

**UNIVERSIDADE FEDERAL DE ALFENAS
INSTITUTO DE CIÊNCIAS EXATAS
BACHARELADO EM CIÊNCIA DA COMPUTAÇÃO**

Edgar Fonseca Franco Júnior

**CLASSIFICAÇÃO DE PADRÕES
PARA O ESCOAMENTO GÁS/LÍQUIDO
VIA MODELOS INTELIGENTES**

Alfenas, 01 de Julho de 2010.

UNIVERSIDADE FEDERAL DE ALFENAS
INSTITUTO DE CIÊNCIAS EXATAS
BACHARELADO EM CIÊNCIA DA COMPUTAÇÃO

CLASSIFICAÇÃO DE PADRÕES
PARA O ESCOAMENTO GÁS/LÍQUIDO
VIA MODELOS INTELIGENTES

Edgar Fonseca Franco Júnior

Monografia apresentada ao Curso de Bacharelado em
Ciência da Computação da Universidade Federal de
Alfenas como requisito parcial para obtenção do Título de
Bacharel em Ciência da Computação.

Orientador: Prof. Dr. Ricardo Menezes Salgado

Alfenas, 01 de Julho de 2010.

Edgar Fonseca Franco Júnior

**CLASSIFICAÇÃO DE PADRÕES
PARA O ESCOAMENTO GÁS/LÍQUIDO
VIA MODELOS INTELIGENTES**

A Banca examinadora abaixo-assinada aprova a monografia apresentada como parte dos requisitos para obtenção do título de Bacharel em Ciência da Computação pela Universidade Federal de Alfenas.

Prof. Dr. Eric Batista Ferreira

Universidade Federal de Alfenas

Prof. Dr. Marcos de Mendonça Passini

Universidade Federal de Alfenas

Prof. Dr. Ricardo Menezes Salgado (Orientador)

Universidade Federal de Alfenas

Alfenas, 01 de Julho de 2010.

*Dedico este **Trabalho de Conclusão de Curso**
a todos que de alguma forma fizeram parte da minha vida.*

*Aos meus **Pais, Edgar e Nailza,**
meus primeiros mestres,
que sempre acreditaram em mim,
inclusive orando por mim.*

*Aos **Familiares,**
que nos momentos de minha ausência,
sempre fizeram entender que o futuro,
é feito a partir da constante dedicação no presente !!!*

*Àquelas pessoas que quando deveriam ser simplesmente **Professores,**
foram **Mestres.**
Que quando deveriam ser **Mestres,**
foram **Amigos.***

*Aos verdadeiros **Amigos,**
que me incentivaram a seguir
quando nem eu mais acreditava que conseguiria.*

*A **Alan Turing,**
o Pai da Inteligência Artificial.*

AGRADECIMENTO

Aos meus pais **Edgar Fonseca Franco** e **Nailza Rosângela Santos Franco**, por terem feito o possível e o impossível para me oferecerem a oportunidade de estudar.

Ao meu primo, **Raphael Franco Gatti**, por ser o principal responsável pela escolha do curso de Ciência da Computação.

À **Universidade Federal de Alfenas** (UNIFAL-MG), juntamente com o **Laboratório de Inteligência Computacional (LInC)** e **Laboratório de Engenharia de Software (LEnS)**, meu reconhecimento e agradecimento pela oportunidade de formar no curso de **Bacharelado em Ciência da Computação (BCC)**.

À **Universidade Estadual de Campinas** (Unicamp), em especial a **Faculdade de Engenharia Mecânica (FEM)** e **Faculdade de Engenharia Elétrica e de Computação (FEEC)**, por incentivar e apoiar este projeto.

Ao Prof. Dr. **Ricardo Salgado Menezes**, pela orientação no desenvolvimento deste projeto e na elaboração desta monografia, guiando de forma sabia os meus passos.

Ao Prof. Dr. **Takaaki Ohishi**, Prof. Dr. **Eugênio Spanó Rosa** e Prof. Dr. **Niederauer Mastelari**, por disponibilizarem seus conhecimentos e pesquisas realizadas em torno deste projeto.

Ao Prof. **Humberto César Brandão de Oliveira** e a Prof^a. **Mariane Moreira de Souza**, por acreditarem e apoiarem o desenvolvimento de trabalhos científicos, incentivando a diversificação do conhecimento, inclusive propondo novos rumos no desenvolvimento deste projeto.

Ao Prof. **Luís Eduardo da Silva** e ao Prof. Dr. **Eric Batista Ferreira** por ajudarem a selecionar e abordar técnicas estatísticas para a validação deste projeto.

A minha prima e Prof^a. **Valéria Barbosa Ferraz**, por auxiliar na correção desta monografia tornando-a mais fácil de ser compreendida.

Aos meus irmãos de República, **Robson Ferreira Mendes**, **Júlio César Dias**, **Neubio Matos Ferreira** e **Jan Victor Soares Santos**, que foram a minha família em Alfenas.

A minha amiga, **Juliana Oliveira Ferreira**, que por tantas vezes fomos companheiros de estudos.

Aos demais **amigos** que, de alguma forma ajudaram na elaboração desta monografia e que enriqueceram meus conhecimentos.

E a você, **amigo leitor**. **Muito Obrigado!!**

O real nos parece um fluxo e no fluxo não há modelos. Daí, a eterna controvérsia dos que admitem, como Heráclito, que o fluxo ou devir é a realidade e dos que entendem, como Parmênides, que o real é imutável e o devir é aparência. Os modelos, portanto, são nossas formas perceptuais e transitórias de apreender, a cada momento, o fluxo. Assim, cada forma perceptual do fluxo só é real em relação ao percebedor no momento da percepção e só se torna aparência ou Maya se prossegue além da percepção.

O real é o agora. O agora é sempre inédito. Quem vê, não precisa de palavras, pois só se fala para aqueles que não viram. E o que se diz, já não é: o presente é mais rápido que o laço da palavra. Por isso, quem fala, não vê, porque, se fala, fala do que já não vê. O eu não existe no presente: surge, quando a experiência já terminou. O eu é o passado.

Cada percepção do real é única e irrepetível. Jamais saberemos o que perdemos, jamais repetiremos o que experimentamos. A riqueza do viver não consiste na acumulação do vivido, mas na capacidade de viver plenamente o momento que passa. Nenhuma experiência deve deixar restos ou saldos, pois eles deformam as novas percepções da realidade.

(Valter da Rosa Borges)

RESUMO

O escoamento gás/líquido é o fluxo de duas substâncias com fases distintas, homogêneas ou heterogêneas, através de uma tubulação, sendo que, para cada tipo de sistema há um conjunto de padrões ou regimes de escoamento que o compõe.

Devido às mudanças contínuas de regimes nas tubulações, há a dificuldade em controlar algumas situações nas linhas de escoamento como: queda de pressão, controle de vazão, gerência de produção, entre outras. Para isso, é necessário um sistema capaz de classificar padrões, em determinados intervalos de tempo, a fim de manter as linhas atualizadas.

Esta pesquisa propôs a construção de um sistema de **Classificação de Padrões para o Escoamento Gás/Líquido via Modelos Inteligentes**. Foram aplicadas técnicas de redes neurais artificiais (RNAs) do tipo multilayer perceptron (MLP), backpropagation com termo momentum (BPM) e método gradiente (GRAD), para a identificação de tais padrões. Na pesquisa houve a utilização da base de dados desenvolvida pelo Laboratório de Fenômenos Multifásicos da Faculdade de Engenharia Mecânica (FEM) da Unicamp, onde foram aplicadas técnicas de transformações estatísticas visando diversificar o conjunto de entradas.

Para avaliação das principais correlações de medição encontradas nos diversos conjuntos de experimentos executados, foram feitos testes estatísticos, visando selecionar as configurações mais prósperas para a classificação de padrões de escoamentos. Sendo que, as configurações mais prósperas obtiveram resultados médios entre 67,73% a 97,57% de acertos.

Palavras-Chave: Classificação de Padrões, Escoamento Gás/Líquido, Modelos Inteligentes, Rede Neural Artificial, Transformações Estatísticas, Análise Estatística.

ABSTRACT

The gas/liquid flow is the flow of two substances with distinct phases, homogeneous or heterogeneous, through a pipe, and, for each type of system is a set of patterns or flow regimes that compose it.

Due to continuous changes of regimes in the pipes, there is the difficulty in controlling some situations such as flow lines: pressure drop, flow control, production management, among others. So, it's necessary a system capable of classifying patterns, in specific time intervals, in order to keep the lines updated.

This research proposed the construction of a system of **Classification of Patterns for Gas/Liquid Flow with Intelligent Models**. Techniques were applied artificial neural networks (ANNs) type multilayer perceptron (MLP), back-propagation with momentum term (BPM) and gradient method (GRAD), for identifying such patterns. In the research was used the database developed by the Laboratory of Multiphase Phenomena of Faculty of Mechanical Engineering (FEM) at Unicamp, where applied techniques of statistics transformations aimed diversifying the set of inputs used in the classification.

For evaluation, the main correlations of measurement found in the different sets of experiments performed were done statistics tests, to select the settings with ate more prosperous for the classification of flow patterns. Since, more prosperous settings obtained mean results between 67.73% to 97.57% accuracy.

Keywords: Classification of Patterns, Gas/Liquid Flow, Intelligent Models, Artificial Neural Networks, Transformations Statistics, Statistical Analysis.

LISTA DE FIGURAS

FIGURA 1 – REGIMES DE ESCOAMENTO GÁS/LÍQUIDO VERTICAL (CORRÊA, 2009, P. 4).	33
FIGURA 2 - REGIMES DE ESCOAMENTO GÁS/LÍQUIDO HORIZONTAL (THERMOPEDIA™, 2006).	34
FIGURA 3 – MAPA DOS PADRÕES DE ESCOAMENTO (JL X JG).	49
FIGURA 4 – VALORES DE TENSÃO DO PADRÃO AGITADO (#39) EM 60S E 3S.....	51
FIGURA 5 – VALORES DE TENSÃO DO PADRÃO ANULAR (#1) EM 60S E 3S.....	51
FIGURA 6 – VALORES DE TENSÃO DO PADRÃO BOLHAS (#2) EM 60S E 3S.....	51
FIGURA 7 – VALORES DE TENSÃO DO PADRÃO CAPA ESFÉRICA (#4) EM 60S E 3S.....	52
FIGURA 8 – VALORES DE TENSÃO DO PADRÃO GOLFADAS (#14) EM 60S E 3S.....	52
FIGURA 9 – VALORES DE TENSÃO DO PADRÃO SEMI-ANULAR (#65) EM 60S E 3S.....	52
FIGURA 10 – DESCRIÇÃO DO <i>BOXPLOT</i>	54
FIGURA 11 – <i>BOXPLOT</i> PARA OS VALORES DE TENSÃO DO PADRÃO AGITADO.....	55
FIGURA 12 – <i>BOXPLOT</i> PARA OS VALORES DE TENSÃO DO PADRÃO ANULAR.....	55
FIGURA 13 – <i>BOXPLOT</i> PARA OS VALORES DE TENSÃO DO PADRÃO BOLHAS.....	55
FIGURA 14 – <i>BOXPLOT</i> PARA OS VALORES DE TENSÃO DO PADRÃO CAPA ESFÉRICA.....	56
FIGURA 15 – <i>BOXPLOT</i> PARA OS VALORES DE TENSÃO DO PADRÃO GOLFADAS.....	56
FIGURA 16 – <i>BOXPLOT</i> PARA OS VALORES DE TENSÃO DO PADRÃO SEMI-ANULAR.....	56
FIGURA 17 – MÉDIA DOS VALORES DE TENSÃO PARA TODOS OS PADRÕES.....	59
FIGURA 18 – DESVIO PADRÃO DOS VALORES DE TENSÃO PARA TODOS OS PADRÕES.....	59
FIGURA 19 – ASSIMETRIA DOS VALORES DE TENSÃO PARA TODOS OS PADRÕES.....	60
FIGURA 20 – CURTOSE DOS VALORES DE TENSÃO PARA TODOS OS PADRÕES.....	61
FIGURA 21 – VALORES MÍNIMOS DOS VALORES DE TENSÃO PARA TODOS OS PADRÕES.....	62
FIGURA 22 – VALORES MÁXIMOS DOS VALORES DE TENSÃO PARA TODOS OS PADRÕES.....	62
FIGURA 23 – MODA DOS VALORES DE TENSÃO PARA TODOS OS PADRÕES.....	63
FIGURA 24 – MEDIANA DOS VALORES DE TENSÃO PARA TODOS OS PADRÕES.....	63
FIGURA 25 – HISTOGRAMA COM 180.000 PONTOS PARA TODOS OS PADRÕES.....	66
FIGURA 26 – HISTOGRAMA COM 424 PONTOS PARA TODOS OS PADRÕES.....	66
FIGURA 27 – HISTOGRAMA COM 300 PONTOS PARA TODOS OS PADRÕES.....	67
FIGURA 28 – HISTOGRAMA COM 58 PONTOS PARA TODOS OS PADRÕES.....	68
FIGURA 29 – HISTOGRAMA COM 18 PONTOS PARA TODOS OS PADRÕES.....	68
FIGURA 30 – PDF COM 180.000 PONTOS PARA TODOS OS PADRÕES.....	70
FIGURA 31 – PDF COM 424 PONTOS PARA TODOS OS PADRÕES.....	70
FIGURA 32 – PDF COM 300 PONTOS PARA TODOS OS PADRÕES.....	71
FIGURA 33 – PDF COM 58 PONTOS PARA TODOS OS PADRÕES.....	72
FIGURA 34 – PDF COM 18 PONTOS PARA TODOS OS PADRÕES.....	72
FIGURA 35 – USO DO TERMO <i>MOMENTUM</i> NO ALGORITMO BP (LACERDA, 2006, P. 71).....	77
FIGURA 36 – ESTRUTURA NEURAL PARA A SAÍDA ÚNICA.....	79
FIGURA 37 - ESTRUTURA NEURAL PARA A SAÍDA MÚLTIPLA.....	80
FIGURA 38 - OSCILAÇÃO DOS PARÂMETROS DE CONFIGURAÇÃO.....	81
FIGURA 39 - ESTAGNAÇÃO DOS PARÂMETROS DE CONFIGURAÇÃO.....	82
FIGURA 40 – <i>BOXPLOT</i> DOS RESULTADOS UTILIZANDO A BASE COM M. AMOSTRAIS.....	94
FIGURA 41 – <i>BOXPLOT</i> DOS RESULTADOS UTILIZANDO A BASE COM HISTOGRAMAS.....	96
FIGURA 42 – <i>BOXPLOT</i> DOS RESULTADOS UTILIZANDO A BASE COM PDFs.....	97
FIGURA 43 – VARIAÇÃO DOS EXPERIMENTOS BPM PARA PARÂMETROS DE CONFIGURAÇÃO.....	101
FIGURA 44 – VARIAÇÃO DOS EXPERIMENTOS BPM PARA PARÂMETROS DE CONFIGURAÇÃO.....	101

FIGURA 45 – PARÂMETROS DE CONFIGURAÇÃO PARA OS RESULTADOS USANDO BPM.....	102
FIGURA 46 – PARÂMETROS DE CONFIGURAÇÃO PARA OS RESULTADOS USANDO GRAD.	102
FIGURA 47 – ESTRUTURA NEURONAL (TATIBANA & KAETSU, 2006).	126
FIGURA 48 – ESTRUTURA DE UM NEURÔNIO ARTIFICIAL (SALGADO, 2009, P. 2).....	127
FIGURA 49 – PRINCIPAIS FUNÇÕES DE ATIVAÇÃO (ARAÚJO, 2008, P. 25).....	127
FIGURA 50 – CAMADAS DE UMA REDE NEURAL ARTIFICIAL (SALGADO, 2009, P. 3).	128

LISTA DE TABELAS

TABELA I - CATEGORIZAÇÃO DOS SINAIS DE ACORDO COM OS PADRÕES DE ESCOAMENTO.....	49
TABELA II – TEMPO DE PROCESSAMENTO E ESPAÇO EM MEMÓRIA PARA AS BASES DE DADOS.	74
TABELA III – DIVISÃO DA BASE DE DADOS (TREINAMENTO E TESTE).....	86
TABELA IV – AGRUPAMENTO DOS EXPERIMENTOS PARA O BPM COM MOMENTOS AMOSTRAIS.	89
TABELA V – AGRUPAMENTO DOS EXPERIMENTOS P/ O GRAD COM MOMENTOS AMOSTRAIS.	90
TABELA VI – AGRUPAMENTO DOS EXPERIMENTOS PARA O BPM COM HISTOGRAMAS.	90
TABELA VII – AGRUPAMENTO DOS EXPERIMENTOS PARA O GRAD COM HISTOGRAMAS.....	91
TABELA VIII- AGRUPAMENTO DOS EXPERIMENTOS PARA O BPM COM PDFs.	91
TABELA IX – AGRUPAMENTO DOS EXPERIMENTOS PARA O GRAD COM PDFs.....	92
TABELA X – ANÁLISE DOS RESULTADOS UTILIZANDO BPM E MOMENTOS AMOSTRAIS.....	93
TABELA XI – ANÁLISE DOS RESULTADOS UTILIZANDO GRAD E MOMENTOS AMOSTRAIS.	93
TABELA XII – ANÁLISE DOS RESULTADOS UTILIZANDO BPM E HISTOGRAMAS.....	95
TABELA XIII – ANÁLISE DOS RESULTADOS UTILIZANDO GRAD E HISTOGRAMAS.	95
TABELA XIV – ANÁLISE DOS RESULTADOS UTILIZANDO BPM E PDFs.	96
TABELA XV – ANÁLISE DOS RESULTADOS UTILIZANDO GRAD E PDFs.	97
TABELA XVI – ANÁLISE DOS TEMPOS DE EXECUÇÃO DO SISTEMA PARA O BPM.	99
TABELA XVII – ANÁLISE DOS TEMPOS DE EXECUÇÃO DO SISTEMA PARA O GRAD.....	99
TABELA XVIII – PARÂMETROS DOS RESULTADOS MAIS PRÓSPEROS.....	100

LISTA DE ABREVIACÕES

ANN	<i>artificial neural network</i> (vide RNA)
BCC	Bacharelado em Ciência da Computação
BP	<i>backpropagation</i>
BPM	<i>backpropagation with momentum term</i> (<i>backpropagation</i> com termo <i>momentum</i>)
DFP	método Davidon-Fletcher-Powell
DIC	delineamento inteiramente casualizado
FDP	função de densidade de probabilidade (vide PDF)
FEEC	Faculdade de Engenharia Elétrica e de Computação
FEM	Faculdade de Engenharia Mecânica (<i>Faculty of Mechanical Engineering</i>)
FR	método Fletcher-Reeves
Gb	gigabytes
GHz	gigahertz
GRAD	<i>gradient method</i> (método gradiente)
h	horas
Hz	hertz
IA	inteligência artificial
JG	velocidade do gás
JL	velocidade do líquido
KHz	quilohertz
LEnS	Laboratório de Engenharia de Software
LInC	Laboratório de Inteligência Computacional
m	metros
m/s	metros por segundo
min	minutos
M5DAC	média, mediana, moda, mínimo, máximo, desvio padrão, assimetria e curtose

MCP	modelo de McCulloch e Pitts
MD	média e desvio padrão
MDAC	média, desvio padrão, assimetria e curtose
MLP	multilayer perceptron
PDF	<i>probability density function</i> (vide FDP)
PR	método Polak-Ribière
PSDF	função de distribuição de probabilidade
RAM	<i>random access memory</i> (memória de acesso randômico)
RBF	funções de base radial
RNA	rede neural artificial (vide ANN)
s	segundos
SE	sistema especialista
TCC	Trabalho de Conclusão de Curso
Unicamp	Universidade Estadual de Campinas
UNIFAL-MG	Universidade Federal de Alfenas

SUMÁRIO

1 INTRODUÇÃO	23
1.1 JUSTIFICATIVA E MOTIVAÇÃO	23
1.2 PROBLEMATIZAÇÃO.....	24
1.3 OBJETIVOS	25
1.3.1 Gerais	25
1.3.2 Específicos	25
1.4 ORGANIZAÇÃO DA MONOGRAFIA.....	26
2 CLASSIFICAÇÃO DE ESCOAMENTO BIFÁSICO	29
2.1 ESCOAMENTO BIFÁSICO GÁS/LÍQUIDO	29
2.1.1 Sentido do Fluxo.....	31
2.1.1.1 Regime Vertical Ascendente.....	32
2.1.1.2 Regime Horizontal.....	33
2.1.2 Aquisição dos Sinais	35
2.2 SINAIS DE ESCOAMENTO	36
2.2.1 Tratamento dos Sinais.....	37
2.3 RECONHECIMENTO DE PADRÕES	38
2.3.1 Modelos de Classificação de Padrões	39
2.3.1.1 Classificação de Padrões Estatísticos.....	40
2.3.1.2 Redes Neurais Artificiais	41
2.3.1.3 Redes Neuro Fuzzy.....	43
2.3.1.4 Sistemas Especialistas.....	44
2.3.2 Classificação de Padrões.....	45
2.3.3 Considerações	46
3 BASE DE DADOS	47
3.1 CONSIDERAÇÕES INICIAIS	47
3.2 AQUISIÇÃO DA BASE DE DADOS	48
3.2.1 Valores de Tensão.....	50
3.3 PRÉ-PROCESSAMENTO DOS SINAIS.....	53
3.3.1 Momentos Amostrais.....	57
3.3.1.1 Base MD	58
3.3.1.2 Base MDAC	60
3.3.1.3 Base M5DAC	61
3.3.2 Histogramas	64
3.3.2.1 Cálculo do Histograma	64
3.3.2.2 Distribuição de Frequência do Histograma.....	65
3.3.3 Função de Densidade de Probabilidade.....	69
3.3.3.1 Cálculo da PDF.....	69
3.4 BASE DE DADOS SELECIONADAS.....	73
4 METODOLOGIA PROPOSTA	75
4.1 PROPOSTA	75
4.1.1 Algoritmos de Treinamento	76
4.1.1.1 <i>Backpropagation</i> com Termo <i>Momentum</i>	76
4.1.1.2 Método Gradiente.....	78

4.1.2 Tipos de Saídas.....	79
4.1.3 Busca Exaustiva de Configurações.....	81
5 ESTUDO DE CASO.....	83
5.1 METODOLOGIA APLICADA.....	83
5.1.1 Configuração do Sistema.....	84
5.1.2 Base de Treinamento e Teste.....	85
5.2 RESULTADOS.....	87
5.2.1 Análise de Variância.....	87
5.2.2 Teste dos Melhores Conjuntos de Configurações.....	93
5.2.2.1 Momentos Amostrais.....	93
5.2.2.2 Histogramas.....	94
5.2.2.3 PDFs.....	96
5.3 CONSIDERAÇÕES FINAIS.....	98
6 CONCLUSÃO E TRABALHOS FUTUROS.....	105
6.1 CONCLUSÃO.....	105
6.2 TRABALHOS FUTUROS.....	107
7 REFERÊNCIAS BIBLIOGRÁFICAS.....	109
8 ANEXOS.....	115
8.1 BACKPROPAGATION COM TERMO MOMENTUM.....	115
8.2 MÉTODO GRADIENTE.....	116
8.3 UNIDIM.....	118
8.4 TESTE DOS PESOS.....	119
8.5 TESTE DE PADRÕES.....	120
8.6 PERMUTAÇÃO DOS PARÂMETROS PARA O BPM.....	121
8.7 PERMUTAÇÃO DOS PARÂMETROS DO GRAD.....	122
9 APÊNDICE.....	125
9.1 REDES NEURAIIS ARTIFICIAIS.....	125
9.1.1 Funcionamento.....	126
9.1.2 Processo de Aprendizado.....	128
9.1.3 Algoritmo de Treinamento.....	129
9.1.4 Aplicação das Redes.....	130

1

Introdução

*A adversidade desperta em nós a capacidades que,
em circunstâncias favoráveis, teriam ficado adormecidas.*
(Horácio)

O capítulo 1 apresenta a motivação e os objetivos para o desenvolvimento deste projeto. A seção 1.1 apresenta os motivos e a justificativa para a escolha do tema. Na seção 1.2, é encontrada a descrição do problema (problematização). Os objetivos, gerais e específicos, são descritos na seção 1.3. A organização do projeto é apresentada ao final do capítulo, na seção 1.4.

1.1 Justificativa e Motivação

Nos últimos anos, tem-se notado um considerável interesse acadêmico e industrial no uso da fluidodinâmica computacional para modelar o sistema de escoamento bifásico em processos químicos.

Segundo Corrêa (2009, p. 1) "*O escoamento bifásico gás-líquido está presente na natureza e em diferentes atividades industriais, entre elas na indústria química, de petróleo e de geração de energia elétrica.*". "*Neste tipo de escoamento, as fases presentes, gás e líquido, podem escoar dispostas em diferentes configurações espaciais no interior do duto, denominadas de padrões ou regimes de escoamento.*" (Corrêa, 2009, p. Resumo).

Devido às mudanças contínuas de regimes nas tubulações, há a dificuldade em controlar algumas situações nas linhas de escoamento como: queda de pressão, controle de vazão, gerência de produção, entre outras. Para isso, é necessário um

sistema capaz de reconhecer e classificar padrões, em determinados intervalos de tempo, a fim de manter as linhas atualizadas.

As ferramentas de reconhecimento e classificação de padrões são importantes em muitas áreas, tais como na determinação de comportamentos típicos e em sistemas de controle de fluxo. Nessa conjuntura, a proposta deste trabalho consiste na classificação de padrões de escoamento bifásico vertical no regime gás/líquido através de redes neurais artificiais (RNAs) do tipo multilayer perceptron (MLP). Serão conduzidos estudos visando a contribuição para um melhor entendimento e capacidade de identificação de fluidos nos sistemas de escoamentos bifásicos.

1.2 Problematização

A classificação de padrões de escoamento é a capacidade de identificar e classificar diferentes regimes (ou padrões) em uma tubulação. Em ambientes reais, a classificação deve ser feita de forma contínua, rápida e com alto grau de confiabilidade.

Nas indústrias químicas e petrolíferas, é grande a importância de reconhecer os regimes de escoamento para que se possa estimar a variação de pressão e de vazão, assim como planejar e gerenciar o fluxo nas linhas de escoamento, dentre outras situações.

A dificuldade em capturar e processar uma grande quantidade de dados, além das mudanças contínuas de padrões, faz com que a tomada de decisão seja um processo complexo, principalmente pela quantidade de características que o problema pode apresentar. Dessa forma, há a dificuldade de selecionar e determinar um conjunto reduzido de características, que mantenha um alto grau de acerto e confiabilidade para o sistema de escoamento.

A utilização de diferentes abordagens para atacar o problema tem o intuito de tornar o sistema de classificação mais robusto e diversificado. Pretende-se trabalhar com a natureza estatística do problema, tanto nos dados de entrada, aplicando métodos estatísticos, quanto com os resultados obtidos através de análises estatísticas, fazendo uso da capacidade das redes neurais em adquirir conhecimento pelo processo de aprendizagem.

Este trabalho está limitado apenas à classificação de padrões para sistema de escoamento gás/líquido vertical, não tendo foco no reconhecimento, captura e transmissão dos dados. Para isso, será utilizada a base de dados desenvolvida pelo Laboratório de Fenômenos Multifásicos da Faculdade de Engenharia Mecânica da Unicamp, para avaliação das principais correlações de medição de escoamentos bifásicos gás/líquido.

1.3 Objetivos

1.3.1 Gerais

Este projeto objetiva desenvolver um sistema que possibilite a identificação e classificação de padrões para sistemas de escoamento bifásicos gás/líquido, pretendendo-se determinar o padrão, em termos da proporção entre a quantidade de líquido e de gás que passa por uma dada seção da tubulação.

Com o intuito de encontrar os resultados mais prósperos, serão feitos diferentes tratamentos na base de dados associados a buscas exaustivas de parâmetros para o algoritmo de reconhecimento de padrões.

1.3.2 Específicos

Os objetivos específicos deste projeto são os seguintes:

- Compreender e analisar os diferentes tipos de padrões de escoamento bifásico vertical no regime gás/líquido;
- Empregar diferentes métricas, calculadas em função da base de dados, para a utilização no sistema;
- Definir, avaliar e aplicar modelos inteligentes para o reconhecimento de padrões;
- Examinar a utilização de uma quantidade reduzida de valores necessários para a identificação de padrões de escoamento;
- Avaliar estatisticamente os resultados obtidos com o intuito de validar o sistema proposto, tornando hábil a divulgação dos conhecimentos obtidos e práticas aplicadas no projeto.

1.4 Organização da Monografia

Este trabalho encontra-se dividido nos seguintes capítulos:

O capítulo 2 apresenta uma contextualização em torno da classificação de padrões para o escoamento bifásico gás/líquido, inserindo seus conceitos e variações clássicas. Após a definição, são apresentadas as diferentes técnicas de classificação de padrões. Também é apresentada uma revisão bibliográfica sobre os principais processos envolvidos para a classificação de padrões de escoamento.

No capítulo 3, é apresentado todo o processo de tratamento da base de dados. Inicialmente é descrito o processo de aquisição da base em função dos valores de tensão captados em uma tubulação. Em seguida, é detalhado todo o processo de tratamento de sinais, onde são aplicados momentos amostrais, cálculos de histograma e PDF (função de densidade de probabilidade), a fim de ampliar e

diversificar o poder de classificação do sistema. Por fim, é feita uma descrição das bases selecionadas e apresentada uma análise comparativa destas.

O capítulo 4 apresenta a proposta principal deste projeto, expondo o sistema de classificação de padrões composto de métodos inteligentes. As redes neurais foram os métodos trabalhados, variando a utilização dos algoritmos de treinamento, *backpropagation* com termo *momentum* (BPM) e método gradiente (GRAD).

Os experimentos realizados com o sistema de classificação são apresentados no capítulo 5. É feita uma descrição da metodologia aplicada e as configurações utilizadas para a execução dos experimentos. Em seguida, são apresentados os resultados obtidos com as diferentes bases utilizadas, que são confrontados com testes estatísticos para a seleção das configurações que obtiveram os resultados mais prósperos. Por fim, são apresentados os tempos de execução e uma análise geral dos resultados.

O capítulo 6 apresenta as conclusões gerais em torno do projeto, as principais contribuições e os possíveis trabalhos futuros.

As referências bibliográficas estão disponíveis no capítulo 7.

No capítulo 8, são apresentados os anexos desta monografia, estando presentes os algoritmos utilizados para o sistema de classificação de padrões

Por fim, o capítulo 9, apêndice, é descrito o funcionamento de uma rede neural artificial (RNA) expondo seus conceitos e processos envolvidos.

2

Classificação de escoamento Bifásico

*As crianças têm mais necessidade de modelos do que de críticas.
(Joseph Joubert)*

O capítulo 2 faz uma descrição dos processos envolvidos para a classificação de sistemas de escoamento bifásico gás/líquido. A seção 2.1 apresenta uma introdução dos sistemas de escoamento multifásicos e bifásicos, assim como a sua importância nas indústrias. Ainda nesta seção são apresentados os diferentes sentidos que as substâncias podem fluir, vertical e horizontal. Os tratamentos aplicados na base de escoamento são apresentados na seção 2.2. Na seção 2.3, é feita uma descrição de sistemas de reconhecimento, e um detalhamento da classificação de padrões e das técnicas de classificação conhecidas.

2.1 Escoamento Bifásico Gás/Líquido

Escoamento bifásico é o fluxo de uma ou mais substâncias com fases distintas, homogêneas ou heterogêneas, através de uma tubulação, sendo que, para cada tipo de sistema, há um conjunto de padrões ou regimes de escoamento que o compõe.

"A definição de fase não é relativa ao estado da matéria (sólido, líquido, gasoso ou plasma), mas sim ao número de interfaces presentes no escoamento multifásico" (Silva, Borges, & Pinheiro, 2000, p. 46). Isto é, um sistema de escoamento pode apresentar apenas uma interface, mas o fluxo de substâncias diferentes (líquidos imiscíveis, líquido-gás, entre outras).

O fluxo de duas ou mais substâncias que escoam, simultaneamente, por uma tubulação, é conhecido como escoamento multifásico. Uma das ramificações do escoamento multifásico é o escoamento bifásico, caracterizado pelo fluxo com apenas duas substâncias, tanto homogêneas quanto heterogêneas, por uma tubulação.

O fluxo simultâneo de diferentes substâncias por uma tubulação pode produzir diversos tipos de regimes de escoamento, diferenciando-se um do outro pela distribuição espacial das fases na mistura.

Whalley (1996) *apud* (Souza, 2009, p. 2) define que o regime de escoamento depende das:

- vazões de gás e de líquido;
- propriedades dos fluídos (massa, viscosidade e tensão superficial);
- condições de operação (pressão, temperatura, gravidade, etc);
- características geométricas do duto (forma, diâmetro e inclinação).

Os sistemas de escoamento são encontrados na natureza e aplicados em diversas áreas industriais. A identificação precisa do regime é crítica para sistemas de escoamento, e dessa forma, essencial para diversos processos em indústrias químicas, nucleares e petrolíferas, tais como caldeiras, condensadores, linhas de transporte de gás e óleo e sistemas de resfriamento de reatores nucleares.

Corrêa (2009, p. 1) faz uma descrição mais detalhada da utilização destes sistemas nas indústrias.

"Na indústria de geração termoelétrica e nuclear, o escoamento bifásico ocorre nos circuitos percorridos pelo fluído de trabalho, entre a geração de calor e a expansão nas turbinas, e, também, nas linhas de retorno. Na indústria química, o escoamento bifásico ocorre em reatores, colunas de destinação e absorção, atomizadores, caldeiras, trocador de calor, reatores e linhas de processo. Na indústria de petróleo, o escoamento bifásico está presente em todo o percurso dos fluídos produzidos, desde a rocha-

reservatório às unidades de separação, passando pela coluna de produção, linha de surgência e riser, além das unidades industriais de refino."

Nesta vertente são encontrados, em diversos processos industriais, sistemas de escoamentos bifásicos que apresentam o fluxo de substâncias gasosas, líquidas e sólidas pelas tubulações.

Dos quatro tipos de escoamento bifásico (gás-líquido, gás-sólido, líquido-líquido e líquido-sólido), o escoamento gás-líquido é o mais complexo, pois combina as características de deformação interfacial e compressão de um dos fluxos. Para o fluxo de duas fases em uma tubulação, a distribuição interfacial gás-líquido possui uma infinidade de possibilidades de fluxos. No entanto, esses fluxos podem ser classificados em diferentes tipos de distribuição interfacial, comumente chamados de regimes de escoamento ou fluxo de padrões.

"O tipo de padrão em cada trecho do duto será dependente do diâmetro e inclinação do duto, das frações volumétricas entre as fases no trecho de interesse e das propriedades das fases" (Corrêa, 2009, p. 1).

2.1.1 Sentido do Fluxo

Segundo Pacheco *et al.* (2007, p. 2) *"devido ao caráter complexo do escoamento multifásico, diversas metodologias foram desenvolvidas com a finalidade de identificar os padrões de escoamento"*.

Freqüentemente é observado o fluxo de substâncias no interior das tubulações, podendo ocorrer em trechos verticais, horizontais ou inclinados. O fluxo dessas substâncias pode alterar de tempos em tempos, mudando os regimes de escoamentos, ou seja, pode haver uma sucessão de padrões diferentes passando pela mesma tubulação.

Os padrões mais comuns são classificados em função do comportamento das fases dentro da tubulação. Diversos pesquisadores classificam os padrões,

baseando-se no sentido do fluxo no duto, sendo os principais o escoamento horizontal e vertical. Detalhes sobre esses padrões são encontrados em (Hewitt, 1982), (Whalley, 1987), (Dukler & Taitel, 1986), (Ishii, 1975), (Mishima & Ishii, 1984) e (Wallis, 1969).

2.1.1.1 Regime Vertical Ascendente

Para sistemas de escoamento bifásico são encontrados o fluxo de substâncias gás-líquido em regimes verticais ascendentes. De acordo com Corrêa (2009, p. 3), os principais padrões formados neste tipo de sistema são:

- **Anular (Annular):** *O líquido flui envolto do duto formando um filme com bolhas dispersas, e o gás escoo no centro da tubulação, carreando gotículas de líquido;*
- **Semi-Anular (Wispy Annular):** *É um sub-padrão entre o anular e o agitado. Onde o fluxo de líquido é maior, a concentração de gotas no núcleo do gás aumenta, levando à formação de grandes fluxos de líquido;*
- **Agitado (Unstable Slug):** *O movimento do líquido é agitado e instável, e constantemente a integridade das bolhas é rompida pela instabilidade do movimento líquido;*
- **Golfadas (Slug):** *É caracterizado por uma sucessão de pistões de líquido aerados seguidos por bolhas de gás alongadas, escoando verticalmente ao redor de um filme de líquido junto à parede;*
- **Capa Esférica (Spherical Cap):** *É caracterizado por bolhas que podem tomar parcialmente ou completamente a seção transversal do tubo na forma de uma "capa" ou "tampa" de gás. Usualmente, este sub-padrão está agrupado com o padrão de bolhas;*
- **Bolhas (Bubble):** *A fase gasosa se encontra dispersa na fase líquida em forma de bolhas discretas, sendo a fase líquida contínua. Esse escoamento ocorre tipicamente para baixas velocidades superficiais de gás.*

Na Figura 1 podem ser vistos todos os padrões de escoamento para o regime vertical descrito acima.

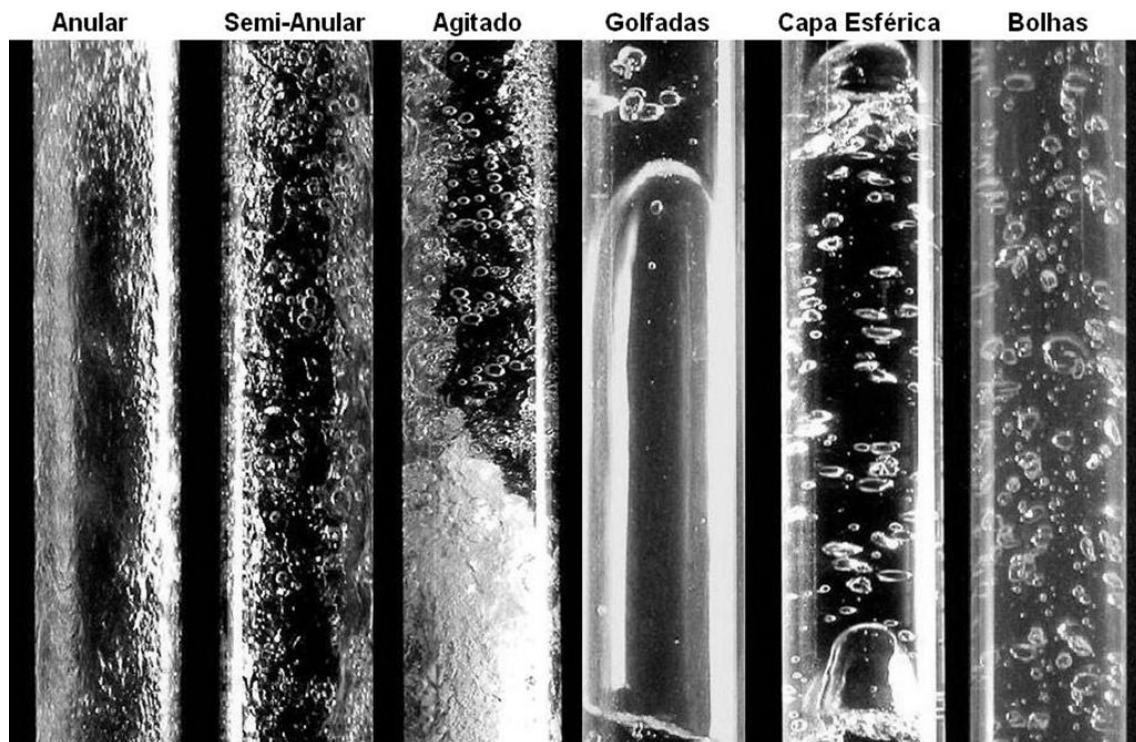


Figura 1 - Regimes de escoamento gás/líquido vertical (Corrêa, 2009, p. 4).

Além destes padrões típicos, consideram-se outros padrões de escoamento. Quando a tubulação está cheia de líquido e vazia de gás, e quando ela está vazia de líquido e cheia de gás, sendo que esta informação é importante para o operador da linha de escoamento.

2.1.1.2 Regime Horizontal

Além dos regimes verticais, também são encontrados sistemas de escoamento de regime horizontal. Souza (2009, p. 3) apresenta os diversos padrões que constituem esse tipo de regime.

- **Borbulhado (Bubble Flow):** *Nesse tipo de regime ocorre a dispersão de bolhas de gás no meio do fluxo contínuo de líquido, sendo que as bolhas ficam concentradas na parte superior;*

- **Estratificado (Stratified Flow):** *As fases fluem separadas, o gás na parte superior e o líquido na parte inferior do tubo. Esse padrão de escoamento está associado a baixas velocidades para os dois fluídos e não é comum de ocorrer;*
- **Ondulado (Wavy Flow):** *As fases fluem separadas, porém com uma interface irregular, apresentando ondulações. Essas ondas estão associadas a velocidades altas do gás;*
- **Tamponado (Plug Flow):** *Os regimes formam bolsões alongados de gás que movem na parte superior do tubo, podendo haver ou não a presença de pequenas bolhas na região imediatamente abaixo dos bolsões;*
- **Pistonado (Slug Flow):** *Ocorre o rompimento da interface líquido-gás em alguns trechos e partes do líquido são suspensas na fase gasosa. Esse escoamento tende a ocorrer à medida que a velocidade do gás aumenta;*
- **Anular (Annular Flow):** *A formação de película de líquido na parede do tubo envolvendo a fase gasosa que escoar como um núcleo, sendo que a espessura do filme líquido na parte inferior do tubo é maior que na superior.*

A Figura 2 representa o fluxo para regimes de escoamento horizontal.

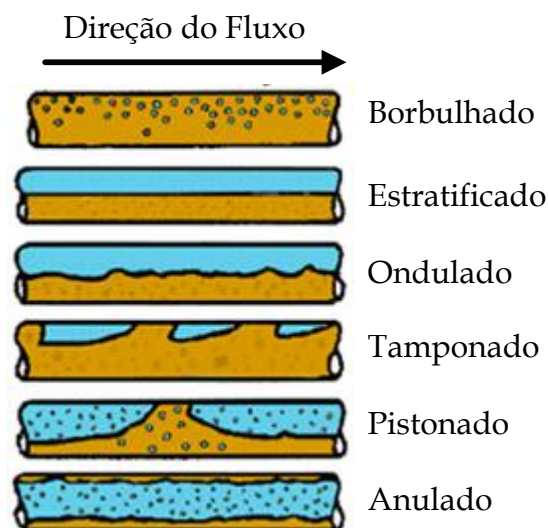


Figura 2 - Regimes de escoamento gás/líquido horizontal (Thermopedia™, 2006).

Nos sistemas de regime horizontal, ocorre a presença dos mesmos padrões básicos presentes no regime vertical. A tubulação cheia de líquido ou cheia de gás.

2.1.2 Aquisição dos Sinais

Para a aquisição, de sinais, foram identificadas técnicas de absorção de radiação, técnicas ópticas, variação de pressão, variação da impedância elétrica, gradiente de pressão e modelos matemáticos.

Jones & Zuber (1975) e Vince & Lahey (1982) utilizaram sistemas de raios-X para a identificação de padrões em função da análise de fração de gás. Esta análise somente é possível de ser realizada com tubulações opacas.

Durst (1976) e Mishima & Hibiki (1996) trabalharam com técnicas ópticas, porém necessitaram que as tubulações fossem transparentes.

Xu & Han (1997) e Warsito *et al.* (1999) fizeram o uso de técnicas de tomografia computadorizada ultra-sônica.

Medições em função da variação de pressão, diferença de pressão entre gás e líquido, foram realizadas por diversos pesquisadores como Tutu (1982), Hulin, Fierfort & Condol (1994), Sun & Zhang (2008).

Uma das técnicas é obtida pela medida da fração do gás sobre uma linha, área ou volume em relação ao total de uma seção restrita da tubulação, descrevendo localmente as proporções, de gás e líquido, de cada fase do escoamento (Corrêa, 2009).

Algumas pesquisas utilizaram a variação da impedância elétrica (Tsoukalas, Ishii & Mi (1997), Mi, Ishii & Tsoukalas (2001) e Lee *et al.* (2004)), onde utilizaram sensores de impedância não intrusivos. Tsoukalas, Ishii & Mi (1997) e Mi, Ishii & Tsoukalas (2001) obtiveram o sinal com uma frequência de amostragem de 200 Hz durante 60 segundos. Lee *et al.* (2004) usaram uma frequência de amostragem de 1 KHz durante um tempo de 60 segundos.

De acordo com Mi *et al.* (1996), técnicas de impedância são também utilizadas para medições de fração de gás, além de que o tempo de resposta dos sensores é feita em tempo real.

Hagedorn & Brown (1965) e Beggs & Brill (1973) fizeram uso do gradiente de pressão para o desenvolvimento de uma correlação empírica de cálculos, que são aplicados sobre qualquer inclinação e orientação das tubulações.

Taitel, Dukler & Barnea (1980) apresentaram um modelo mecanicista para previsão da transição entre os padrões de escoamento em regimes verticais.

Independentemente do fluxo de diferentes substâncias ou mesmo do regime de escoamento, o processo de aquisição de conhecimento é vital para o sucesso de qualquer sistema. Em função disto, várias técnicas vêm sendo desenvolvidas para a identificação de padrões, onde o tratamento de sinais é feito a partir de certas propriedades do escoamento.

2.2 Sinais de Escoamento

"Embora haja algumas regularidades no processo de aquisição de conhecimento, ainda não foi possível estabelecer um consenso sobre um método eficaz. O que existe de concreto são recomendações e linhas gerais de atuação" (Rezende, 2005, p. 54).

Como não existem métodos gerais para a aquisição de conhecimento, algumas técnicas de aquisição podem auxiliar nesse processo.

Uma das técnicas envolve a tecnologia de aprendizado de máquina, *"cujo objetivo é o desenvolvimento de técnicas computacionais sobre o aprendizado bem como a construção de sistemas capazes de adquirir conhecimento de forma automática"* (Rezende, 2005, p. 89).

Outra técnica é caracterizada pela aquisição do conhecimento de forma manual, onde o conhecimento é adquirido por meio do acompanhamento de

especialistas, sendo que o uso de ferramentas que minimizem o problema com ruídos (erros na aquisição dos dados) torna o processo semi-automático.

Uma das práticas utilizadas para a aquisição de conhecimento é através da fluidodinâmica computacional que simula, através de modelos da mecânica dos fluidos, os processos físicos e/ou químicos presentes em sistemas de escoamento.

Souza (2009, p. 7) diz que, com a fluidodinâmica computacional

"apenas alguns padrões de escoamentos bifásicos líquido-gás podem ser simulados com bons resultados. Não existe ainda uma representação matemática suficientemente geral, que descreva a interface líquido-gás para todos os padrões de escoamento conhecidos. Essas interfaces são na verdade contornos físicos dinâmicos de topologia extremamente complexa, que se movem e deformam ao longo da tubulação."

Entretanto, há outras práticas que fazem a aquisição de conhecimento sem o uso de modelos matemáticos. Estas outras têm características de técnicas de aquisição manuais ou mesmo semi-automáticas.

2.2.1 Tratamento dos Sinais

Independentemente da forma como os sinais são adquiridos, a maioria dos trabalhos não os utiliza no estado bruto. Isto por causa da dificuldade em identificar os padrões a partir dos sinais. Assim, muitas pesquisas fazem uso de transformações dos sinais adquiridos para a utilização nas metodologias propostas.

Uma das transformações é com o cálculo da função de densidade de probabilidade (PDF) que representa de forma contínua a distribuição de probabilidade de um conjunto de dados, sendo que o número de frequência de dados varia entre pesquisas. Esta transformação foi utilizada em Tutu (1982), Matsui (1984), Mi *et al.* (1996), Tsoukalas, Ishii & Mi (1997) e Rosa *et al.* (2010).

Lee *et al.* (2004) utilizam a função de distribuição de probabilidade (PSDF), que é a ordenação dos sinais adquiridos em função da magnitude em determinados períodos de tempo.

Momentos estatísticos são utilizados por Mi, Ishii & Tsoukalas (1998) e Corrêa (2009), os quais calculam a média e o desvio padrão dos sinais.

A transformada de Gabor sobre os sinais tem o objetivo de realizar uma análise em tempo-frequência e foi utilizado por Selli & Selegim (2004). Bin *et al.* (2006) e Ding *et al.* (2007) utilizaram também um método de análise de tempo-frequência, a transformada de Hilbert Huang.

Tendo em vistas as várias técnicas de aquisição de conhecimento para sistemas de escoamento, é importante ter em mente que cada uma se aplica melhor a determinados sistemas e contextos. Pois, *"todo conhecimento reflete um padrão, mas nem todo padrão pode ser considerado conhecimento. Um padrão deve ser previsível para se tornar conhecimento. A sabedoria vem com o uso bem sucedido do conhecimento"* (Rezende, 2005, p. 85).

2.3 Reconhecimento de Padrões

Segundo Cera (2005, p. 4) *"Os seres humanos são capazes de reconhecer padrões com grande rapidez. Faz parte da natureza humana, ao observar um objeto, fazer uma coleta de informações, as quais são comparadas com as propriedades e comportamentos conhecidos e armazenados em sua mente."* Através dessa comparação, os seres humanos são capazes de reconhecer o alvo de sua observação.

Este conceito de reconhecimento é simples e familiar para todos em um ambiente real, mas no contexto da inteligência artificial (IA), o reconhecimento não é uma tarefa trivial. *"Não existe nenhuma máquina ou software de inteligência artificial*

que seja capaz de igualar-se à capacidade de reconhecimento do cérebro humano" (Jesan, 2005) *apud* (Cera, 2005, p. 4).

O reconhecimento de padrões computacional envolve três níveis de processamento: filtragem da entrada, extração de características e classificação. Jesan (2005) *apud* (Cera, 2005, p. 4) descreve esses processos da seguinte maneira:

- **Filtragem da entrada:** *tem o objetivo de eliminar dados desnecessários ou distorcidos fazendo com que a entrada apresente apenas dados relevantes para o reconhecimento do objeto em análise;*
- **Extração de Características:** *consiste da análise dos dados de entrada a fim de extrair e derivar informações úteis para o processo de reconhecimento;*
- **Classificação:** *é o estágio final do reconhecimento de padrões onde, através da análise das características da entrada de dados, é declarado como pertencente ou não a uma determinada categoria.*

Tendo em vista que este projeto visa somente à classificação de padrões, serão apresentados, a seguir, mais detalhes sobre diferentes técnicas envolvidas neste processo.

2.3.1 Modelos de Classificação de Padrões

Um classificador de padrões consiste, basicamente, em uma técnica na qual objetiva armazenar em um computador o conhecimento necessário para, futuramente, classificar determinado objeto, baseando-se em determinadas características.

Esta técnica utiliza uma base de dados que contém características de vários objetos e a classe a qual este pertence. Trabalhando nesses dados, o sistema localiza uma forma de separá-los em suas diferentes classes, de acordo com essas características disponíveis, permitindo assim a classificação de novos padrões. Os sistemas classificadores começam sem nenhum conhecimento, e através de

algoritmos de treinamento, são capazes de aprender, de modo que são capazes de buscarem novas soluções para classificar os objetos.

Atualmente, há diversos sistemas de classificação de padrões, cada qual com características diferentes e dessa forma se adaptam melhor a determinados tipos de problemas, a seguir são apresentados os que mais destacam na classificação de padrões.

2.3.1.1 Classificação de Padrões Estatísticos

Segundo Bishop (1995), a forma mais geral e natural de formular soluções para a identificação de padrões é a classificação de padrões estatísticos, por meio da qual é classificada a natureza estatística tanto da informação que se quer representar quanto dos resultados.

Schalkoff (1992) diz que a classificação de padrões estatísticos, como sugere o próprio nome, assume uma base estatística para os algoritmos de classificação. Um conjunto de medidas, que denotam as características, é extraído dos dados de entrada e usado para associar cada vetor de características a uma entre n classes. Presume-se que as características são geradas por um estado da natureza, e portanto, existe um modelo subordinado a um estado com um conjunto de probabilidades e/ou funções de densidade de probabilidade (PDF) correspondente, passível de ser usado para representá-lo.

O classificador toma decisões baseando-se no aprendizado realizado a partir de um conjunto de treinamento, o qual contém exemplos de padrões de todas as classes existentes no sistema. A classificação é realizada utilizando estimativas de distribuições de probabilidade. O classificador de padrões é avaliado através de um conjunto de testes, preferencialmente composto por padrões de todas as classes, mas que não estejam no conjunto de treinamento.

2.3.1.2 Redes Neurais Artificiais

Bilobrovec, Marçal & Pilatti (2004, p. 2) dizem que as redes neurais consistem em:

“um método de solucionar problemas de inteligência artificial, construindo um sistema que tenha circuitos que simulem o cérebro humano, inclusive seu comportamento, ou seja, aprendendo, errando e fazendo descobertas. São mais que isso, são técnicas computacionais que apresentam um modelo inspirado na estrutura neural de organismos inteligentes e que adquirem conhecimento através da experiência.”

Uma rede neural artificial é um modelo de grafo orientado disposto em uma rede em que os nós representam neurônios artificiais e as arestas orientadas são as conexões entre as entradas e saídas dos neurônios. As redes neurais podem ser vistas como máquinas massivamente paralelas com muitos processadores simples e muitas interconexões (Jain, 2000). As redes neurais podem ser hábeis a aprender relações entre entradas e saídas complexas e não lineares. Dentre outras características, inclui-se a possibilidade de usar procedimentos seqüenciais de treinamento e o poder de adaptação aos dados.

O processo de aprendizado envolve atualização da arquitetura da rede e dos pesos das conexões, de forma que a rede pode efetuar tarefas específicas de classificação e análise de agrupamentos (*clustering*).

As redes neurais possuem a capacidade de se adaptarem a problemas específicos utilizando diferentes tipos de aprendizagem. Uma delas é a aprendizagem supervisionada, onde as saídas desejadas são fornecidas junto com os padrões para as entradas. As saídas são supervisionadas por um agente humano, que coleta os dados e faz uma análise desses. O algoritmo de treinamento verifica o erro para cada padrão e altera os pesos da rede a fim de minimizá-lo.

A popularidade das redes neurais cresceu devido ao fato que, aparentemente, elas têm uma baixa dependência a um específico domínio, de

forma que o mesmo tipo de rede pode ser aplicado em vários problemas, o que não ocorre com as abordagens baseadas em modelos ou em regras.

Outro fator importante na popularidade é o fato de existirem muitos algoritmos eficientes de aprendizado com redes neurais. As redes neurais disponibilizam conjuntos de algoritmos não lineares para extração de características (usando camadas escondidas) e classificação (multilayer perceptrons - MLP). Além disso, algoritmos de extração de características podem ser mapeados em arquiteturas de redes neurais para implementação eficiente em hardware.

Do ponto de vista estrutural, as redes neurais podem ser classificadas dentro de dois grandes grupos: redes neurais estáticas ou não recorrentes, e redes neurais dinâmicas ou recorrentes.

As redes neurais estáticas são aquelas cujas saídas dependem somente do valor atual das entradas e não de valores passados nem futuros. *"A estrutura estática mais utilizada na literatura é sem dúvida a Rede Multicamadas com Função de Ativação Sigmoidal, isto devido a sua capacidade de aproximação universal e processamento paralelo"* (Scarselli & Tsoi, 1998). Outro modelo estático importante é denominado rede neural com funções de ativação de base radial (RBF) onde possui apenas uma única camada oculta, e os neurônios possuem função de ativação gaussiana, em vez de sigmoidal.

As redes neurais recorrentes são estruturas de processamento capazes de representar uma grande variedade de comportamentos dinâmicos. *"A presença de realimentação de informação permite a criação de representações internas e dispositivos de memória capazes de processar e armazenar informações temporais e sinais seqüenciais"* (Lima, 2000, p. 24).

Apesar de possuírem aparentemente princípios diferentes, a maioria dos modelos de redes neurais populares são implicitamente equivalentes ou similares a métodos clássicos de classificação estatística de padrões. Entretanto, as redes neurais oferecem várias vantagens, como abordagens unificadas para extração e

classificação, além de procedimentos flexíveis para encontrar boas soluções não lineares.

2.3.1.3 Redes Neuro Fuzzy

Segundo Oliveira & Belderrain (2006, p. 2), a lógica fuzzy consiste em:

"aproximar a decisão computacional da decisão humana, tornando as máquinas mais capacitadas a seu trabalho. Isto é feito de forma que a decisão de uma máquina não se resume apenas a um 'sim' ou um 'não', mas também tenha decisões 'abstratas', do tipo 'um pouco mais', 'talvez sim', e outras tantas variáveis que representem as decisões humanas. É um modo de interligar inerentemente processos analógicos que deslocam-se através de uma faixa contínua para um computador digital que podem ver coisas com valores numéricos bem definidos (valores discretos)."

"A rede neuro fuzzy é uma tecnologia da inteligência artificial que gera regras diretamente de dados numéricos, e associa probabilidades a estas regras" (Buragohain & Mahanta, 2008) *apud* (Nose, Lotufo, & Lopes, 2008, p. 2). A rede neuro fuzzy combina a lógica fuzzy com o aprendizado e capacidades adaptativas das redes neurais. Portanto, o usuário pode inicializar a rede neural usando conhecimento apurado expressado com regras fuzzy, e então treiná-la usando dados numéricos. Após treinar a rede, as regras podem ser avaliadas e editadas para sobrepor qualquer deficiência nos dados treinados.

Uma das principais potencialidades da lógica fuzzy é que suas bases de conhecimento, as quais estão no formato de regras de produção, são fáceis de examinar e entender. Este formato de regra também torna fácil a manutenção e a atualização da base de conhecimento.

"As redes neurais e os sistemas fuzzy têm sido intensamente aplicados a muitos problemas incluindo identificação de sistema, predição, classificação e controle" (Nose, Lotufo, & Lopes, 2008, p. 2). Os sistemas fuzzy têm sido desenvolvidos para gerenciar conhecimento de maneira mais natural.

2.3.1.4 Sistemas Especialistas

Segundo Nikolopoulos (1997) *apud* (Dapper, Strauss, Leiria, & Furlanetto, 2003, p. 1) "*um sistema especialista é um programa de computador inteligente que usa o conhecimento e procedimentos de inferência para resolver problemas que são difíceis o bastante para requerer um especialista humano para sua solução.*".

Os sistemas especialistas resolvem problemas através do conhecimento de humanos especialistas, dessa forma não é possível resolver problemas em geral. De acordo com Dapper *et al.* (2003, p. 3), a arquitetura básica dos sistemas especialistas consiste em quatro processos:

- **Aquisição do Conhecimento:** consiste no entendimento do problema através de informações passadas por especialistas. Com base nessas informações, é formulado o domínio do problema;
- **Base de Conhecimento:** contém a representação do conhecimento, geralmente através de regras lógicas (se - então - senão), para a resolução do problema;
- **Máquina de Inferência:** "*A máquina de inferência contém algoritmos gerais, os quais estão aptos a manipular o conhecimento armazenado na base de conhecimento para a solução de problemas.*" Dapper *et al.* (2003, p. 5);
- **Interface de entrada e saída:** é o modo como o usuário interage com os dados utilizados e os resultados gerados pelo sistema especialista.

"O Conhecimento envolvido para solucionar problemas complexos do mundo real, são geralmente incertos, imprecisos e incompletos." Nikolopoulos (1997) *apud* (Dapper, Strauss, Leiria, & Furlanetto, 2003, p. 7). Apesar do desenvolvimento de um sistema especialista ser trabalhoso, este é justificado quando o conhecimento sobre determinado assunto é escasso, ou quando a presença de um especialista não é possível devido à dificuldade em alocá-lo em lugares hostis (Corrêa, 2009, p. 28).

2.3.2 Classificação de Padrões

Diversas metodologias para a classificação de padrões foram propostas para os sistemas de escoamento, tais como: lógica fuzzy, redes neurais, redes neuro fuzzy e sistemas especialistas. Assim, são apresentados algumas pesquisas e os resultados obtidos por essas.

Utilizando de um conjunto de regras lógicas (se - então - senão), Dailiang *et al.* (2004) trabalharam com a lógica fuzzy, e obtiveram a identificação dos padrões em tempo real, variando de 85,2% a 96,4% de acertos.

Diversos pesquisadores utilizam as redes neurais artificiais (RNAs) para a classificação de padrões devido a sua simplicidade estrutural e funcional.

Mi, Ishii & Tsoukalas (2001) utilizaram valores estatísticos de impedância como entrada das RNAs e identificaram em tempo real os padrões de escoamentos.

A topologia de Kohonen foi adotado por Cai *et al.* (1994), enquanto Lee *et al.* (2004) utilizou das topologias multilayer perceptron (MLP) e Kohonen.

A RNA com a arquitetura de função de ativação base radial foi adotada por Dong *et al.* (2003) e Malyeri, Muller-Steinhagen & Smith (2003). Já Rosa *et al.* (2010) utilizaram redes MLP, onde os resultados foram de 92%.

Wu, Zhou & Wu (2001) fizeram uso do algoritmo backpropagation (BP) para a identificação de padrões, alcançando 92% de acertos.

Mi *et al.* (1996) usaram a combinação das redes neurais com a lógica fuzzy, conhecida como redes neuro fuzzy, e obtiveram até 100% de acertos.

Corrêa (2009) desenvolveu um sistema especialista, o qual através de um conjunto de regras lógicas é capaz de fazer a identificação de padrões, sendo que seus resultados variaram de 70% a 100% dependendo do tipo de teste adotado.

Salgado *et al.* (2010) utilizaram técnicas de agrupamento para criar a base de dados de forma otimizada, e obteve de 70% a 91% de acertos com o uso de sistemas especialistas.

2.3.3 Considerações

Dentro os modelos de classificação de padrões citados acima, as redes neurais artificiais do tipo multilayer perceptron é as que serão utilizadas neste projeto. A escolha desse modelo está no fato da sua grande utilização para resolver problemas de classificação, inclusive de escoamento.

As redes são facilmente aplicadas, pois podem ter várias entradas e saídas. Além disso, elas trabalham com dados imprecisos, onde possuem a capacidade de generalização dos resultados que não estão presentes na sua base de conhecimento.

O poder de adaptação e aprendizagem das redes neurais é um fator muito atrativo, levando assim ao avanço tecnológico dos hardwares voltados para as redes neurais, nos quais agregam maior poder computacional.

3

Base de Dados

*Em todas as coisas o sucesso depende de uma preparação prévia,
e sem tal preparação o fracasso é certo.
(Confúcio)*

O capítulo 3 tem por finalidade descrever todo o processo de construção da base de dados. A seção 3.1 define de forma geral o conceito de base de dados. Já a seção 3.2 explica como foi feita a aquisição desta base, assim como a categorização dos padrões adquiridos. Na seção 3.3, são apresentados os diferentes tipos de tratamentos feitos em cima da base de dados, tais como momentos amostrais, histogramas e função de densidade de probabilidade. Por fim, a seção 3.4 faz uma análise geral dos tratamentos das bases a serem utilizadas.

3.1 Considerações Iniciais

Uma base de dados pode ser entendida como uma forma matemática de estruturar e codificar o conhecimento sobre uma determinada aplicação, ou seja, contém uma abstração do mundo descrita por um formalismo processável computacionalmente.

"Obter e representar todo o conhecimento de um determinado domínio é extremamente difícil" (Rezende, 2005, p. 47). *"A representação do conhecimento é um dos problemas cruciais de inteligência artificial (IA)"* (Araribóia, 1987) *apud* (Rezende, 2005, p. 30).

"Para resolver problemas nas áreas em que, até o momento, o ser humano é mais bem sucedido que as máquinas, as máquinas resolvidoras de problemas precisam saber o que

os resolvedores humanos sabem sobre o assunto" ((Rich & Knight, 1993) apud (Rezende, 2005, p. 14)).

Dessa forma, com o intuito de aprimorar os conhecimentos e agregar novos dados para os modelos de escoamento, os pesquisadores da Faculdade de Engenharia Mecânica (FEM) da Universidade Estadual de Campinas (Unicamp) exploraram de forma exaustiva os limites do Laboratório de Escoamento, a fim de obterem novos sinais para diversos padrões de escoamento.

3.2 Aquisição da Base de Dados

O processo de aquisição da base de dados foi baseado no escoamento bifásico de duas substâncias imiscíveis, gás e líquido, que fluem por uma tubulação no regime vertical ascendente.

Através de um sistema supervisionado, foram produzidos diversos padrões de escoamento, variando a velocidade de líquido e de gás (JL e JG) que passam pela tubulação. Após o padrão de escoamento entrar em regime estacionário, inicia-se o processo de aquisição, através do instrumento de medição. A identificação do padrão é feita por um humano especialista que observa o fluxo por uma secção transparente da tubulação. Corrêa (2009, p. 37) apresenta mais detalhes sobre como são feitos os processos de aquisição da base de dados.

Na Figura 3 é possível analisar o gráfico, na base logarítmica, com as diferentes velocidades de líquido e gás empregadas para a aquisição dos dados. A velocidade do líquido varia até 3,08 m/s enquanto a do gás atinge até 28,75 m/s.

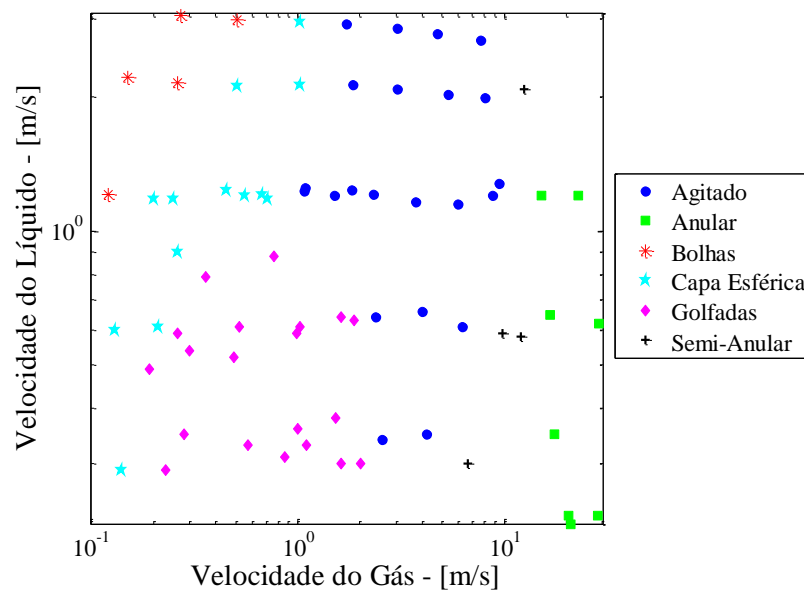


Figura 3 - Mapa dos padrões de escoamento (JL x JG).

Durante a aquisição dos sinais, o especialista humano catalogou 73 diferentes padrões produzidos na tubulação, como pode ser visto na Tabela I. Estes padrões foram divididos em seis categorias, de acordo com o tipo de escoamento trabalhado. E, com o intuito de facilitar a visualização dos gráficos subseqüentes, foram atribuídas cores idênticas para os padrões de mesma categoria.

Tabela I - Categorização dos sinais de acordo com os padrões de escoamento.

Padrão	Sinais	Cor
Agitado	#39, #40, #41, #44, #45, #48	Azul
Anular	#1, #27, #29, #30, #31, #37, #38, #42, #43, #46, #47, #50, #51	Verde
Bolhas	#2, #3, #5, #6, #19, #20, #21, #22, #23, #24, #25, #26, #28, #32, #33, #34, #35, #36, #52, #53	Vermelho
Capa Esférica	#4, #7, #8, #9, #10, #11, #12, #13, #15, #16, #18, #49, #54, #55, #56, #57, #58, #59, #60, #62, #63, #64	Cyan
Golfadas	#14, #17, #61, #70	Magenta
Semi-Anular	#65, #66, #67, #68, #69, #71, #72, #73	Preto

Simultaneamente a esse processo, foram feitas análises, através de sensores, do fluxo de gás e líquido que escoam por uma tubulação. Tais análises resultaram na representação do fluxo por meio de valores de tensão, de modo que seja possível a utilização destes dados para a identificação de padrões de escoamento.

3.2.1 Valores de Tensão

A informação sobre o fluxo gás/líquido em uma dada seção é obtida através de uma sonda capacitiva, que apresenta um sinal em termos de tensão (volts). Quando a tubulação está cheia de líquido, a tensão é de 5 volts, e quando cheia de gás o valor é de 1 volt. Esses sinais variam bastante devido ao processo de escoamento, já que se torna difícil identificar o padrão diretamente a partir desse sinal.

Esses valores foram adquiridos com uma frequência de 3 KHz durante 60 segundos, totalizando 180.000 pontos para cada um dos 73 padrões da base. Dessa forma, é necessária uma análise dos dados, em um espaço de tempo, que seja capaz de diferenciar os vários padrões possíveis de serem encontrados.

A fim de facilitar a visualização da dinâmica que os valores de tensão comportam para os diferentes padrões, são apresentados dois tipos de figuras, uma com 60 segundos representando todos os pontos adquiridos, e a outra com uma segmentação desta primeira, com um intervalo de 3 segundos, sendo que somente a primeira com 60 segundos foi utilizada neste projeto.

Esta comparação entre os diferentes tempos de aquisição, 60 e 3 segundos, pode ser visualizada da Figura 4 à Figura 9, onde são comparados o tempo em função da tensão relativa e os diferentes padrões de escoamentos adquiridos.

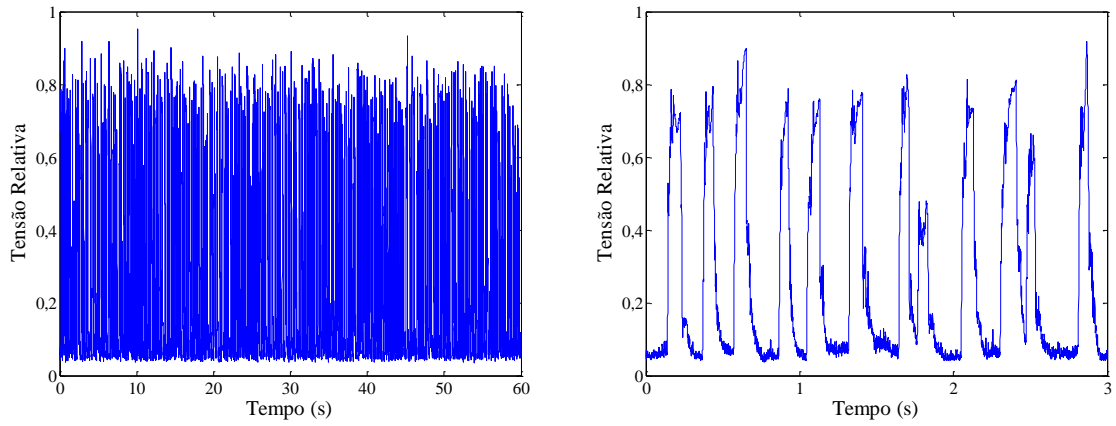


Figura 4 - Valores de tensão do padrão agitado (#39) em 60s e 3s.

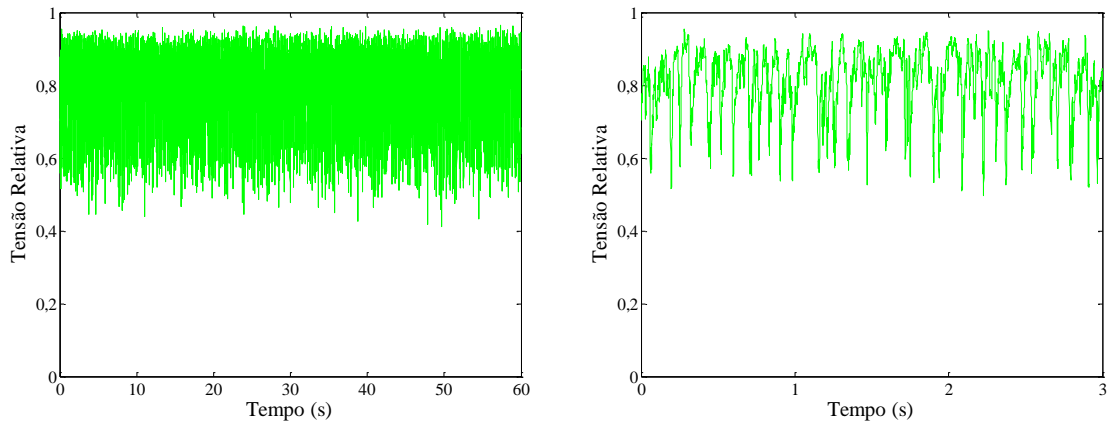


Figura 5 - Valores de tensão do padrão anular (#1) em 60s e 3s.

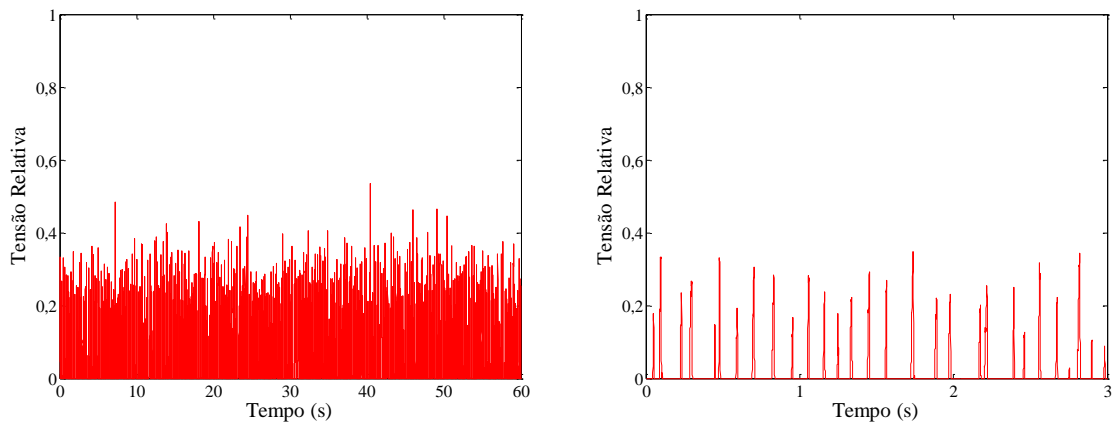


Figura 6 - Valores de tensão do padrão bolhas (#2) em 60s e 3s

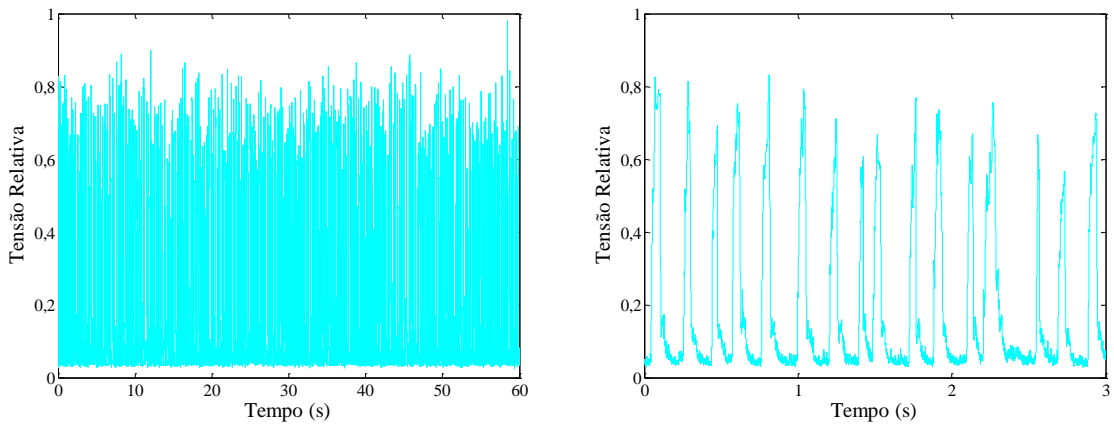


Figura 7 - Valores de tensão do padrão capa esférica (#4) em 60s e 3s.

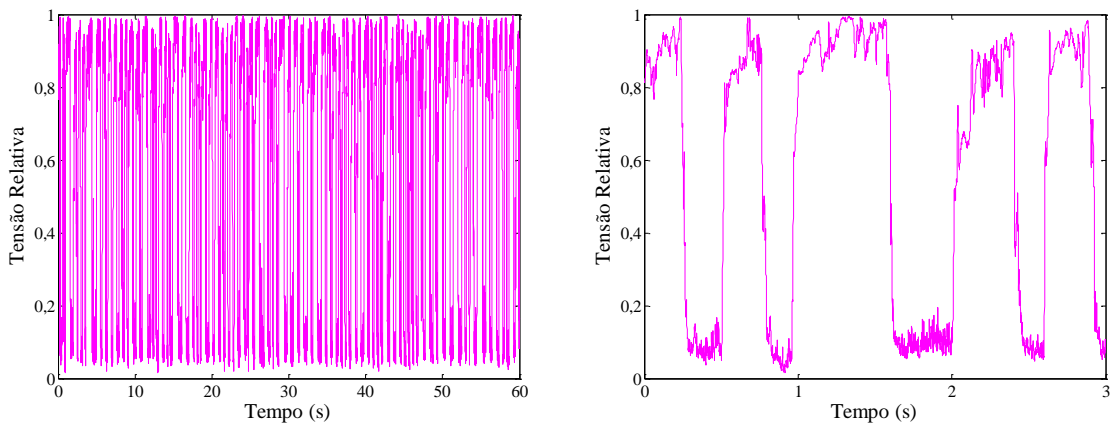


Figura 8 - Valores de tensão do padrão golfadas (#14) em 60s e 3s.

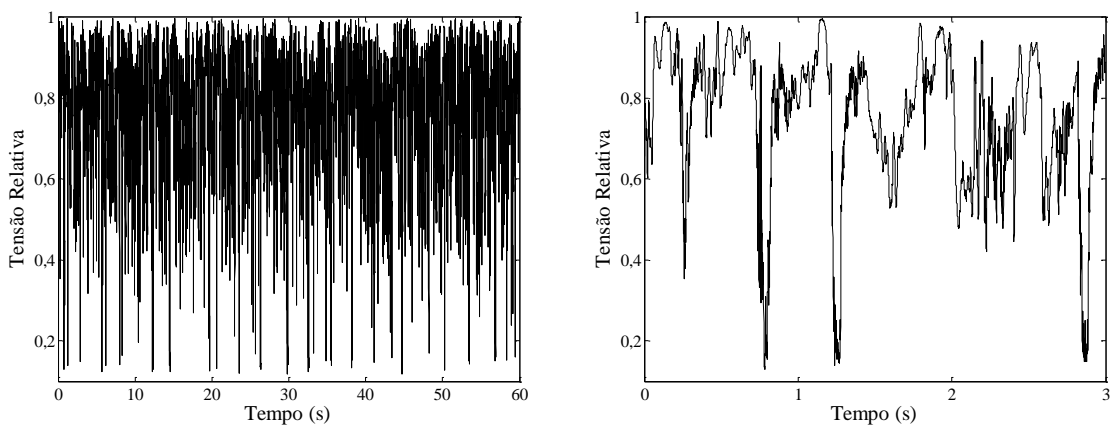


Figura 9 - Valores de tensão do padrão semi-anular (#65) em 60s e 3s.

Segundo Rosa *et al.* (2010, p. 8)

"Um fator de decisão que está atrelado ao problema de reconhecimento de padrões de uma maneira geral é como representar cada sinal de forma que o modelo de reconhecimento de padrões seja capaz de enxergar as diferenças existentes em cada sinal apresentado. Uma maneira direta de representação de um determinado padrão em um problema de reconhecimento de padrões é utilizar o próprio sinal (completo) sem nenhum tratamento ou manipulação para representar o respectivo padrão. No problema de reconhecimento de padrão de escoamento bifásico, tratado neste trabalho, a representação direta de um padrão de escoamento poderia ser feita através da apresentação de cada sinal de 180.000 pontos. O inconveniente desta abordagem é a quantidade de dados, gerando uma alta dimensão na entrada do modelo. Por exemplo, se for considerado um modelo baseado em redes neurais multicamadas, a arquitetura deste modelo será formada por 180.000 neurônios na camada de entrada, podendo causar a lentidão no processo de ajuste dos parâmetros da rede e até mesmo a não convergência do modelo em determinados casos.

Para contornar este problema de dimensionalidade uma alternativa encontrada é realizar um pré-processamento do sinal original buscando estimar um novo sinal com dimensão reduzida e que mantenha a representatividade do sinal original. Neste sentido, é possível encontrar na literatura alguns trabalhos científicos que utilizam procedimentos estatísticos ou transformadas específicas para realizar a etapa de pré-processamento do sinal."

3.3 Pré-Processamento dos Sinais

"Diante da deficiência para analisar e compreender grandes volumes de dados, diversos estudos têm sido direcionados ao desenvolvimento de tecnologias de extração automática de conhecimento de base de dados" (Rezende, 2005, p. 309).

Com o intuito de fazer uma análise estatística dos dados, utilizou-se de um recurso chamado "Diagrama de Caixas" ou "*BoxPlot*". Este recurso é uma maneira rápida de examinar um ou mais conjuntos de dados graficamente.

Em um gráfico *BoxPlot* são apresentadas diversas informações: quartil inferior (Q1), mediana (Q2), quartil superior (Q3), o valor adjacente inferior (*lower whisker*), e valor adjacente superior (*upper whisker*), sendo que os *whiskers* são os valores que estão dispersos em até 1,5 vezes de qualquer um dos quartis (Q1 e Q3). O *BoxPlot* também pode indicar que alguns dados são *outliers*, isto é, dados externos aos *whiskers*, inferior e superior. A Figura 10 apresenta a posição das informações citadas acima.

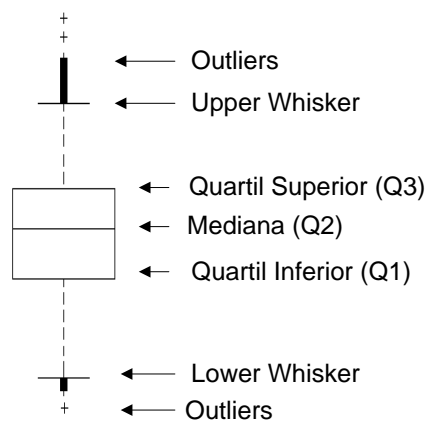


Figura 10 - Descrição do *BoxPlot*.

O *BoxPlot* pode parecer mais primitivo do que um histograma ou estimativa de densidade, mas ele tem algumas vantagens. Ele ocupa menos espaço e é, portanto, particularmente útil para a comparação entre as distribuições de vários grupos ou conjuntos de dados. Pode-se observar, a partir da Figura 11 à Figura 16, os diferentes resultados obtidos com o *BoxPlot* para os diversos padrões obtidos.

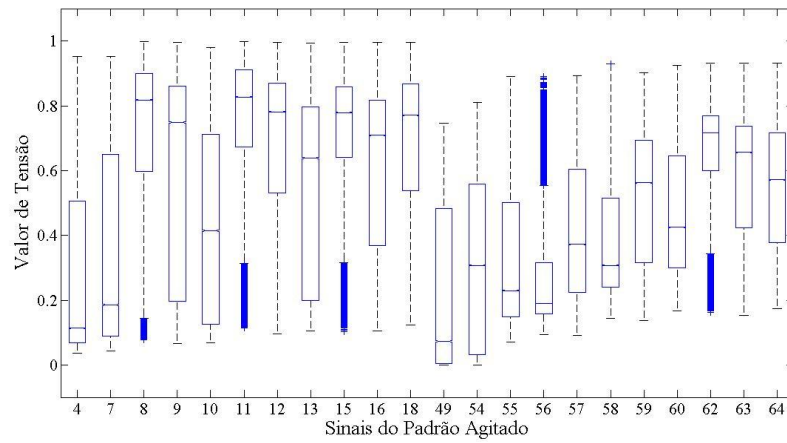


Figura 11 - BoxPlot para os valores de tensão do padrão agitado.

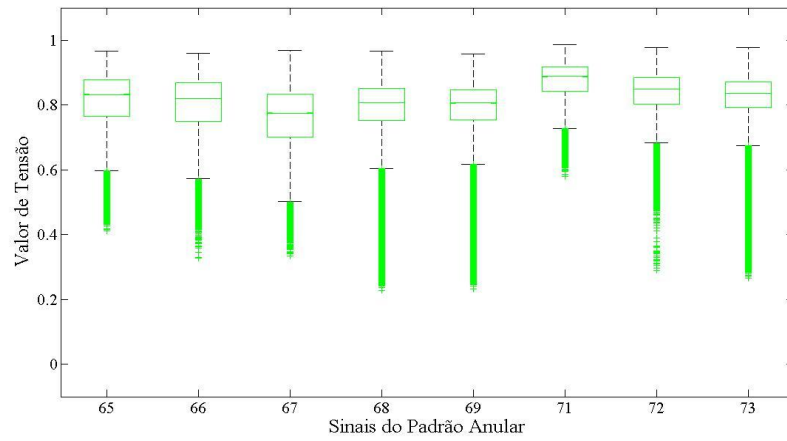


Figura 12 - BoxPlot para os valores de tensão do padrão anular.

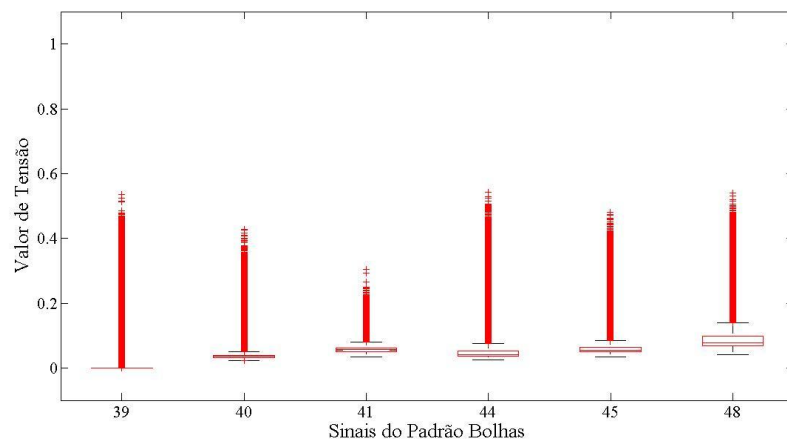


Figura 13 - BoxPlot para os valores de tensão do padrão bolhas.

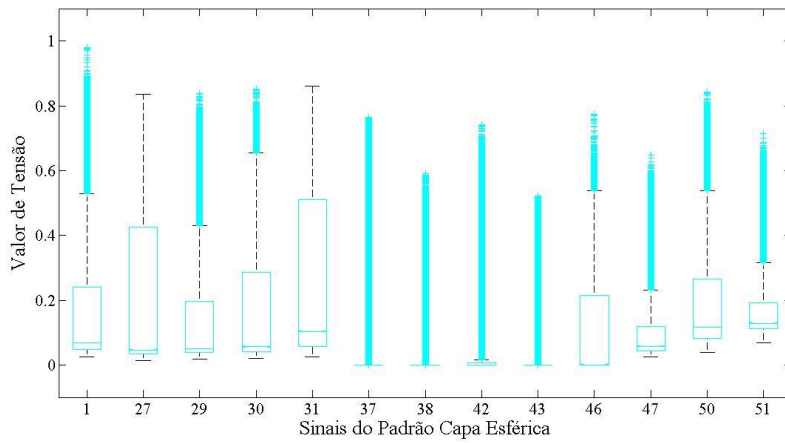


Figura 14 - BoxPlot para os valores de tensão do padrão capa esférica.

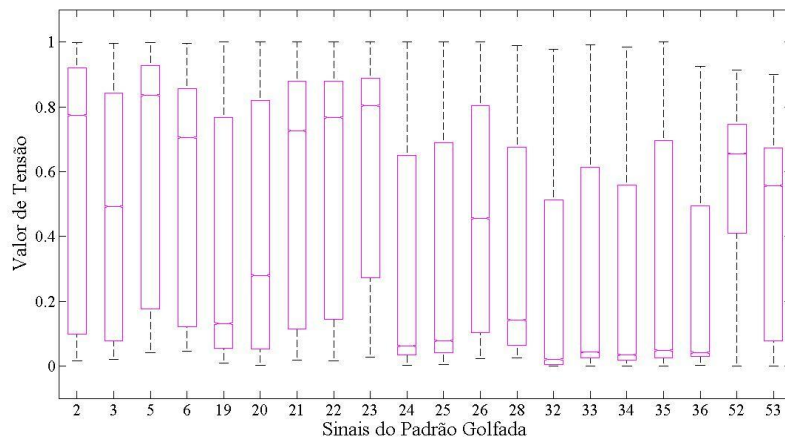


Figura 15 - BoxPlot para os valores de tensão do padrão golfadas.

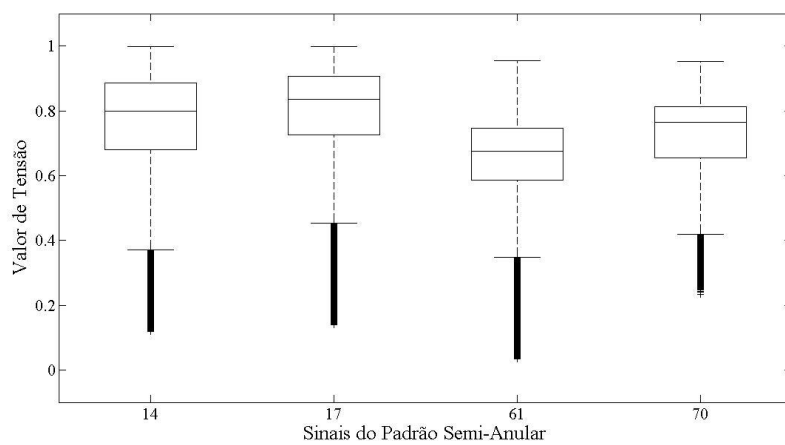


Figura 16 - BoxPlot para os valores de tensão do padrão semi-anular.

O *BoxPlot* mostra as diferenças entre as populações, sem fazer quaisquer suposições da distribuição estatística. A caixa central inclui os 50% dos dados centrais, os espaçamentos entre as diferentes partes da caixa ajudam a indicar o grau de dispersão (*spread*) e a simetria nos dados, visto que é relativamente fácil comparar grupos de populações.

A partir dessas informações foi possível estabelecer alguns métodos para o pré-processamento dos dados, que incluem o tratamento e redução do volume de dados. Segundo Rezende (2005, p. 317) "*o pré-processamento deve ser realizado criteriosamente e com o devido cuidado, uma vez que é fundamental garantir que as informações presentes nos dados brutos continuem presentes nas amostras geradas, para que os modelos finais sejam representativos da realidade expressa nos dados brutos.*".

Foram selecionados três diferentes tipos de tratamentos para processar e representar os dados: momentos amostrais, histogramas e função de densidade de probabilidade (PDF). A seguir são apresentados tais tratamentos.

3.3.1 Momentos Amostrais

Os dados foram analisados pela estatística descritiva, caracterizando seus comportamentos e calculando os momentos amostrais. Os momentos podem ser caracterizados como quantidades numéricas, calculadas a partir de um conjunto de dados, e que são utilizadas para fornecer uma descrição resumida da base utilizada.

Para a geração das bases de dados foram utilizadas três classes de medidas:

- **Medidas de Posição:** representa uma série de dados que orientam quanto à posição da distribuição em relação ao eixo horizontal do gráfico da curva de frequência. As medidas mais utilizadas são: média aritmética, moda, mediana, valor mínimo e máximo;

- **Medidas de Dispersão:** "*expressam a maior ou menor concentração de partículas em torno da média*" (Dias, 2004, p. 31). Determina a variabilidade ou dispersão dos dados, em relação à medida de localização do centro da amostra. Foi utilizado o desvio padrão;
- **Medidas de Assimetria:** proporcionam a descrição e compreensão completa da distribuição de frequências. A assimetria e curtose foram as medidas utilizadas.

Baseando nas medidas acima, foram geradas três bases de dados, utilizando diferentes combinações. A descrição mais detalhada das bases é feita a seguir.

3.3.1.1 Base MD

A primeira base foi construída utilizando duas medidas: Média aritmética e desvio padrão (medida de posição e dispersão respectivamente). A nomenclatura foi baseada na primeira inicial de cada medida (M – média; D – desvio padrão).

A média aritmética é a razão entre a soma total dos valores pela quantidade de valores. Ela representa a quantidade total expressa pelos dados, sendo que ao multiplicar a média pelo total de elementos, obtém a quantidade pretendida.

O desvio padrão é a raiz quadrada da média aritmética dos quadrados dos desvios. "*Indica a dispersão dos dados dentro da amostra, isto é, o quanto os dados em geral diferem da média. Quanto menor o desvio padrão, mais parecidos são os valores da série estatística*" (Ferrari, 2004, p. 25).

Os resultados dos cálculos da média e desvio padrão podem ser vistos nas Figura 17 e Figura 18, respectivamente, sendo estes agrupados pelos padrões.

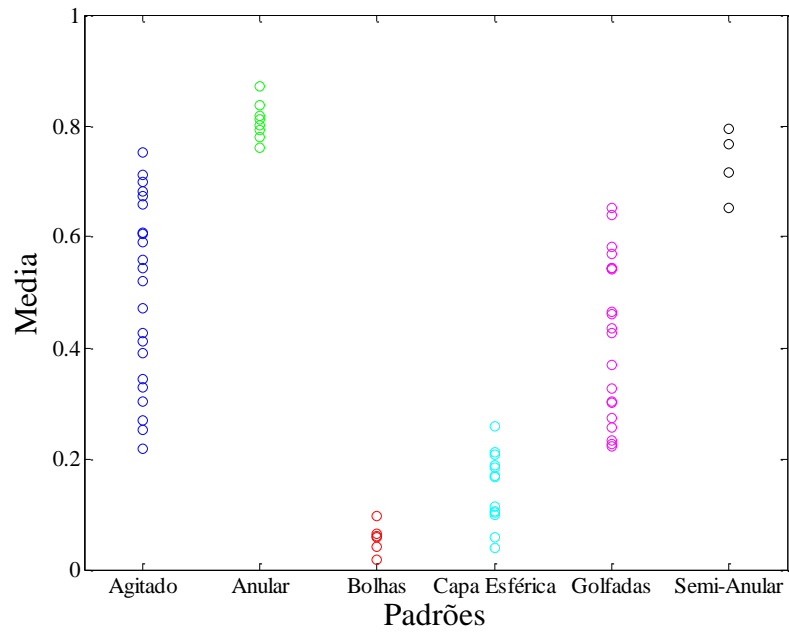


Figura 17 - Média dos valores de tensão para todos os padrões.

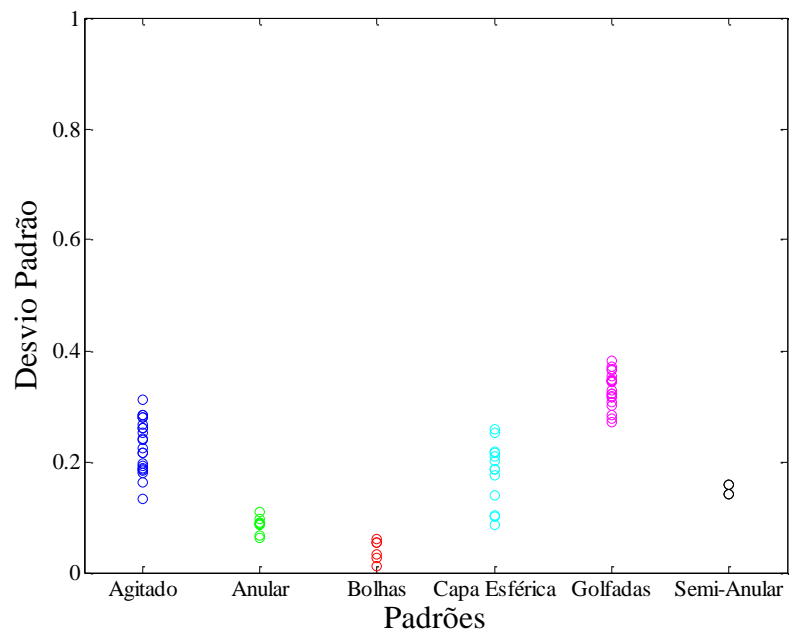


Figura 18 - Desvio padrão dos valores de tensão para todos os padrões.

3.3.1.2 Base MDAC

A segunda base foi construída utilizando as duas medidas anteriores (média aritmética e desvio padrão) e mais duas medidas de assimetria: assimetria e curtose (M - média; D - desvio padrão; A - assimetria; C - curtose).

A assimetria é quanto a média, mediana e a moda apresentam valores distintos. O deslocamento desses valores pode ser para a direita ou para a esquerda, variando o grau de deformação.

A curtose mede o grau de achatamento de uma distribuição em relação a uma distribuição padrão ou normal.

A Figura 19 e Figura 20 apresentam os valores encontrados da assimetria e curtose para a base de dados.

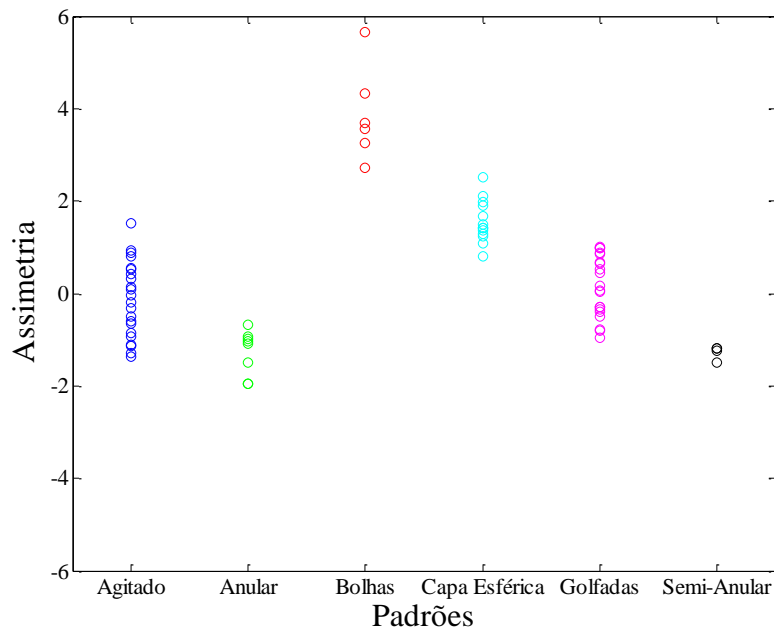


Figura 19 - Assimetria dos valores de tensão para todos os padrões.

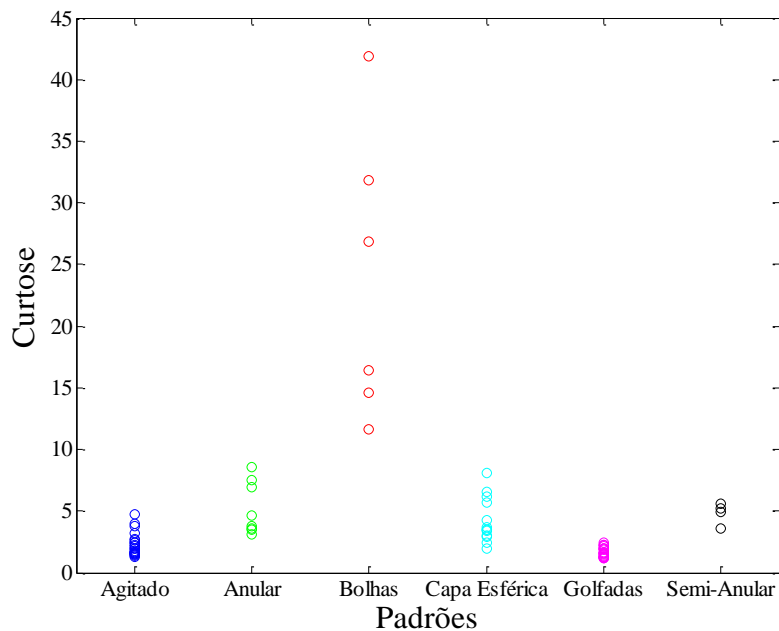


Figura 20 - Curtose dos valores de tensão para todos os padrões.

3.3.1.3 Base M5DAC

A última base, utilizando momentos amostrais, faz uso de todas as medidas estatísticas citadas, e ainda inclui a mediana, moda, valor mínimo e máximo. A nomenclatura foi baseada nos mesmos princípios anteriores, a letra inicial das medidas, mas substituindo todos os 'M' por 'M5' (M5 - média, mediana, moda, valor mínimo e máximo; D - desvio padrão; A - assimetria; C - curtose).

Os valores mínimos e máximos são os valores extremos (inferior e superior) presentes na base. A mediana é o valor situado de tal forma no conjunto, que o separa em dois subconjuntos de mesmo número de elementos, enquanto moda é o valor que ocorre com maior frequência em uma série de valores.

Como nas medidas anteriores é apresentado um gráfico para cada nova medida apresentada, valor mínimo (Figura 21), valor máximo (Figura 22), moda (Figura 23) e mediana (Figura 24).

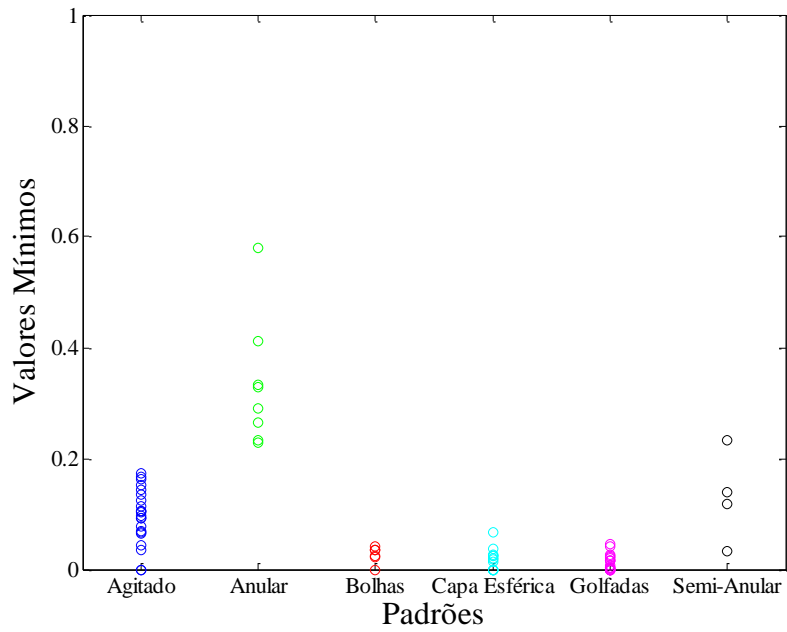


Figura 21 - Valores mínimos dos valores de tensão para todos os padrões.

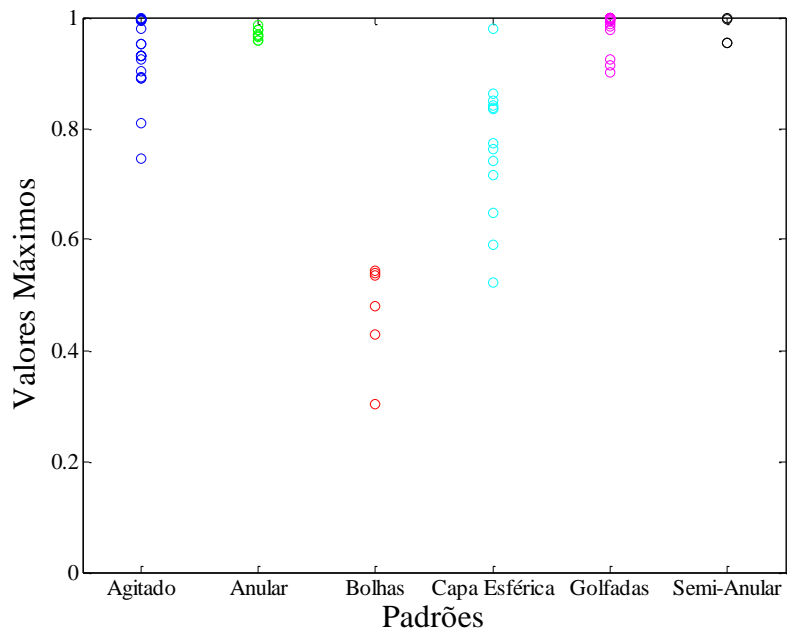


Figura 22 - Valores máximos dos valores de tensão para todos os padrões.

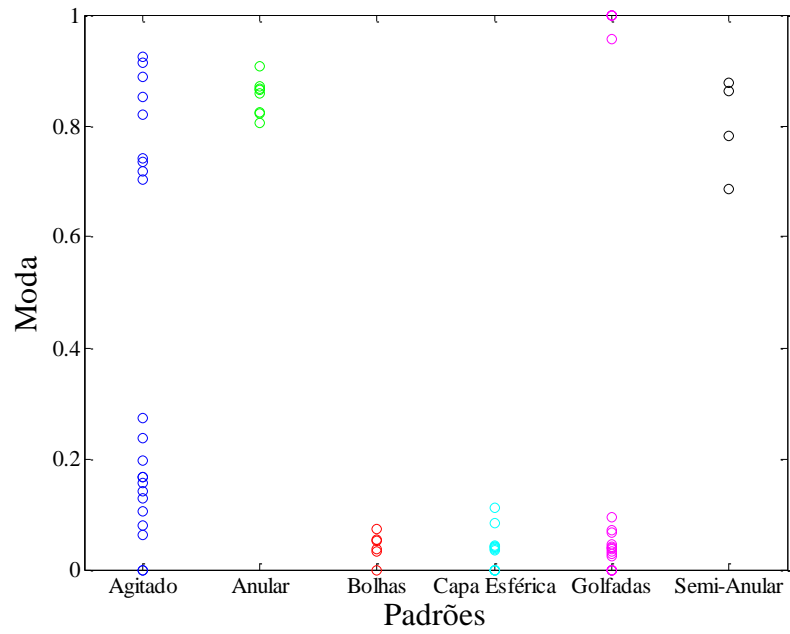


Figura 23 - Moda dos valores de tensão para todos os padrões.

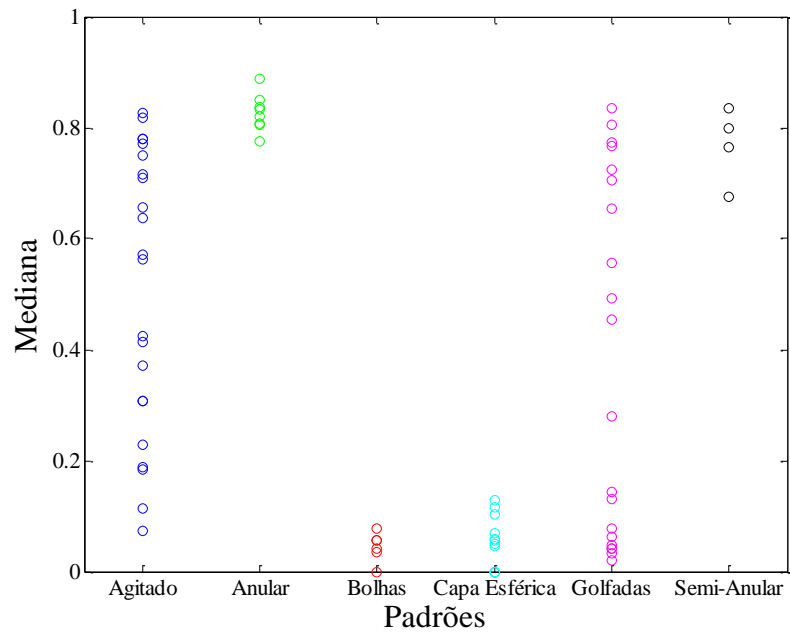


Figura 24 - Mediana dos valores de tensão para todos os padrões.

3.3.2 Histogramas

"O Histograma é formado por um conjunto de retângulos, de tal modo que seus pontos médios coincidam com os pontos médios dos intervalos de classe. A área é proporcional à soma das frequências simples ou absolutas" (Ferrari, 2004, p. 18).

"Através dos histogramas é fácil determinar as principais características do sedimento, designadamente qual é a classe granulométrica mais abundante, se a amostra é bem ou mal calibrada, se existem assimetrias entre as quantidades de partículas grosseiras e finas, etc" (Dias, 2004, p. 28).

3.3.2.1 Cálculo do Histograma

Há várias ferramentas que implementam as funções para o cálculo e construção de histogramas, as quais também suportam construção de gráficos. A seguir é apresentada uma breve descrição de como gerar um histograma:

- obter uma amostra dos dados;
- calcular a amplitude ($R = \text{valor máximo} - \text{valor mínimo}$);
- determinar o número de classes (parametrizável);
- calcular a amplitude das classes;
- determinar os limites das classes;
- construir a tabela;
- traçar o diagrama.

"Em virtude das restrições de espaço em memória ou tempo de processamento, pode ser necessária a redução dos dados. A redução deve ser feita a fim de manter as características do conjunto de dados originais", de acordo com Glymour et al. (1997) apud (Rezende, 2005, p. 315).

Em razão dessas restrições, é possível, através da variação do "número de classes", controlar o tamanho e o tempo de processamento para gerar os dados do

histograma. O número de classes influencia diretamente o resultado do histograma, podendo agrupar ou separar a frequência dos valores. Tal processo pode ser controlado pela "distribuição de frequência".

3.3.2.2 Distribuição de Frequência do Histograma

A distribuição de frequência é um método para agrupar dados em classes de modo a fornecer a quantidade de dados em cada classe. Com isso, pode-se resumir e visualizar um conjunto de dados sem precisar levar em conta os valores individuais.

Com o intuito de gerar diferentes bases a partir do histograma, foram construídas cinco distribuições de frequência, visando abranger diferentes pontos de configurações para o "número de classes". Todos os cálculos são baseados na base de dados brutos, isto é, nos valores de tensão, onde estes apresentam 180.000 pontos para cada padrão de escoamento.

3.3.2.2.1 Base com Todos os Pontos

O primeiro histograma foi gerado mantendo todos os pontos dos dados brutos. Este foi apenas um processo com o intuito de verificar o comportamento real do histograma, não sendo viável a sua utilização em um sistema de classificação de padrões¹. O histograma com "Todos os Pontos" pode ser visualizado na Figura 25.

3.3.2.2.2 Base Raiz Quadrada

O segundo histograma utilizou da "Raiz Quadrada" (Eq. 4.1) para calcular a distribuição de frequência (descrito em Ferrari (2004, p. 16)), onde n é a quantidade de pontos e k é a distribuição de frequência (número de classes).

$$k \cong \sqrt{n} \quad (4.1)$$

lxv|lxv_____

¹ Vide seção 3.2.1, página 55, citação de Rosa *et al.* (2010, p. 8).

Para a "Raiz Quadrada", foi encontrado o valor de k igual aproximadamente a 424. Com base nesse resultado, foi construído o histograma utilizando esse valor no número de classes (Figura 26).

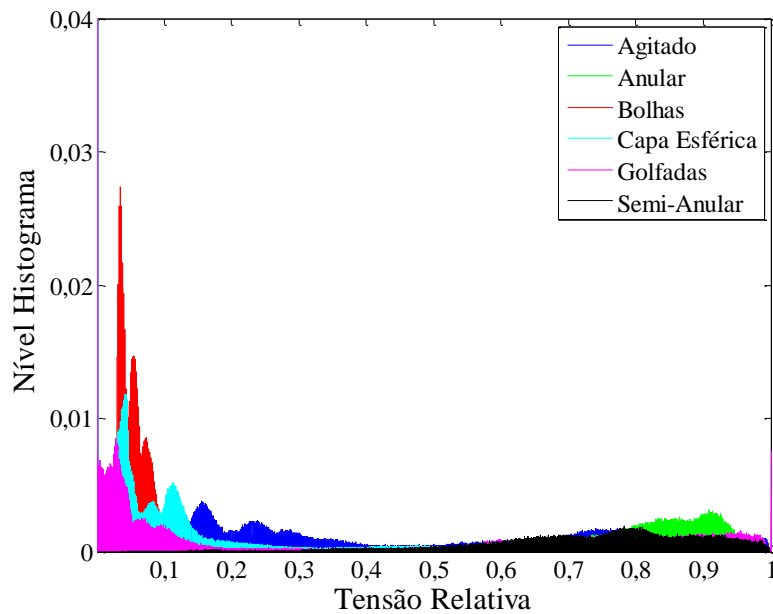


Figura 25 - Histograma com 180.000 pontos para todos os padrões.

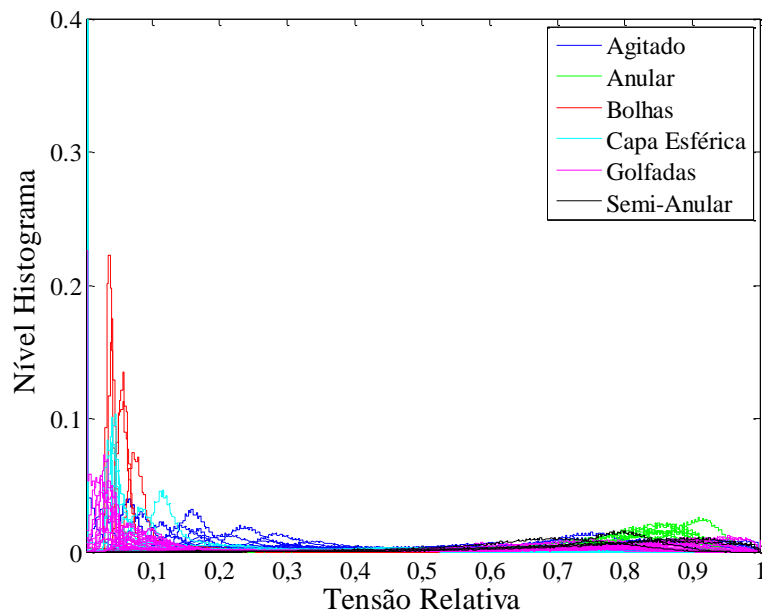


Figura 26 - Histograma com 424 pontos para todos os padrões.

3.3.2.2.3 Base Empírica

Outra configuração selecionada foi a com o número de classes igual a 300. Esse valor foi proposto empiricamente depois de diversos testes executados pelo Laboratório de Escoamento da Unicamp. Esses testes podem ser vistos em Rosa *et al.* (2010, p. 32). A análise do histograma é apresentada na Figura 27.

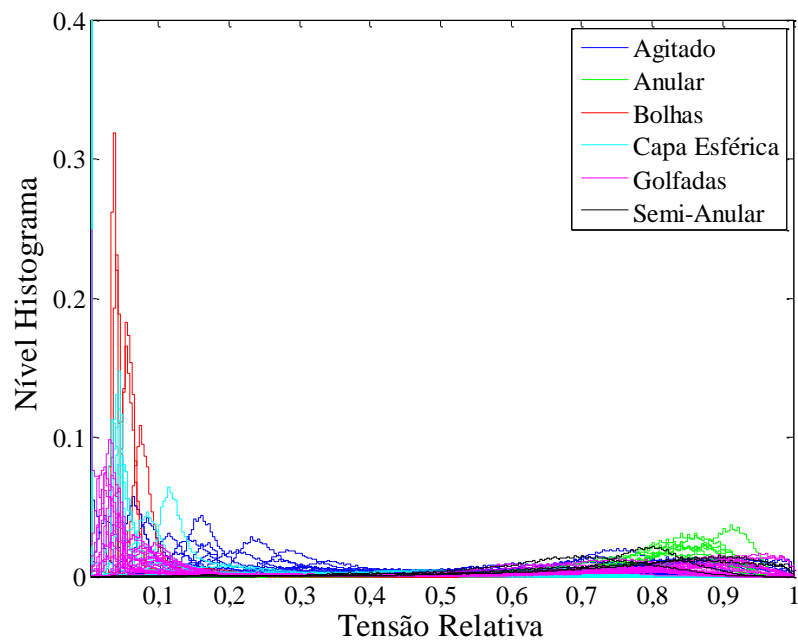


Figura 27 – Histograma com 300 pontos para todos os padrões.

3.3.2.2.4 Base Sturge2 e Base Sturges10

Por último, foram geradas mais duas bases, ambas utilizando a Regra de Sturges. A diferença entre elas está na base logarítmica utilizada, uma na base 2 e a outra na base 10 (Ferrari, 2004, p. 16), sendo a base 2 uma variação proposta em função da base 10, como pode ser visto nas Eq. (4.2) e Eq. (4.3) respectivamente.

$$k \cong 1 + 3,3 \cdot \log_2 n \quad (4.2)$$

$$k \cong 1 + 3,3 \cdot \log_{10} n \quad (4.3)$$

Para a base 2, foi encontrado o valor aproximado de k igual a 58, enquanto na base 10 o valor foi de aproximadamente 18. A Figura 28 e Figura 29 apresentam o comportamento do histograma para os dois testes executados.

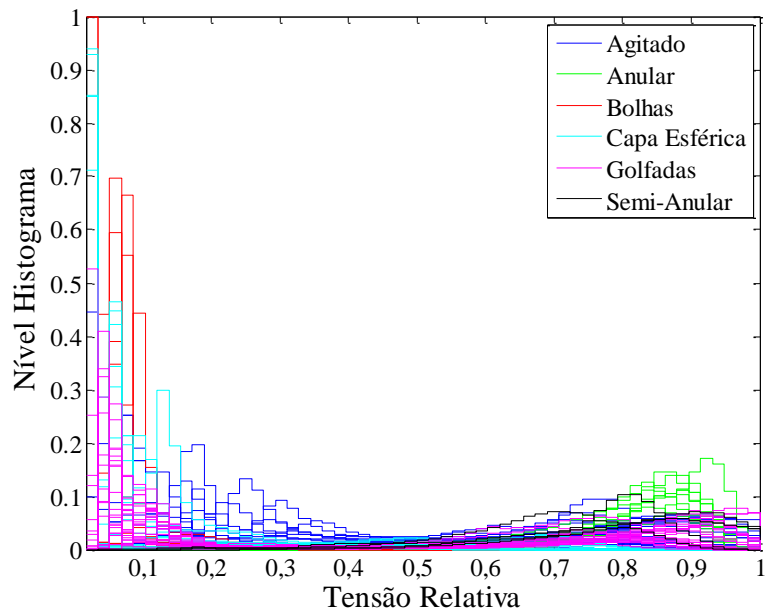


Figura 28 - Histograma com 58 pontos para todos os padrões.

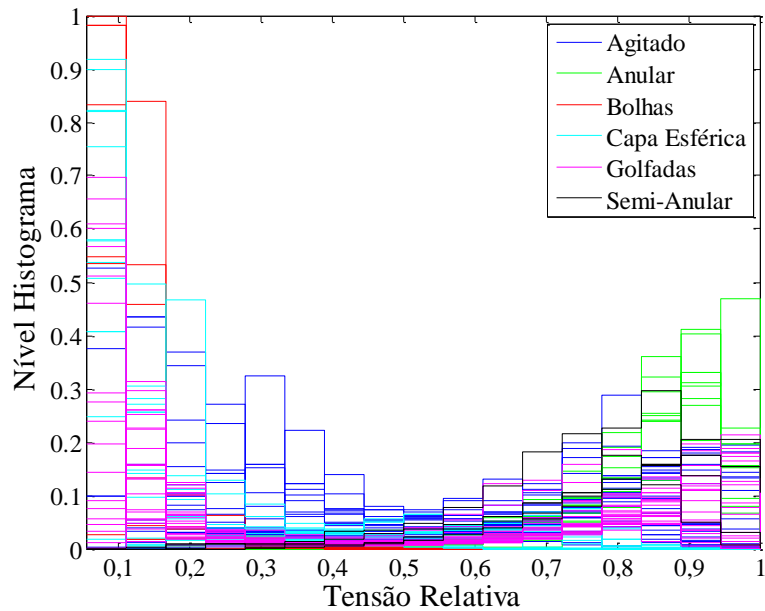


Figura 29 - Histograma com 18 pontos para todos os padrões.

3.3.3 Função de Densidade de Probabilidade

A PDF é uma função que associa a cada valor assumido pela variável aleatória a probabilidade do evento correspondente. Esta função descreve a forma da curva de distribuição da probabilidade de ocorrência de cada valor da variável aleatória contínua.

"A faixa de valores que os dados apresentam é dividida em um determinado número de células de mesmo 'comprimento'. A 'altura' de cada célula é dada pela contagem do número de dados cujos valores estão na faixa correspondente a essa célula" (Rosa & Ohishi, 2008, p. 33).

3.3.3.1 Cálculo da PDF

Como a PDF indica a faixa de valores observados com maior ocorrência, o seu formato é bastante semelhante ao seu histograma. Mais detalhes quanto ao cálculo da função de densidade de probabilidade pode ser encontrado em Rosa & Ohishi (2008) e Rosa *et al.* (2010).

Seguindo os mesmos princípios do cálculo do histograma (seção 3.3.2.2), foram utilizados os mesmos testes para a construção da função de densidade probabilidade. Mais detalhes a seguir.

3.3.3.1.1 Base com Todos os Pontos

Foram geradas as PDFs para todos os padrões de escoamento, mantendo os 180.000 pontos adquiridos a partir dos valores de tensão. A Figura 30 apresenta o gráfico das PDFs geradas, porém da mesma forma, essa base não é viável para utilização em um sistema de classificação de padrões.

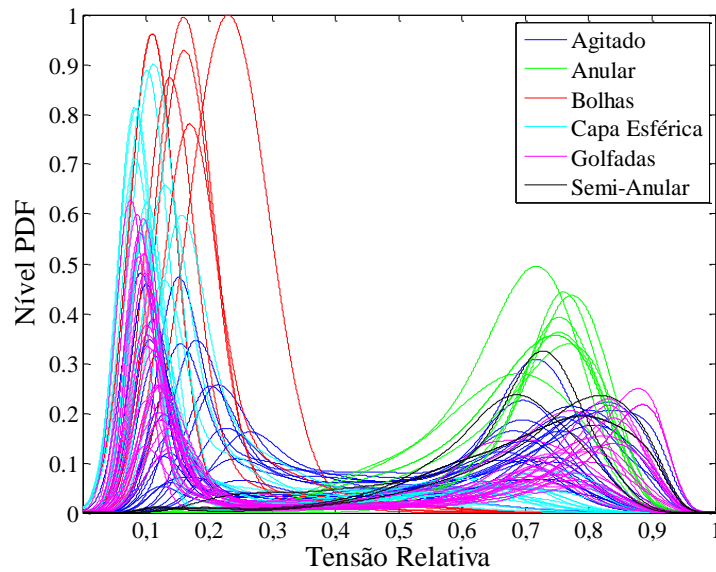


Figura 30 – PDF com 180.000 pontos para todos os padrões.

3.3.3.1.2 Base Raiz Quadrada

A segunda base foi calculada com "Raiz Quadrada", sendo que o valor encontrado para o número de pontos foi de 424. Na Figura 31, é apresentada a PDF com 424 pontos, o que não difere visualmente da PDF com 180.000 pontos (Figura 30).

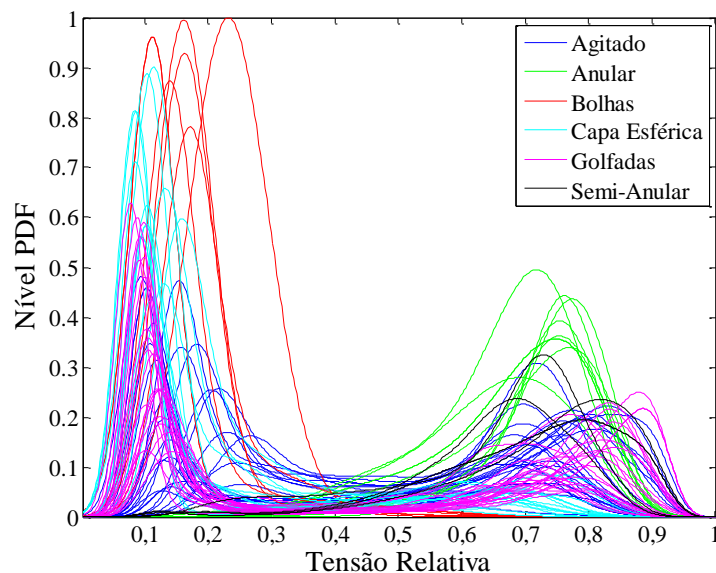


Figura 31 – PDF com 424 pontos para todos os padrões.

3.3.3.1.3 Base Empírica

A quantidade atribuída para o teste empírico foi de 300 pontos (Figura 32), número que foi encontrado empiricamente.

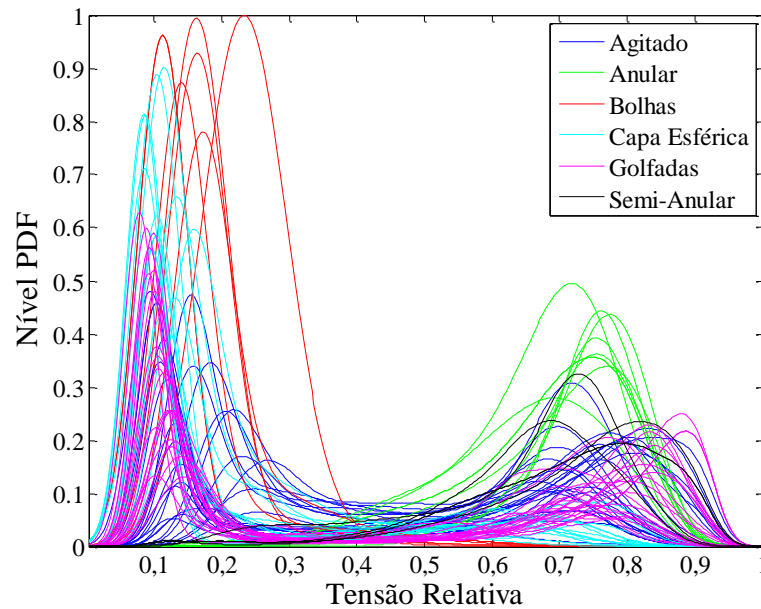


Figura 32 - PDF com 300 pontos para todos os padrões.

3.3.3.1.4 Base Sturges2 e Base Sturges10

Por fim, a Regra de Sturges obteve 58 pontos para a base 2 (Figura 33) e 18 pontos para a base 10 (Figura 34), tais bases apresentaram maiores diferenças para os outros testes, tornando visível a influência que os valores atribuídos à densidade de frequência podem acarretar na construção da base.

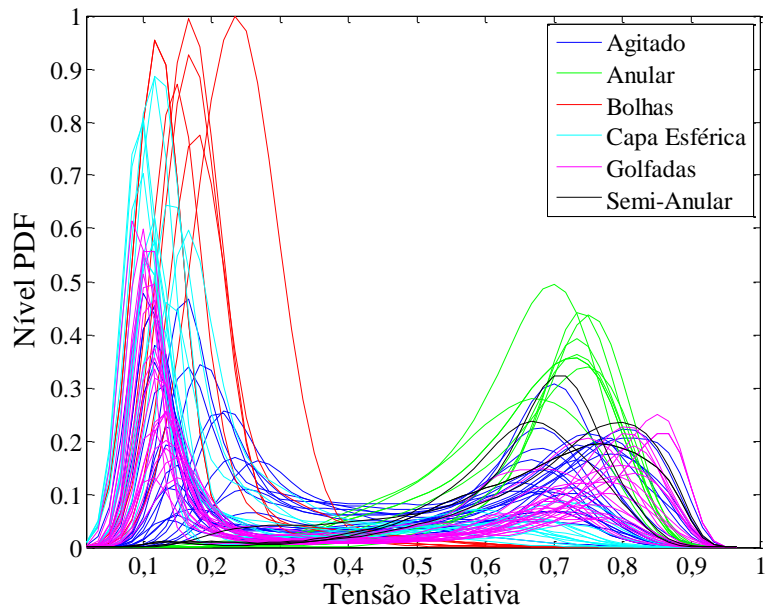


Figura 33 - PDF com 58 pontos para todos os padrões.

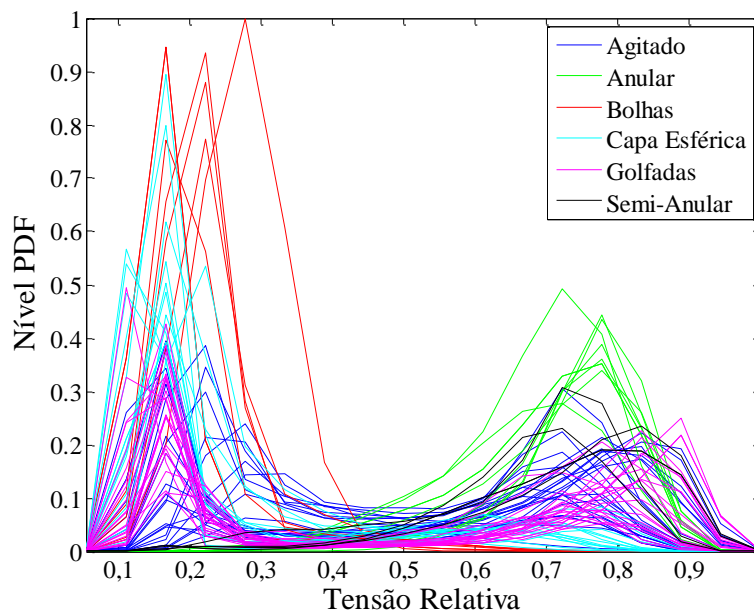


Figura 34 - PDF com 18 pontos para todos os padrões.

3.4 Base de Dados Seleccionadas

Após a descrição de todo o processo de coleta e geração de diferentes bases de dados, foi possível obter uma gama de conjunto de dados que são passíveis de serem utilizados para a classificação de padrões.

Foram selecionados 11 conjuntos de dados gerados a partir dos dados brutos (valores de tensão), sendo eles:

- 3 conjuntos de dados tratados com momentos amostrais (base MD, base MDAC e base M5DAC);
- 4 conjuntos de dados tratados com histogramas (base raiz quadrada, base empírica, base Sturges2 e base Sturges10), e
- 4 conjuntos de dados tratados com PDFs (base raiz quadrada, base empírica, base Sturges2 e base Sturges10).

Ao gerar as bases, foram coletados alguns dados importantes quanto ao processo de construção da base, tempo de processamento e espaço em memória. A Tabela II apresenta os dados quanto ao tempo de processamento e o espaço de memória ocupado. Para efeitos de comparação, foram colocados também a base de dados gerada com "Todos os Pontos" tanto para o histograma quanto para a PDF.

Esse árduo processo de geração de bases tem o intuito de atacar de diferentes formas os problemas de espaço de memória e tempo de processamento, visto que quanto mais pontos são processados e armazenados, mais oneroso pode tornar-se o processo de classificação de padrões.

Da mesma forma, deve-se ter em mente que a redução, de forma drástica, da base de dados, pode levar a uma queda no desempenho do sistema por suprimir os dados da base.

Dessa maneira, é preciso estabelecer um conjunto de teste para averiguar quais tratamentos melhor se aplicam ao problema proposto.

Tabela II - Tempo de processamento e espaço em memória para as bases de dados.

<i>Base de Dados</i>		<i>Tempo</i>	<i>Espaço (Bytes)</i>
Dados Brutos (Valores de Tensão)		1min	105.120.000
Momentos Amostrais	MD	1s	1168
	MDAC	10s	2336
	M5DAC	14s	4672
Histogramas	Todos os Pontos	4s	105.120.000
	Raiz Quadrada	2s	247.616
	Empírico	2s	175.200
	Sturges2	1s	33.872
	Sturges10	1s	10.512
Função de Densidade de Probabilidade	Todos os Pontos	21h 7min	105.120.000
	Raiz Quadrada	3min	247.616
	Empírico	2min	175.200
	Sturges2	24s	33.872
	Sturges10	10s	10.512

Todo o processo de geração da base de dados foi executado em um computador com Sistema Operacional Linux Ubuntu 8.04, com processador Core2Quad de 2,0 GHz e com 4Gb de memória RAM. O Código-M foi a linguagem utilizada para o desenvolvimento na plataforma MATLAB© R2008a.

4

Metodologia Proposta

*Você precisa fazer aquilo que pensa que não é capaz de fazer.
(Eleanor Roosevelt)*

No capítulo 4 será descrito toda a metodologia proposta para o desenvolvimento do sistema de classificação. Onde a seção 4.1, apresenta o sistema utilizando redes neurais, sendo descrito os algoritmos de treinamento e o método de configuração dos parâmetros da rede. O capítulo 9.1, apêndice, apresenta os principais conceitos envolvidos sobre redes neurais, tais como o seu funcionamento, processos de aprendizado, algoritmos de treinamento e classes de problemas aplicados.

4.1 Proposta

Segundo Pacheco *et al.* (2007, p. 3) "*o uso de RNAs na indústria de petróleo é recente, e seu potencial não foi completamente explorado. Apesar disso, muitos pesquisadores têm aplicado redes neurais artificiais para resolver vários problemas associados ao escoamento multifásico.*".

Em função deste potencial ramo das redes neurais para solucionar problemas de escoamento, o sistema aqui proposto tem o objetivo de averiguar e selecionar as melhores configurações para diferentes algoritmos de treinamento, já que o grande desafio na utilização de redes neurais é selecionar um algoritmo de treinamento, e conseqüentemente ajustar seus parâmetros, a fim de obter os pesos que melhor representem a função que classifica o sistema de escoamento (Rezende, 2005, p. 145).

É importante salientar que para este tipo de problema de classificação são encontrados dois tipos de bases: base de treinamento e base de teste. A base de treinamento é utilizada para o aprendizado da rede, enquanto a base de teste tem a função de determinar a quantidade de acertos que a rede é capaz obter.

4.1.1 Algoritmos de Treinamento

Os algoritmos de treinamento adotados foram o *backpropagation* com termo *momentum* (BPM) e o método gradiente (GRAD), sendo que ambos utilizam a tangente hiperbólica como função de ativação. A seguir são dados mais detalhes sobre o funcionamento, assim como os parâmetros de configuração utilizados.

4.1.1.1 *Backpropagation* com Termo *Momentum*

Um dos algoritmos utilizados para o treinamento de redes MLP é o *backpropagation*, no qual são ajustados os pesos das camadas intermediárias.

"Como não existem valores de saída desejados para camadas intermediárias da rede, somente para as saídas, o ajuste dos pesos das camadas intermediárias é feito através da propagação para trás do erro da camada de saída." (Rezende, 2005, p. 152).

O treinamento pode ser dividido em duas fases. Na primeira fase, *forward*, os valores de entrada são propagados para frente, até a saída da rede, onde é obtido o valor da saída da rede que é comparado com o valor desejado. Caso o resultado não seja o esperado, o ajuste dos pesos é feito para trás (*backward*). Esses ajustes são feitos até que a rede obtenha um resultado satisfatório.

Um problema do *backpropagation* é que o seu treinamento pode convergir para regiões de mínimos locais e gradientes nulo, isto é, a aprendizagem da rede fica presa em determinados pontos e não consegue sair deles, causando assim baixa taxa de aprendizado da rede.

Uma forma encontrada para corrigir essa deficiência foi através do uso do backpropagation com termo *momentum* (BPM), o qual utiliza dos ajustes anteriores para aumentar a velocidade de treinamento e estabilizar a convergência.

"Se o treinamento passa por um mínimo local, o termo *momentum* forçará os ajustes na mesma direção dos ajustes anteriores à passagem pelo mínimo." (Rezende, 2005, p. 155). Essa situação pode ser visualizada na Figura 35, onde é possível averiguar a velocidade de treinamento com a utilização do termo *momentum* e sem ele.

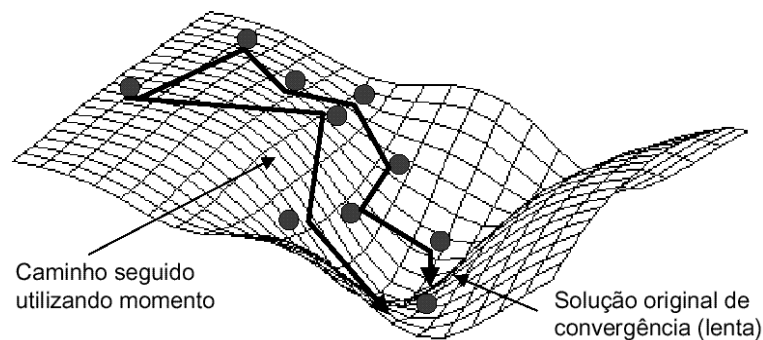


Figura 35 - Uso do termo *momentum* no algoritmo BP (Lacerda, 2006, p. 71).

Com base no entendimento do funcionamento do BPM, foi utilizado o algoritmo codificado por Silva & Zuben (1998, p. 9), utilizando o Código-M, o qual pode ser visualizado no Anexo 8.1. Neste algoritmo, estão presentes alguns parâmetros de configuração como:

- **Valores do conjunto de treinamento:** características selecionadas do problema a ser classificado (base de dados);
- **Saída do conjunto de treinamento:** classes de padrões associados ao conjunto sinais;
- **Número de neurônios:** quantidade de neurônios presentes nas camadas intermediárias;
- **Taxa de aprendizagem:** capacidade do neurônio em aprender a cada interação. O valor deve ser pequeno para evitar saturação;

- **Termo momentum:** influência do termo *momentum* na atualização dos pesos;
- **Erro mínimo desejado:** taxa de erro esperado para a parada do algoritmo;
- **Número máximo de épocas:** número de interações para a atualização dos pesos da rede;
- **Aleatoriedade:** fator de aleatoriedade do algoritmo.

E os retornos deste algoritmo são;

- **Pesos da camada intermediária:** pesos de ativação referentes a camada intermediária;
- **Pesos da camada de saída:** pesos da camada de saída;

4.1.1.2 Método Gradiente

Outra técnica de treinamento de uma rede MLP é conhecida como método gradiente (GRAD). O GRAD é um dos métodos mais simples para a obtenção dos pesos da rede.

De acordo com Silva (1998, p. 57), o GRAD apresenta "*um comportamento monotonicamente decrescente da função de custo (erro), percebe-se que a taxa de aprendizagem oscila (aumenta e diminui) e o número de épocas para convergência é menor do que o do algoritmo BP, quando o mesmo critério de parada é adotado.*".

Silva & Zuben (1998, p. 11) codificaram também o algoritmo GRAD utilizando Código-M, este pode ser visualizado nos Anexos 8.2 e 8.3. E a única diferença quanto aos parâmetros do BPM, é que o GRAD não apresenta a taxa de aprendizado.

4.1.2 Tipos de Saídas

Todo sistema de classificação deve apresentar um conjunto de valores que representam os padrões desejados para as entradas utilizadas. Tais valores são chamados de classes de padrões.

As classes de padrões são utilizadas para verificar a taxa de acertos da rede. Essa verificação é feita utilizando-se da base de teste com os pesos obtidos pela rede neural. O algoritmo responsável por essa verificação está no Anexo 8.4.

Da mesma maneira com que a rede neural deve receber o conjunto de características de forma estruturada, os valores de saídas devem seguir o mesmo princípio. Desse modo foram selecionadas duas formas de representar as classes de padrões, sendo que as duas representam os mesmos dados, podendo ser visualizado no Anexo 8.5 o algoritmo que verifica os padrões de escoamento.

A primeira é a "saída única" na qual as classes são enumeradas e colocadas em um vetor. Todas as classes contêm um número único que as representam, podendo dessa forma diferenciar uma das outras. A Figura 36 representa a estrutura neural com saída única, sendo que n é a quantidade de características passadas na entrada, e m o número de neurônios na camada oculta.

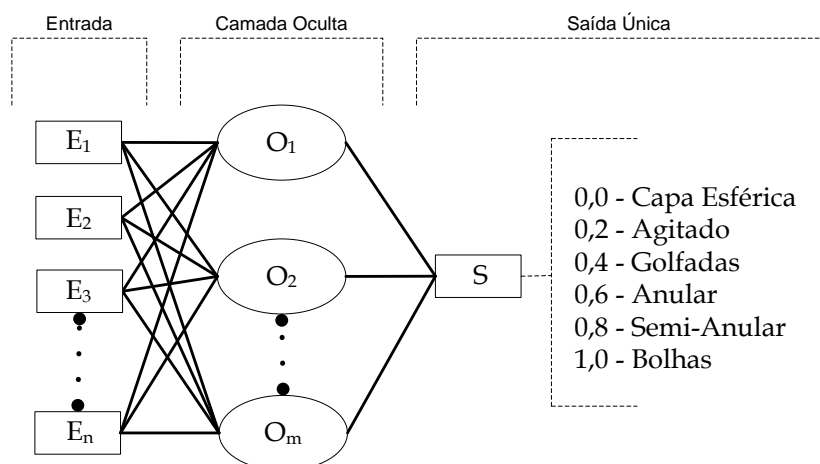


Figura 36 - Estrutura neural para a saída única.

A outra forma é a "saída múltipla", onde cada classe é representada por uma t-upla de um vetor. A posição da t-upla que representa a classe, recebe o valor 1 e as demais recebem o valor 0. Assim, o conjunto de sinais é representado pela classe onde o valor 1 estiver associado.

Para uma melhor compreensão da "saída múltipla" é apresentado na Figura 37 a estrutura neural com múltiplas saídas.

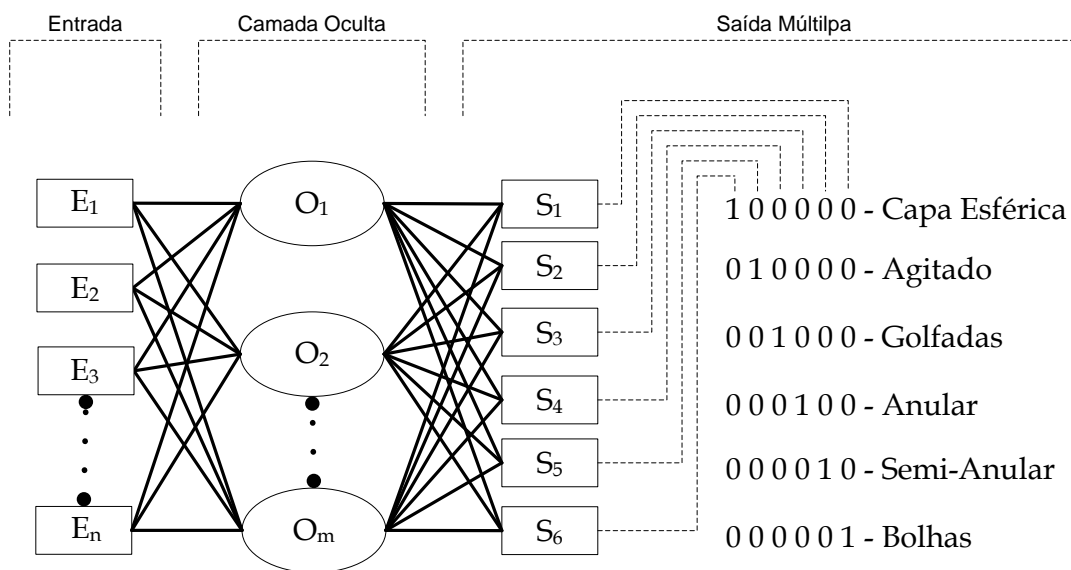


Figura 37 - Estrutura neural para a saída múltipla.

Apesar do tipo da saída representar as mesmas informações, ela pode influenciar a rede, devido ao fato dos valores utilizados não retornarem de forma exata às classes desejadas, isto é, o valor pode variar de 0 a 1 para os dois tipos de saídas, sendo que a múltipla acontece com todas as t-uplas do vetor.

Esse fato é corrigido para a saída única através da aproximação do valor obtido com os valores correspondentes e para a saída múltipla, é selecionado o padrão representado pela t-upla que apresenta o maior valor.

Como exemplo, caso seja obtido o valor (0,25) para a saída única ou para a saída múltipla $(0,1\ 0,9\ 0,5\ 0,6\ 0,8\ 0,3)^T$, os valores serão transformados respectivamente em (0,2) e $(0\ 1\ 0\ 0\ 0\ 0)^T$, que representam o Padrão Agitado.

4.1.3 Busca Exaustiva de Configurações

Para a seleção dos parâmetros de configuração, foi adotado o processo de busca exaustiva (força bruta), o qual consiste no teste de todas as configurações. Esse processo foi adotado devido ao fato da complexidade em escolher tais valores de forma arbitrária.

Apesar do demasiado tempo gasto para varrer todo o conjunto de parâmetros, esse processo se mostrou bastante útil, pois o torna automático e independente do fator humano.

A Figura 38 e Figura 39 demonstram diferentes simulações de problemas, utilizando várias combinações de parâmetros de configuração. É possível averiguar o comportamento da rede para diferentes combinações, e também as oscilações e estagnações presentes.

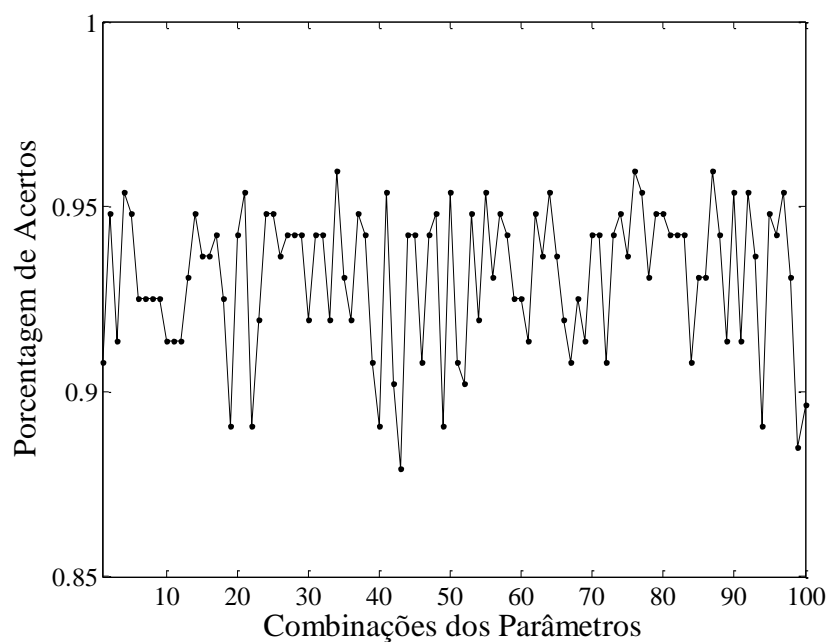


Figura 38 - Oscilação dos parâmetros de configuração.

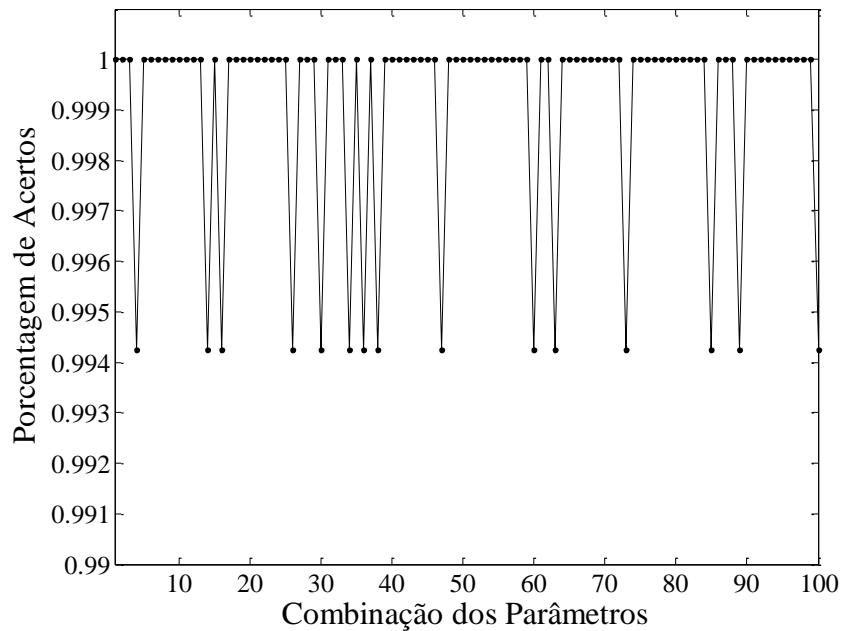


Figura 39 - Estagnação dos parâmetros de configuração.

Suponha que haja os parâmetros, número de neurônios (nh) e termo *momentum* (cm), onde $nh \in \{1; 2; \dots; 10\}$ e $cm \in \{0,1; 0,2; \dots; 1,0\}$. É feita uma enumeração da permutação desses parâmetros ($nh; cm$), como pode ser visto abaixo.

- Combinação 1 - (1; 0,1);
- Combinação 2 - (1; 0,2);
- ...
- Combinação n-1 - (10; 0,9);
- Combinação n - (10; 1,0);

Desse modo, o esforço humano é determinar os valores que devem fazer parte do conjunto de parâmetros configuráveis, para que o sistema realize de forma exaustiva a busca por resultados que melhor classifiquem o problema tratado.

5

Estudo de caso

*Determinação, coragem e auto-confiança são fatores decisivos para o sucesso.
Se estamos possuídos por uma inabalável determinação conseguiremos superá-los.
Independentemente das circunstâncias, devemos ser sempre humildes,
recatados e despidos de orgulho.
(Dalai Lama)*

O capítulo 5 apresenta todo o procedimento para a execução dos experimentos. Na seção 5.1, há uma breve descrição da metodologia aplicada para a execução dos experimentos, onde foram evidenciados os conjuntos de parâmetros utilizados assim como a divisão da base de dados em base de treinamento e de teste. A seção 5.2 apresenta os resultados mais relevantes para todos os experimentos, juntamente com uma análise estatística dos resultados. Por fim, na seção 5.3 são apresentados os tempos de execução assim como as considerações sobre a execução dos experimentos.

5.1 Metodologia Aplicada

A busca por sistemas que obtenham resultados prósperos é almejada por todos os pesquisadores, dessa forma, diferentes linhas de pesquisas são estudadas a fim de estabelecer novos resultados para diversos tipos de problemas.

Em virtude desse objetivo, este projeto tem o intuito de fazer uma busca exaustiva de configurações para as redes neurais, de forma que se possa encontrar as melhores configurações que geram os resultados mais prósperos para o sistema de escoamento bifásico gás/líquido, uma vez que o grande desafio na utilização de redes neurais é selecionar um algoritmo de treinamento, e conseqüentemente,

ajustar seus parâmetros. A fim de obter os pesos que melhor representem a função que classifica o sistema de escoamento (Rezende, 2005, p. 145).

5.1.1 Configuração do Sistema

Primeiramente, foram definidos os algoritmos de treinamentos a serem utilizados: backpropagation com termo *momentum* (BPM) e método gradiente (GRAD). Estes dois algoritmos foram escolhidos devido a sua rápida execução e ampla divulgação na literatura. Mais detalhes sobre estes algoritmos podem ser encontrados nas subseções 4.1.1.1 e 4.1.1.2, respectivamente.

Em seguida, após analisar o comportamento dos algoritmos, com diferentes parâmetros, foram determinados quais os intervalos de variações que cada algoritmo é mais apto a executar.

Para o algoritmo BPM, foram definidos os seguintes conjuntos de configuração para cada parâmetro utilizado: número de neurônios, nh , (Eq. 6.1), taxa de aprendizagem, $alfa$, (Eq. 6.2) e termo *momentum*, cm , (Eq. 6.3).

$$neuronios = \left\{ \begin{array}{l} nh \mid nh \in \mathbb{N} \text{ e } 2 \leq nh \leq 20 \text{ e} \\ (\exists x)(x \in \mathbb{N} \text{ e } nh = 2x) \end{array} \right\} \quad (6.1)$$

$$aprendizado = \left\{ \begin{array}{l} alfa \mid alfa \in \mathbb{R} \text{ e } 0,001 \leq alfa \leq 0,01 \text{ e} \\ (\exists x)(x \in \mathbb{N} \text{ e } alfa = 0,001 + 0,0015x) \end{array} \right\} \quad (6.2)$$

$$momentum = \left\{ \begin{array}{l} cm \mid cm \in \mathbb{R} \text{ e } 0,1 \leq cm \leq 1,0 \text{ e} \\ (\exists x)(x \in \mathbb{N} \text{ e } cm = 0,1 + 0,15x) \end{array} \right\} \quad (6.3)$$

Enquanto para o GRAD, os conjuntos utilizados para os parâmetros foram: número de neurônios (Eq. 6.4) e taxa de aprendizagem (Eq. 6.5).

$$neurônios = \{nh \mid nh \in \mathbb{N} \text{ e } 1 \leq nh \leq 20\} \quad (6.4)$$

$$aprendizado = \left\{ \begin{array}{l} \text{alfa} \mid \text{alfa} \in \mathbb{R} \text{ e } 0,1 \leq \text{alfa} \leq 1,0 \text{ e} \\ (\exists x)(x \in \mathbb{N} \text{ e } \text{alfa} = 0,1x \end{array} \right\} \quad (6.5)$$

Sendo que, para os dois algoritmos, BPM e GRAD, foram utilizados três parâmetros com valores constantes: erro mínimo desejado, *minerro*, (Eq. 6.6), número máximo de épocas, *maxep*, (Eq. 6.7) e aleatoriedade, *val*, (Eq. 6.8).

$$erro = \{minerro = 1^{-7}\} \quad (6.6)$$

$$iterações = \{maxep = 5000\} \quad (6.7)$$

$$aleatoriedade = \{val = 0,5\} \quad (6.8)$$

Desta forma, o número de possibilidades de permutações para cada algoritmo foi de 490 para o BPM e 200 para o GRAD. Nos Anexos 8.6 e 8.7, podem ser observados os algoritmos que são responsáveis pela permutação do conjunto de parâmetros e conseqüentemente, a execução dos experimentos.

Definidos os conjuntos de parâmetros, fez-se necessária a definição da base de dados com a qual serão realizados experimentos.

5.1.2 Base de Treinamento e Teste

Um sistema de classificação de padrões necessita de que a base de dados seja separada em dois grupos: base de treinamento e base de teste. Essa divisão tem a finalidade de distinguir duas fases do processo de classificação, onde a primeira fase utiliza a base de treinamento no algoritmo de classificação com o intuito de treinar a rede neural, e a segunda tem a finalidade de testar o resultado obtido pelo treinamento da rede com a base de teste.

Essa divisão não segue nenhum padrão, cada pesquisador separa a base de acordo com sua experiência, sendo que as divisões mais comuns encontradas na literatura fazem a divisão na seguinte proporção: 2/3 para a base de treinamento e

1/3 para a base de teste. A seleção dos padrões que irão pertencer a cada base normalmente é feita de forma aleatória.

É importante salientar que alguns pesquisadores utilizam toda a base de dados para o treinamento, e para os testes eles fazem a utilização parcial da mesma base. Essa prática não é considerada uma boa técnica, devido ao fato dos testes serem feitos com padrões que foram treinados previamente.

Este projeto seguiu as mesmas divisões propostas por Corrêa (2009) e Rosa *et al.* (2010). A base de dados foi dividida em aproximadamente 70% para o treinamento e 30% para o teste, sendo que, dos 73 conjuntos de sinais, 50 são de treinamento e 23 de teste. Na Tabela III é possível verificar essa divisão, inclusive a proporção para todas as classes de padrões.

Tabela III - Divisão da base de dados (treinamento e teste).

<i>Padrões</i>	<i>Treinamento</i>	<i>Teste</i>	<i>Total</i>
Agitado	15 - 20,55 %	7 - 9,58 %	22 - 30,13 %
Anular	5 - 6,85 %	3 - 4,11 %	8 - 10,96 %
Bolhas	4 - 5,48 %	2 - 2,74 %	6 - 8,22 %
Capa Esférica	9 - 12,33 %	4 - 5,48 %	13 - 17,81 %
Golfadas	14 - 19,18 %	6 - 8,22 %	20 - 27,40 %
Semi-Anular	3 - 4,11 %	1 - 1,37 %	4 - 5,48 %
Total	50 - 68,50 %	23 - 31,50 %	73 - 100 %

Selecionados os algoritmos de treinamentos, juntamente com o conjunto de configurações e determinada a divisão da base nos conjuntos de treinamento e teste, é possível, assim, executar toda uma gama de experimentos a fim de validar o sistema proposto.

5.2 Resultados

A fase de experimentação tem a finalidade de obter os resultados de diversas execuções do sistema. Esse procedimento visa testar repetidamente o sistema com o intuito de analisar e validar toda a proposta envolvida. Assim, para cada conjunto de configuração dos algoritmos de treinamento e base de dados trabalhada, foram rodados 50 vezes os testes.

Uma combinação de análises será feita, de tal forma que seja possível averiguar a influência da diversificação dos conjuntos de parâmetros no treinamento da rede, e conseqüentemente, identificar e analisar as configurações mais prósperas encontradas.

5.2.1 Análise de Variância

Com o intuito de realizar uma análise mais profunda em todos os conjuntos de configurações, foram feitos testes não-paramétricos (teste de Kruskal-Wallis), e quando possível aplicado o teste de Scott-Knott para verificar quais os tratamentos que possuem médias próximas.

O Teste de Kruskal-Wallis é um teste estatístico muito utilizado pelos pesquisadores, e visa verificar se existe uma diferença significativa entre as médias. A primeira etapa consiste na formalização da hipótese (Eq. 6.9).

$$\begin{cases} H_0 : \tau_1 = \tau_2 = \dots = \tau_n = 0 \\ H_1 : \tau_i \neq 0 \text{ para pelo menos um } i \end{cases} \quad (6.9)$$

A hipótese H_0 diz que não existe diferença entre os tratamentos dos experimentos, enquanto a hipótese H_1 diz que há pelo menos um tratamento diferente dos demais.

Para testar a hipótese H_0 , utiliza-se a análise de variância. Tal análise é realizada utilizando o delineamento inteiramente casualizado (DIC), pois todas as variáveis do sistema são controladas. O teste foi ajustado para o nível de 5% de significância, tendo o número de repetições igual a 50 vezes e o número de tratamentos, sendo 490 para o BPM e 200 para o GRAD.

Caso a hipótese H_0 não seja verdadeira, é interessante aplicar um teste de agrupamento de médias, teste de Scott-Knott, para a avaliação de múltiplos resultados. O teste de Scott-Knott compara toda e qualquer diferença entre duas médias dos tratamentos.

Dessa forma o teste de Scott-Knott tem a função de agrupar os diferentes tratamentos em grupos similares. Por exemplo, um experimento possui n diferentes tipos de tratamentos (sendo $n = 7$, onde: $T1, T2, \dots, T7$), onde cada um tem m repetições. Ao aplicar esse teste, os tratamentos que possuem médias significativamente similares serão agrupados em até n grupos, ordenados pelas maiores médias. Assim, os tratamentos seriam separados nos seguintes grupos:

- grupo A (T1, T5 e T7);
- grupo B (T3);
- grupo C (T2 e T4);
- grupo D (T6).

Tais agrupamentos implicam que, os tratamentos T1, T5 e T7 são significativamente iguais, apenas variando seus resultados pelo acaso. O mesmo acontece para os demais grupos.

Essa seqüência de testes foi realizada para todos os conjuntos de experimentos a fim de averiguar quais possuem o conjunto de parâmetros mais próximos. Para quase todos os experimentos, a hipótese H_0 foi rejeitada, levando assim ao teste de Scott-Knott, com a significância de 5%, onde foi possível determinar quais conjuntos de parâmetros similares. A hipótese H_0 foi aceita para

os experimentos que utilizaram o GRAD com as bases raiz quadrada e empírica, sendo as duas tratadas com o histograma e com saída única.

Devido à grande quantidade de conjuntos de configurações selecionadas para fazer parte dos experimentos, não é possível apresentar todos os agrupamentos encontrados no teste de Scott-Knott. Desta forma, optou-se por apresentar apenas os três grupos com maiores resultados na média e um quarto campo com os demais agrupamentos.

Em cada campo foram apresentadas duas informações. Na parte superior, há a quantidade de tratamentos (experimentos) do grupo, sendo que no campo "*Demais Grupos*" é indicado, entre parênteses, a quantidade de grupos a mais que foram encontrados. A parte inferior apresenta a variação das médias, sendo o intervalo destas no grupo com os valores ([mínimo máximo]).

Os resultados foram separados de acordo com o algoritmo de treinamento e a base de dados, que podem ser visualizadas nas Tabela IV e Tabela XVII.

Tabela IV - Agrupamento dos experimentos para o BPM com momentos amostrais.

<i>Base de Dados</i>	<i>Saída</i>	<i>Grupo A</i>	<i>Grupo B</i>	<i>Grupo C</i>	<i>Demais Grupo</i>
MD	Única	154 [78,60 82,78] %	123 [74,08 78,52] %	24 [71,39 73,91] %	189 (13) [1,47 71,04] %
	Multi	302 [85,91 90,61] %	24 [83,21 85,73] %	23 [80,00 83,04] %	141 (13) [0,00 79,91] %
MDAC	Única	64 [91,04 97,57] %	131 [86,52 90,86] %	36 [80,08 86,34] %	259 (11) [0,00 79,04] %
	Multi	124 [88,34 93,83] %	41 [83,39 88,17] %	9 [78,26 82,86] %	316 (11) [0,00 76,78] %
M5DAC	Única	172 [84,00 91,13] %	71 [75,82 83,91] %	6 [72,26 74,78] %	241 (10) [0,00 69,65] %
	Multi	125 [90,00 95,22] %	28 [85,56 89,82] %	11 [83,21 84,95] %	326 (10) [0,00 80,00] %

Tabela V - Agrupamento dos experimentos p/ o GRAD com momentos amostrais.

<i>Base de Dados</i>	<i>Saída</i>	<i>Grupo A</i>	<i>Grupo B</i>	<i>Grupo C</i>	<i>Demais Grupo</i>
MD	Única	136 [82,26 84,35] %	46 [80,78 82,17] %	7 [79,56 80,52] %	11 (2) [52,17 77,73] %
	Multi	90 [89,91 90,87] %	80 [88,26 89,82] %	10 [81,73 83,04] %	20 (2) [47,82 69,82] %
MDAC	Única	180 [91,13 94,70] %	3 [88,26 89,04] %	7 [86,52 87,65] %	10 (2) [64,17 67,73] %
	Multi	137 [88,69 91,22] %	14 [87,30 88,60] %	9 [85,56 87,04] %	40 (5) [44,43 84,08] %
M5DAC	Única	180 [86,26 88,87] %	53 [84,69 86,17] %	16 [82,78 84,60] %	20 (2) [74,69 82,52] %
	Multi	91 [93,91 95,39] %	27 [92,52 93,82] %	13 [91,56 92,43] %	69 (10) [51,73 91,47] %

Tabela VI - Agrupamento dos experimentos para o BPM com histogramas.

<i>Base de Dados</i>	<i>Saída</i>	<i>Grupo A</i>	<i>Grupo B</i>	<i>Grupo C</i>	<i>Demais Grupo</i>
Raiz Quadrada	Única	25 [68,00 70,43] %	38 [66,00 67,82] %	278 [62,78 65,91] %	149 (9) [3,82 62,69] %
	Multi	50 [77,73 80,26] %	54 [75,65 77,47] %	277 [72,43 75,56] %	109 (13) [0,00 72,26] %
Empírico	Única	66 [68,26 72,08] %	103 [65,56 68,17] %	235 [62,26 65,47] %	86 (8) [11,30 62,17] %
	Multi	58 [77,47 80,08] %	58 [75,73 77,30] %	296 [68,69 75,65] %	78 (11) [0,00 68,34] %
Sturges2	Única	261 [72,00 77,13] %	167 [68,17 71,91] %	34 [64,17 68,00] %	28 (7) [9,04 62,86] %
	Multi	302 [83,56 89,30] %	51 [80,43 83,39] %	56 [76,26 80,26] %	81 (11) [0,00 76,08] %
Sturges10	Única	238 [74,08 76,78] %	106 [71,82 74,00] %	76 [69,65 71,73] %	70 (11) [4,52 69,47] %
	Multi	206 [92,86 95,73] %	137 [89,47 92,78] %	41 [85,47 89,30] %	106 (11) [0,00 85,30] %

Tabela VII - Agrupamento dos experimentos para o GRAD com histogramas.

<i>Base de Dados</i>	<i>Saída</i>	<i>Grupo A</i>	<i>Grupo B</i>	<i>Grupo C</i>	<i>Demais Grupo</i>
Raiz Quadrada	Única	200 [70,69 74,08] %	-	-	-
	Multi	115 [77,30 78,69] %	41 [76,60 77,21] %	14 [75,56 76,52] %	30 (4) [47,82 70,78] %
Empírico	Única	200 [65,52 72,08] %	-	-	-
	Multi	170 [76,52 78,17] %	10 [68,86 69,91] %	10 [56,17 57,21] %	10 (1) [47,91 48,52] %
Sturges2	Única	106 [72,52 73,82] %	94 [71,21 72,43] %	-	-
	Multi	124 [86,95 89,21] %	34 [85,13 86,86] %	12 [83,30 84,95] %	30 (3) [47,82 81,65] %
Sturges10	Única	80 [76,17 77,82] %	120 [74,34 76,08] %	-	-
	Multi	151 [94,34 96,78] %	9 [93,21 94,08] %	10 [90,95 92,86] %	30 (3) [41,39 86,17] %

Tabela VIII- Agrupamento dos experimentos para o BPM com PDFs.

<i>Base de Dados</i>	<i>Saída</i>	<i>1º Grupo</i>	<i>2º Grupo</i>	<i>3º Grupo</i>	<i>Demais Grupo</i>
Raiz Quadrada	Única	85 [67,56 75,56] %	227 [60,43 67,30] %	40 [56,78 60,26] %	138 (9) [0,00 56,43] %
	Multi	161 [88,00 93,47] %	103 [76,95 87,91] %	9 [70,43 76,17] %	217 (10) [0,00 69,04] %
Empírico	Única	55 [71,13 75,73] %	43 [67,39 71,04] %	184 [62,78 67,30] %	208 (11) [0,00 62,69] %
	Multi	182 [88,17 93,47] %	105 [78,17 88,00] %	10 [72,08 76,52] %	193 (10) [0,00 69,91] %
Sturges2	Única	79 [73,56 78,69] %	88 [68,78 73,47] %	181 [64,17 68,52] %	142 (9) [0,00 64,08] %
	Multi	228 [86,86 91,82] %	157 [76,60 86,86] %	51 [62,86 71,13] %	54 (8) [0,00 61,73] %

Sturges10	Única	123 [77,13 80,78] %	53 [74,00 76,86] %	156 [70,43 73,91] %	158 (10) [2,78 70,34] %
	Multi	293 [84,78 89,91] %	117 [77,82 84,69] %	10 [70,52 73,73] %	70 (8) [0,00 66,78] %

Tabela IX - Agrupamento dos experimentos para o GRAD com PDFs.

<i>Base De Dados</i>	<i>Saída</i>	<i>Grupo A</i>	<i>Grupo B</i>	<i>Grupo C</i>	<i>Demais Grupo</i>
Raiz Quadrada	Única	158 [63,30 67,73] %	15 [619,1 63,04] %	27 [58,86 61,04] %	-
	Multi	96 [92,00 93,47] %	43 [90,26 91,91] %	10 [89,13 90,00] %	51 (5) [48,00 88,78] %
Empírico	Única	80 [65,91 68,00] %	72 [63,91 65,82] %	14 [62,08 63,65] %	34 (1) [59,47 61,82] %
	Multi	94 [92,26 93,65] %	38 [90,78 92,17] %	13 [89,47 90,43] %	55 (5) [48,26 89,13] %
Sturges2	Única	10 [67,30 69,56] %	112 [63,13 65,56] %	78 [61,13 63,04] %	-
	Multi	117 [89,47 91,47] %	31 [88,34 89,39] %	22 [86,95 88,26] %	30 (5) [47,82 78,95] %
Sturges10	Única	15 [70,43 72,08] %	69 [68,26 70,26] %	62 [66,86 68,17] %	54 (1) [64,86 66,78] %
	Multi	82 [86,78 88,52] %	88 [85,04 86,69] %	10 [76,00 77,82] %	20 (2) [47,82 65,56] %

Em praticamente todas as análises, houve uma distribuição das configurações entre os grupos pertencentes aos seus experimentos. Contudo, o primeiro grupo, o que apresenta as maiores médias, teve relativamente uma grande quantidade de configurações próximas, indicando dessa forma, que tais configurações são semelhantes e que as suas variações foram meros acasos gerados de forma aleatória.

Assim, qualquer configuração presente no primeiro grupo está propensa a alcançar resultados tão prósperos quanto qualquer outra configuração presente

neste mesmo grupo. Com base nesses dados e visando avaliar o desempenho das melhores configurações encontradas, foram feitas análises para a melhor configuração de cada experimento.

5.2.2 Teste dos Melhores Conjuntos de Configurações

Nesta análise, foram feitos alguns testes com o intuito de aprofundar o entendimento dos resultados mais prósperos. Foram calculados a média, desvio padrão, valor mínimo e máximo para cada experimento. E assim foram selecionadas as configurações que obtiveram as melhores médias.

5.2.2.1 Momentos Amostrais

Abaixo são apresentadas as análises feitas com os experimentos, onde foram executados os algoritmos de treinamento BPM (Tabela X) e GRAD (Tabela XI), utilizando as bases geradas com momentos amostrais.

Tabela X - Análise dos resultados utilizando BPM e momentos amostrais.

<i>Base de Dados</i>	<i>Tipo de Saída</i>	<i>Média</i>	<i>Desvio Padrão</i>	<i>Valor Mínimo</i>	<i>Valor Máximo</i>
MD	Única	82,78 %	2,32 %	73,91 %	86,96 %
	Multi	90,61 %	1,61 %	86,96 %	91,30 %
MDAC	Única	97,57 %	3,19 %	86,96 %	100,00 %
	Multi	93,83 %	2,17 %	91,30 %	95,65 %
M5DAC	Única	91,13 %	4,12 %	82,61 %	100,00 %
	Multi	95,22 %	1,32 %	91,30 %	95,65 %

Tabela XI - Análise dos resultados utilizando GRAD e momentos amostrais.

<i>Base de Dados</i>	<i>Tipo de Saída</i>	<i>Média</i>	<i>Desvio Padrão</i>	<i>Valor Mínimo</i>	<i>Valor Máximo</i>
MD	Única	84,35 %	2,32 %	78,26 %	86,96 %
	Multi	90,87 %	1,32 %	86,96 %	91,30 %

MDAC	Única	94,70 %	3,20 %	86,96 %	100,00 %
	Multi	91,22 %	2,56 %	82,61 %	95,65 %
M5DAC	Única	88,87 %	4,04 %	78,26 %	100,00 %
	Multi	95,39 %	1,04 %	91,30 %	95,65 %

Para o algoritmo BPM a melhor média encontrada foi com a base MDAC com saída única. Já para o GRAD, a base M5DAC com saída múltipla obteve a melhor média. Uma melhor comparação, entre os diversos experimentos apresentada acima, pode ser visualizada na Figura 40, sendo assim possível averiguar a dispersão dos resultados.

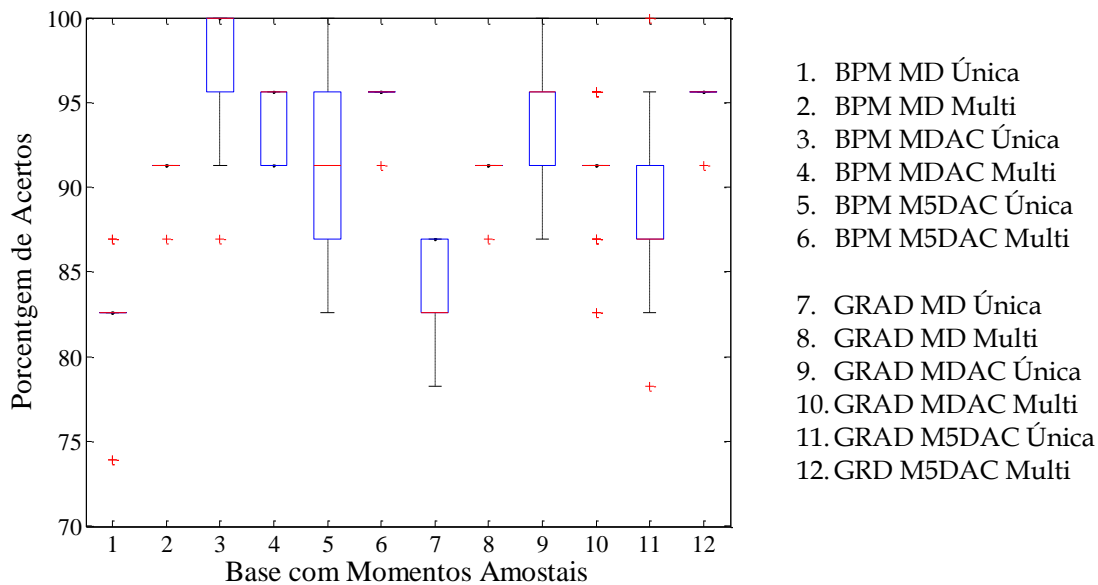


Figura 40 - BoxPlot dos resultados utilizando a base com m. amostrais.

5.2.2.2 Histogramas

Foram executados os algoritmos de treinamento BPM (Tabela XII) e GRAD (Tabela XIII), utilizando as bases geradas com histogramas.

Tabela XII – Análise dos resultados utilizando BPM e histogramas.

<i>Base de Dados</i>	<i>Tipo de Saída</i>	<i>Média</i>	<i>Desvio Padrão</i>	<i>Valor Mínimo</i>	<i>Valor Máximo</i>
Raiz Quadrada	Única	70,43 %	4,30 %	60,86 %	78,26 %
	Multi	80,26 %	3,74 %	73,91 %	91,30 %
Empírico	Única	72,08 %	4,73 %	60,86 %	82,60 %
	Multi	80,08 %	5,05 %	69,56 %	91,30 %
Sturges2	Única	77,13 %	6,44 %	60,86 %	82,60 %
	Multi	89,30 %	3,64 %	82,60 %	95,65 %
Sturges10	Única	76,78 %	2,25 %	69,56 %	78,26 %
	Multi	95,73 %	2,84 %	86,95 %	100,00 %

Tabela XIII – Análise dos resultados utilizando GRAD e histogramas.

<i>Base de Dados</i>	<i>Tipo de Saída</i>	<i>Média</i>	<i>Desvio Padrão</i>	<i>Valor Mínimo</i>	<i>Valor Máximo</i>
Raiz Quadrada	Única	74,08 %	4,29 %	65,21 %	82,60 %
	Multi	78,69 %	3,19 %	73,91 %	86,95 %
Empírico	Única	72,08 %	4,31 %	65,21 %	82,60 %
	Multi	78,17 %	1,07 %	73,91 %	82,60 %
Sturges2	Única	73,82 %	2,40 %	65,21 %	78,26 %
	Multi	89,21 %	3,07 %	82,60 %	100,00 %
Sturges10	Única	77,82 %	3,64 %	69,56 %	91,30 %
	Multi	96,78 %	2,29 %	91,30 %	100,00 %

Para ambos os algoritmos BPM e GRAD, as melhores médias encontradas foram com a base Sturges10 com saída múltipla. Como no caso anterior, uma melhor comparação entre os experimentos pode ser visualizada na Figura 41.

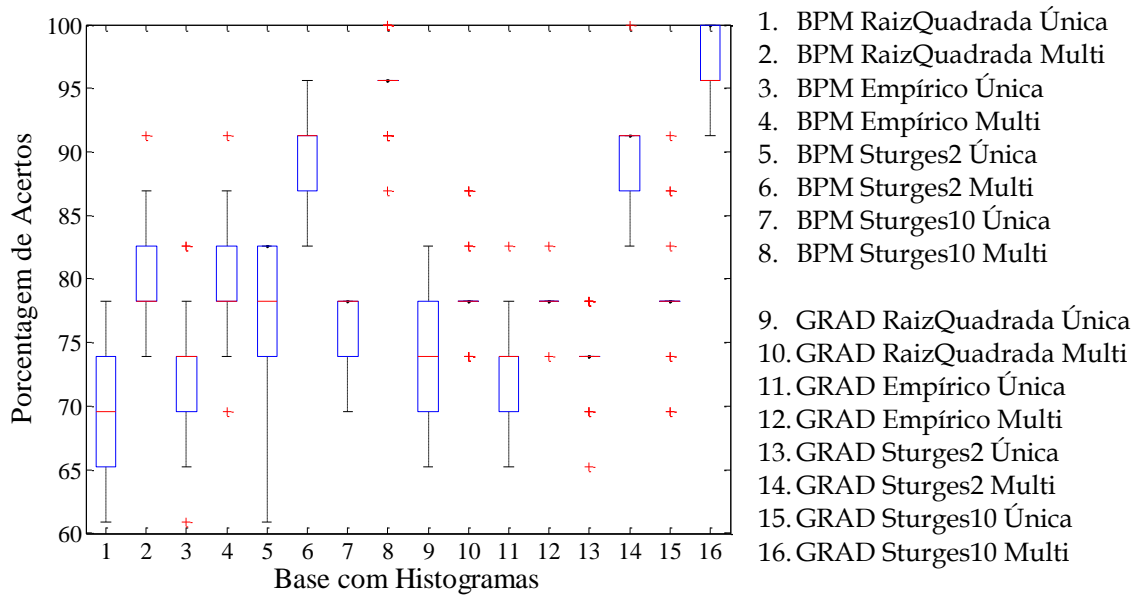


Figura 41 - *BoxPlot* dos resultados utilizando a base com histogramas.

5.2.2.3 PDFs

Foram executados os algoritmos de treinamento BPM (Tabela XIV) e GRAD (Tabela XV), utilizando as bases geradas com PDFs.

Tabela XIV - Análise dos resultados utilizando BPM e PDFs.

<i>Base de Dados</i>	<i>Tipo de Saída</i>	<i>Média</i>	<i>Desvio Padrão</i>	<i>Valor Mínimo</i>	<i>Valor Máximo</i>
Raiz Quadrada	Única	75,56 %	6,68 %	56,52 %	86,95 %
	Multi	93,47 %	3,19 %	86,95 %	100,00 %
Empírico	Única	75,73 %	6,75 %	56,52 %	86,95 %
	Multi	93,47 %	3,95 %	82,60 %	100,00 %
Sturges2	Única	78,69 %	4,66 %	69,56 %	91,30 %
	Multi	91,82 %	4,36 %	78,26 %	100,00 %
Sturges10	Única	80,78 %	2,92 %	73,91 %	86,95 %
	Multi	89,91 %	4,06 %	73,91 %	95,65 %

Tabela XV – Análise dos resultados utilizando GRAD e PDFs.

<i>Base de Dados</i>	<i>Tipo de Saída</i>	<i>Média</i>	<i>Desvio Padrão</i>	<i>Valor Mínimo</i>	<i>Valor Máximo</i>
Raiz Quadrada	Única	67,73 %	5,20 %	52,17 %	78,26 %
	Multi	93,47 %	3,07 %	86,95 %	100,00 %
Empírico	Única	68,00 %	5,94 %	56,52 %	82,60 %
	Multi	93,65 %	3,06 %	82,60 %	100,00 %
Sturges2	Única	69,56 %	3,92 %	60,86 %	73,91 %
	Multi	91,47 %	2,47 %	86,95 %	100,00 %
Sturges10	Única	72,08 %	2,16 %	69,56 %	73,91 %
	Multi	88,52 %	4,37 %	78,26 %	96,65 %

Para ambos os algoritmos, BPM e GRAD, as melhores médias encontradas foram com a base empírica com saída múltipla, porém no BPM a base raiz quadrada com saída múltipla obteve a mesma média. Na Figura 42, é apresentada a comparação entre os experimentos.

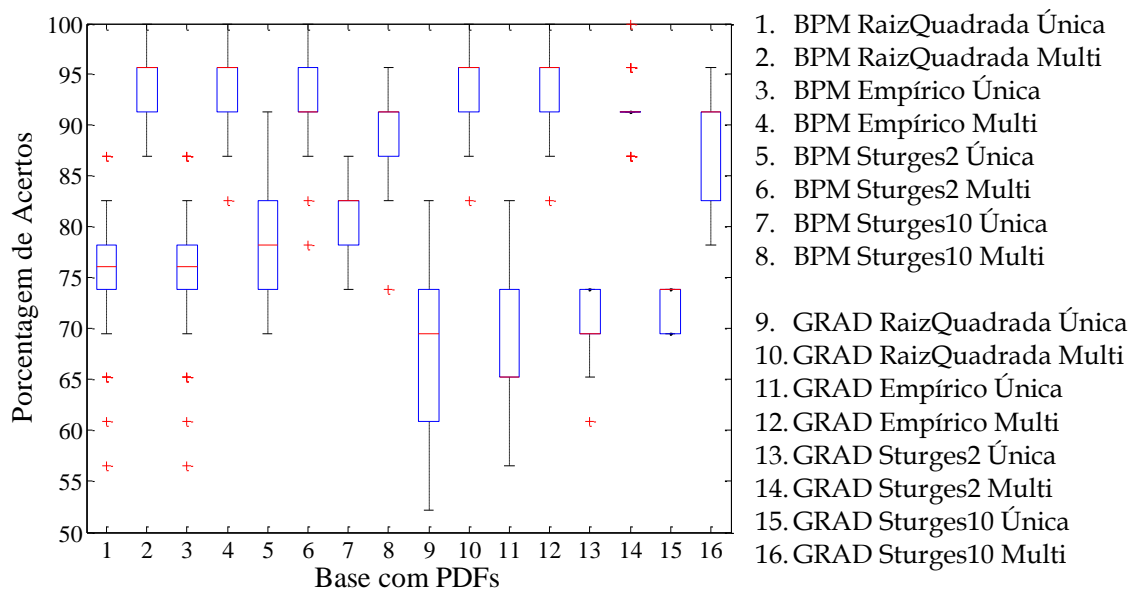


Figura 42 - BoxPlot dos resultados utilizando a base com PDFs.

5.3 Considerações Finais

Todo o sistema proposto foi desenvolvido na linguagem conhecida como Código-M, e foi utilizado na fase de elaboração e execução, a plataforma de desenvolvimento MATLAB© R2008a, e para os testes estatísticos foi utilizado o Software Estatístico R, com alguns pacotes desenvolvidos pelo Prof. Dr. Eric B. Ferreira². Os experimentos foram executados em um computador com Sistema Operacional Linux Ubuntu 8.04, com processador Core2Quad de 2,0GHz e 4Gb de memória RAM.

Durante todos os experimentos, foram armazenados os tempos de execução, cujo controle tem o intuito de comparar a diferença entre as diversas configurações que o sistema possui, além de servir como um comparativo para outros trabalhos.

É possível averiguar o tempo de execução em três níveis do algoritmo. O primeiro, "Total", refere-se ao tempo gasto para executar todos os 50 experimentos para cada conjunto de parâmetros (490×50 para o BPM, e 200×50 para o GRAD). O segundo, "Parcial", informa o tempo médio de cada um dos experimentos (490 para o BPM, e 200 para o GRAD). E o último, "Único", é o tempo médio que a rede neural leva para treinar e testar uma configuração do conjunto de parâmetros.

Esses valores podem ser vistos em duas tabelas, que apresentam de forma discriminada, os algoritmos de treinamento e as bases utilizadas. A Tabela XVI apresenta os experimentos com o algoritmo BPM e a Tabela XVII com o GRAD.

xcviii—

² Prof. Dr. Eric Batista Ferreira possui Pós-Doutorado em Estatística Multivariada, e faz parte do Instituto de Ciências Exatas da UNIFAL-MG como Professor Adjunto Nível I.

Tabela XVI - Análise dos tempos de execução do sistema para o BPM.

<i>Base de Dados</i>		<i>Total</i>	<i>Parcial</i>	<i>Único</i>
Momentos Amostrais	MD	3h 41min	4min 26s	0,39s
	MDAC	2h 1min	2min 25s	0,42s
	M5DAC	2h 2min	2min 27s	0,42s
Histogramas	Raiz Quadrada	36h 39min	43min 59s	1,93s
	Empírico	28h 10min	33min 48s	1,78s
	Sturges2	13h 43min	16min 27s	0,55s
	Sturges10	11h 22min	13min 38s	0,41s
Função de Densidade de Probabilidade	Raiz Quadrada	33h 14min	39min 53s	1,85s
	Empírico	28h 57min	34min 45s	1,59s
	Sturges2	12h 20min	14min 48s	0,52s
	Sturges10	4h 13min	5min 3s	0,45s

Tabela XVII - Análise dos tempos de execução do sistema para o GRAD.

<i>Base de Dados</i>		<i>Total</i>	<i>Parcial</i>	<i>Único</i>
Momentos Amostrais	MD	7h 48min	9min 22s	0,72s
	MDAC	11h 30min	13min 48s	0,69s
	M5DAC	11h 50min	14min 12s	0,69s
Histogramas	Raiz Quadrada	21h 45min	26min 6s	0,94s
	Empírico	16h 35min	19min 55s	0,86s
	Sturges2	8h 8min	9min 45s	0,72s
	Sturges10	6h 55min	8min 18s	0,70s
Função de Densidade de Probabilidade	Raiz Quadrada	22h 32min	27min 3s	0,94s
	Empírico	17h 11min	20min 37s	0,86s
	Sturges2	8h 16min	9min 55s	0,74s
	Sturges10	6h 54min	8min 17s	0,69s

Apesar do alto custo computacional gasto para executar todos os experimentos e configurações, a execução de apenas uma configuração é muito rápida, em média menos de 1 segundo, o que a torna muito viável para a utilização em sistemas de classificação de tempo real.

De modo geral, os experimentos que obtiveram os melhores resultados apresentaram, quase na totalidade, o tipo de saída múltipla. Quanto às bases utilizadas: para os momentos amostrais foi a base M5DAC; para os histogramas, foi a base Sturges10 e para os PDFs, foi a raiz quadrada. Curiosamente, o algoritmo de treinamento não apresentou muitas variações entre eles, BPM e GRAD.

As duas melhores configurações encontradas, uma para o BPM e outra para o GRAD, obtiveram respectivamente 97,57% e 96,78% de acertos. Tais valores representam mais de 22 acertos, do total de 23 conjuntos de sinais utilizados para os testes.

O BPM alcançou esse valor com o experimento que utiliza a base tratada com momentos amostrais MDAC, possuindo o tipo de saída única. Enquanto o GRAD obteve tal resultado utilizando a base tratada com histogramas Sturges10, possuindo saídas múltiplas. Os parâmetros que obtiveram tais resultados foram os seguintes conforme a Tabela XVIII, sendo que o erro, o número de iterações e a aleatoriedade são constantes com valores de 1^{-7} , 5000 e 0,5 respectivamente.

Tabela XVIII - Parâmetros dos resultados mais prósperos.

<i>Base de Dados</i>	<i>Neurônios</i>	<i>Aprendizado</i>	<i>Momentum</i>
BPM - momentos amostrais (MDAC) - única	20	0,001	0,1
GRAD - histogramas (Sturges10) - múltiplas	10	0,5	-

A seguir, são apresentadas as variações de resultados para todos os conjuntos de configurações (490 configurações), ocorrida nos 50 experimentos do algoritmo BPM utilizando a base MDAC com saída única. É possível averiguar os pontos mínimos e máximos, além da ocorrência média de resultados, na Figura 43.

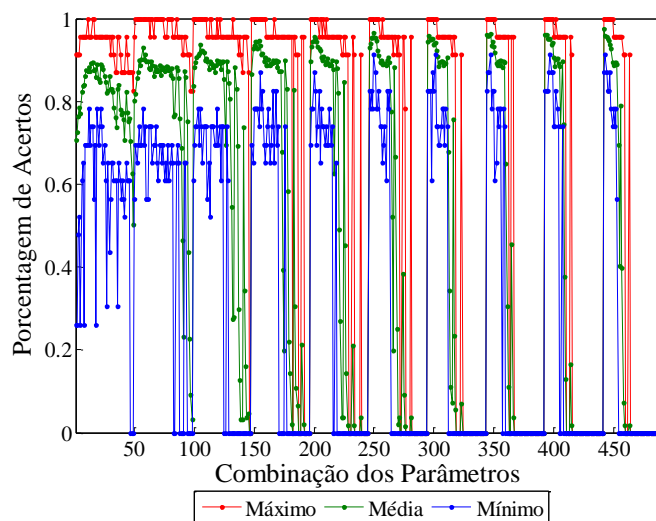


Figura 43 - Variação dos experimentos BPM para parâmetros de configuração.

Enquanto, para o algoritmo GRAD, a Figura 44 apresenta as mesmas variações de resultados para o conjunto de 200 configurações, sendo a base Struges10 através da transformação usando histograma, com múltiplas saídas.

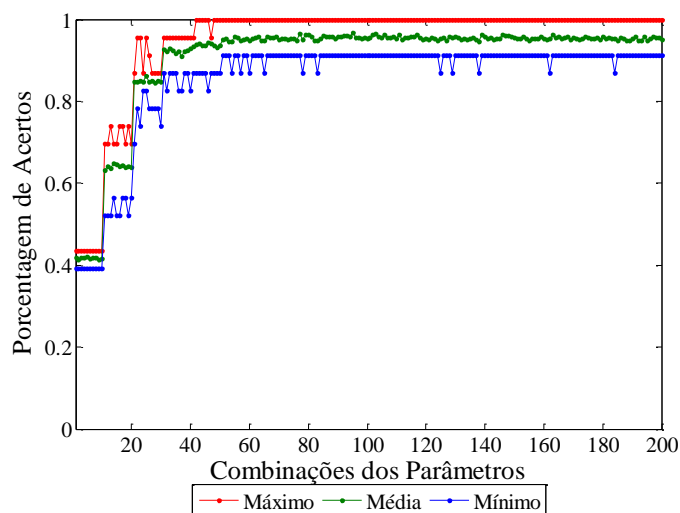


Figura 44 - Variação dos experimentos BPM para parâmetros de configuração.

Por fim, é possível verificar os parâmetros de configuração para todos os melhores resultados encontrados nos experimentos, sendo estes divididos de

acordo com o tratamento feito na base. A Figura 45 apresenta, em três dimensões, todos os três parâmetros de configuração para o algoritmo de treinamento BPM.

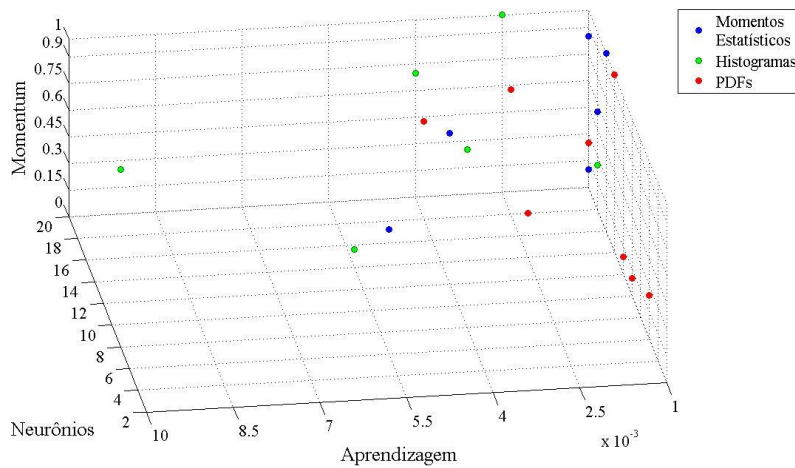


Figura 45 – Parâmetros de configuração para os resultados usando BPM.

Já para o GRAD, que possui dois parâmetros configuráveis, são apresentados os parâmetros de configuração distribuídos como segue a Figura 46.

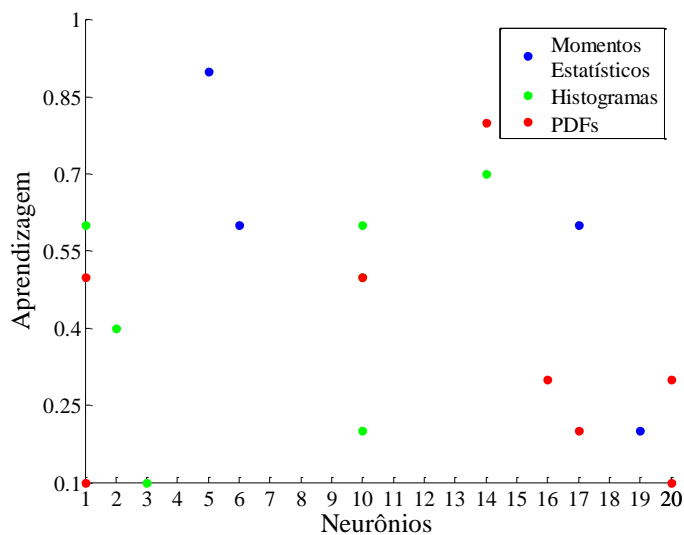


Figura 46 – Parâmetros de configuração para os resultados usando GRAD.

É importante, salientar, que os experimentos não foram executados para a base de dados bruta, assim como para os tratamentos com histograma e PDF que

utilizam "Todos os Pontos". O tempo gasto, para treinar as RNAs, é inviável com entradas muito grandes, 180.000 características neste caso. Além disso, pode haver, no treinamento, uma paralisia e estagnação dos pesos em mínimos locais, inviabilizando a convergência dos resultados.

Com base em toda a análise feita, é de grande valor que haja uma busca por um conjunto de configurações a fim de determinar quais destas possuem uma maior propensão em obter resultados que satisfaçam o problema a ser tratado, que neste caso é a classificação de padrões de escoamento para sistemas bifásico gás/líquido.

6

Conclusão e Trabalhos Futuros

*O sucesso é uma consequência e não um objetivo.
(Gustave Flaubert)*

O capítulo 6 apresenta algumas discussões envolvidas neste projeto de Trabalho de Conclusão de Curso. Na seção 6.1, é apresentada a conclusão, onde é descrita de forma sucinta a proposta envolvida, assim como as dificuldades encontradas e os resultados obtidos pelo sistema. A subseção 6.2 traz os possíveis trabalhos futuros seguindo esta mesma linha de pesquisa ou atuando em outros problemas de classificação de padrões.

6.1 Conclusão

Este projeto objetivou tratar um dos problemas mais comuns da fluidodinâmica utilizando técnicas computacionais. Em sistemas de escoamento de fluídos, há a dificuldade em se identificar e assim classificar os padrões de escoamento que passam por um duto. Com o intuito de sanar esse problema, foi proposto um sistema capaz de classificar tais padrões de escoamento, utilizando uma base de dados previamente capturada.

Especificamente, este projeto trabalhou com uma base onde o sistema de escoamento flui no sentido vertical, com substâncias gás e líquido. Utilizou-se a base desenvolvida pela FEM da Unicamp, sendo que esta foi tratada neste projeto

com diferentes técnicas (momentos amostrais, histogramas e função de densidade de probabilidade) a fim de ampliar o potencial de classificação.

Para a classificação, foram utilizadas as rede neurais do tipo MLP (treinada com os algoritmos BPM e GRAD), com o propósito de averiguar, de forma ampla, o potencial da rede para a classificação de padrões. Foram feitos testes exaustivos com um grande conjunto de parâmetros e diferentes algoritmos de treinamento.

Finalmente, houve uma análise estatística dos resultados, na qual foi possível verificar todos os experimentos realizados e assim indicar quais os parâmetros de configurações mais adequados para o problema de classificação de padrões de escoamento bifásico gás/líquido.

Inicialmente, as maiores dificuldades encontradas foram quanto à natureza do problema, tanto para os sistemas de escoamento quanto para as técnicas de classificação e esse fato deu-se, devido a pouca experiência para resolver o problema. Outra dificuldade encontrada foi perante aos tratamentos estatísticos que deveriam ser utilizados, uma vez que poucos são os trabalhos na literatura que utilizam esse tipo de análise para problemas computacionais.

Superadas as dificuldades, foi aplicada toda a proposta descrita acima, da qual se obteve resultados satisfatórios para o problema. Foi possível analisar o comportamento da rede, assim como averiguar que muitos conjuntos de configurações apresentam resultados "similares", sendo diferenciados pelo acaso.

Ainda assim, podem-se destacar os resultados obtidos com maior ganho, pois estes alcançaram médias de 97,57% e 96,78%, o que representa mais de 22,26 acertos em 23 padrões de teste em média. Outro fator destacável é que o maior desvio padrão foi cerca de 7%, que devido à quantidade de padrões utilizados para teste é pouco significativo (menos de dois padrões), além de muitos experimentos alcançarem 100% de acertos.

Após a seleção dos parâmetros, a execução e identificação dos padrões de escoamento acontecem de forma rápida, menos de 1 segundo, podendo ser utilizadas em sistemas de classificação em tempo real.

6.2 Trabalhos Futuros

Em termos de trabalhos futuros, podem-se destacar duas ramificações a serem seguidas. Uma é continuar com o trabalho em sistemas de escoamento, no qual destacam-se tais propostas:

- classificar padrões de escoamento que utilizam diferentes meios de aquisição de sinais, o que inclui também o tempo para essa captura;
- executar o sistema com um conjunto maior de sinais na base de dados;
- aplicar os parâmetros de configuração em sistemas reais;

A outra ramificação é quanto ao sistema de classificação:

- aplicar o sistema para outros problemas de classificação;
- avaliar diferentes configurações dos parâmetros, buscando uma faixa de valores que atendam aos diversos sistemas de classificação;
- utilizar outros algoritmos de treinamento;
- reduzir o tempo de execução, mas mantendo ou superando o poder de classificação;

7

Referências Bibliográficas

- Araribóia, G. (1987). *Inteligência Artificial: Um Curso Prático*. Rio de Janeiro, RJ, Brasil: Livros Técnicos e Científicos Editora Ltda.
- Araújo, W. R. (2008). *Aplicação de Técnicas de Inteligência Artificial Visando Eficiência Energética e Estimação de Parâmetros em Sistemas Motrizes Industriais*. Dissertação de Mestrado, UFMS, Engenharia Elétrica, Campo Grande - MS.
- Beggs, H. D., & Brill, J. P. (Maio de 1973). *Study of Two-phase Flow in Inclined Pipes*. *Journal of Petroleum Technology* , 607-617.
- Bilobrovec, M., Marçal, R. F., & Pilatti, L. A. (novembro de 2004). *Aplicações de redes neurais artificiais em simulações*. XI SIMPEP .
- Bin, S., Hongjian, Z., Lu, C., & Yuxiao, Z. (2006). *Flow regime identification of gas-liquid two-phase flow based on hht*. *Chinese Journal Chemical Engineering* , 14 , 1, 24-30.
- Bishop, C. M. (1995). *Neural Networks for Pattern Recognition*. Oxford University press.
- Buragohain, M., & Mahanta, C. (2008). *A Novel Approach for ANFIS Modelling based on Full Factorial Design*. *Applied Soft Computing* , 8, 609-625.
- Cai, S., Toral, H., Qiu, J., & Archer, J. S. (1994). *Neural network based objective flow regime identification in air-water two-phase flow*. *Canadian Journal of Chemical Engineering* , 72 , 440-445.
- Cera, M. C. (2005). *Uso De Redes Neurais Para O Reconhecimento de Padrões*. Trabalho da Disciplina Arquiteturas Especiais de Computadores.
- Corrêa, F. C. (2009). *Sistema Embarcado Para Medidas De Escoamento Bifásico Gás-líquido*. Dissertação (Mestrado em Engenharia Mecânica), Unicamp, Faculdade de Engenharia Mecânica, Campinas.
- Dailiang, X., Haifeng, J., Zhiyao, H., & Haiqing, L. (2004). *An Online Flow Pattern Identification System For Gas-oil Two-phase Flow Using Electrical Capacitance Tomography*. *Instrumentation and Measurement - Technology Conference* , 2320-2325.
- Dapper, C. D., Strauss, L. M., Leiria, R., & Furlanetto, T. M. (2003). *Introdução ao Estudo dos Sistemas Especialistas*. Universidade do Vale do Rio dos Sinos - UNISINOS, Análise de Sistemas.
- Dias, J. A. (2004). *A Análise Sedimentar do Conhecimento dos Sistemas Marinhos*. Versão Preliminar.

- Ding, H., Huang, Z., Song, Z., & Yan, Y. (2007). *Hilbert-huang transform based signal analysis for the characterization of gas-liquid two-phase flow*. Chinese Journal Chemical Engineering , 18 , 37-46.
- Dong, F., Jiang, Z. X., Qiao, X. T., & Xu, L. A. (2003). *Application Of Electrical Resistance Tomography To Two-phase Pipe Flow Parameters Measurement*. Flow Measurement and Instrumentation , 14, 183-192.
- Dukler, A. E., & Taitel, W. (1986). *Flow Pattern Transitions in Gas-Liquid Systems: Measurement and Modelling*. *Multiphase Science and Technology*. Hemisphere Publishing Corporation.
- Durst, F. (1976). *Principles and practice of laser-doppler anemometry*. Londres, Inglaterra: London: Academic Press.
- Ferrari, F. (2004). *Estatística Básica*. Baseado em Estatística Fácil de Antônio Arnot Crespo, Editora Saraiva, 1999.
- Garcia, A. L. (1989). *Probability and Random processes for Electrical Engineering*. 583. New York, E.U.A.: Addison-Wesley Publishing Compan.
- Glymour, C. D., Madigan, D., Pregibon, D., & Smyth, P. (1997). *Statistical Themes and Lessons for Data Mining*. Data Mining and Knowledge Discovery , 1 , 1, 11-28.
- Hagedorn, A. R., & Brown, K. E. (abril de 1965). *Experimental Study of Pressure Gradients Occumng During Continuous Two-Phase Flow in Small-Diameter Vertical Conduits*. Journal of Petroleum Technology , 475-484.
- Hewiit, G. F. (1982). *Handbook of Multiphase Systems*. Hemisphere Publishing Corporation.
- Hulin, J. P., Fierfort, C., & Condol, R. (1994). *Experimental Study of Vortex Emission Behind Bluff Obstacles in a Gas Liquid Vertical Two Phase Flow*. International Journal Multiphase Flow , 8 , 475-90.
- Ishii, M. (1975). *Thermo-Fluid Dynamic Theory of Two-Phase Flow*. Collection de la Direction des Etudes et Recherches d'Electricite de France.
- Jain, A. K. (2000). *Statistical pattern recognition: A review*. IEEE Transactions on Pattern Analysis and Machine Intelligence , 22 (1), 4-37.
- Jesan, J. P. (2005). *The Neural Approach To Pattern Recognition*. Acesso em 01 de 03 de 2010, disponível em Ubiquity: IT Magazine and forum: http://www.acm.org/ubiquity/views/v5i7_jesan.html
- Jones, O. C., & Zuber, N. (1975). *The interrelation between void fraction fluctuactions and flow patterns in two phase flow*. International Journal Multiphase Flow , 2 , 273-306.
- Lacerda, W. S. (2006). *Redes Neurais Artificiais*. http://algor.dcc.ufla.br/~lacerda/download/palestras/rna/introducao_RNA_2006.ppt, Universidade Federal de Lavras, Departamento de Ciência da Computação.

- Lee, J. Y., Paranjape, S. S., Kim, N. S., & Ishii, M. (2004). *Instantaneous And Objective Flow Regime Identification Method*. 5th International Conference on Multiphase Flow , 204.
- Lima, C. A. (2000). *Emprego de Teoria de Agentes no Desenvolvimento de Dispositivos Neurocomputacionais Híbridos e Aplicação ao Controle e Identificação de Sistemas Dinâmicos*. Tese de Mestrado, Unicamp, FEEC.
- Malyeri, M. R., Muller-Steinhagen, H., & Smith, J. M. (2003). *Neural Network Analysis Of Void Fraction In Air/Water Two-phase Flows At Elevated Temperatures In A Gas-liquid Vertical Two-phase Flow*. Chemical Engineering and Processing , 42, 587-597.
- Matsui, G. (1984). *Identification of Flow Regimes in Vertical Gas-Liquid Two-Phase Flow Using Differential Pressure Fluctuations*. International Journal Multiphase Flow , 711-720.
- Mi, Y., Ishii, M., & Tsoukalas, L. H. (2001). *Flow Regime Identification Methodology With Neural Networks And Two-phase Flow Models*. Nuclear and Engineering Design , 87-100.
- Mi, Y., Ishii, M., & Tsoukalas, L. H. (1998). *Vertical Two-Phase Flow Identification Using Advanced Instrumentation and Neural Networks*. Nuclear Engineering and Design , 409-420.
- Mi, Y., Ishii, M., Tsoukalas, L. H., Li, M., & Xiao, Z. (1996). *Hybrid Fuzzy-Neural Flow Identification Methodology*. 5th IEEE International Conference on Fuzzy Systems .
- Mishima, K., & Hibiki, T. (1996). *Visualization and Measurements of Two-Phase Flow in Metallic Ducts Using Neutrons as Microscopic Probes*. JSME B .
- Mishima, K., & Ishii, M. (1984). *Flow Regime Transition Criteria For Upward Two-Phase Flow in Vertical Tubes*. International Journal Heat Mass Transfer .
- Moreira, F. C. (2002). *Reconhecimento e Classificação de Padrões de Imagens de Núcleos de Linfócitos do Sangue Periférico Humano com a Utilização de Redes Neurais Artificiais*. Dissertação de Mestrado, Universidade Federal de Santa Catarina, Ciência da Computação, Florianópolis.
- Nikolopoulos, C. (1997). *Expert Systems: Introduction to First and Second Generation and Hybrid Knowledge Based Systems*. Dekker, New York, USA.
- Nose, K. F., Lotufo, A. D., & Lopes, M. L. (maio de 2008). *Utilização de Redes Neurais Artificiais e Redes Neuro Fuzzy para Previsão de Cargas Elétrica*. 7th Brazilian Conference on Dynamics, Control and Applications .
- Oliveira, C. A., & Belderrain, M. C. (2006). *Classificação de Unidades Móveis no Ambiente Marítimo*. Instituto Tecnológico de Aeronáutica, São José dos Campos, SP.
- Pacheco, F., Cavalcante, C. C., Mendes, J. R., Bannwart, A. C., & Serapião, A. B. (21-24 de outubro de 2007). *Identificação de Padrões de escoamento Multifásico e Determinação da Perda de Carga com o Auxílio de Redes Neurais*. 4º PDPETRO , p. 10.

- Rezende, S. O. (2005). *Sistemas Inteligentes: Fundamentos e Aplicações* (1ª ed. 1ª reimpressão ed.). Barueri, SP, Brasil: Editora Manole Ltda.
- Rich, E., & Knight, K. (1993). *Inteligência Artificial* (2ª ed. ed.). Makron Books.
- Rosa, E. S., & Ohishi, T. (2008). *Desenvolvimento de Sensor para Detecção de Golfadas para Operação em Gas-Lift Intermitente*. 4º Relatório de Atividades, Unicamp, DE/FEM; DENSIS/FEEC, Campinas.
- Rosa, E., Salgado, R., Ohishi, T., & Mastelari, N. (2010). *Performance Comparison of Artificial Neural Networks and Expert Systems Applied to Flow Pattern Identification in Vertical Ascendant Gas-Liquid Flows*. *International Journal of Multiphase Flow* .
- Salgado, R. M. (2009). *Métodos de Treinamento, Redes Neurais - Perceptron MLP*. Data Show, Universidade Estadual de Campinas, Faculdade de Engenharia Elétrica e de Computação, Campinas.
- Salgado, R. M., Ohishi, T., Rosa, E. S., Mastelari, N., & de Oliveira, H. C. (2010). *Algoritmos de Agrupamento e Sistema Especialistas Aplicados no Reconhecimento de Padrões de Escoamentos Gás/Líquido*. Simpósio Brasileiro de Pesquisa Operacional .
- Scarselli, F., & Tsoi, A. C. (1998). *Universal Approximation Using Feedforward Neural Networks: a Survey of Some Existing Methods, and Some New Results*. *Neural Networks* , 11 (1), 15–37.
- Schalkoff, R. J. (1992). *Pattern Recognition: Statistical, Structural and Neural Approches*. John Wiley & Sons.
- Selli, M. F., & Selegim, P. (2004). *On-line Identification Of Horizontal Two-phase Flow Regimes Through Gabor Transform And Neural Network Processing*. 10th Congresso Brasileiro de Ciências Térmicas e Engenharia (ENCIT) .
- Shepherd, A. J. (1995). *Stopped Training and Other Remedies for Overfitting*. Proceeding of the Symposium on the Interface .
- Silva, C. B., Borges, M. J., & Pinheiro, J. A. (2000). *Medição de Vazão e Propriedades em Escoamento Multifásico: Solução Econômica para Diferentes Atividades Industriais*. Bol. téc. PETROBRAS, Petrobrás, SEPROT, SEPROJ, GERPRO, Rio de Janeiro.
- Silva, L. N. (1998). *Análise e Síntese de Estratégias de Aprendizado para Redes Neurais Artificiais*. Dissertação de Mestrado, Unicamp, Faculdade de Engenharia Elétrica e de Computação, Campians -SP.
- Silva, L. N., & Zuben, F. J. (1998). *Optimised Traininig Techniques for Feedforward Neural Networks*. Relatório Técnico, Unicamp, Faculdade de Engenharia Elétrica e de Computação, Campinas - SP.
- Souza, R. G. (2009). *Escoamento Bifásico Líquido-Gás: Previsão de Gradientes de Pressão com a Correlação de Lockhart & Martinelli e Fluidodinâmica Computacional*. Rio de Janeiro, RJ.

- Sun, Z., & Zhang, H. (2008). *Neural Networks Approach for Prediction of Gas-Liquid Two-Phase Flow Pattern Based on Frequency Domain Analysis of Vortex Flowmeter Signal*. *Measurement Science and Technology*, 19.
- Taitel, Y., Dukler, A. E., & Barnea, D. (1980). *Modeling Flow Pattern Transitions for Steady Upward Gas-Liquid Flow in Vertical Tubes*. *AIChE Journal*, 26, 345-354.
- Tatibana, C. Y., & Kaetsu, D. Y. (2006). *Seja bem vindo à homepage de Redes Neurais*. (W. Designer, Editor) Acesso em 10 de 06 de 2010, disponível em Uma Introdução às Redes Neurais: <http://www.din.uem.br/ia/neurais/>
- Thermopedia™. (2006). *A-to-Z Guide to Thermodynamics, Heat & Mass Transfer, and Fluids Engineering Online*. (B. House, Editor) Acesso em 10 de 06 de 2010, disponível em Thermopedia: http://www.thermopedia.com/toc/chapt_g/GAS-LIQUID_FLOW.html
- Tsoukalas, L. H., Ishii, M., & Mi, Y. A. (1997). *Neuro-fuzzy Methodology For Impedance-based Multiphase Flow*. *Engineering Application of Artificial Intelligence*, 10, 545-555.
- Tutu, N. (1982). *Pressure Fluctuations and Flow Pattern Recognition in Vertical Two Phase Gas-Liquid Flows*. *International Journal Multiphase Flow*, 8, 443-447.
- Vince, M. A., & Lahey, R. T. (1982). *On the Development of an Objective Flow Regime Indicator*. *International Journal Multiphase Flow*, 8, 93-124.
- Wallis, G. B. (1969). *One-dimensional Two-phase Flow*. New York.
- Warsito, Ohkawa, M., Kawata, N., & Uchida, S. (1999). *Cross-Sectional Distributions of Gas and Solid Holdups in Slurry Bubble Column Investigated by Ultrasonic Computed Tomography*. *Chem Eng Sc*, 54, 4711-4728.
- Whalley, P. B. (1987). *Boiling and Condensation and Gas-Liquid Flow*. Oxford: Clarendon Press.
- Whalley, P. B. (1996). *Two-Phase Flow and Heat Transfer*. Oxford University Press.
- Wu, H., Zhou, F., & Wu, Y. (2001). *Intelligent Identification System Of Flow Regime Of Oil-gas-water Multiphase Flow*. *International Journal of Multiphase Flow*, 27, 459-475.
- Xu, L., & Han, Y. (1997). *Application of Ultrasonic Tomography to Monitoring Gas/Liquid Flow*. *Chem Eng Sci*, 52, 13, 2171-2183.

8

Anexos

8.1 Backpropagation com Termo Momentum

```

% BPM
% Main Program (function)
% MLP net with Backprop training
% Standard BP with Momentum
% Off-line Updating
% Author: Leandro Nunes de Castro
% Unicamp, January 1998

% Entrada
% P = Valores do Conjunto de Entrada
% T = Saída do Conjunto de Entrada
% nh = Numero de Neurônios
% alfa = Taxa de aprendizagem
% cm = Termo Momentum
% minerro = erro mínimo desejado
% maxep = numero máximo de épocas
% val = valor para incrementar a aleatoriedade
% Saída
% w1 = Pesos da Camada Intermediária
% w2 = Pesos da Camada de Saída
% y = Saída da Função Linear

function [w1,w2,y] = Treino_BPM(P,T,nh,alfa,cm,minerro,maxep,val)
%-----
% Definição e Inicialização dos parâmetros
%-----
P0 = P; %Histórico de Treinamento
[np,ni] = size(P0);
[no] = size(T,2); % Saída desejada no treinamento
ep = 0;
w1 = 2.* val.*rand(ni+1,nh) - val;
w2 = 2.* val.*rand(nh+1,no) - val;

%-----
% Treinamento da Rede
%-----
sse = 10; sseant = sse; veter = [];
P = [ones(np,1) P0];

```

```

while (ep < maxep & sse > minerro)
    sseant = sse; sse = 0;
    gdw1=zeros(ni+1,nh); gdw2=zeros(nh+1,no);

    %-----
    % Forward
    %-----
    z0 = tanh(P*w1);
    z = [ones(np,1) z0];
    y = z*w2; % Linear output

    %-----
    % Correção e Calculo dos Erros
    %-----
    dk = (T-y); gdw2 = z'*dk; % Linear output
    w20 = reshape(w2(2:nh+1,:),nh,no);
    dj = (dk*w20').*(1-z0.^2);
    gdw1 = P'*dj;
    verr = (T-y); verr = reshape(verr,np*no,1);
    sse = verr'*verr;

    %-----
    % Atualização do Termo Momentum
    %-----
    w1a = w1; w2a = w2;
    w1 = w1 + alfa*gdw1;
    w2 = w2 + alfa*gdw2;
    w1 = w1 + cm *(w1-w1a);
    w2 = w2 + cm *(w2-w2a);
    ep = ep + 1;
    vgrad = [reshape(gdw1,(ni+1)*nh,1);
    reshape(gdw2,(nh+1)*no,1)];
    ngrad = norm(vgrad);
    veter = [veter sse];
end;
end

```

8.2 Método Gradiente

```

% GRAD
% Main Program (function)
% MLP net with Backprop training
% Gradient method
% Secondary functions: UNIDIM
% Off-line Updating
% Author: Leandro Nunes de Castro
% Unicamp, January 1998

```

```

% Entrada
%     P = Valores do Conjunto de Entrada
%     T = Saída do Conjunto de Entrada
%     nh = Numero de Neurônios
%     cm = Termo Momentum
%     minerro = erro mínimo desejado
%     maxep = numero máximo de épocas
%     val = valor para incrementar a aleatoriedade
% Saída
%     w1 = Pesos da Camada Intermediária
%     w2 = Pesos da Camada de Saída
%     y = Saída da Função Linear

function [w1, w2, y] = Treino_GRAD(P,T,nh,cm,minerr,maxep,val)
%-----
% Definição e Inicialização dos parâmetros
%-----
warning off
P0 = P;
[np,ni] = size(P0);
[no] = size(T,2);

ep = 0; alfa = 0.001;
w1 = 2.* val.*rand(ni+1,nh) - val;
w2 = 2.* val.*rand(nh+1,no) - val;

%-----
% Treinamento da Rede
%-----
sse = 10; sseant = sse;
P = [ones(np,1) P0]; veter = []; vetalfa = [];
%fini = flops; t0 = clock;
val = 0;

while (ep < maxep & sse > minerr)
    sseant = sse; sse = 0;
    gdw1=zeros(ni+1,nh); gdw2=zeros(nh+1,no);

    %-----
    % Forward
    %-----
    z0 = tanh(P*w1);
    z = [ones(np,1) z0];
    y = z*w2; % linear output

    %-----
    % Correção e Calculo dos Erros
    %-----
    dk = (T-y); gdw2 = z'*dk;
    w20 = reshape(w2(2:nh+1,:),nh,no);
    dj = (dk*w20').*(1-z0.^2);
    gdw1 = P'*dj;
    verr = (T-y); verr = reshape(verr,np*no,1);
    sse = verr'*verr;

```

```

vgrad = [reshape(gdw1, (ni+1)*nh,1); reshape(gdw2, (nh+1)*no,1)];
ngrad = norm(vgrad);
alfa = Treino_UNIDIM(gdw1,gdw2,w1,w2,alfa,cm,sse,sseant,T,P);

%-----
% Atualização do Termo Momentum
%-----
w1a = w1; w2a = w2;
w1 = w1 + alfa*gdw1/ngrad;
w2 = w2 + alfa*gdw2/ngrad;
w1 = w1 + cm *(w1-w1a);
w2 = w2 + cm *(w2-w2a);
ep = ep + 1;
veter = [veter sse]; vetalfa = [vetalfa alfa];
end;
end

```

8.3 Unidim

```

% UNIDIM
% Secondary function
% Inexact line search
% Off-line Updating
% Author: Leandro Nunes de Castro
% Unicamp, January 1998

function alfa = Treino_UNIDIM(gdw1,gdw2,w1,w2,alfa,cm,sse,sseant,T,P)
% Line search
[np,ni] = size(P); ni = ni-1;
[nh,no] = size(w2); nh = nh-1;

if sse < sseant
    w1a = w1; w2a = w2;
    w1 = w1 + alfa*gdw1;
    w2 = w2 + alfa*gdw2;
    w1 = w1 + cm *(w1-w1a);
    w2 = w2 + cm *(w2-w2a);
    if alfa < .5, % Maximum value limit
        alfa = 1.2*alfa;
    end;
else
    aux = 1;

    while (sse >= sseant & aux < 5)
        aux = aux + 1;
        ssep=sse; sse=0;
        alfa = .618*alfa;
        w1a = w1; w2a = w2;
        w1p = w1 + alfa*gdw1;

```



```

w2p = w2 + alfa*gdw2;
w1p = w1p + cm *(w1p-w1a);
w2p = w2p + cm *(w2p-w2a);

%-----
% Forward pass
%-----
z = tanh(P*w1p);
z = [ones(np,1) z];
y = z*w2p; % Linear output
verr = (T-y); verr = reshape(verr,np*no,1);
sse = verr'*verr;
end;
end;
end

```

8.4 Teste dos Pesos

```

%Previsor via RNA Estatica
% Backpropagation Classico (bmp)
% Backpropagation Classico-Gradiente descentedente

% Entrada
% arquivo_teste = base de teste
% w1 = peso da camada intermediaria
% w2 = peso da camada de saida
% Saída
% Valor_previsto = valores obtidos para a base de teste

function [Valor_previsto]=Testa_RNA(arquivo_teste, w1, w2)
arquivo_teste0 = arquivo_teste; %Historico de Treinamento
[np,ni] = size(arquivo_teste0);
arquivo_teste = [ones(np,1) arquivo_teste0];

%-----
% Forward
%-----
z0 = tanh(arquivo_teste*w1);
z = [ones(np,1) z0];
Valor_previsto = z*w2;

end

```

8.5 Teste de Padrões

```

%Função Testa os Resultados
%Entrada
% obtido - Resultado Obtido pelo teste
% desejado - Resultado esperado pela base de Teste
%Saída
% padroes - retorna os valores dos padroes aproximados
% acertos - quantidade de acertos obtidos

function [padroes, acertos] = Testa_Padroes(obtido, desejado)
    [LObtido, CObtido] = size(obtido);

    %Única saída
    if (CObtido == 1)
        %Percorre todas as linhas
        for linha = 1:LObtido
            %Padrao Capa Esferica - ].. 0.1[
            if (obtido(linha) < 0.1)
                aproximacao = 0.0;
            %Padrao Agitado - [0.1 0.3[
            elseif (obtido(linha) < 0.3)
                aproximacao = 0.2;
            %Padrao Golfadas - [0.3 0.5[
            elseif (obtido(linha) < 0.5)
                aproximacao = 0.4;
            %Padrao Anular - [0.5 0.7[
            elseif (obtido(linha) < 0.7)
                aproximacao = 0.6;
            %Padrao Semi Anular - [0.7 0.9[
            elseif (obtido(linha) < 0.9)
                aproximacao = 0.8;
            %Padrao Bolhas - [0.9 ..[
            else
                aproximacao = 1.0;
            end

            %Retorna o Padrao Mais Proximo
            padroes(linha) = aproximacao;
            acertos(linha) = (padroes(linha) == desejado(linha));
        end

    %Multiplas Saídas
    else
        for linha = 1:LObtido
            [x,y] = find(obtido(linha,:) == max(obtido(linha,:)), 1);
            padroes(linha) = y;
            [x, y] = find(desejado(linha, :) == 1);
            acertos(linha) = (padroes(linha) == y);
        end
    end
end

```

8.6 Permutação dos Parâmetros para o BPM

```

%Entrada
%   repeticoes - numero de interacoes do algoritmo
%   treinoIN - matriz de treinamento entrada
%   treinoOUT - matriz de treinamento saida
%   testeIN - matriz de teste entrada
%   testeOUT - matriz de teste saida
%Saida
%   configuracao - retorna o valor dos teste

function [resultado, parametros] = Alg_Config_BPM(repeticoes,
treinoIN, treinoOUT, testeIN, testeOUT)
    [Lteste, Cteste] = size(testeIN);

    %variaveis da Rede
    erro = 1e-7;
    iteracoes = 5000;
    aleatoriedade = 0.5;

    %Iteracoes do algortimo
    for repeticao = 1 : 1 : repeticoes
        cont = 1;
        Tempo_Config = cputime;

        %Numero de neuronios
        for neuronio = 2 : 2 : 20
            %Taxa de aprendizado
            for aprendizado = 0.001 : 0.0015 : 0.01
                %termo momentum
                for momentum = 0.1 : 0.15 : 1.0

                    [w1, w2] = Treino_BPM(treinoIN , treinoOUT ,
neuronio, aprendizado, momentum, erro, iteracoes, aleatoriedade);
                    ValoresPrevistos = Testa_RNA(testeIN, w1, w2);

                    if (sum(isnan(ValoresPrevistos)) < Lteste)
                        [padroes, acertos] =
Testa_Padros(ValoresPrevistos, testeOUT);

                        resultado.Configuracao(repeticao, cont) =
(sum(acertos) / Lteste);
                        Configuracao_BPM = {repeticao neuronio
aprendizado momentum sum(acertos)}
                    else
                        resultado.Configuracao(repeticao, cont) = 0;
                        Configuracao_BPM = {repeticao neuronio
aprendizado momentum 'ERROR!!'}
                    end
                end
            end
        end
    end

```

```

        parametros(:, cont) = [cont, neuronio,
aprendizado, momentum, erro, iteracoes, aleatoriedade];

        cont = cont + 1;

    end
end
end

    resultado.Tempo(repeticao, :) = cputime - Tempo_Config;
    resultado.Analise = [mean(resultado.Configuracao);
std(resultado.Configuracao); min(resultado.Configuracao);
max(resultado.Configuracao)]

end
end

```

8.7 Permutação dos Parâmetros do GRAD

```

%Entrada
%   repeticoes - numero de interacoes do algoritmo
%   treinoIN - matriz de treinamento entrada
%   treinoOUT - matriz de treinamento saida
%   testeIN - matriz de teste entrada
%   testeOUT - matriz de teste saida
%Saida
%   configuracao - retorna o valor dos teste

function [resultado, parametros] = Alg_Config_GRAD(repeticoes,
treinoIN, treinoOUT, testeIN, testeOUT)
    [Lteste, Cteste] = size(testeIN);

    %variaveis da Rede
    erro = 1e-7;
    iteracoes = 5000;
    aleatoriedade = 0.5;

    %Numero de neuronios
    for neuronio = 1 : 1 : 20
        %Taxa de aprendizado
        for aprendizado = 0.1 : 0.1 : 1.0

            [w1, w2] = Treino_GRAD(treinoIN ,treinoOUT , neuronio,
aprendizado, erro, iteracoes, aleatoriedade);
            ValoresPrevistos = Testa_RNA(testeIN, w1, w2);

            if (sum(isnan(ValoresPrevistos)) < Lteste)
                [padroes,acertos]=Testa_Padros (ValoresPrevistos,
testeOUT);
            end
        end
    end
end

```

```

                                resultado.Configuracao(repeticao, cont) =
(sum(acertos) / Lteste);
                                Configuracao_GRAD = {repeticao neuronio
aprendizado sum(acertos)}
                                else
                                resultado.Configuracao(repeticao, cont) = 0;
                                Configuracao_GRAD = {repeticao neuronio
aprendizado 'ERROR!!'}
                                end

                                parametros(:, cont) = [cont, neuronio, aprendizado, 0,
erro, iteracoes, aleatoriedade];

                                cont = cont + 1;

                                end
                                end

                                resultado.Tempo(repeticao, :) = cputime - Tempo_Config;
                                resultado.Analise = [mean(resultado.Configuracao);
std(resultado.Configuracao); min(resultado.Configuracao);
max(resultado.Configuracao)]

                                end
                                end

```


9

Apêndice

9.1 Redes Neurais Artificiais

A rede neural artificial (RNA) é uma técnica da inteligência artificial (IA), sendo um dos ramos da IA que mais se desenvolveu ultimamente. *"As redes neurais artificiais configuram um modelo matemático-computacional com base na estrutura neural de seres vivos inteligentes e que possuem a capacidade de adquirir conhecimento através da experiência."* (Araújo, 2008, p. 22).

O sistema nervoso humano é constituído por uma rede de neurônios. Cada neurônio é capaz de receber, processar e transmitir sinais a outros neurônios, compreendendo desta forma, um sistema de comunicação cerebral.

A comunicação cerebral é feita através da conexão de vários neurônios por meio dos dendritos e dos axônios, que são responsáveis por conduzir os estímulos sinápticos ao núcleo celular. Esses estímulos capacitam o cérebro a armazenarem informações e conseqüentemente a aprender. A estrutura de um neurônio pode ser vista na Figura 47.

Com base no sistema nervoso, diversas pesquisas foram produzidas com o intuito de compreender o funcionamento do raciocínio humano. Dessa maneira foram desenvolvidos sistemas computacionais com bases matemáticas que simulam redes neurais.

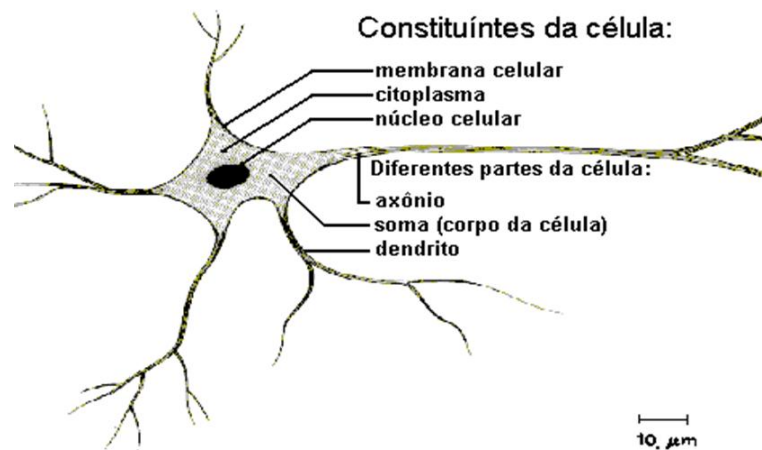


Figura 47 - Estrutura neuronal (Tatibana & Kaetsu, 2006).

9.1.1 Funcionamento

"Assim como na natureza, um neurônio artificial é uma estrutura relativamente simples, que responde a estímulos de outros neurônios conectados a ele. Essa arquitetura dá as redes neurais características marcantes de intenso paralelismo e robustez." (Moreira, 2002, p. 32).

O modelo geral de neurônio é uma generalização do modelo de McCulloch e Pitts (modelo MCP). Um neurônio é uma estrutura, o qual recebe de entrada, estímulos de um conjunto de valores $X = [x_1, x_2, \dots, x_n]^T$ contendo n elementos. Para cada valor x_i é atribuído um peso w_i , onde $w_i \in W$, sendo que $W = [w_1, w_2, \dots, w_n]^T$.

A soma dos produtos dos valores x_i pelos pesos w_i é chamada de saída linear u , onde $u = \sum_{i=0}^n (x_i \cdot w_i)$. A saída linear u é utilizada no cálculo da função de ativação, $y = f(u)$, onde y é a saída do neurônio (saída de ativação).

Em alguns modelos de neurônios, é encontrado um parâmetro denominado bias, w_b , que têm o papel de aumentar ou diminuir a influência dos valores das entradas. A estrutura de um neurônio pode ser visualizada na Figura 48.

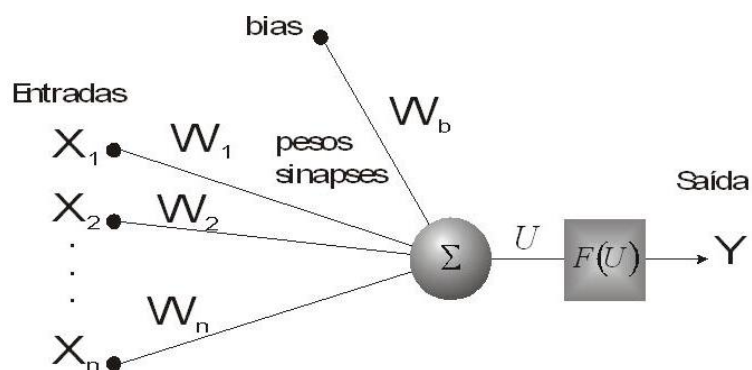


Figura 48 - Estrutura de um neurônio artificial (Salgado, 2009, p. 2).

Um fator importante é a escolha da função de ativação, que tem o propósito de normalizar a saída linear u . As principais funções de ativação (Figura 49) encontradas na literatura são: identidade, degrau, rampa, sigmoideal e tangente hiperbólica.

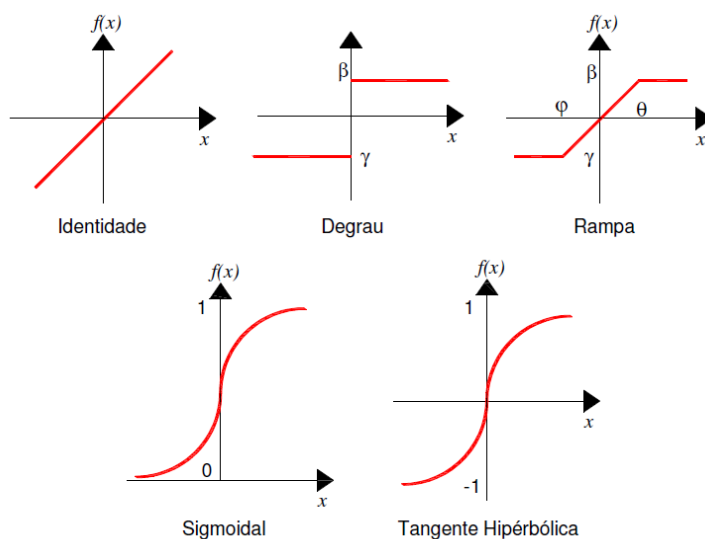


Figura 49 - Principais funções de ativação (Araújo, 2008, p. 25).

Segundo Bilobrovec, Marçal & Pilatti (2004, p. 3) "*Uma rede neural é formada por um conjunto de unidades de processamento simples que se comunicam enviando sinais uma para a outra através de conexões ponderadas.*"

Existem vários tipos de redes neurais artificiais e dentre essas redes, a mais usada é a rede neural do tipo multilayer perceptron (MLP), devido a sua fácil

implementação e por sua simplicidade. Uma rede neural MLP pode ser composta de várias camadas de processamento, sendo classificadas em três grupos (Figura 50):

- **Camada de entrada:** passagem dos valores de entrada para a rede;
- **Camadas intermediárias ou escondidas:** conjunto de neurônios responsáveis pelo aprendizado da rede;
- **Camada de saída:** resultados obtidos no treinamento da rede.

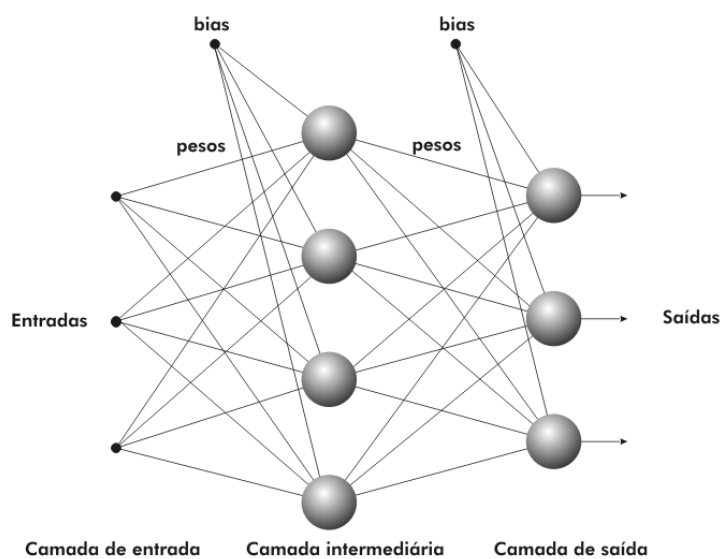


Figura 50 – Camadas de uma rede neural artificial (Salgado, 2009, p. 3).

De acordo com Rezende (2005, p. 144), o aprendizado de uma RNA consiste no ajuste sucessivo dos pesos $w_i \in W$, de modo que a superfície de decisão atenda aos critérios de parada.

9.1.2 Processo de Aprendizado

O processo de aprendizado envolve atualização da arquitetura da rede e dos pesos das conexões (w_i), de maneira que a rede pode efetuar tarefas específicas de

classificação e análise de agrupamentos. As RNAs possuem a capacidade de se adaptarem a problemas específicos, utilizando diferentes tipos de aprendizagem:

- **Aprendizagem Supervisionada:** as saídas desejadas são fornecidas junto com os padrões para as entradas. As saídas são supervisionadas por um agente humano, que coleta os dados e faz uma análise desses. O algoritmo de treinamento verifica o erro para cada padrão e altera os pesos da rede, a fim de minimizá-lo;
- **Aprendizagem Não Supervisionada:** a rede recebe somente os padrões como entrada. Dessa forma, o objetivo é classificá-los de acordo com suas características comuns, não obedecendo a requisitos pré-estabelecidos. Os pesos da rede são alterados até esta atingir um ponto de estabilidade;
- **Aprendizagem Por Reforço:** é uma mescla entre o treinamento supervisionado e o não supervisionado. Em lugar do supervisor, há um crítico no qual retorna uma penalidade ou um sinal de reforço para última ação da rede.

"O algoritmo de aprendizado deve ser responsável pela adaptação dos parâmetros da rede, de maneira que, em um número finito de iterações do algoritmo, haja convergência para uma solução." (Rezende, 2005, p. 148).

O processo de aprendizado tem como objetivo ajustar os pesos (w_i), de modo a causar uma mudança no comportamento da rede resultando em uma melhoria no seu desempenho. Esses ajustes são utilizados para que a convergência seja atingida pelo algoritmo de treinamento (Rezende, 2005, p. 149).

9.1.3 Algoritmo de Treinamento

O algoritmo de treinamento tem o objetivo de atualizar e corrigir os pesos das redes neurais. Seu processo é feito através de funções matemáticas que modificam

os valores dos pesos e essa atualização é realizada diversas vezes, buscando a convergência da solução.

"O treinamento de redes neurais com várias camadas pode ser entendido como um caso especial de aproximação de funções, onde não é levado em consideração nenhum modelo explícito dos dados." (Shepherd, 1995) *apud* (Silva L. N., 1998, p. 48).

Diversos algoritmos de treinamento foram desenvolvidos. Eles se diferem principalmente por como os pesos são atualizados, sendo os mais conhecidos:

- Backpropagation (BP);
- Backpropagation com Termo *Momentum* (BPM);
- Método Gradiente (GRAD);
- Método Fletcher-Reeves (FR);
- Método Polak-Ribière (PR);
- Método Davidon-Fletcher-Powell (DFP), e entre outros.

Diante de tais informações e do poder de aprendizagem das redes neurais, diversos problemas foram atrelados às redes neurais, visto o fato desta ser capaz de resolver diferentes categorias de problemas, modificando apenas alguns parâmetros e configurações.

9.1.4 Aplicação das Redes

A utilidade das redes neurais está no fato delas serem utilizadas para resolver diversos problemas que apresentam um alto grau de complexidade, uma vez que as redes são capazes de serem adaptadas a diferentes problemas.

As RNAs são aplicadas na identificação e controle de sistemas, jogo de interpretação, tomada de decisão, reconhecimento de padrões, reconhecimento de seqüências, diagnóstico médico, aplicações financeiras, de mineração de dados

entre outras áreas. Os problemas podem ser divididos nas seguintes categorias, como descreve (Rezende, 2005, p. 145):

- **Predição:** consiste na previsão de situações futuras de um sistema, baseando em informações dos estados anteriores. Sua aplicação é utilizada para prever cotações financeiras e também séries temporais;
- **Classificação:** a partir de um conjunto de classes conhecidas, é atribuída uma classe correspondente ao padrão testado. São passados para o treinamento da rede os valores de entrada (características do problema) assim como as classes definidas (resultados obtidos). Sua utilização é encontrada em problemas de reconhecimento de padrões;
- **Categorização:** tem a finalidade de dividir e agrupar diferentes conjuntos de dados de acordo com suas características. Desse modo, somente os valores de entrada são conhecidos, cabendo à rede classificar os dados em grupos de padrões.