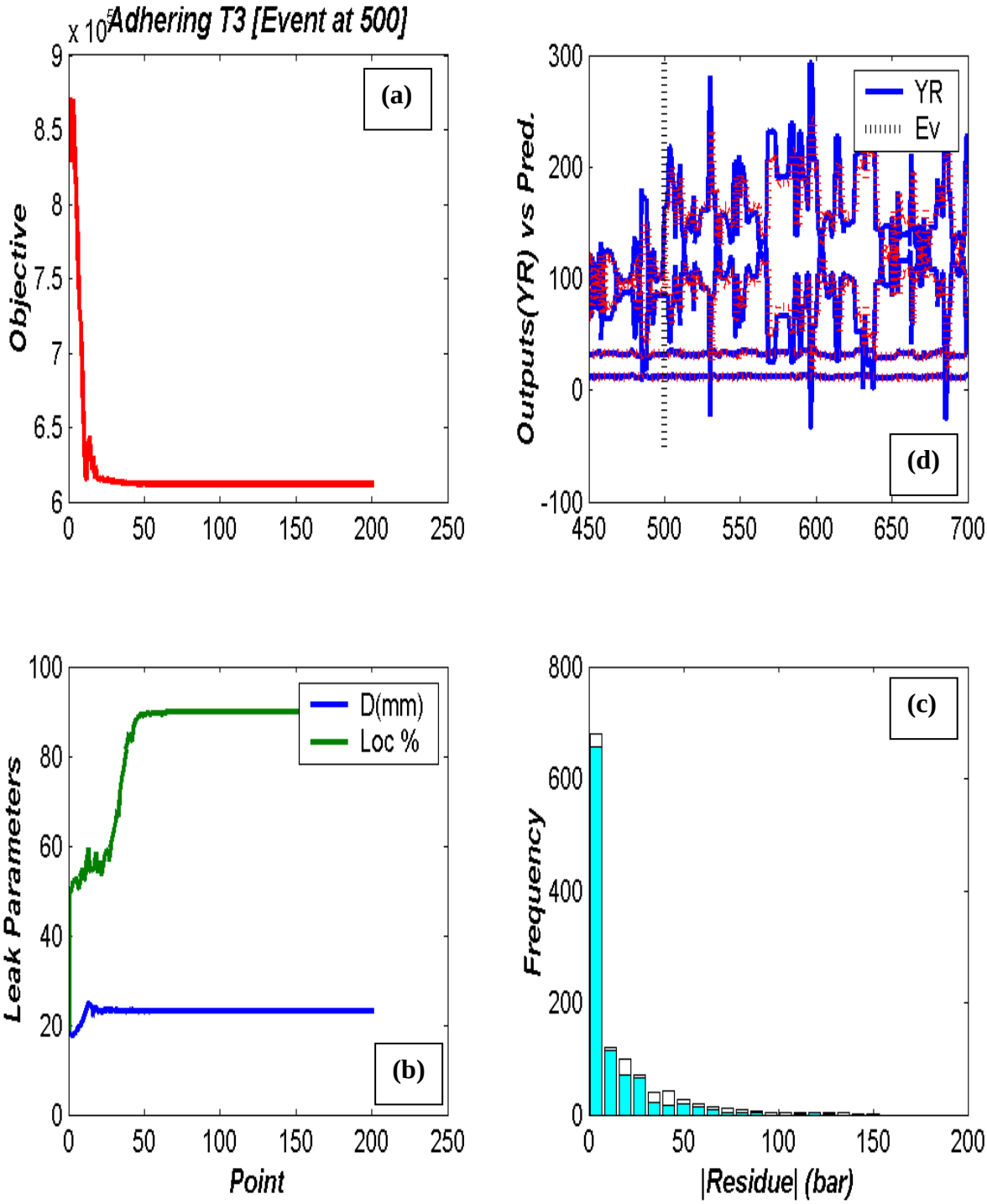


Figura 7.26: Busca de Localização de Vazamento para *Tubo* 3
(a) Trajeto do Objetivo; (b) Trajeto de Parâmetros; (c) Histograma de Resíduos Finais;
(d) Recorte de Amostragens (Y_R) versus Predições (Vermelho)

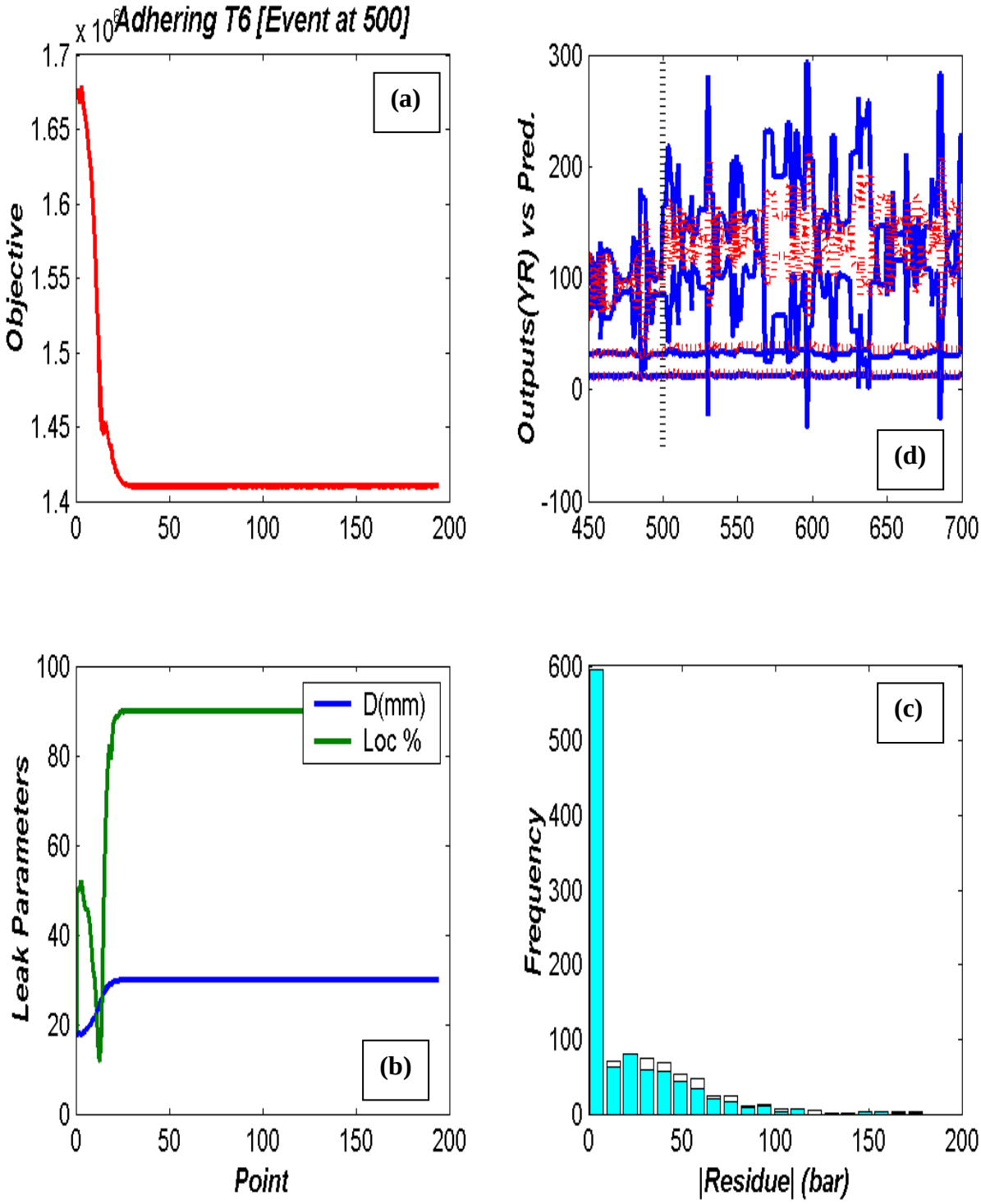


Figura 7.27: Busca de Localização de Vazamento para *Tubo* 6
(a) Trajeto do Objetivo; (b) Trajeto de Parâmetros; (c) Histograma de Resíduos Finais;
(d) Recorte de Amostragens (Y_R) versus Predições (Vermelho)

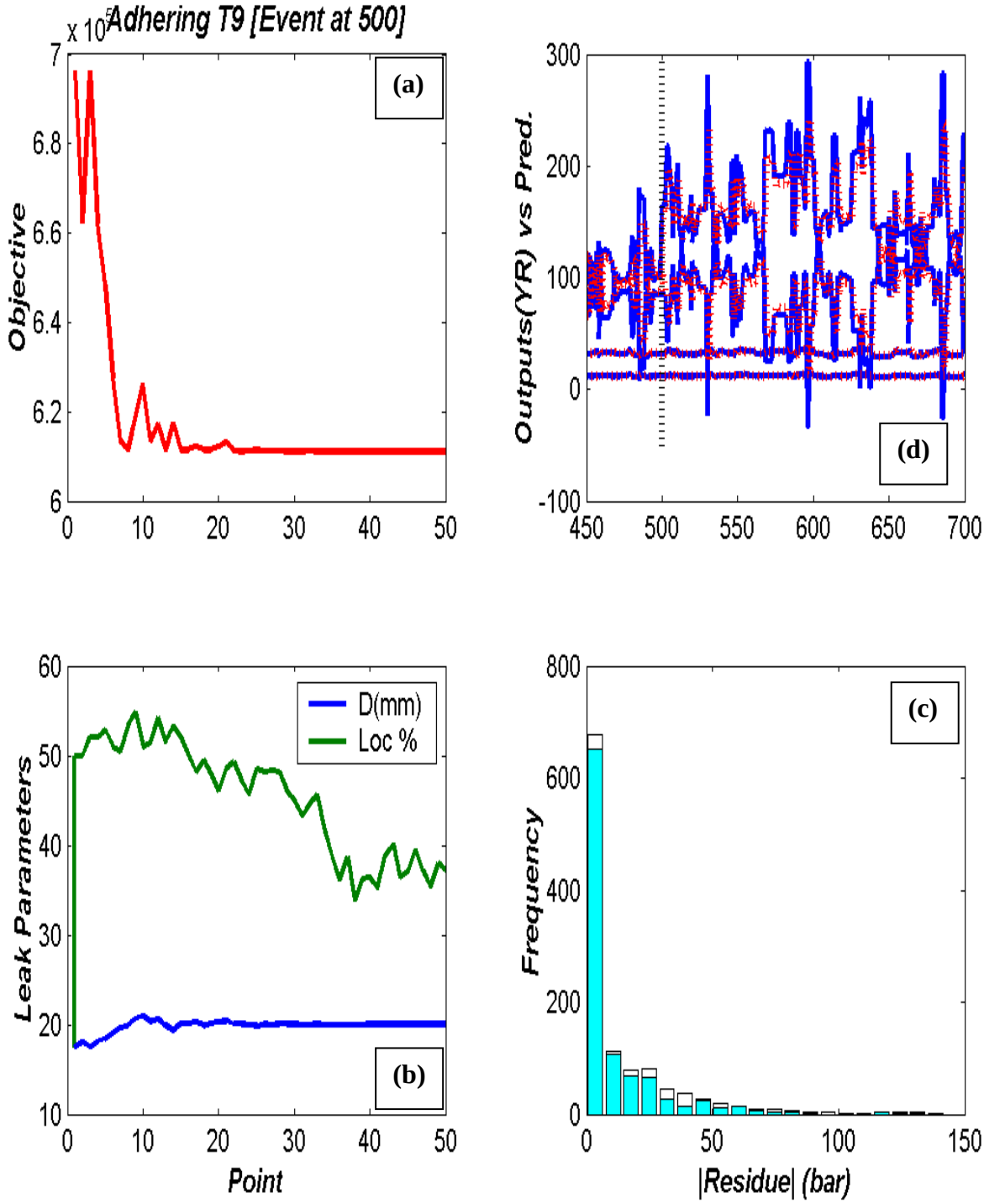


Figura 7.28: Busca de Localização de Vazamento para *Tubo* 9
(a) Trajeto do Objetivo; (b) Trajeto de Parâmetros; (c) Histograma de Resíduos Finais;
(d) Recorte de Amostragens (Y_R) versus Predições (Vermelho)

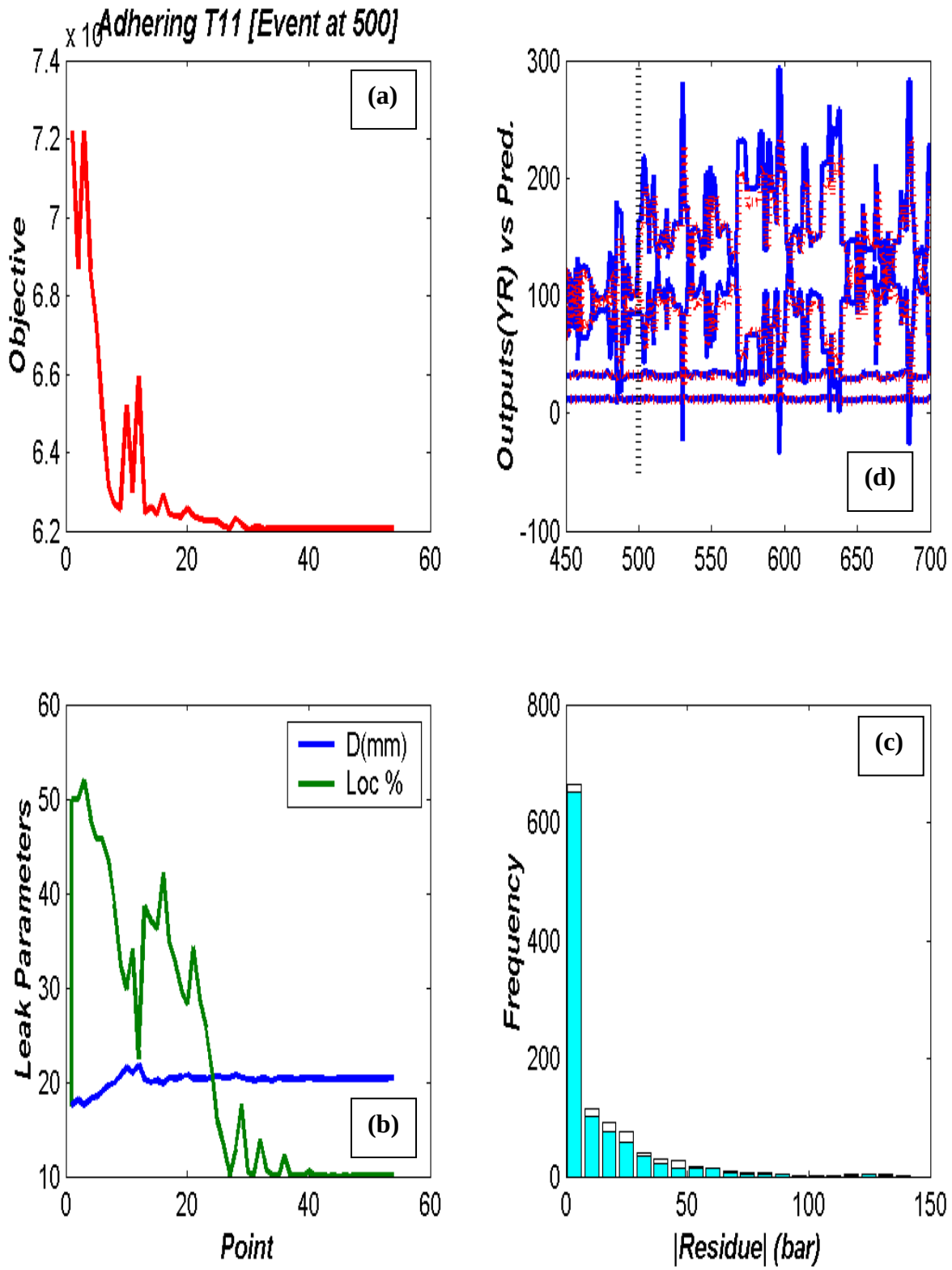


Figura 7.29: Busca de Localização de Vazamento para *Tubo* 11
(a) Trajeto do Objetivo; (b) Trajeto de Parâmetros; (c) Histograma de Resíduos Finais;
(d) Recorte de Amostragens (Y_R) versus Predições (Vermelho)

Tabela 7.2 : Resultados Finais de Localização de Vazamento para R1LP7P15W1W5

Valores Finais	***Tubo* 3**	***Tubo* 6**	***Tubo* 9**	***Tubo* 11**
Objetivo	$6{,}199.10^5$	$14{,}1.10^5$	$6{,}111.10^5$	$6{,}204.10^5$
$D_F(mm)$	23,1	30	20,0	20,3
L_F/L_k	90%	90%	37,2%	10%
Resíduos Pequenos	~650	~600	~650	~650

A Tabela 7.2 mostra que a busca de vazamento para o tubo 9 foi a que atingiu melhor valor de objetivo, embora bastante próximo daqueles encontrados para as buscas do tubo 3 e tubo 11. Os diagramas de predição versus amostragens (Figuras 7.27D, 7.28D e 7.29D) mostram aderências similares para as três buscas de localização, o que justifica os valores próximos de objetivo. A otimização no tubo 6, por outro lado, não apresentou nenhuma chance de ser competitiva (provavelmente, o vazamento situado no centro do NAFTA4 inviabiliza qualquer tentativa de ajuste em trechos terminais do NAFTA3).

No entanto é possível desconfiar-se que as buscas no tubo 3 e tubo 11 levaram a resultados suspeitos por atingirem, ambas, os limites de busca referentes à localização da ruptura, respectivamente, superior e inferior. Como a ruptura encontra-se no tubo 9, a posição do tubo 3 mais próxima deste é realmente a extremidade direita do domínio de busca, a saber a 90%. Analogamente, como o tubo 11 encontra-se a jusante do ponto de vazamento faz sentido que a otimização tenha convergido para locar o vazamento na sua extremidade esquerda do domínio de busca.

A menos destes detalhes, o fato importante é que o vazamento foi, afinal, corretamente localizado no tubo 9 com diâmetro de ruptura de 20.0mm e posição percentual de 37.2%. As coordenadas corretas do vazamento são 20mm de diâmetro e 40% de posição percentual.

Os resultados são promissores indicando que a metodologia tem potencial para utilização prática, especialmente quando integrada as demais abordagens aqui descritas.

8 Conclusões e Recomendações

Com o desenvolvimento da sociedade e sua demanda crescente pelos mais diferentes bens, cresce também a estrutura industrial necessária para atender esse consumo. O crescimento do setor químico e petroquímico, e com ele o aumento no uso de dutovias, é um reflexo do progresso que hoje vivemos. É importante ressaltar que juntamente com as vantagens operacionais que as dutovias proporcionam, existe também a possibilidade de ocorrerem acidentes graves envolvendo operações com este tipo de equipamento.

Apresentam-se as principais metodologias atualmente em uso na detecção de anomalias em dutovias. Enquanto determinadas técnicas são baseadas em hardware, buscando detectar sons característicos de vazamentos ou mesmo a emissão de vapores ou gases tóxicos, outras se baseiam apenas na análise dos dados gerados a partir da operação normal do sistema. É nesta classe que se enquadram as metodologias aqui propostas.

Para teste das ferramentas de detecção, emprega-se dutovia baseada em uma estrutura em operação no Brasil. O simulador foi previamente validado, sendo fonte confiável de dados para o estudo. Procedeu-se cenários de detecção, identificação e quantificação de anomalias operacionais com simulador do processo alvo. Uma série de metodologias foram então propostas e analisadas, compreendendo três grandes grupos: ferramentas baseadas em Análise de Componentes Principais (PCA); Classificadores Hierárquicos; e ferramentas baseadas em Modelos Auto-regressivos (ARX).

Um estudo preliminar, reduzindo a base de dados para um sistema de apenas duas dimensões, mostrou a viabilidade do uso da metodologia dos componentes principais na detecção de anomalias do tipo falhas em sensores ou vazamentos. Anomalias severas foram facilmente detectadas mesmo considerando apenas os dois primeiros componentes principais. Enquanto a análise visual dos resultados gerados pelo modelo PCA é suficiente para detectar as falhas de maior severidade, para anomalias menores torna-se necessário o uso de ferramentas estáticas mais precisas. As métricas de erro “Q” e “SPE” foram utilizadas com relativo sucesso,

sendo capazes de detectar grande parcela das anomalias as quais a rede foi submetida. Apesar do êxito, a eficiência dessas duas abordagens mostrou-se influenciada pela severidade da falha e sua disposição temporal na janela de análise.

Visando uma maior capacidade detectora foram propostas ferramentas baseadas nas métricas de similaridade Spca, Sdist e SF, esta última uma combinação das duas primeiras. Enquanto as abordagens por Spca e SF não obtiveram sucesso, o emprego da métrica Sdist elevou consideravelmente a eficiência de detecção. Esse sucesso foi especialmente significante quando a fase de teste tornou-se móvel, reduzindo a influência do instante temporal da anomalia sobre a capacidade do método. Esta abordagem, denominada “Sdist dinâmico”, mostrou-se igualmente eficaz na detecção dos dois tipos de anomalias aqui empregadas: falhas em sensores e vazamentos.

Além das análises por componentes principais, outra abordagem, totalmente distinta também foi aplicada: a Classificação Hierárquica. Os resultados encontrados a partir das técnicas baseadas em PCA em geral, e Sdist dinâmica em particular, mostraram elevada capacidade detectora de falha em uma janela temporal. Porém, mostraram-se incapazes de especificar o instante exato de início do problema operacional. Com este objetivo foram testadas ferramentas baseadas em classificação hierárquicas. Essas ferramentas separam os instantes amostrais em dois grandes conjuntos: “com” e “sem” falha. Apesar de apresentar elevado grau de sucesso quando empregada na análise de cenários com falhas de grande severidade, esta abordagem mostrou-se pouco eficiente para situações de menor gravidade. Deste modo, o emprego de tal técnica em ferramentas detectoras parece vinculado ao uso conjunto com outras metodologias que apresentem maior eficiência detectora, compondo um sistema híbrido. Não sendo adequado o uso isolado da classificação hierárquica na detecção de anomalias. Por outro lado, os classificadores hierárquicos mostram-se fundamentais na estimativa da severidade da falha encontrada, possibilitando correlacionar o afastamento entre os clusters formados e a gravidade da falha.

Além de detectar a presença de uma anomalia, a ferramenta deve também ser capaz de identificar a origem da falha. No caso de falhas em sensores foram propostas duas abordagens distintas: “SVI-D” e “Protótipos”. Para a rede empregada, composta por 16 sensores, o índice SVI-D foi capaz de indicar com sucesso o sensor em falha. A metodologia baseada em protótipos, por outro lado, apresentou desempenho bastante inferior. Para cenários de vazamento a eficiência na indicação de sua origem se inverteu. Enquanto o índice SVI-D não apresentou sucesso, a abordagem via protótipos revelou a localização física da ruptura da dutovia em um número significativo de casos.

A quantificação da anomalia é outra tarefa do sistema de detecção de falhas. Tal função foi efetuada a partir do afastamento encontrado entre os clusters formados na classificação hierárquica.

Tão importante quanto detectar, localizar e quantificar uma anomalia, é identificar o seu tipo. No caso de redes de escoamento os dois tipos de falhas mais comuns são: “falhas em sensores” e “vazamento”. O uso de classificadores hierárquicos acoplados com método dos protótipos permitiu identificar com sucesso a anomalia ocorrida.

Buscando melhorar ainda mais desempenho da ferramenta diante de cenários de vazamentos, testou-se então uma terceira abordagem: modelos Auto-Regressivos. O motivo de focar tal metodologia apenas para lidar com cenários de vazamento deveu-se a maior gravidade desse tipo de situação na operação de dutovias.

A ferramenta de detecção proposta consiste de um preditor auto-regressivo com entrada externa (ARX), montado em um contexto MIMO (multi-input, multi-output) com ordens (3,3) nos seus dois filtros de processamento. Tal ferramenta mostrou-se bastante sensível e precisa em aplicações em detecção de vazamentos, desde que seja devidamente abastecida por dados de entrada e saída do processo capazes de cobrir as principais características operacionais. Desta forma é necessário avaliar se a estrutura de coleta de medições do processo é capaz de permitir que o detector desenvolva uma boa performance.

Foram conduzidos diversos testes com detectores ARX-MIMO. Através deles pode-se concluir que sensores de vazão são mais relevantes do que sensores de pressão para montagem de estruturas de detecção de vazamento. O crescimento do diâmetro da ruptura tende a facilitar enormemente o processo de detecção; no entanto esta relação não é linear: maiores impactos na facilidade de detecção ocorrem quando o tamanho da ruptura passa de pequeno a médio do que quando passa de grande a muito grande. Quanto maior o diâmetro da linha em questão, piores as condições para detecção rápida e segura de vazamentos.

Neste trabalho também foi proposta uma estratégia de localização e quantificação de vazamentos baseada em preditores ARX. O Localizador/Quantificador é um preditor ARX-MIMO-DUAL munido de dois filtros apropriados a dois tipos de entradas:

- As entradas dinâmicas ordinárias do processo; e
- Os sinais de ruptura referentes ao calibre e à posição da ruptura em um tubo.

Os parâmetros do Localizador/Quantificador devem ser estimados via sintonia com séries temporais de respostas da rede sob a ação de vazamentos em múltiplas posições, tubos e calibres. Somente após a assimilação deste "aprendizado" em reconhecer vazamentos e seus efeitos é que a ferramenta torna-se capaz de cumprir sua missão. Neste trabalho não testamos o desempenho dela frente a dois ou mais vazamentos concomitantes.

A abordagem foi testada com sucesso na busca de um vazamento de médias proporções na rede de escoamento industrial estudada neste trabalho. O tubo danificado foi corretamente identificado dentre um conjunto de 4 tubos rivais e os parâmetros do vazamento em questão foram estimados com boa precisão.

Entendendo que os principais desafios da ferramenta proposta são detectar, identificar, localizar e quantificar situações de falha, pode-se dizer que as abordagens aqui propostas atendem aos objetivos. Ressalta-se que, apesar de terem sido descritas separadamente, o uso das diferentes metodologias deve ser

feito de modo integrado, visando a elaboração de abordagens híbridas. A integração permite aproveitar as características mais favoráveis de cada método, resultando numa ferramenta final com desempenho muito superior aquela que poderia ser desenvolvida a partir de uma única metodologia.

9. Referências Bibliográficas

ABHULIMEN, K.E., SUSU, A.A., Liquid Pipeline Leak Detection System: Model Development and Numerical Simulation. Chemical Engineering Journal, 97, p. 47-67, 2004.

ABHULIMEN, K.E., SUSU, A.A., Modelling Complex Pipelin Network Leak Detection Systems. Process Safety and Environmental Protection, 85, p. 579-598, 2007.

ALASKA DEPARTMENT OF ENVIRONMENTAL CONSERVATION, Technical Review of Leak Detection Technologies – Volume I – Crude Oil Transmission Pipelines, 1996.

AKAL, M., Forecasting Turkey's Tourism Revenues by ARMAX Model. Tourism Management, 25, p. 565-580, 2004.

AMAND, T., HEYEN, G., KALITVENTZEFF, B., Plant Monitoring and Fault Detection: Synergy Between Data Reconciliation and Principal Component Analysis, Computers and Chemical Eng., p. 501 – 507, 2001.

BIMPAS, M., AMDITIS, A., UZUNOGLU, N., Detection of Water Leaks in Supply Pipes Using Continuous Wave Sensor Operating at 2.45 GHz, Journal of Applied Geophysics, v.70, p. 226 – 236, 2010.

BOX, G. E. P., JENKINS, G. M., REINSEL, G. C., Time Series Analysis Forecasting and Control, 3ª ed., Editora Prentice Hall, 1994.

COLOMBO, A. F., LEE, P., KARNEY, B.W., A Selective Literature Review of Transient-based Leak Detection Methods, Journal of Hydro-environment Research 2, v. 2, p. 212-227, 2009.

CONAMA, Resolução nº 1, de 23 de janeiro de 1986, http://www.mma.gov.br/conama/res/
res86/ res0186.html

CONAMA, Resolução nº 293, de 12 de dezembro de 2001,
http://www.mma.gov.br/port/conama/res/ res01/res29301.html

CONNER, J.S., SEBORG, D.E., Assessing the Need for Process Re-identification, Ind Eng Chem Res, Vol 44, p 2767-2775, 2005.

COSTA, A.L.H, MEDEIROS, J.L., ARAÚJO, O.Q.F., A Time Series Approach for Pipe Network Simulation, Proceedings of IPC-2002 International Pipeline Conference, Calgary, Canada, Paper IPC02-27145, 2002

DASH, S., RENGASWAMY, R., VENKATASUBRAMANIAN, V., Fuzzy-logic Trend Classification for Fault Diagnosis of Chemical Processes, Computers and Chemical Engineering, Vol. 27, p. 347-362, 2003.

DUNIA, R., QIN, J., EDGAR, T., McAVOY, T. Identification of Faulty Sensors Using Principal Component Analysis. AIChE Journal, v.42, n.10, p.2797-2812, 1996 (a).

DUNIA, R., QIN, J., EDGAR, T., McAVOY, T., Use of principal component analysis for sensor fault identification. Computers and Chemical Engineering, 20, p. 713-718, 1996 (b).

FRAUSTO, H. U.; PIETERS, J. G.; DELTOUR, J. M., Modelling Greenhouse Temperature by Means of Auto Regressive Models. Biosystems Engineering, 84(2), p. 147-157, 2003.

FUKUSHIMA, K.; MAESHIMA, R.; KINOSHITA, A.; SHIRAISHI, H.; KOSHIJIMA, I., Gás Pipeline Leak Detection System Using the Online Simulation Method. Computers and Chemical Engineering, 24, p. 453-456, 2000.

GAO, X., FAN, H., HUANG, T., WANG, X., BAO, J., LI, X., HUANG, W., ZHANG, W., Natural Gas Pipeline Leak Detector Based on NIR Diode Laser Absorption Spectroscopy. Spectrochimica Acta Part A, 65, p.133-138, 2006.

GAO, Y., BRENNAN, M.J., JOSEPH, P.F., MUGGLETON, J.M., HUNAIDI, O., On the Selection of Acoustic/vibration Sensors for Leak Detection in Plastic Water Pipes. Journal of Sound and Vibration, 283, p. 927-941, 2005.

GESSING, R.; BLACHUTA, M., Set Point and Indentifiability in the Closed Loop with a Minimum-variance Controller. Control Eng. Practice, 4, nº 5, p. 665-670, 1996.

GUSMÃO, A. C. F.; DE MARTINI, L.C., Gestão Ambiental na Indústrial, Editora SMS Digital, 2009

HART, P., DUDA, R.O. Pattern Classification. 2ª ed. John Wiley Pro., 2000.

KAM, S. I., Mechanistic Modeling of Pipeline Leak Detection at Fixed Inlet Rate, Journal of Petroleum Science and Engineering, v. 70, p. 145-156, 2010.

KANO, M., NAGAO, K., OHNO, H., HASEBE, S., HASHIMOTO, I., Dissimilarity of Process data for statistical process monitoring. Proceedings of IFAC Symposium on Advanced Control of Chemical Processes, Vol 1, p. 231-236, 2000.

KRZANOWSKI, W.J., Between Groups Comparison of Principal Components, Journal of American Statistical Association, Vol 74, 1979.

LÓPEZ, R.M., SPIRIN, V.V., MIRIDONOV, S.V., SHLYAGIN, M.G., BELTRÁN, G., KUZIN, E.A., Fiber Optic Distributed Sensor for Hydrocarbon Leak Localization Based on Transmission/reflection Measurement. Optics & Laser Technology, 34, p. 465-469, 2002.

LOQUASTO. F., SEBORG, D.E.,Model Predictive Controller Monitoring Based on Pattern Classification and PCA, Proceedings of American Control Conference, 2003.

MADDISON, D., Air Pollution and Hospital Admissions an ARMAX Modelling Approach. Journal of Environmental Economics and Management, 49, p. 116-131, 2005.

MARQUES, J.M., MARQUES, M.A.M. As Componentes Principais no Descarte de Variáveis em um Modelo de Regressão Múltipla. Revista FAE, Vol 8, p. 93-101, 2005.

MEDEIROS, J.L., XAVIER, G.M., DUARTE M.V.E., ARAÚJO, O.Q.F., Stochastic Identification of a Process with Multiple Stationary States, Latin American Applied Research, 31: 367-372, 2001

MORETTIN, P.A.; TOLOI, C.M.C., Análise de Séries Temporais. Editora Edgard Blucher Ltda, 2004.

RIDDER, D., Pekalska, E., Duin, R., The Economics of Classification: Error vs. Complexity, IEEE, 2002

RIVERA, D. E., GAIKWAD, S. V., Digital PID Controller Design Using ARX Estimation. Computers Chem. Engineering, 20, p. 1317-1334, 1996.

SADORSKY, P, Time-varying Risk Premiums in Petroleum Futures Prices. Energy Economics, 24, p. 539-556, 2002.

SINGHAL, A., SEBORG, S. Pattern Matching in Multivariate Time Series Databases Using a Moving-Window Approach. Ind. Eng. Chem. Res. n.41, p.3822-3838, 2002.

SILVA, R. A.; BUIATTI, C. M.; CRUZ, S. L.; PEREIRA, J. A. F. R., Pressure Wave Behaviour and Leak Detection in Pipelines. Computers Chem. Engineering, 20, p. S491-S496, 1996.

SMITH, S., A Tutorial on Principal Components Analysis, 2006

SOUZA, J.N.M., Desenvolvimento de Algoritmos para Detecção, Localização e Quantificação de Vazamentos em Redes de Dutos. Tese de M,Sc. pelo Programa de Tecnologia de Processos Químicos e Petroquímicos – EQ / UFRJ , 2004.

VANDAELE, W., Applied time series and Box-Jenkins Models. Academic Press, 1983.

VAZ JUNIOR, C.A. Detecção, Localização e Quantificação de Vazamentos: uma Abordagem em Séries Temporais". Dissertação de Mestrado - Escola de Química da UFRJ, 2006.

VAZ JUNIOR, C. A., BARBABELA, B. A. Análise Exploratória da Aplicação de Modelos Auto-Regressivos na Previsão do Comportamento dos Preços de Derivados de Petróleo, IX SEMEAD - USP, São Paulo 2006

VAZ JUNIOR, C. A., ARAUJO, O. Q. F., MEDEIROS, J. L.
Classificadores Hierárquicos e PCA na Detecção e Identificação de Falhas em Redes de Escoamento, 4 Congresso Brasileiro de P&D em Petróleo e Gás, 2007, Rio de Janeiro.

VAZ JUNIOR, C. A., BARBABELA, B. A. Previsão do Comportamento dos Preços de Derivados de Petróleo: uma Abordagem por Séries Temporais, Congreso de Computación Aplicada CAIP, 2007, Assunção, Paraguai.

VAZ JUNIOR, C. A., MEDEIROS, J. L., ARAUJO, O. Q. F.
Ferramentas de Detecção de Anomalias Baseadas em Classificadores Hierárquicos e Teoria dos Protótipos, XVII Congresso Brasileiro de Engenharia Química, 2008, Recife.

VAZ JUNIOR, C. A., ARAUJO, O. Q. F., MEDEIROS, J. L. Failure Diagnostics Using Data Mining Tools. Computer-Aided Chemical Engineering (PSE). , v.27, p.1539 - 1544, 2009.

VAZ JUNIOR, C. A., ARAUJO, O. Q. F., MEDEIROS, J. L. ARX Modeling Approach to Leak Detection and Diagnosis. Journal of Loss Prevention in the Process Industries, **IN PRESS,** http://dx.doi.org/10.1016/j.jlp.2010.03.001, 2010.

VENKATASUBRAMANNIAN, V., RENGASWAMY, R., KAVURI, S., YIN, K. A Review of Process Fault Diagnosis – Part III: Process History Based Methods, Computers and Chemical Engnieering, Vol. 27, p.327-346, 2003.

WARNES, M. R., GLASSEY, J., MONTAGUE, G. A., KARA, B., On Data-based Modelling Techniques for Fermentation Process. Process Biochemistry, 31, nº 2, p. 147-155, 1996.

WERNER, L., RIBEIRO, J. L. D., Previsão de Demanda: uma Aplicação Box-Jenkins na Área de Assistência Técnica de Computadores Pessoais. Gestão & Produção, 10, nº1, p. 47-67, 2003.

WISE, B., GALLAGHER, N. The Process Chemometrics Approach to Process monitoring and faulty detection. J. Proc. Cont. v.6, n.6, p.329-348, 1996.

WU, C. S., LIN, J. L., TIAO, G. C., CHO, D. D., Is Money Demand in Taiwan Stable? Economic Modelling, 22, p. 327-346, 2005.

ZHANG, L., QIN, X. Y., WANG, Z., LIANG, W., Designing a Reliable Leak Detection System for West Products Pipeline. Journal of Prevention in the Process Industries v.22, p. 981-989, 2009.

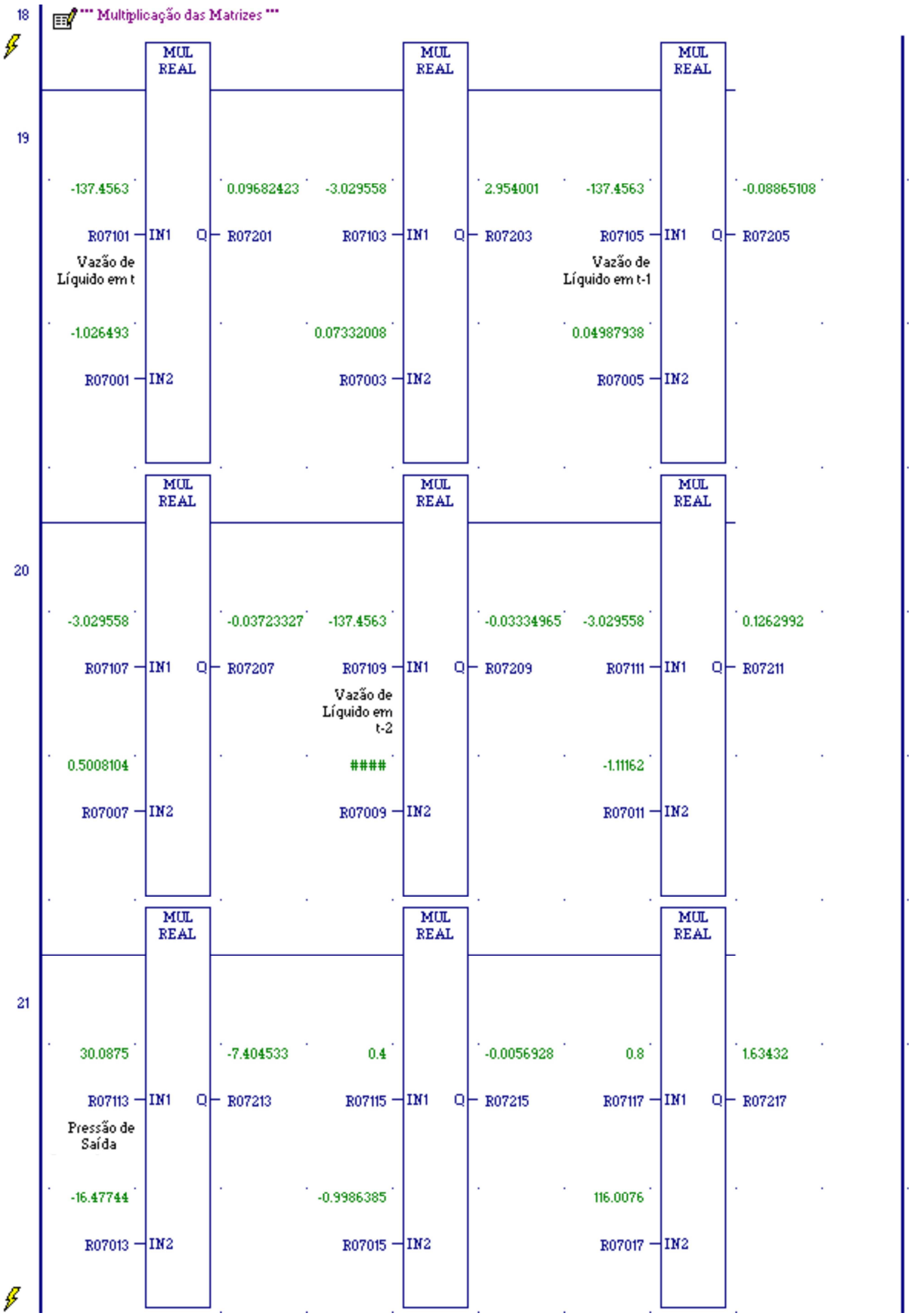
ZHANG, J., Designing a Cost Effective and Reliable Pipeline Leak Detection System. Pipeline Reliability Conference, Houston, USA, 1996.

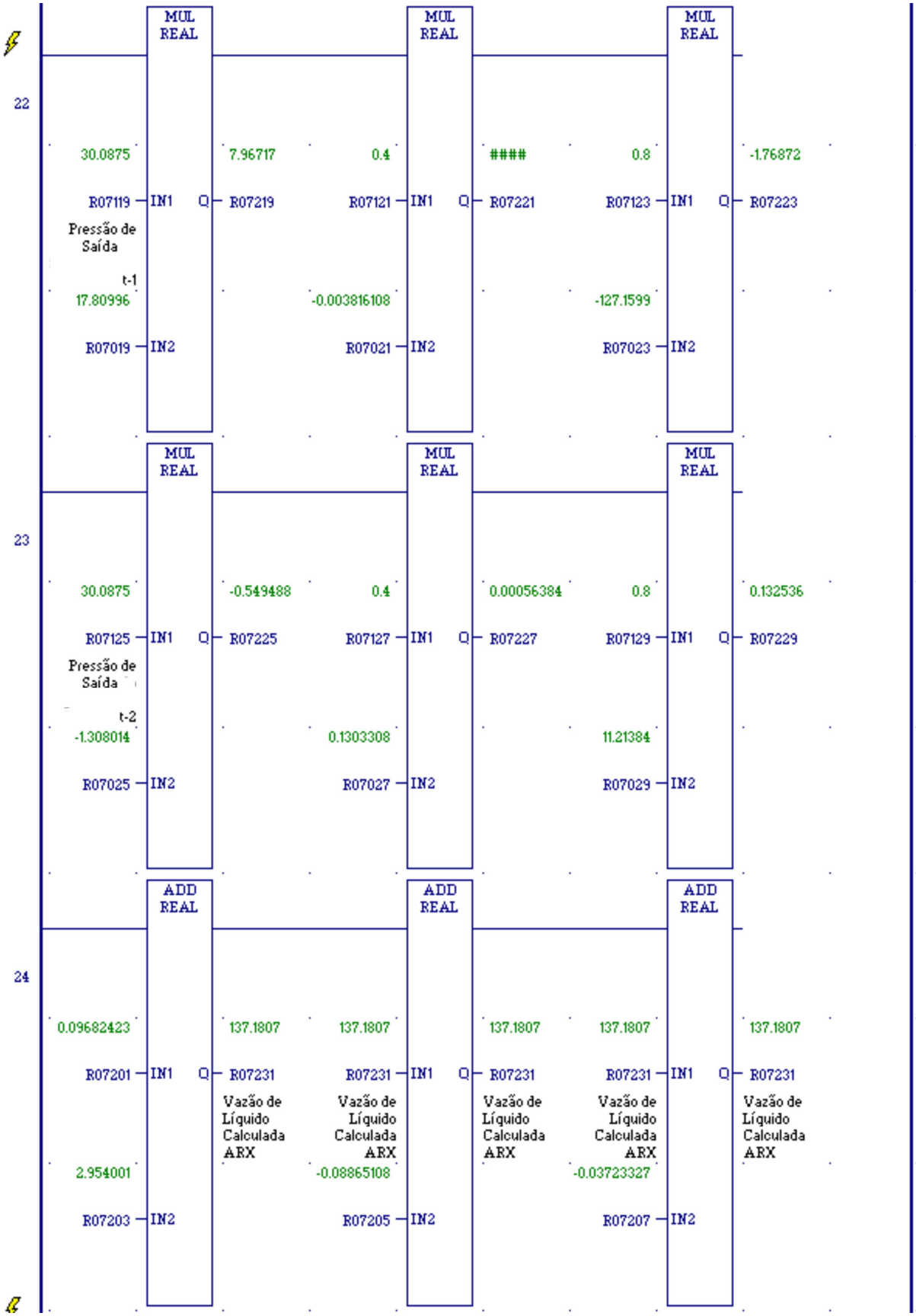
ZHOU, L., ZENG, Y., Automatic Alignment of Infrared Vídeo Frames for Equipment Leak Detection. Analytica Chimica Acta, v. 584, p. 223-227, 2006.

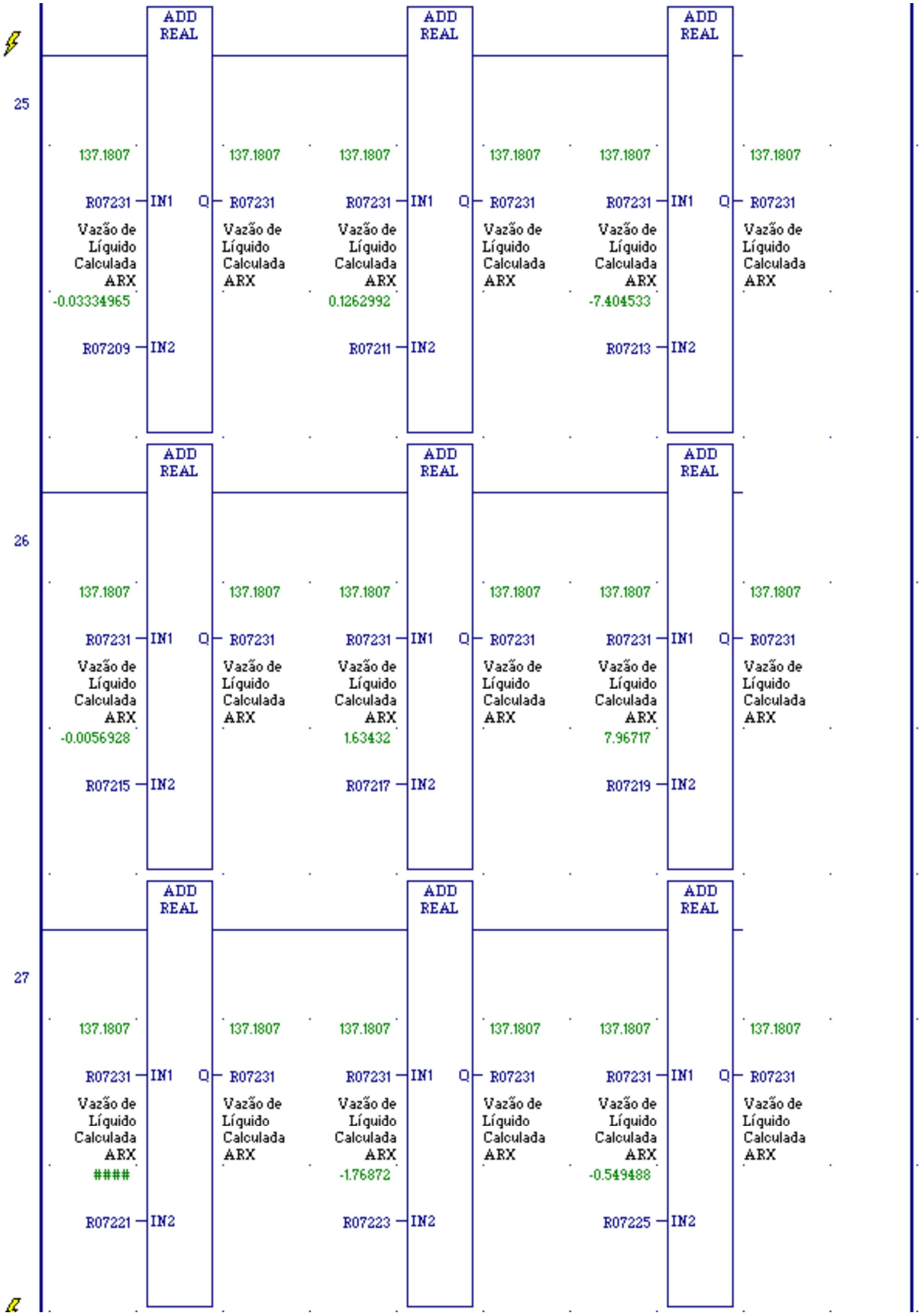
ANEXO I – Implementação de preditor ARX em Linguagem LADDER

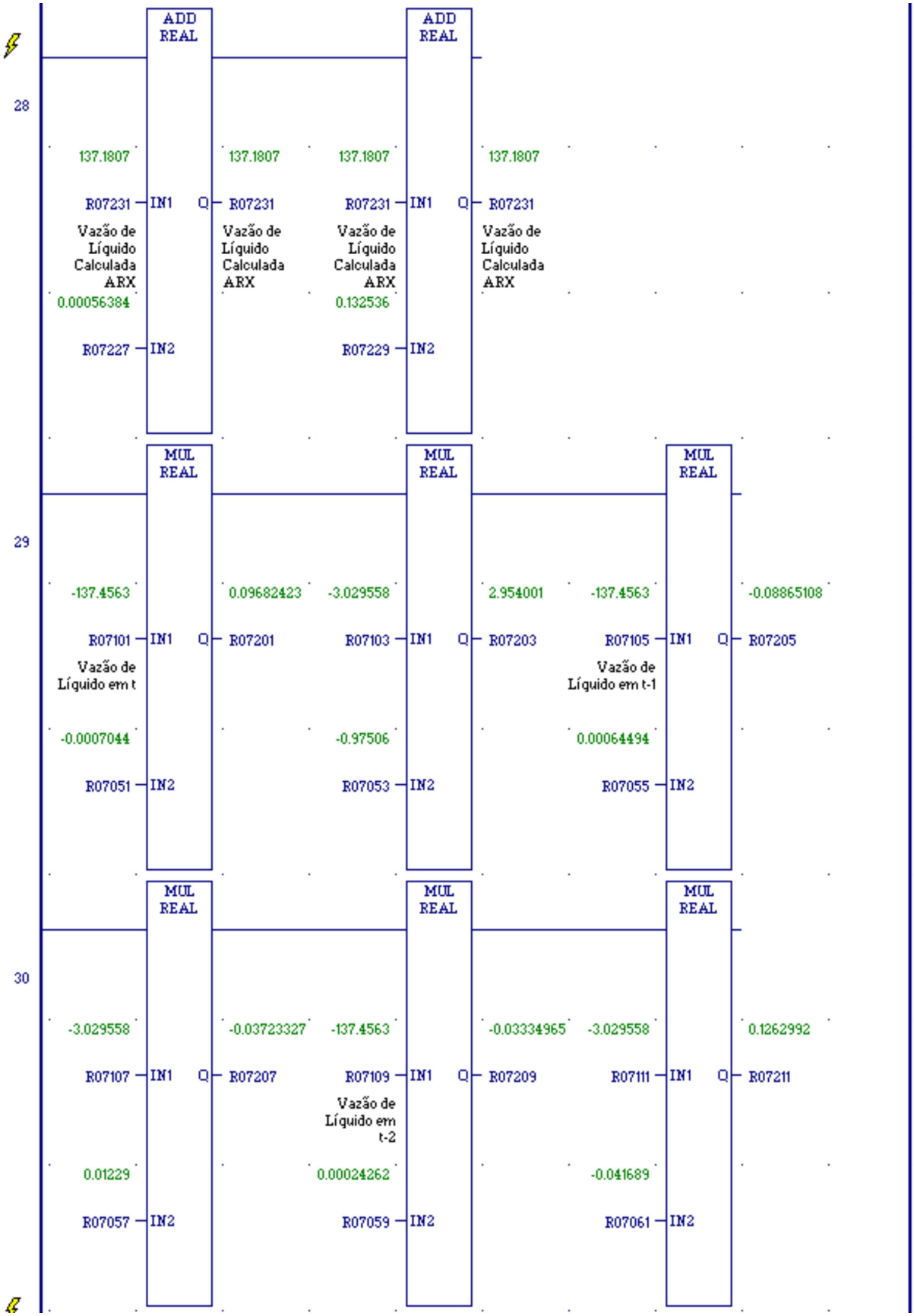
As figures abaixo descrevem a implementação de um preditor ARX em ambiente LADDER do controlador lógico programável GE Fanuc família 90:30 (CPU 374).

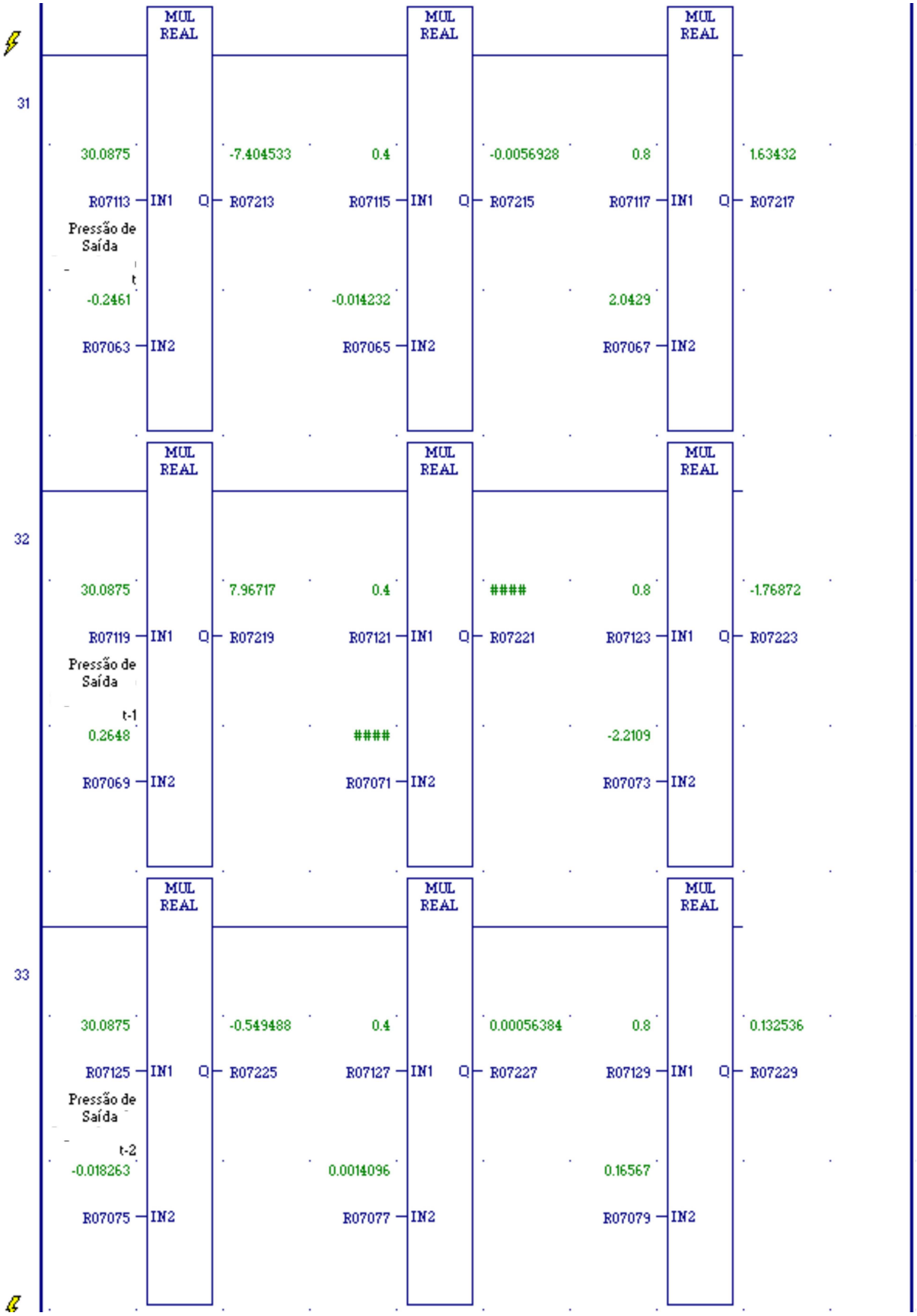
A linguagem LADDER não processa matrizes, de modo que todos os cálculos matriciais tiveram de ser abertos. Foi criada uma subrotina destinada à implementação dos cálculos do modelo ARX para a previsão das vazões de líquido no instante t + 1, baseado em uma série temporal t, t – 1 e t – 2. Ou seja, o modelo ARX utiliza-se dos valores obtidos em três instantes de tempo passado para então prever o valor seguinte.

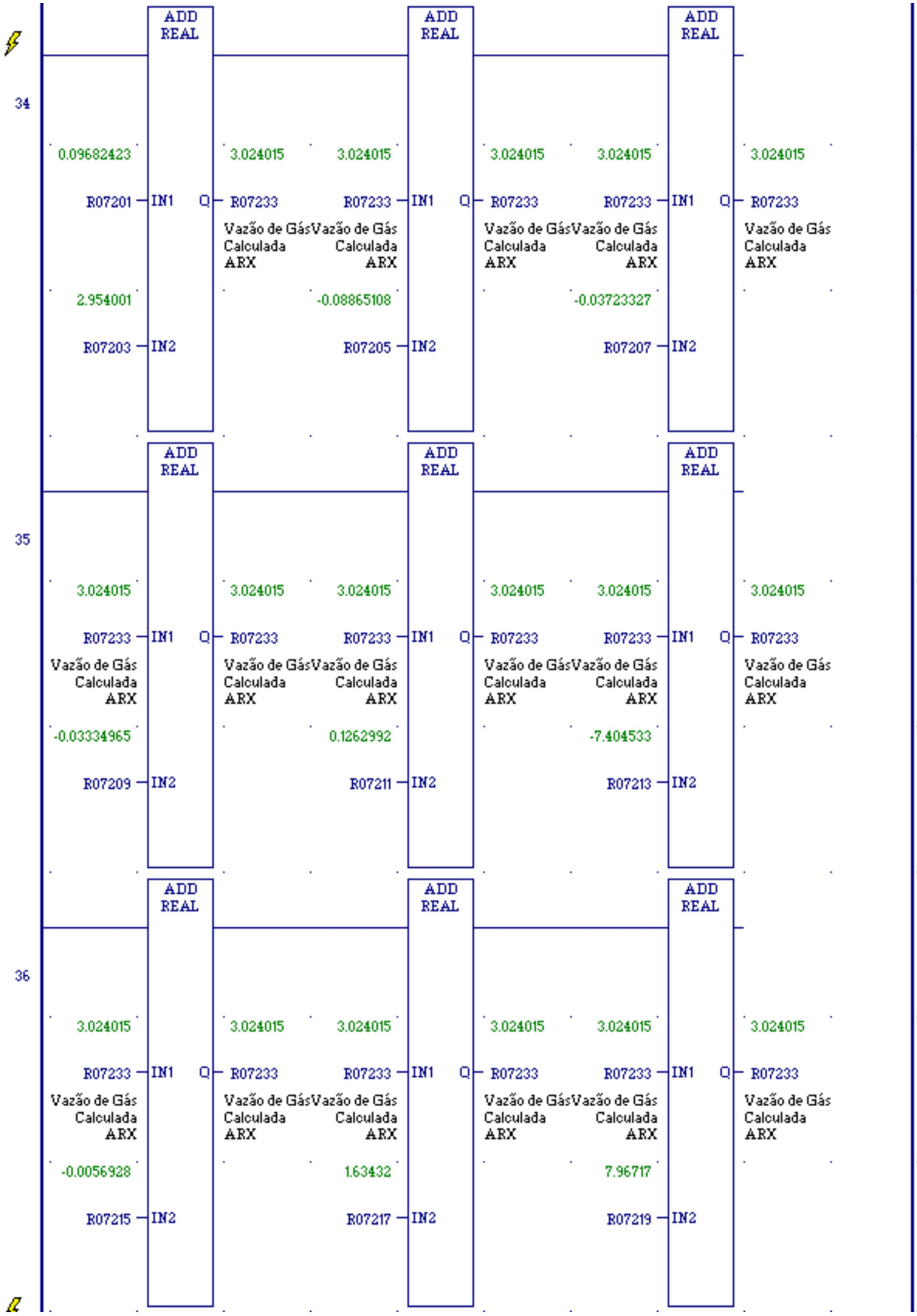

18
*** Multiplicação das Matrizes ***
MUL REAL
19
-137.4563
R07101
Vazão de Líquido em t
IN1
Q
0.09682423
R07201
-3.029558
R07103
2.954001
R07203
-137.4563
R07105
Vazão de Líquido em t-1
-0.08865108
R07205
-1.026493
R07001
IN2
0.07332008
R07003
0.04987938
R07005
20
-3.029558
R07107
0.5008104
R07007
-0.03723327
R07207
-137.4563
R07109
Vazão de Líquido em t-2
####
R07009
-0.03334965
R07209
-3.029558
R07111
-1.11162
R07011
0.1262992
R07211
21
30.0875
R07113
Pressão de Saída
-7.404533
R07213
0.4
R07115
-0.0056928
R07215
0.8
R07117
1.63432
R07217
-16.47744
R07013
-0.9986385
R07015
116.0076
R07017

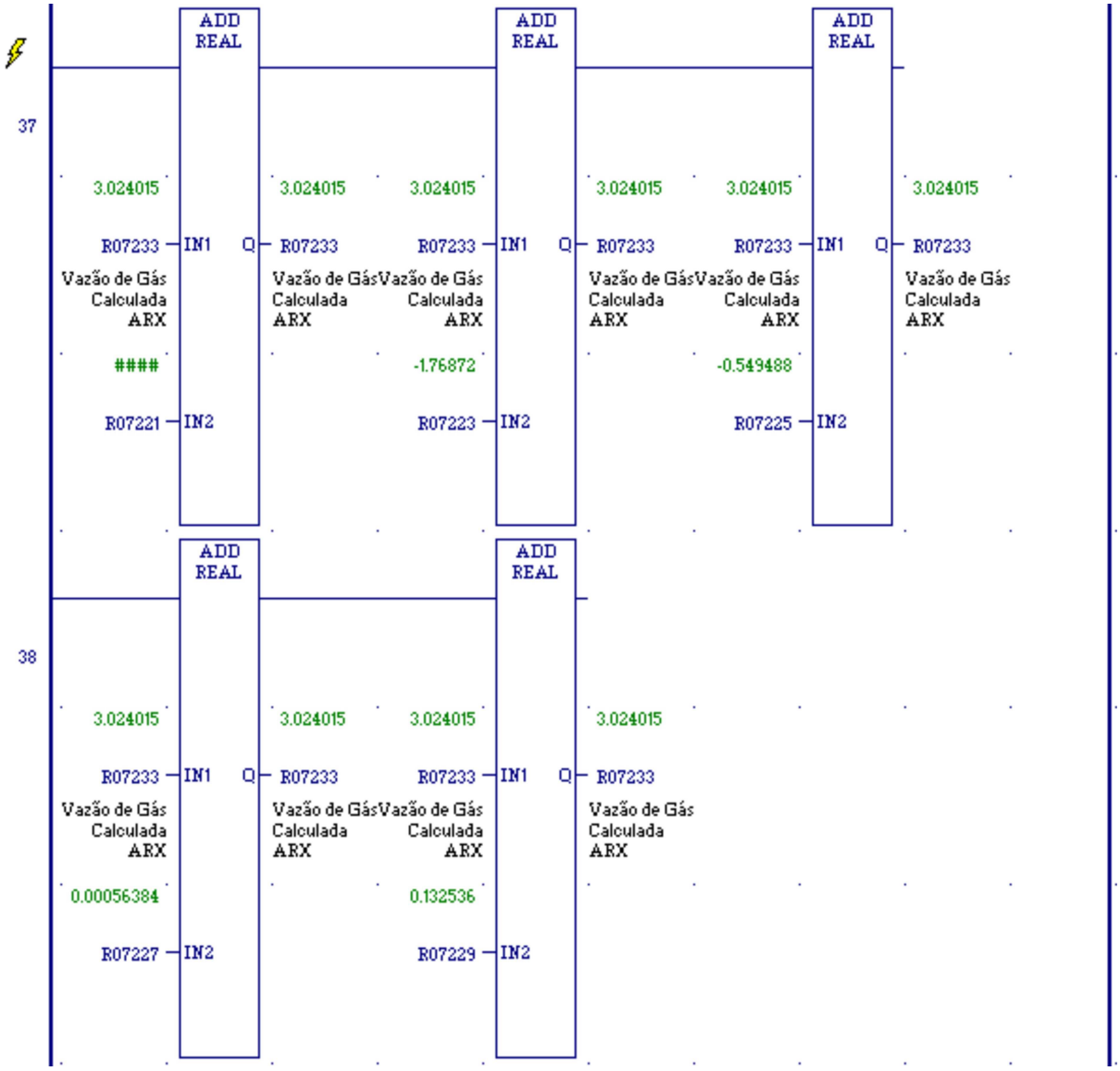

22
MUL REAL
30.0875
R07119
IN1
Q
R07219
7.96717
Pressão de Saída
t-1
17.80996
R07019
IN2
R07121
0.4
-0.003816108
R07021
MUL REAL
IN1
Q
R07221
####
R07123
0.8
-127.1599
R07023
MUL REAL
IN1
Q
R07223
-1.76872
23
MUL REAL
30.0875
R07125
IN1
Q
R07225
-0.549488
Pressão de Saída
t-2
-1.308014
R07025
IN2
R07127
0.4
0.1303308
R07027
MUL REAL
IN1
Q
R07227
0.00056384
R07129
0.8
11.21384
R07029
MUL REAL
IN1
Q
R07229
0.132536
24
ADD REAL
0.09682423
R07201
IN1
Q
R07231
137.1807
Vazão de Líquido Calculada ARX
2.954001
R07203
IN2
R07231
137.1807
Vazão de Líquido Calculada ARX
-0.08865108
R07205
ADD REAL
IN1
Q
R07231
137.1807
Vazão de Líquido Calculada ARX
R07231
137.1807
Vazão de Líquido Calculada ARX
-0.03723327
R07207
ADD REAL
IN1
Q
R07231
137.1807
Vazão de Líquido Calculada ARX


ADD REAL
25
137.1807
R07231
IN1
Q
Vazão de Líquido Calculada ARX
-0.03334965
R07209
IN2
137.1807
R07231
Vazão de Líquido Calculada ARX
137.1807
R07231
IN1
Q
Vazão de Líquido Calculada ARX
0.1262992
R07211
IN2
137.1807
R07231
Vazão de Líquido Calculada ARX
137.1807
R07231
IN1
Q
Vazão de Líquido Calculada ARX
-7.404533
R07213
IN2
137.1807
R07231
Vazão de Líquido Calculada ARX
26
-0.0056928
R07215
1.63432
R07217
7.96717
R07219
27
####
R07221
-1.76872
R07223
-0.549488
R07225


28
ADD REAL
137.1807
R07231
IN1
Q
R07231
Vazão de Líquido Calculada ARX
0.00056384
R07227
IN2
137.1807
137.1807
R07231
IN1
Q
R07231
Vazão de Líquido Calculada ARX
0.132536
R07229
IN2
137.1807
R07231
Vazão de Líquido Calculada ARX
29
MUL REAL
-137.4563
R07101
IN1
Q
R07201
Vazão de Líquido em t
-0.0007044
R07051
IN2
0.09682423
-3.029558
R07103
IN1
Q
R07203
-0.97506
R07053
IN2
2.954001
-137.4563
R07105
IN1
Q
R07205
Vazão de Líquido em t-1
0.00064494
R07055
IN2
-0.08865108
30
MUL REAL
-3.029558
R07107
IN1
Q
R07207
0.01229
R07057
IN2
-0.03723327
-137.4563
R07109
IN1
Q
R07209
Vazão de Líquido em t-2
0.00024262
R07059
IN2
-0.03334965
-3.029558
R07111
IN1
Q
R07211
-0.041689
R07061
IN2
0.1262992


31
MUL REAL
30.0875
R07113
IN1
Q
R07213
-7.404533
R07115
IN1
Q
R07215
0.4
-0.0056928
R07117
IN1
Q
R07217
0.8
1.63432
Pressão de Saída
t
-0.2461
R07063
IN2
-0.014232
R07065
IN2
2.0429
R07067
IN2
32
MUL REAL
30.0875
R07119
IN1
Q
R07219
7.96717
R07121
IN1
Q
R07221
0.4
####
R07123
IN1
Q
R07223
0.8
-1.76872
Pressão de Saída
t-1
0.2648
R07069
IN2
####
R07071
IN2
-2.2109
R07073
IN2
33
MUL REAL
30.0875
R07125
IN1
Q
R07225
-0.549488
R07127
IN1
Q
R07227
0.4
0.00056384
R07129
IN1
Q
R07229
0.8
0.132536
Pressão de Saída
t-2
-0.018263
R07075
IN2
0.0014096
R07077
IN2
0.16567
R07079
IN2


34
ADD REAL
0.09682423
R07201
IN1
Q
2.954001
R07203
IN2
3.024015
R07233
Vazão de Gás Calculada ARX
3.024015
R07233
Vazão de Gás Calculada ARX
-0.08865108
R07205
3.024015
R07233
Vazão de Gás Calculada ARX
3.024015
R07233
Vazão de Gás Calculada ARX
-0.03723327
R07207
3.024015
R07233
Vazão de Gás Calculada ARX
35
3.024015
R07233
Vazão de Gás Calculada ARX
-0.03334965
R07209
3.024015
R07233
Vazão de Gás Calculada ARX
3.024015
R07233
Vazão de Gás Calculada ARX
0.1262992
R07211
3.024015
R07233
Vazão de Gás Calculada ARX
3.024015
R07233
Vazão de Gás Calculada ARX
-7.404533
R07213
3.024015
R07233
Vazão de Gás Calculada ARX
36
3.024015
R07233
Vazão de Gás Calculada ARX
-0.0056928
R07215
3.024015
R07233
Vazão de Gás Calculada ARX
3.024015
R07233
Vazão de Gás Calculada ARX
1.63432
R07217
3.024015
R07233
Vazão de Gás Calculada ARX
3.024015
R07233
Vazão de Gás Calculada ARX
7.96717
R07219
3.024015
R07233
Vazão de Gás Calculada ARX


37
ADD REAL
3.024015
R07233
IN1
Q
Vazão de Gás Calculada ARX
####
R07221
IN2
3.024015
R07233
Vazão de Gás Calculada ARX
3.024015
R07233
Vazão de Gás Calculada ARX
-1.76872
R07223
IN2
3.024015
R07233
Vazão de Gás Calculada ARX
3.024015
R07233
Vazão de Gás Calculada ARX
-0.549488
R07225
IN2
3.024015
R07233
Vazão de Gás Calculada ARX
38
ADD REAL
3.024015
R07233
Vazão de Gás Calculada ARX
0.00056384
R07227
IN2
3.024015
R07233
Vazão de Gás Calculada ARX
3.024015
R07233
Vazão de Gás Calculada ARX
0.132536
R07229
IN2
3.024015
R07233
Vazão de Gás Calculada ARX