DESENVOLVIMENTO DE MODELOS NEURAIIS PARA O PROCESSAMENTO DE
SINAIS ACÚSTICOS VISANDO A MEDIÇÃO DE PROPRIEDADES TOPOLÓGICAS
EM ESCOAMENTOS MULTIFÁSICOS

ÉRICA REGINA FILLETTI NASCIMENTO

*Tese apresentada à Escola de Engenharia de
São Carlos, da Universidade de São Paulo,
como parte dos requisitos para obtenção do
título de Doutor em Engenharia Mecânica.*

Orientador: PROF. ASSOC. PAULO SELEGHIM JÚNIOR

*Ao meu marido Marcelo
e ao meu filho Lucas,
com amor.*

*“Se procurar bem, você acaba encontrando
não a explicação (duvidosa) da vida,
mas a poesia (inexplicável) da vida”.*

(Carlos Drummond de Andrade)

Agradecimentos^{*}

Agradeço a Deus por suas bênçãos.

Ao Prof. Paulo Seleglim Júnior, pelo apoio e orientação durante este trabalho.

Ao Marcelo, meu marido, por todo amor, carinho, força e incentivo que me deu em todos os momentos que passamos juntos.

Ao meu filho Lucas, a maior bênção que eu já recebi, pela imensa alegria que trouxe à minha vida.

Aos meus pais Nelson e Neusa, pelo amor incondicional em todos os momentos e aos meus irmãos Evandro, Elis e Éllen pela amizade e companheirismo.

Aos colegas de curso, funcionários e docentes do Departamento de Engenharia Mecânica da EESC-USP.

De modo especial, agradeço ainda à toda minha família e a todos os amigos que, direta ou indiretamente, contribuíram para a realização deste trabalho.

^{*}Este trabalho foi parcialmente financiado pelo PROSET CT-Petro CNPq 01/2002 (processo n^o 500101/2002-7)

Resumo

FILLETTI, E. R. (2007). *Desenvolvimento de modelos neurais para o processamento de sinais acústicos visando a medição de propriedades topológicas em escoamentos multifásicos*. Tese (Doutorado) - Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2007.

Uma nova metodologia para a medida não intrusiva da fração volumétrica e da área interfacial é proposta neste trabalho, com base em redes neurais para processar respostas obtidas de sinais acústicos. A distribuição geométrica das fases dentro do escoamento é mapeada pela velocidade local de propagação acústica, considerada na equação diferencial que governa o problema. Esta equação é resolvida numericamente pelo método de diferenças finitas com as condições de contorno reproduzindo a estratégia de pulso/eco. Um número significativo de distribuições das velocidades de propagação foi considerado na solução da equação diferencial para construir uma base de dados, da qual os parâmetros da rede podem ser ajustados. Especificamente, o modelo neural é construído para mapear características extraídas dos sinais obtidos de quatro sensores acústicos, localizados no contorno externo do domínio de sensoriamento, estimando a fração volumétrica e a área interfacial correspondentes. Estas características correspondem às amplitudes e aos tempos de chegada dos três maiores picos da onda acústica. Os resultados numéricos mostram que o modelo neural pode ser treinado em um tempo computacional razoável e é capaz de estimar os valores da fração volumétrica e da área interfacial dos exemplos do conjunto de teste.

Palavras-chave: escoamento bifásico, monitoração acústica, fração volumétrica, área interfacial, redes neurais artificiais.

Abstract

FILLETTI, E. R. (2007). *Development of neural models for the processing of acoustic signals aiming at the measurement of topological properties in multi-phase flow*. Ph.D. Thesis - Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2007.

A new methodology for measuring the volumetric fraction and interfacial area in two-phase flows is proposed in this work, based on neural network for processing the responses obtained from an acoustic interrogation signal. The geometrical distribution of the phases within the flow is mapped by the local acoustic propagation velocity which is considered in the governing differential equation. This equation is solved numerically by the finite difference method with boundary conditions reproducing the pulse/echo strategy. A significant number of propagation velocities distributions were considered in the solution of the differential equation in order to construct a database from which the neural model parameters could be adjusted. Specifically, the neural model is constructed to map the features extracted from the signals delivered by four acoustic sensors, placed on the external boundary of the sensing domain, into the corresponding volumetric fraction and interfacial area. These features correspond to the amplitudes and the times of arrival on the three first peaks of the acoustic wave. Numerical results showed that the neural model can be trained in a reasonable computational time and it is capable of estimating the values of the volumetric fraction and the interfacial area of examples of the set of test.

Keywords: two-phase flow, acoustic monitoring, volumetric fraction, interfacial area, artificial neural network.

Lista de Figuras

3.1	Esboço do sensor proposto.	23
3.2	Posição dos elementos de medida na geometria quadrada.	24
3.3	Posição dos elementos de medida na geometria circular.	24
3.4	Sinal acústico e as respectivas características extraídas.	25
3.5	Geração aleatória de gotas.	27
3.6	Célula computacional da geometria quadrada.	32
3.7	Células computacionais relativas: (a) ao canto inferior esquerdo e (b) à borda inferior do quadrado.	33
3.8	Células computacionais referentes: (a) ao canto inferior direito e (b) à borda lateral esquerda do quadrado.	34
3.9	Células computacionais referentes: (a) ao interior e (b) à borda lateral direita do quadrado.	34
3.10	Células computacionais relativas: (a) ao canto superior esquerdo e (b) à borda superior do quadrado.	35
3.11	Célula computacional concernente ao canto superior direito do qua- drado.	35
3.12	Célula computacional da geometria circular.	38
3.13	Malha em coordenadas polares.	41
3.14	Célula computacional relativa ao ponto central do círculo.	41
3.15	Célula computacional relativa aos pontos pertencentes ao semi-eixo positivo das abscissas que não têm intersecção com o círculo. . . .	42
3.16	Célula computacional relativa ao ponto de intersecção do semi-eixo positivo das abscissas com o círculo.	43

3.17 Célula computacional relativa aos pontos do interior do círculo para $i = 1, \dots, N - 1$ e $j = 2, \dots, M - 1$	43
3.18 Célula computacional concernente aos pontos do contorno do círculo para $j = 2, \dots, M - 1$	44
3.19 Célula computacional relativa aos pontos do interior do círculo para $i = 1, \dots, N - 1$ e $j = M$	44
3.20 Célula computacional relativa ao ponto do círculo onde $i = N, j = M$	45
4.1 Neurônio de McCulloch e Pitts.	50
4.2 Modelo de neurônio não-linear.	51
4.3 Exemplos de funções de ativação: (a) Degrau, (b) Linear por Partes e (c) Sigmóide.	53
4.4 Rede neural <i>feed-forward</i> com uma única camada.	55
4.5 Rede neural <i>feed-forward</i> com múltiplas camadas.	55
4.6 Rede neural recorrente.	56
4.7 Rede neural 5 - 5 - 4 - 3.	58
4.8 Ilustração da propagação para frente dos sinais funcionais e da retropropagação dos sinais de erro em um MLP.	59
4.9 Função de ativação utilizada nas camadas intermediárias das redes implementadas neste trabalho.	73
4.10 Função de ativação utilizada na camada de saída das redes implementadas neste trabalho.	74
5.1 Simulação numérica da propagação da energia acústica em uma seção quadrada com uma inclusão localizada no centro do domínio.	76
5.2 Simulação numérica da propagação da energia acústica em uma seção circular com uma inclusão à esquerda do centro do domínio.	76
5.3 Malha utilizada no domínio quadrado.	77
5.4 Sinais acústicos obtidos nos testes da geometria quadrada.	78
5.5 Sinais acústicos filtrados obtidos nos testes da geometria quadrada.	79

5.6	Histogramas da maior amplitude dos sinais de pressão acústica obtidos nos elementos (a) 1, (b) 2, (c) 3 e (d) 4.	82
5.7	Histogramas do tempo de chegada da onda acústica dos sinais de pressão acústica obtidos nos elementos (a) 1, (b) 2, (c) 3 e (d) 4.	82
5.8	Malha utilizada no domínio circular.	83
5.9	Sinais acústicos obtidos nos testes da geometria circular.	84
5.10	Sinais acústicos filtrados obtidos nos testes da geometria circular.	85
5.11	Decaimento do erro durante o treinamento da rede para estimar a fração volumétrica e a área interfacial para os dados obtidos na geometria quadrada.	89
5.12	Histogramas da fração volumétrica e da área interfacial calculadas para os exemplos de treinamento e de teste, respectivamente, na geometria quadrada.	91
5.13	Comparação entre os valores estimados pela rede neural e os valores reais para a fração volumétrica (geometria quadrada).	92
5.14	Intervalo de confiança dos resultados fornecidos pela RNA para os valores da fração volumétrica na geometria quadrada.	94
5.15	Comparação entre os valores estimados pela rede neural e os valores reais para a área interfacial (geometria quadrada).	95
5.16	Intervalo de confiança dos resultados fornecidos pela RNA para os valores da área interfacial na geometria quadrada.	96
5.17	Decaimento do erro durante o treinamento da rede para estimar a fração volumétrica e a área interfacial para os dados obtidos na geometria circular.	97
5.18	Histogramas da fração volumétrica e da área interfacial calculadas para os exemplos de treinamento e de teste, respectivamente, na geometria circular.	99
5.19	Comparação entre os valores estimados pela rede neural e os valores reais para a fração volumétrica (geometria circular).	99
5.20	Intervalo de confiança dos resultados fornecidos pela RNA para os valores da fração volumétrica na geometria circular.	100

5.21	Comparação entre os valores estimados pela rede neural e os valores reais para a área interfacial (geometria circular).	102
5.22	Intervalo de confiança dos resultados fornecidos pela RNA para os valores da área interfacial na geometria circular.	102

Lista de Tabelas

5.1	Características extraídas dos sinais de resposta para diferentes testes numéricos em uma tubulação quadrada.	81
5.2	Características extraídas dos sinais de resposta para diferentes testes numéricos em uma tubulação circular.	87
5.3	Respostas da rede neural para a fração volumétrica em alguns testes realizados na geometria quadrada.	90
5.4	Resposta da rede neural para a área interfacial em alguns testes realizados na geometria quadrada.	90
5.5	Resposta da rede neural para a fração volumétrica no caso da geometria circular.	97
5.6	Respostas da rede neural para a área interfacial no caso da geometria circular.	98

Lista de Símbolos

Formulação do Problema

- A - coeficiente da equação discretizada em coordenadas cartesianas.
- a - coeficiente da equação discretizada em coordenadas polares.
- A_1 - maior amplitude do sinal de pressão acústica.
- A_2 - segunda maior amplitude do sinal de pressão acústica.
- A_3 - terceira maior amplitude do sinal de pressão acústica.
- $A_i(t)$ - área instantânea de todas as gotas no domínio de simulação.
- $\tilde{A}$ - matriz do sistema linear da geometria quadrada.
- $\widetilde{A}_1$ - matriz do sistema linear da geometria circular.
- $\hat{A}$ - matriz pré-condicionada.
- $\overline{AB}$ - segmento de reta.
- B - coeficiente da equação discretizada em coordenadas cartesianas.
- b - coeficiente da equação discretizada em coordenadas polares.
- $\tilde{b}$ - vetor independente do sistema linear na caso da geometria quadrada.
- $\tilde{b}_1$ - vetor independente do sistema linear na caso da geometria circular.
- C - coeficiente da equação discretizada em coordenadas cartesianas.
- c - coeficiente da equação discretizada em coordenadas polares.
- $\overline{CD}$ - segmento de reta.
- $c(\vec{x})$ - velocidade local de propagação acústica.
- c_0 - velocidade de referência.
- $c_{\text{água}}$ - velocidade de propagação acústica na água.
- $c_{\text{óleo}}$ - velocidade de propagação acústica no óleo.
- $c_{i,j}$ - velocidade local de propagação acústica no ponto (i, j) da malha.

- D - coeficiente da equação discretizada em coordenadas cartesianas.
- d - coeficiente da equação discretizada em coordenadas polares.
- E - coeficiente da equação discretizada em coordenadas cartesianas.
- e - coeficiente da equação discretizada em coordenadas polares.
- F - coeficiente da equação discretizada em coordenadas cartesianas.
- f - coeficiente da equação discretizada em coordenadas polares.
- G - coeficiente da equação discretizada em coordenadas cartesianas.
- g - coeficiente da equação discretizada em coordenadas polares.
- i - coordenada nas direções x e r das malhas cartesiana e polar, respectivamente.
- j - coordenada nas direções y e θ das malhas cartesiana e polar, respectivamente.
- L_{exc} - elemento responsável pela excitação.
- L_k - k -ésimo elemento de medida, $k = 1, \dots, 4$.
- L_x - comprimento do domínio de sensoramento na direção x .
- L_y - comprimento do domínio de sensoramento na direção y .
- M - número de pontos das malhas cartesiana e polar, nos intervalos $[0, L_y]$ e $[0, 2\pi]$, respectivamente.
- N - número de pontos das malhas cartesiana e polar, nos intervalos $[0, L_x]$ e $[0, R]$, respectivamente.
- $\vec{n}$ - vetor normal.
- P - pressão acústica.
- P^* - pressão adimensional.
- P_0 - pressão de referência
- $P_i(t)$ - perímetro instantâneo de todas as gotas no domínio de simulação.
- $P_{i,j}$ - pressão acústica no ponto (i,j) da malha.
- $P_{i,j}^k$ - pressão acústica no ponto (i,j) da malha, no instante k .
- $P_{i,j}^{k-1}$ - pressão acústica no ponto (i,j) da malha, no instante $k - 1$.
- $P_{i,j}^{k-2}$ - pressão acústica no ponto (i,j) da malha, no instante $k - 2$.
- r^* - raio adimensional.
- R - raio do domínio circular.
- $R_k(t)$ - sinal de resposta obtido no k -ésimo elemento de medida.
- s - coordenada curvilínea que percorre o contorno do domínio do problema.

t - tempo.

t_0 - tempo de referência.

t^* - tempo adimensional.

T - tempo de chegada da onda acústica.

T_1 - tempo de chegada de A_1 .

T_2 - tempo de chegada de A_2 .

T_3 - tempo de chegada de A_3 .

V - volume de controle.

$V_b(t)$ - volume ocupado pelas gotas no volume V .

x^* - variável adimensional.

x^k - vetor de soluções no instante k .

x^{k-1} - vetor de soluções no instante $k - 1$.

x^{k-2} - vetor de soluções no instante $k - 2$.

y^* - variável adimensional.

$\alpha(t)$ - fração volumétrica.

Δr - tamanho do intervalo da malha na direção r .

Δt - tamanho do intervalo de tempo.

Δx - tamanho do intervalo da malha na direção x .

Δy - tamanho do intervalo da malha na direção y .

$\Delta \theta$ - tamanho do intervalo da malha na direção θ .

$\Gamma(t)$ - área interfacial volumétrica.

λ - constante que aparece na equação adimensionalizada em coordenadas cartesianas.

Ω - domínio do problema.

$\partial\Omega$ - contorno do domínio do problema.

Conceitos Básicos de Redes Neurais Artificiais

A_1 - maior amplitude do sinal de pressão acústica.

A_2 - segunda maior amplitude do sinal de pressão acústica.

- A_3 - terceira maior amplitude do sinal de pressão acústica.
- C - conjunto de todos os neurônios da camada de saída da rede neural.
- $d_j(n)$ - resposta desejada para o neurônio j , na iteração n .
- $d_k(n)$ - resposta desejada para o neurônio k , na iteração n .
- $e_j(n)$ - sinal de erro na saída do neurônio j , na iteração n .
- $e_k(n)$ - sinal de erro na saída do neurônio k , na iteração n .
- $E(n)$ - erro total.
- E_{med} - erro médio quadrático.
- l - l -ésima camada de neurônios da rede neural.
- m - número total de entradas (excluindo o *bias*) aplicadas ao neurônio j .
- n - iteração onde o n -ésimo padrão de treinamento é apresentado à rede neural.
- N - número total de exemplos contidos no conjunto de treinamento.
- p - número de entradas do neurônio de McCulloch e Pitts e do neurônio não linear.
- R - número de entradas de uma rede neural de múltiplas camadas.
- S^i - número de neurônios em cada uma das camadas intermediárias e de saída.
- T - tempo de chegada da onda acústica.
- T_1 - tempo de chegada de A_1 .
- T_2 - tempo de chegada de A_2 .
- T_3 - tempo de chegada de A_3 .
- u_k - saída do somador devido aos sinais de entrada.
- v - potencial de ativação.
- $v_j(n)$ - potencial de ativação do neurônio j , na iteração n .
- $v_k(n)$ - potencial de ativação do neurônio k , na iteração n .
- w_j - peso sináptico associado à entrada x_j .
- w_{j0} - *bias* aplicado ao neurônio j .
- w_{k0} - *bias* aplicado ao neurônio k .
- $w_{ji}(n)$ - peso sináptico conectando a saída do neurônio i à entrada do neurônio j , na iteração n .
- $w_{kj}(n)$ - peso sináptico conectando a saída do neurônio j à entrada do neurônio k , na iteração n .
- x_j - j -ésimo sinal de entrada do neurônio.

y - sinal de saída do neurônio.

y_0 - entrada correspondente ao *bias*.

$y_i(n)$ - sinal funcional na saída do neurônio i , na iteração n .

$y_j(n)$ - sinal funcional na saída do neurônio j , na iteração n .

$y_k(n)$ - sinal funcional na saída do neurônio k , na iteração n .

α - constante de momento.

$\delta_j(n)$ - gradiente local do neurônio j , na iteração n .

$\delta_k(n)$ - gradiente local do neurônio k , na iteração n .

$\Delta w_{ji}(n)$ - correção aplicada ao peso sináptico $w_{ji}(n)$.

η - parâmetro da taxa de aprendizado.

$\varphi(\cdot)$ - função de ativação.

$\varphi_j(\cdot)$ - função de ativação associada ao neurônio j .

$\varphi_k(\cdot)$ - função de ativação associada ao neurônio k .

θ - *threshold*, *bias*.

$\theta_k(n)$ - *bias* aplicado ao neurônio k , na iteração n .

Resultados

A_1 - maior amplitude do sinal de pressão acústica.

A_2 - segunda maior amplitude do sinal de pressão acústica.

A_3 - terceira maior amplitude do sinal de pressão acústica.

b - intersecção da reta $y = mx + b$ com o eixo y .

e_i - erro local da fração volumétrica e da área interfacial.

i - coordenada nas direções x e r das malhas cartesiana e polar, respectivamente.

j - coordenada nas direções y e θ das malhas cartesiana e polar, respectivamente.

K - número total de pares (x_i, y_i) .

M - número de pontos das malhas cartesiana e polar, nas direções y e θ , respectivamente.

m - grau de inclinação da reta $y = mx + b$.

N - número de pontos das malhas cartesiana e polar, nas direções x e r , respec-

tivamente.

R - raio da geometria circular.

r - coeficiente de correlação.

T - tempo de chegada da onda acústica.

T_1 - tempo de chegada de A_1 .

T_2 - tempo de chegada de A_2 .

T_3 - tempo de chegada de A_3 .

x_i - valor real da fração volumétrica e da área interfacial.

y_i - valor estimado da fração volumétrica e da área interfacial pela RNA.

α - fração volumétrica.

Γ - área interfacial.

σ_e - desvio padrão de e_i .

Lista de Siglas

ADALINE - *Adaptative Linear Element*
CNPq - Conselho Nacional de Desenvolvimento Científico e Tecnológico
EESC - Escola de Engenharia de São Carlos
ERT - Tomografia de Resistência Elétrica
FAPESP - Fundação de Amparo à Pesquisa do Estado de São Paulo
FEM - Faculdade de Engenharia Mecânica
MADALINE - *Multiple Adaline*
MATLAB - *Matrix Laboratory*
MCP - Neurônio de McCulloch e Pitts
MDF - Método de Diferenças Finitas
MLP - *Multilayer Perceptrons*
NETeF - Núcleo de Engenharia Térmica e Fluidos
RNAs - Redes Neurais Artificiais
SOM - *Self-Organizing Maps*
PBCG - Método dos Gradientes Bi-Conjugados Pré-Condicionado
PDP - Sistema de Processamento Paralelo Distribuído
TOFD - *Time of Flight Diffraction*
TSVD - *Truncated Singular Value Decomposition*
UCT - Tomografia Computadorizada por Ultra-som
UNICAMP - Universidade Estadual de Campinas
USP - Universidade de São Paulo
VOF - *Volume of Fluid*

Sumário

Resumo	iii
Abstract	v
1 Introdução	1
2 Revisão Bibliográfica	7
2.1 Escoamentos Multifásicos	7
2.2 Ultra-som	11
2.2.1 Histórico	11
2.2.2 Pesquisas Recentes Envolvendo Ultra-Som	12
2.3 Redes Neurais Artificiais	15
2.3.1 Histórico	15
2.3.2 Novos Trabalhos Envolvendo Redes Neurais	18
3 Formulação do Problema	21
3.1 Propagação Acústica em Meios Heterogêneos	21
3.1.1 Sinais Acústicos	23
3.1.2 Fração Volumétrica e Área Interfacial	25
3.2 Construção do Modelo Numérico	27
3.2.1 Adimensionalização	27
3.2.2 Discretização da Equação Diferencial Parcial	30
3.2.3 Armazenamento	46
3.2.4 Resolução do Sistema Linear	47

4	Conceitos Básicos de Redes Neurais Artificiais	49
4.1	O que é uma Rede Neural Artificial?	49
4.2	Modelos Matemáticos de Neurônios	50
4.3	Funções de Ativação	53
4.4	Arquitetura de Redes Neurais	54
4.5	Tipos de Aprendizagem	56
4.6	<i>Perceptrons</i> de Múltiplas Camadas	58
4.7	O Algoritmo de Retropropagação	60
4.7.1	O Neurônio j é um Nó da Camada de Saída	64
4.7.2	O Neurônio j é um Nó em uma Camada Intermediária	64
4.8	Treinamento por Lote	66
4.9	A Taxa de Aprendizagem	68
4.10	Conjunto de Teste e Generalização	68
4.11	Normalização dos Dados	70
4.11.1	Normalização através da Divisão pelo Máximo Valor Absoluto	70
4.11.2	Normalização através da Média e do Desvio Padrão	71
4.11.3	Análise das Componentes Principais	71
4.12	Modelo Neural	72
5	Resultados	75
5.1	Resultados das Simulações Numéricas	75
5.1.1	Simulações em Tubulação de Seção Quadrada	77
5.1.2	Simulações em Tubulação de Seção Circular	83
5.2	Resultados das Redes Neurais Artificiais	88
5.2.1	Geometria Quadrada	88
5.2.2	Geometria Circular	96
6	Conclusões	105
	Referências Bibliográficas	107
	Índice Remissivo	117

Introdução

A monitoração acústica de processos industriais surge como uma técnica extremamente interessante, sobretudo em aplicações envolvendo o transporte e a manipulação de escoamentos multifásicos. Em específico, tanto podem ser mensurados parâmetros físicos (vazões, densidade, viscosidade, fração volumétrica, temperatura e outros), quanto é possível obter informações mais completas sobre a fenomenologia do processo, normalmente expressas através de imagens montadas diretamente ou reconstruídas numericamente a partir de sinais de resposta acústica (LYNCH, 1991). Dentre as principais vantagens da monitoração acústica (robustez, imunidade a ruídos elétricos, baixo custo e outras), a possibilidade de se fazer o sensoriamento de forma não-intrusiva e não-invasiva é fundamental para algumas aplicações em que não são possíveis intervenções mais diretas. No caso do sensoriamento acústico, a propriedade que o ultra-som tem de atravessar diferentes materiais permite que monitoração seja realizada por simples contato com as paredes da tubulação.

Este trabalho vincula-se a projetos de relevância, não somente para a indústria petroquímica, mas também para a segurança e a preservação de recursos naturais. Em particular, este trabalho está vinculado aos projetos "Instrumentação e controle de escoamentos multifásicos industriais", em curso no NETeF - EESC - USP (proc. FAPESP 03/13819-6) e "Produção e transporte multifásico de óleos pesados", em curso na FEM - UNICAMP sob a coordenação do professor A. C. Bannwart (proc. CNPq/CT-Petro 360654/02-9).

Isso posto, é possível concluir que o presente trabalho tem grande importância

nestas pesquisas, uma vez que um dos seus objetivos é o desenvolvimento de uma ferramenta numérica para o monitoramento acústico de escoamentos multifásicos. Em específico, deseja-se determinar e quantificar como os campos de sensoria-mento acústico se estabelecem, tendo em vista sua interação com o escoamento, assim como com a geometria da tubulação. Este é, de fato, um problema bas-tante importante dada a multiplicidade de interfaces bifásicas entre diferentes impedâncias acústicas, o que gera fenômenos também múltiplos de reflexão, ab-sorção e espalhamento, de forma que a estrutura do sinal obtido como resposta à excitação pode ser bastante complexa e seu processamento extremamente deli-cado.

Tal simulador numérico para estudo da propagação acústica em meios hete-rôneos foi desenvolvido possibilitando a realização de vários testes numéricos com o objetivo de formar uma base de dados para o treinamento de redes neurais artificiais. Especificamente, pretende-se implementar redes neurais que, a partir da distribuição de variáveis acústicas (tempos de chegada da onda acústica, am-plitudes dos sinais acústicos e outros), sejam capazes de fornecer parâmetros do escoamento bifásico, dentre os quais a fração volumétrica e a área interfacial são certamente os mais importantes.

Conforme se sabe da literatura, assim como dos resultados numéricos obtidos neste trabalho, o escoamento de fases com diferentes propriedades acústicas torna extremamente complexa a propagação de ondas acústicas e, conseqüentemente, a interpretação de seus sinais de resposta é igualmente bastante complicada. Esta interpretação pode ser realizada segundo duas abordagens distintas:

- (i) através de reconstrução tomográfica, ou seja, cálculo da distribuição in-terna da impedância acústica a partir de medidas de intensidade acústica (SCHLABERG, PODD e HOYLE 2000);
- (ii) utilização de redes neurais artificiais ajustadas a partir de uma base de dados de treinamento de forma a mapear o conjunto de possíveis medidas de variáveis acústicas sobre os correspondentes parâmetros topológicos do escoamento.

Em ambos os casos trata-se da solução de um problema inverso e intrinsecamente mal condicionado. Portanto, para obter respostas em condições experimentais realistas, quanto ao nível de ruídos por exemplo, é imprescindível o emprego de técnicas de regularização numérica. A abordagem tomográfica, sendo geralmente implementada *off-line*, permite aplicar algumas técnicas baseadas em medidas redundantes (ROLNIK e SELEGHIM, 2002; TICHONOV, 1963) ou TSVD (*Truncated Singular Value Decomposition* - (PRESS *et al.*, 1992)). Já a utilização de redes neurais tem o mal condicionamento associado ao ajuste de seus parâmetros internos, podendo então ser feito *a priori*, ainda que isso demande métodos de regularização semelhantes. Conseqüentemente, uma vez treinada, a rede neural "aprende" a regularizar o problema inverso e pode então ser implementada em tempo real.

O objetivo principal deste trabalho é verificar a possibilidade, através de simulações numéricas, de desenvolver uma metodologia para a medida não intrusiva da fração volumétrica e da área interfacial de escoamentos bifásicos líquido-líquido, utilizando sensoramento acústico e resolvendo o problema inverso associado. Neste sentido, o problema foi investigado através do desenvolvimento e treinamento de redes neurais para a solução do problema inverso.

As redes são treinadas com dados sintéticos oriundos de simulações numéricas. Para a realização de tais simulações, foi desenvolvida uma ferramenta numérica capaz de calcular a propagação acústica em meios heterogêneos. Este simulador foi utilizado para construir uma base de dados constituída de distribuições de contraste acústico * geradas aleatoriamente. Para cada teste numérico realizado, foram obtidas a fração volumétrica e a área interfacial, além de quatro sinais acústicos, dos quais extraiu-se os tempos de chegada dos pulsos acústicos, a amplitude dos três maiores picos de cada sinal, bem como o tempo de chegada destes picos. Após concluída esta etapa, realizou-se a implementação de redes neurais e estudos visando obter arquiteturas otimizadas relativamente ao problema tratado. Em seguida, fez-se o treinamento dos melhores modelos com exemplos extraídos da base de dados, buscando determinação de parâmetros como a fração volumétrica e

* Razão entre as impedâncias de propagação acústica.

a área interfacial. Em particular, verificou-se que estes modelos são treináveis em um tempo computacional razoável e os intervalos de confiança para as respostas das redes neurais são compatíveis com futuras aplicações práticas.

Sendo assim, as seguintes metas intermediárias podem ser enumeradas:

- (i) Construção de um modelo físico de propagação acústica, adaptado às características do escoamento multifásico e da geometria da tubulação, capaz de reproduzir os fenômenos de reflexão, absorção e espalhamento;
- (ii) Desenvolvimento do modelo numérico correspondente tendo como base o MDF (Método de Diferenças Finitas);
- (iii) Implementação numérica desse modelo em plataforma computacional adequada, neste caso, programação direta em Fortran, para cálculos de maior esforço computacional;
- (iv) Construção de um banco de dados de resposta acústica para tubulações de seção quadrada e circular a partir da geração aleatória de inclusões características de escoamentos dispersos (gotas) óleo-água. Em específico, a excitação consiste de impulsos (Diracs) e a resposta, determinada pelo simulador, é constituída das distribuições dos respectivos tempos de chegada da onda acústica, além das amplitudes dos três maiores picos do sinal e dos tempos de chegada destes picos;
- (v) Implementação de redes neurais artificiais visando a otimização de suas arquiteturas, principalmente em termos do números de camadas e de neurônios em cada uma delas;
- (vi) Treinamento dos modelos com exemplos extraídos da base de dados buscando determinação de importantes parâmetros do escoamento, no caso, a fração volumétrica e a área interfacial.

O presente texto está organizado da seguinte forma:

- (i) No capítulo 2, encontra-se a revisão bibliográfica dos assuntos envolvidos neste trabalho: escoamentos multifásicos, ultra-som e redes neurais artificiais. São expostos os principais acontecimentos que marcaram o progresso do ultra-som e o avanço das pesquisas na área de redes neurais artificiais.
- (ii) O capítulo 3 está dividido em duas seções: na primeira, é apresentada a formulação matemática do problema de propagação acústica, incluindo as interações com as paredes da tubulação e as diferentes fases do escoamento. Na segunda seção, é descrita a construção do modelo numérico para o monitoramento acústico de escoamento bifásico em meios heterogêneos.
- (iii) A finalidade do capítulo 4 é apresentar os conceitos básicos de redes neurais artificiais, tais como funções de ativação, arquiteturas, aprendizagem, treinamento e generalização. Este capítulo traz ainda, uma descrição detalhada do algoritmo de retropropagação do erro, utilizando a regra da cadeia do cálculo.
- (iv) No capítulo 5 são expostos os resultados das simulações numéricas realizadas a partir da ferramenta construída para o monitoramento acústico de escoamentos multifásicos. Também são apresentados os resultados do treinamento e da generalização das redes neurais.
- (v) Por fim, no capítulo 6 são discutidas as conclusões obtidas no presente trabalho.

Revisão Bibliográfica

Este capítulo apresenta a revisão bibliográfica dos temas abordados neste trabalho. A revisão foi dividida da seguinte forma: escoamentos multifásicos, ultrassom e redes neurais artificiais.

2.1 Escoamentos Multifásicos

Um fluido multifásico pode escoar segundo diversas estruturas topológicas conhecidas como padrões de escoamento ou configurações de escoamento. O surgimento desses padrões é determinado pelas vazões e propriedades físicas de seus componentes, além de parâmetros geométricos, como o diâmetro e a inclinação da tubulação. As duas fases de uma mistura gás-líquido, por exemplo, podem escoar horizontalmente de forma dispersa (regimes a bolhas) ou segregada (regimes anular ou estratificado).

Em 1970, a indústria do petróleo começou a identificar mecanismos básicos para diagnosticar padrões de escoamento. Nesta época, Estados Unidos, França e Japão passaram a investir maciçamente na geração termonuclear de energia, fazendo com que o desenvolvimento de técnicas de instrumentação adequadas para escoamentos multifásicos tivesse um grande impulso.

Escoamentos bifásicos gás-líquido e líquido-líquido estão presentes em diversos setores industriais, incluindo o petroquímico, o metalúrgico, o farmacêutico entre outros. Escoamento trifásicos também são muito encontrados na prática,

mas na literatura as referências são escassas. Uma das principais dificuldades para a evolução das técnicas de produção e transporte dos líquidos está relacionada à medida de vazão em escoamentos multifásicos. Atualmente há uma falta de sistemas de medidas que combinam características simples com baixo custo, confiabilidade e segurança. Com base nisto, Almeida, Castro e Rodriguez (2005) propuseram uma sonda capacitiva não-intrusiva e não-invasiva capaz de medir a fração volumétrica em escoamentos bifásicos água-óleo e ar-água, como ponto de partida para a medição em escoamentos trifásicos líquido-líquido-gás.

Um dos parâmetros mais importantes para descrever regimes de escoamento bifásico é, sem dúvida, a fração volumétrica. A importância da medição desse parâmetro apresenta um crescimento significativo em muitos processos industriais, tais como físicos, químicos e petrolíferos. Por esta razão, diversas técnicas de medida da fração volumétrica têm sido desenvolvidas por pesquisadores nos últimos anos. Alguns trabalhos nesta área são citados a seguir:

- (i) Dong *et al.* (2003) empregaram um sistema de tomografia de resistência elétrica (ERT) para calcular a fração volumétrica em escoamentos bifásicos. Os resultados obtidos indicaram que o método é factível para diferentes condições de escoamento e demonstraram que é possível um monitoramento, em tempo real, de processos de transportes de escoamentos bifásicos ar-água a partir do sistema de ERT;
- (ii) Huang, Wang e Li (2003) basearam-se na técnica de tomografia por capacitância elétrica para o desenvolvimento de um novo método de medição de fração volumétrica para escoamentos bifásicos. Os resultados experimentais mostraram uma precisão satisfatória e concluiu-se que o método proposto é apropriado para medir a fração volumétrica em diversos tipos de escoamentos bifásicos;
- (iii) Daidzic *et al.* (2005) estimaram a fração volumétrica em escoamentos bifásicos gás-líquido em tubos verticais através de um sistema de imageamento por ressonância magnética, concluindo que esta é uma técnica não-invasiva e não-intrusiva promissora para a caracterização de escoamentos bifásicos;

- (iv) Prasser, Misawa e Tiseanu (2005) mediram a fração volumétrica utilizando um sensor de fio elétrico e também através de tomografia de raio-x. Os resultados obtidos pelas duas técnicas foram satisfatórios, sendo que o sensor de fio elétrico apresenta melhor resolução no tempo e no espaço e a tomografia possui a vantagem de ser um método não-invasivo.

Outro parâmetro significativo dos escoamento bifásicos é a área interfacial volumétrica, muito utilizada nos modelos com transferência interfacial, e importante para descrever a estrutura geométrica dos escoamento bifásicos. Alguns trabalhos relacionados à área interfacial são mencionados a seguir:

- (i) Ishii e Kim (2001) mediram a área interfacial de escoamentos bifásicos a bolhas transportados em tubulações circulares, utilizando uma sonda com quatro micro sensores e compararam os resultados com a equação

$$\text{Transferência interfacial} = \Gamma(t) \times \text{força motriz},$$

onde $\Gamma(t)$ é a área interfacial, verificando que os resultados obtidos com a sonda possuem um erro de $\pm 10\%$;

- (ii) Euh *et al.* (2001) desenvolveram um método de medida da área interfacial em escoamentos bifásicos utilizando uma sonda de condutividade com cinco sensores. Neste trabalho, foi descrito o modelo matemático e foram realizadas simulações numéricas para vários formatos de bolhas, afim de avaliar o novo método. Os resultados obtidos com este método são mais precisos quando comparados aos resultados alcançados com uma sonda de condutividade com quatro sensores.
- (iii) Dongjian *et al.* (2005) investigaram a área interfacial experimentalmente usando um sistema composto por uma câmera digital de alta velocidade e por uma sonda de condutividade com dois sensores, afim de calcular este parâmetro de escoamentos ar-água.

Sobre a utilização de técnicas tomográficas para visualização ou identificação de regimes, Kim *et al.* (2001) empregaram tomografia de impedância elétrica para visualizar escoamentos bifásicos gás-líquido, fazendo um estudo comparativo entre alguns métodos de regularização utilizados na resolução do problema inverso; Lehner e Wirth (1999) utilizaram um sistema de tomografia de raio-x para obter a fração volumétrica de sólido em um reator gás-sólido; Ma *et al.* (2001) aplicaram uma técnica de tomografia de resistência elétrica para monitorar escoamentos bifásicos em uma tubulação horizontal.

A respeito de escoamentos gás-líquido em tubulações horizontais, Crivelaro e Seleglim (2002) empregaram um modelo neural com aprendizado supervisionado para processar os sinais de uma sonda de visualização direta e diagnosticar os regimes de escoamento horizontal gás-líquido, produzidos em um circuito experimental. Para tanto, os sinais produzidos por uma sonda de visualização direta foram utilizados para constituir uma imagem qualitativa do escoamento e, em seguida, foram enviados à rede neural para a identificação dos regimes. As respostas obtidas pela rede foram satisfatórias, alcançando uma porcentagem média de acerto de 82,8%. Al-Sarkhi e Hanratty (2002) analisaram o efeito do diâmetro da tubulação no tamanho das gotas de água em um escoamento gás-líquido anular horizontal. Klein (2004) utilizou a análise tempo-frequência da transformada de Gabor para investigar a existência de sub-regimes em escoamentos intermitentes, comprovando o grande potencial desta técnica.

Com relação aos escoamentos bifásicos líquido-líquido, encontrados particularmente na indústria de petróleo, onde o óleo e a água freqüentemente são produzidos e transportados juntamente, Angeli e Hewitt (2000) identificaram os padrões de escoamento óleo-água em tubos horizontais através de dois métodos, a saber gravação em vídeo com alta velocidade e determinação da fração volumétrica local utilizando uma sonda de impedância de alta frequência. Gao, Gu e Guo (2003) fizeram um estudo numérico sobre escoamento turbulento estratificado bifásico água-óleo em um tubo horizontal usando um modelo VOF.

Sobre escoamento bifásico óleo-água em tubos verticais, Jin *et al.* (2003a) utilizaram as características da velocidade de propagação da onda cinemática

para distinguir os regimes de escoamento e Jin *et al.* (2003b) diagnosticaram as configurações de escoamento pela análise de séries temporais não-lineares.

Neste trabalho, o escoamento de interesse é o escoamento bifásico líquido-líquido, em que a fase discreta (gotas) é óleo e a fase contínua é água.

2.2 Ultra-som

O ultra-som, em sua definição mais simples, refere-se à energia associada a ondas acústicas com uma frequência de 20.000 Hz ou mais (BUTZ e TAUSCHER, 2002). O ouvido humano só é capaz de perceber sons de frequências compreendidas entre 16 Hz e 20.000 Hz, aproximadamente. Desse modo, os ultra-sons não são captados por nosso ouvido, mas são percebidos por alguns animais, como os cães, que ouvem sons de 25.000 Hz, e os morcegos, que chegam a ouvir sons de até 50.000 Hz.

A seguir, encontra-se um breve histórico dos principais acontecimentos que marcaram o progresso da utilização tecnológica do ultra-som.

2.2.1 Histórico

Desde o início do século passado, muitas pesquisas têm sido realizadas no campo de ultra-som. A energia das vibrações ultra-sônicas encontraram aplicações práticas rapidamente em diversas áreas.

Em 1881, Jacques e Pierre Curie deram uma contribuição valiosa para o estudo do ultra-som, quando publicaram os resultados obtidos ao experimentar a aplicação de um campo elétrico sobre cristais de quartzo e turmalina, os quais produziram ondas sonoras de frequências muito altas (DÍAZ, 2005; LEITE, 2005).

De acordo com Blitz (1967), uma das primeiras utilizações do ultra-som verificou-se em 1883, quando Galton mediu o limite das frequências de sons audíveis pelo ser humano.

Em abril de 1912, pouco depois do naufrágio do Titanic, L. F. Richardson sugeriu a utilização de ecos ultra-sônicos para detectar objetos submersos, a fim

de impedir ocorrências futuras deste tipo (WADE, 2000).

A primeira aplicação importante do ultra-som aconteceu durante a Primeira Guerra Mundial (1914-18), quando o físico francês Paul Langevin propôs o uso do ultra-som para detectar submarinos alemães. Ele emitiu pulsos ultra-sônicos na água e mediu o tempo decorrido entre a emissão do pulso e o sinal de eco refletido a partir da embarcação inimiga (BLITZ, 1967).

Em 1917, Paul Langevin e Chilowsky produziram o primeiro gerador piezoelétrico de ultra-som, cujo cristal servia também como receptor e gerava mudanças elétricas ao receber vibrações mecânicas. O aparelho foi utilizado para estudar o fundo marinho, com uma sonda ultra-sônica para avaliar a profundidade dos mares, conforme cita Díaz (2005).

Segundo Blitz (1967), um acontecimento importante durante os anos 30 foi o trabalho pioneiro do cientista russo Sergei Sokolov, que propôs o uso do ultra-som para detectar fendas em metal.

Após a Segunda Guerra Mundial (1939-45) houve um aumento muito grande de aplicações do ultra-som nos mais diversos campos. A descoberta do radar conduziu ao desenvolvimento da técnica do pulso-eco com aplicações em diagnósticos médicos e em testes não-destrutivos de materiais, que surgiram em 1942, com F. A. Firestone (PEREIRA, 1985). Ao mesmo tempo, avanços consideráveis foram alcançados em aplicações de alta intensidade em processos industriais e médicos.

Esta área de desenvolvimento abrange um universo maior do que se imaginava. O conhecimento já conquistado e o potencial de novas possibilidades são inúmeros, indo desde aplicações simples, como apitos ultra-sônicos para "espan-tar" pragas, podendo chegar até à aplicações extremamente complexas como armas ultra-sônicas de última geração (ALTMANN, 2001).

2.2.2 Pesquisas Recentes Envolvendo Ultra-Som

Conforme mencionado na seção anterior, o ultra-som é utilizado em diversas áreas como engenharia, física, química, medicina e navegação entre outras. Com os dispositivos ultra-sônicos existentes atualmente, pode-se medir a densidade de

um líquido (PÜTTMER *et al.* 1998), detectar rachaduras em concreto (BOND, KEPLER e FRANGOPOL 2000), encontrar defeitos em soldas (VEIGA *et al.* 2005), caracterizar alguns tipos de alimentos (LÉTANG 2001), controlar o processo de cristalização da lactose (BUND e PANDIT 2007), entre outros.

O ultra-som pode ser empregado com intensidade bastante alta. As aplicações do ultra-som de alta intensidade têm como propósito, produzir alteração do meio através do qual a onda se propaga. Como exemplo citam-se homogeneização de materiais e solda. Já as aplicações de baixa intensidade têm como objetivo transmitir a energia através de um meio e com isso obter informações do mesmo. Como exemplo dessas aplicações tem-se: ensaio não destrutivo de materiais, medida das propriedades elásticas dos materiais e diagnóstico médico (PÉCORA e GUERISOLI, 2005).

Püttmer *et al.* (1999) utilizaram sensores de densidade ultra-sônicos para determinar o coeficiente de reflexão do ultra-som entre um material de referência com parâmetros conhecidos e o líquido investigado e analisaram os erros devido às camadas finas de depósito na superfície dos sensores. Estes sensores determinam a impedância acústica com a técnica pulso-eco utilizando a amplitude do pulso acústico que foi refletido. Os resultados mostraram que as camadas finas de depósitos na interface de tais sensores têm grande influência nas medidas, afetando menos a medida do tempo (velocidade do som) do que as amplitudes (coeficiente de reflexão).

Henning *et al.* (2000) desenvolveram um sistema ultra-sônico que consiste em um arranjo miniaturizado de multi-sensores para a medida contínua e simultânea de propriedades acústicas de misturas líquidas para o controle de processos. Este sistema de sensores mede a velocidade do som, o coeficiente de impedância acústica, o coeficiente de atenuação e a temperatura. Uma característica particularmente interessante deste sistema é a boa precisão para a medida da densidade dos líquidos. Atualmente, este sistema de sensores é utilizado com sucesso em várias aplicações de processos industriais, por exemplo para medir rapidamente, de forma precisa e confiável as concentrações das substâncias ou para monitorar, de forma contínua, o curso do processo.

O ultra-som também é muito utilizado na área tomográfica, tanto em tomografia médica quanto em tomografia de processos industriais.

A aplicação de tomografia computadorizada por ultra-som (UCT) em problemas industriais é uma técnica importante capaz de reconstruir a distribuição de escoamentos bifásicos gás-líquido em uma seção transversal de um tubo. Xu *et al.* (1997) investigaram a aplicação da UCT para monitorar tais escoamentos, estabelecendo um modelo matemático que descreve as características acústicas dos mesmos, e desenvolveram um algoritmo para a reconstrução da imagem. Xu, L. A. e Xu L. J. (1998) avaliaram o desempenho da tomografia por ultra-som para a identificação de regimes de escoamentos bifásicos gás-líquido e para a medição da fração de vazio em uma seção transversal de um tubo.

A tomografia acústica também pode ser utilizada para reconstruir campos internos de temperatura de um corpo ou escoamento. Filletti (2002) fez um estudo nesse sentido, baseado no fato de que a distribuição interna de impedância acústica depende fortemente da temperatura. Mais especificamente, o problema inverso foi modelado matematicamente a partir da equação de Helmholtz modificada e de um funcional de erro quantificando a sensibilidade dos perfis de pressão acústica externa relativamente a variações na distribuição interna de impedância acústica. Foram testadas duas técnicas de excitação, a clássica, tipo Dirac e uma estratégia otimizada segundo um perfil triangular. Os resultados mostraram que a reconstrução tomográfica segundo a técnica otimizada de sensoriamento tem maiores chances de sucesso em condições práticas de aplicação do que com a estratégia clássica. Fujii e Zhang (2001) realizaram simulações numéricas e testes experimentais na tentativa de reconstruir a distribuição interna de temperatura em materiais sólidos utilizando tomografia computadorizada por ultra-som, comparando dois métodos de reconstrução da distribuição de velocidade do som: o método de convolução e um método de interpolação linear.

Em aplicações médicas, Cloutier, Chen e Durand (2001) realizaram uma avaliação exaustiva do desempenho de técnicas da análise tempo-frequência para aplicações relacionadas à medida da turbulência do fluxo sanguíneo utilizando ultra-som. Malinen, Huttunen e Kaipio (2002) propuseram um controlador de

aquecimento induzido de um tumor cujo objetivo principal é obter a distribuição de temperatura com um campo de ultra-som pré-focalizado, a fim de destruir o tecido cancerígeno.

Muitas pesquisas na área de ultra-som estão ainda em aperfeiçoamento. Várias universidades espalhadas pelo mundo estão empenhadas no desenvolvimento de novas aplicações, na pesquisa de novas topologias construtivas e de novos materiais. Além das universidades, muitos outros grandes centros tecnológicos estão investindo em laboratórios e fornecendo recursos para pesquisas nesta área.

2.3 Redes Neurais Artificiais

As Redes Neurais Artificiais (RNAs), também referenciadas na literatura como sistemas conexionistas ou sistemas de processamento paralelo distribuído (PDP), são modelos computacionais inspirados na estrutura e no funcionamento do cérebro humano.

A principal semelhança entre as RNAs e os circuitos formados no cérebro está no intenso processamento paralelo apresentado por ambos. Uma outra semelhança é que as operações nos dois circuitos não são dependentes de qualquer elemento (neurônio) isolado, mas sim de uma função do conjunto dos elementos. Assim, se um pequeno número de elementos forem danificados, a rede ainda poderia responder com uma mínima degradação de desempenho.

A seguir, encontra-se um breve histórico sobre o avanço das pesquisas na área de redes neurais.

2.3.1 Histórico

Primeiramente identificado pelo cientista Santiago Ramón y Cajal em 1911, o neurônio constitui a base do sistema nervoso (HAYKIN 1999).

Em 1943, Warren McCulloch (psiquiatra e neuroanatomista) e Walter Pitts (matemático) propuseram o primeiro modelo matemático de um neurônio artificial, que pode ser visto como uma simplificação do que já havia sido descoberto

a respeito do neurônio biológico (BRAGA, CARVALHO e LUDERMIR 2000). Este neurônio ficou conhecido como nó MCP (iniciais dos pesquisadores). Contudo, o neurônio desenvolvido por McCulloch e Pitts ficou restrito à descrição de um modelo artificial que representasse, computacionalmente, o neurônio biológico e não o processo de aprendizagem que ocorria no cérebro.

O primeiro trabalho sobre aprendizado foi descrito pelo neuropsicólogo Donald Hebb em 1949 em seu livro *The Organization of Behavior*. Hebb propôs uma teoria para explicar o aprendizado em neurônios biológicos, baseada no reforço das ligações sinápticas entre neurônios excitados.

Em 1958, Rosenblatt (1958), apresentou um modelo de rede denominada *perceptron*, que consistia de uma rede formada por nós MCP acrescidos de pesos sinápticos ajustáveis por uma regra de aprendizado. Em 1962, Rosenblatt demonstrou o teorema de convergência do *perceptron*, mostrando que uma rede *perceptron* sempre converge para problemas linearmente separáveis, ou seja, problemas onde a solução pode ser encontrada dividindo-se o espaço de entrada através de uma reta (BRAGA, CARVALHO e LUDERMIR 2000).

Rosenblatt criou muitas variações do *perceptron*. Uma das mais simples foi uma rede de uma única camada cujos pesos e *bias* poderiam ser treinados para produzir um vetor de saída correto, a partir do vetor de entradas correspondente. A técnica do treinamento usada é chamada de regra de aprendizagem do *perceptron*. O grande interesse gerado pelo *perceptron* foi devido à sua habilidade de generalizar seus dados do treinamento e de aprender a partir de conexões inicialmente aleatórias. *Perceptrons* são apropriados para problemas simples de classificação de padrões. São redes rápidas e de confiança. Além disso, uma compreensão das operações do *perceptron* fornece uma boa base para o entendimento de redes mais complexas (DEMUTH e BEALE, 2000).

Quase em paralelo com Rosenblatt, em 1960, Widrow e Hoff apresentaram uma nova regra de aprendizagem, que ficou conhecida como regra delta, sendo esta utilizada na formulação do modelo neural linear ADALINE (*Adaptive Linear Element*). A diferença entre o *perceptron* e o ADALINE está no procedimento de aprendizagem. Posteriormente, Widrow e seus alunos, em 1962, desenvolveram a

rede MADALINE (*Multiple-Adaline*) que é um modelo multidimensional da rede ADALINE (HAYKIN 1999).

Entretanto, em 1969, Minsky e Pappert apresentaram algumas limitações do *perceptron*, mostrando que este era capaz de resolver apenas uma classe limitada de problemas. Além disso, em seu livro publicado em 1969, Minsky e Pappert afirmaram que mesmo o *perceptron* de múltiplas camadas, possivelmente, não seria capaz de superar as limitações apresentadas pelo *perceptron* de uma única camada (HAYKIN 1999).

Após a publicação de Minsky e Papert, muitos pesquisadores redirecionaram suas pesquisas para outras áreas. Como consequência, a pesquisa no campo das redes neurais ficou estagnada por toda a década de 70 e início da década de 80, restando apenas um pequeno número de pesquisadores contribuindo com a área. Entre eles podem ser destacados: Teuvo Kohonen (memórias associativas e auto-organizadas) na Finlândia, Igor Aleksander (redes neurais sem pesos) na Inglaterra, Kunihiko Fukushima (*cognitron* e *neocognitron*) no Japão e Steven Grossberg (sistemas auto-adaptativos) nos Estados Unidos (BRAGA, CARVALHO e LUDERMIR 2000).

No início da década de 80, a publicação do trabalho desenvolvido pelo pesquisador John Hopfield (1982), utilizando redes recorrentes para construção de memórias associativas, provocou um ressurgimento do interesse pela pesquisa em redes neurais. Este modelo, conhecido como redes de Hopfield, se caracterizou por sua capacidade de associar modelos de entrada com modelos previamente armazenados na memória da rede (HAYKIN 1999).

Em 1982, o pesquisador Teuvo Kohonen apresentou o modelo de rede neural denominado Mapas Auto-Organizáveis (*SOM - Self-Organizing Maps*). Este modelo se diferenciou dos demais por ser inspirado na neurofisiologia do córtex cerebral (KOHONEN 2001).

Porém, apenas no final da década de 80, com o desenvolvimento do algoritmo de retropropagação do erro (*error back-propagation*) pelos pesquisadores Rumelhart, Hinton e Williams (1986), o assunto ganhou credibilidade novamente. Este algoritmo é visto como uma resposta às críticas feitas por Minsky e Papert

nos anos 60 (HAGAN, DEMUTH e BEALE 1996). Redes neurais de múltiplas camadas podem, sem dúvida, resolver problemas complexos. Além disso, houve o desenvolvimento de novos *hardwares* e novas capacidades de processamento, reabrindo-se então uma nova era nas pesquisas em redes neurais artificiais.

2.3.2 Novos Trabalhos Envolvendo Redes Neurais

Atualmente vários artigos têm sido publicados apresentando resultados interessantes envolvendo redes neurais nas mais diversas linhas de pesquisa, como por exemplo, os trabalhos de Pina Filho, Dutra e Lima (2006) que simularam um robô bípede subindo uma rampa com a ajuda de redes neurais para determinar o comportamento das pernas do robô; Parlak *et al.* (2005) sobre a aplicação de modelos neurais com o algoritmo de retropropagação para prever o consumo de combustível e a temperatura de exaustão específicos para um motor diesel para vários tempos de injeção; Kara, Kemaloğlu e Güven (2005) sobre o diagnóstico de obstrução da artéria femoral com a ajuda de uma rede treinada com o algoritmo de Levenberg-Marquart; Ince (2004) sobre como prever parâmetros de rachaduras em concreto com o auxílio de uma rede neural artificial; Gismondi, Almeida e Infantosi (2002) sobre a investigação do uso de redes neurais para modelar a mortalidade infantil, considerando condições sócio-econômicas, demográficas e ambientais; Hilbert e Ostendorf (2001) sobre a modelagem da distribuição da vegetação em diferentes condições climáticas através de modelos neurais.

Como já foi dito anteriormente, o algoritmo de retropropagação (regra delta generalizada) proposto por Rumelhart, Hinton e Williams (1986) fornece uma regra de aprendizagem para redes neurais com múltiplas camadas. Muitas aplicações deste algoritmo têm sido mostradas recentemente. Todavia, foram realizadas poucas pesquisas teóricas sobre a potencialidade das redes de Rumelhart, Hinton e Williams. O trabalho de Funahashi (1989) prova que qualquer mapeamento contínuo pode ser realizado através das redes neurais de Rumelhart, Hinton e Williams com pelo menos uma camada oculta cujas funções de ativação da camada de saída são funções sigmóides. O ponto inicial para provar o caso de uma

camada oculta é uma fórmula integral proposta por Irie e Miyake (1988)* apud Funahashi, 1989, e a partir disto, o caso geral (para qualquer número de camadas ocultas) pode ser provado por indução.

Yang, Kuo e Ling (2003) implementaram redes neurais com algoritmo de retropropagação, utilizando um conjunto de treinamento obtido de resultados numéricos de um modelo teórico a partir das equações de Navier-Stokes, para demonstrar sua utilidade na determinação em tempo real de propriedades de escoamento com alta pressão. Os dados de entrada das redes neurais são a pressão e a distância axial. As redes neurais adotadas predizem as distribuições de fração de vazio e de velocidades do líquido e do gás, em um escoamento bifásico turbulento vertical a bolhas no interior de um tubo circular. Nesta formulação, a fração de vazio é estimada diretamente através das equações que governam o problema. Comparações entre os resultados das redes e os resultados desejados mostraram erros menores que 4.59%.

Yan, Liu Y. e Liu C. (2004) propuseram a identificação de regimes de escoamentos bifásicos, utilizando redes neurais supervisionadas e um sensor de tomografia de capacitância. Dez parâmetros característicos do escoamento são extraídos das medidas de capacitância e introduzidos na rede neural supervisionada para identificar os regimes de escoamento. Resultados das simulações mostraram que o novo método tem uma boa precisão e alta velocidade.

Com o objetivo de aumentar a confiabilidade na classificação de descon- tinuidades detectadas por técnicas de ultra-som, Veiga *et al.* (2005) apresentaram o desenvolvimento e operação de redes neurais artificiais no reconhecimento de padrões de sinais ultra-sônicos obtidos na inspeção de soldas pelas técnicas pulso- eco e TOFD (*Time of Flight Diffraction*) referindo-se a quatro condições da junta soldada: sem defeito, falta de penetração, falta de fusão e porosidade. Foi utilizado o *software* MATLAB para a implementação de uma rede neural *feed-forward* supervisionada com algoritmo de retropropagação. Os sinais ultra-sônicos adquiri- dos pelas duas técnicas foram introduzidos separadamente na rede neural, com

*IRIE, B. and MIYAKE, S., (1988). Capabilities of three-layered Perceptrons, *IEEE Interna- tional Conference on Neural Networks*, v.1, p.641-648.

e sem pré processamento. Os resultados mostraram um desempenho superior a 72% de sucesso para os testes. Apesar do pré processamento dos sinais melhorar consideravelmente o desempenho da classificação para os sinais adquiridos através da técnica TOFD, o mesmo não acontece para os sinais obtidos pela técnica pulso-eco.

O grande número de artigos publicados recentemente comprova a importância destas técnicas em aplicações industriais nas mais diversas áreas.

Formulação do Problema

Neste capítulo, encontra-se a formulação matemática do problema de propagação acústica, incluindo as interações com as paredes da tubulação e as diferentes fases do escoamento. Este capítulo traz ainda, a descrição do simulador numérico construído para monitoramento acústico de escoamento bifásico em meios heterogêneos.

3.1 Propagação Acústica em Meios Heterogêneos

De uma maneira geral, as propriedades físicas de uma mistura multifásica (densidade aparente, viscosidade aparente, área interfacial local, fração volumétrica) influenciam fortemente a interação acústica entre o fluido e suas interfaces, definidas por suas superfícies de contato (tubulação, conexões) (ESBENSEN *et al.*, 1999).

Sendo assim, a propagação de ondas acústicas no interior de um volume de sensoramento Ω pode ser modelada com auxílio da equação (KINSLER e FREY, 1962)

$$\frac{\partial^2 P(t, \vec{x})}{\partial t^2} - c^2(\vec{x}) \nabla^2 P(t, \vec{x}) = 0, \quad \text{para } (t, \vec{x}) \in \mathbb{R}_+ \times \Omega \quad (3.1)$$

onde P é a pressão acústica, $c(\vec{x})$ é a velocidade local do som. É importante enfatizar que a velocidade local do som "mapeia" a distribuição da fase dentro do volume de medida Ω . A interação entre Ω e o ambiente externo se dá através de

seu contorno $\partial\Omega$ e é definida por relações representando as variáveis de excitação. As condições inicial e de contorno que modelam a estratégia de medida são dadas por

$$P(0, \vec{x}) = \begin{cases} 1, & \text{se } \vec{x} \in L_{exc} \\ 0, & \text{se } \vec{x} \in \Omega - L_{exc} \end{cases} \quad (3.2)$$

$$\vec{n} \cdot \vec{\nabla} P(t, \vec{x}) = 0, \quad \text{para } (t, \vec{x}) \in \mathbb{R}_+^* \times \partial\Omega \quad (3.3)$$

L_{exc} na equação (3.2) denota o elemento responsável pela excitação.

Considerando a geometria do problema, sobretudo a disposição dos elementos de medida, justifica-se desprezar as variações axiais de forma a obter uma equação bidimensional. Portanto, explicitando os termos da equação (3.1) para um problema bidimensional, onde $P = P(t, x, y)$, obtém-se

$$\frac{\partial^2 P}{\partial t^2} - c_0^2 c^2(x, y) \left(\frac{\partial^2 P}{\partial x^2} + \frac{\partial^2 P}{\partial y^2} \right) = 0, \quad \text{em } \Omega \quad (3.4)$$

onde P é a pressão acústica, $c(x, y)$ é a velocidade local do som, t é o tempo e c_0 é a velocidade de referência, Ω é o domínio do problema e (x, y) representa um ponto no interior de Ω .

Um esboço do sensor acústico proposto neste trabalho é mostrado na figura 3.1. A idéia básica é aplicar um pulso de excitação através de um elemento piezoelétrico e capturar as respostas correspondentes medidas diretamente nos elementos piezoelétricos. Como mencionado antes, apesar do pulso da excitação que é simples na forma, um delta de Dirac por exemplo, as respostas medidas têm geralmente uma forma complexa devido aos ecos múltiplos e às difrações associados com as paredes da tubulação e com a mistura multifásica.

Figura 3.1: Esboço do sensor proposto.

3.1.1 Sinais Acústicos

Os sinais de pressão acústica adquiridos neste trabalho através das simulações numéricas têm uma estrutura muito complexa, mesmo em resposta a um único pulso de excitação tipo Dirac, aplicado em um dos vértices da seção transversal do quadrado, e no caso da geometria circular, aplicado no ponto de intersecção do semi-eixo positivo das abscissas com o contorno do domínio. Essa complexidade deve-se aos ecos múltiplos nas paredes da tubulação e também às reflexões e dispersão da onda acústica. Outro fator que influencia na forma dos sinais de resposta é a configuração do escoamento, uma vez que a velocidade de propagação acústica é diferente na fase contínua (água) e na fase dispersa (óleo).

Neste trabalho foram obtidos 4 sinais de resposta $R_k(t)$, $k = 1, \dots, 4$, para cada teste numérico realizado, dados pela equação (3.5):

$$R_k(t) = \oint_{s \in L_k} P(t, \vec{x}) ds, \quad \text{para } L_k \in \partial\Omega \quad (3.5)$$

onde $R_k(t)$ e L_k representam, respectivamente, o k -ésimo sinal de resposta e o k -ésimo elemento de medida (que pode corresponder a L_{exc}).

As posições dos elementos de medida $L_k \in \partial\Omega$, $k = 1, \dots, 4$ podem ser vistas

nas figuras 3.2 e 3.3 nas geometrias quadrada e circular, respectivamente. Pode-se observar que no domínio quadrado, os elementos de medida $L_k, k = 1, \dots, 4$ são responsáveis apenas por obter os sinais acústicos enquanto que no domínio circular, o elemento de medida L_1 é responsável pela aplicação do Dirac e também pela obtenção de um sinal de resposta e os elementos $L_k, k = 2, 3, 4$ têm somente a incumbência de adquirir sinais acústicos.

Figura 3.2: Posição dos elementos de medida na geometria quadrada.

Figura 3.3: Posição dos elementos de medida na geometria circular.

Foram extraídas algumas características destes sinais, tais como o tempo de chegada, T , da onda acústica em cada um dos elementos de medida, as amplitudes dos três maiores picos de cada sinal, A_1 , A_2 e A_3 , bem como o tempo de chegada desses picos, indicados por T_1 , T_2 e T_3 na figura 3.4, a fim de caracterizar a complexidade do sinal em um nível suficiente para a inferência da área interfacial e da fração volumétrica.

Figura 3.4: Sinal acústico e as respectivas características extraídas.

3.1.2 Fração Volumétrica e Área Interfacial

De um ponto de vista físico, a estrutura geométrica de um escoamento disperso pode ser caracterizada pela área interfacial volumétrica, definida como

$$\Gamma(t) = \frac{A_i(t)}{V} \quad (3.6)$$

onde $A_i(t)$ é a área interfacial instantânea de todas as gotas circunscritas em um volume de controle V . Aceita-se geralmente que $\Gamma(t)$ governa trocas interfaciais com o seguinte princípio (REVANKAR e ISHII, 1993; TAN e ISHII, 1990):

$$\text{Transferência interfacial} = \Gamma(t) \times \text{força motriz} \quad (3.7)$$

Conseqüentemente, $\Gamma(t)$ é relacionado fortemente à transferência de massa, momento e energia. Um conhecimento exato desta variável é crucial para uma operação eficiente de um processo industrial de escoamentos multifásicos.

Um outro parâmetro importante na caracterização de escoamento disperso é a fração volumétrica, definida por

$$\alpha(t) = \frac{V_b(t)}{V} \quad (3.8)$$

onde $V_b(t)$ representa o volume ocupado pelas gotas contidas no volume V . A fração volumétrica é fortemente relacionada à distribuição da fase e é um parâmetro requerido para projetos hidrodinâmico e térmico de várias aplicações. Por exemplo, a coalescência das gotas tende a aumentar significativamente em um reator de coluna para $\alpha > 0,3$, com conseqüente redução na taxa da reação.

Na formulação bidimensional, a área interfacial e a fração volumétrica são calculadas, respectivamente, por

$$\Gamma(t) = \frac{\sum P_i(t)}{A} \quad (3.9)$$

$$\alpha(t) = \frac{\sum A_i(t)}{A} \quad (3.10)$$

onde P_i e A_i são o perímetro e a área instantâneos de todas as gotas circunscritas na seção transversal da tubulação, respectivamente, e A é a área total da seção transversal.

Como a distribuição geométrica das fases no escoamento é mapeada pela velocidade local de propagação acústica, considerada na equação diferencial que governa o problema, utilizou-se a velocidade de propagação acústica na água, ou seja, $c_{\text{água}} = 1,48 \frac{Km}{s}$ e no óleo, $c_{\text{óleo}} = 1,7 \frac{Km}{s}$ (óleo de motor SAE 30), uma vez que o escoamento simulado neste trabalho é um escoamento óleo-água em que a fase contínua é água e a fase dispersa (gotas) é óleo.

As gotas são geradas aleatoriamente em cada teste numérico de modo que não haja intersecção entre elas. Outra imposição é que a fração volumétrica esteja no intervalo $[0; 0,3]$, já que o escoamento de interesse é um escoamento disperso. Alguns casos são ilustrados na figura 3.5, considerando-se uma geometria quadrada.

Figura 3.5: Geração aleatória de gotas.

3.2 Construção do Modelo Numérico

Conforme mencionado anteriormente, um simulador numérico capaz de calcular a propagação acústica em escoamentos bifásicos foi desenvolvido, possibilitando a realização de vários testes numéricos, com o objetivo de formar a base de dados para o treinamento das redes neurais. Para a construção desse simulador considerou-se a interação dos campos de sensoriamento acústicos com a geometria da tubulação. Neste trabalho, foram realizadas simulações numéricas em tubos de seção quadrada e em tubos de seção circular.

3.2.1 Adimensionalização

Tubo de Seção Quadrada

Para a realização de simulações numéricas em tubos de seção quadrada, a equação (3.4) foi escrita em coordenadas cartesianas. No contexto deste trabalho, propõe-se a seguinte adimensionalização:

$$P^* = \frac{P}{P_0} \quad (3.11)$$

$$x^* = \frac{x}{L_x} \quad (3.12)$$

$$y^* = \frac{y}{L_y} \quad (3.13)$$

$$t^* = \frac{t}{t_0}, \quad (3.14)$$

onde L_x é o comprimento do domínio de sensoriamento na direção x , L_y é o comprimento na direção y (neste caso, $L_x = L_y = 1$, ou seja, o domínio Ω é um quadrado de lados iguais a 1), P_0 é a pressão de referência e t_0 é dado por

$$t_0 = \frac{L_x}{c_0}. \quad (3.15)$$

Reescrevendo a equação (3.4), obtém-se

$$\frac{\partial^2(P^*P_0)}{\partial t^2} - c_0^2 c^2(x, y) \left[\frac{\partial^2(P^*P_0)}{\partial x^2} + \frac{\partial^2(P^*P_0)}{\partial y^2} \right] = 0. \quad (3.16)$$

Utilizando a regra da cadeia e fazendo as substituições adequadas, tem-se

$$\frac{\partial^2(P^*P_0)}{\partial t^2} = \frac{P_0 c_0^2}{L_x^2} \frac{\partial^2 P^*}{\partial t^{*2}} \quad (3.17)$$

$$\frac{\partial^2(P^*P_0)}{\partial x^2} = \frac{P_0}{L_x^2} \frac{\partial^2 P^*}{\partial x^{*2}} \quad (3.18)$$

$$\frac{\partial^2(P^*P_0)}{\partial y^2} = \frac{P_0}{L_y^2} \frac{\partial^2 P^*}{\partial y^{*2}}. \quad (3.19)$$

Substituindo as equações (3.17), (3.18), (3.19) em (3.16), encontra-se

$$\frac{P_0 c_0^2}{L_x^2} \frac{\partial^2 P^*}{\partial t^{*2}} - c_0^2 c^2(x, y) \left[\frac{P_0}{L_x^2} \frac{\partial^2 P^*}{\partial x^{*2}} + \frac{P_0}{L_y^2} \frac{\partial^2 P^*}{\partial y^{*2}} \right] = 0. \quad (3.20)$$

Multiplicando a equação acima por $\frac{L_x^2}{P_0 c_0^2}$ e definindo $\lambda = \frac{L_x^2}{L_y^2}$, obtém-se

$$\frac{\partial^2 P^*}{\partial t^{*2}} - c^2(x, y) \frac{\partial^2 P^*}{\partial x^{*2}} - \lambda c^2(x, y) \frac{\partial^2 P^*}{\partial y^{*2}} = 0, \quad (3.21)$$

que é a forma adimensionalizada da equação (3.4), para a tubulação de geometria quadrada.

Tubo de Seção Circular

Para realizar simulações numéricas em tubos de seção circular de raio R , a equação (3.4) foi escrita em coordenadas polares, r e θ , onde $x = r \cos \theta$ e $y = r \sin \theta$, obtendo-se

$$\frac{\partial^2 P}{\partial t^2} - c_0^2 c^2(r, \theta) \left[\frac{\partial^2 P}{\partial r^2} + \frac{1}{r} \frac{\partial P}{\partial r} + \frac{1}{r^2} \frac{\partial^2 P}{\partial \theta^2} \right] = 0, \quad (3.22)$$

onde P é a pressão acústica, $c(r, \theta)$ é a velocidade acústica variando com r e θ , t é o tempo e c_0 é a velocidade de referência.

A equação (3.22) foi adimensionalizada, propondo-se

$$P^* = \frac{P}{P_0} \quad (3.23)$$

$$r^* = \frac{r}{R} \quad (3.24)$$

$$t^* = \frac{t}{t_0}, \quad (3.25)$$

sendo que R é o raio da circunferência do domínio de sensoriamento e t_0 é dado por

$$t_0 = \frac{R}{c_0}. \quad (3.26)$$

Reescrevendo a equação (3.22), tem-se

$$\frac{\partial^2(P^*P_0)}{\partial t^2} - c_0^2 c^2(r, \theta) \left[\frac{\partial^2(P^*P_0)}{\partial r^2} + \frac{1}{r} \frac{\partial(P^*P_0)}{\partial r} + \frac{1}{r^2} \frac{\partial^2(P^*P_0)}{\partial \theta^2} \right] = 0. \quad (3.27)$$

Analogamente ao item anterior, obtém-se

$$\frac{\partial^2 P^*}{\partial t^{*2}} - c^2(r, \theta) \left[\frac{\partial^2 P^*}{\partial r^{*2}} + \frac{1}{r^*} \frac{\partial P^*}{\partial r^*} + \frac{1}{r^{*2}} \frac{\partial^2 P^*}{\partial \theta^2} \right] = 0, \quad (3.28)$$

que é a forma adimensionalizada da equação (3.22), para a tubulação de geometria circular.

3.2.2 Discretização da Equação Diferencial Parcial

As equações de propagação acústica foram discretizadas em coordenadas cartesianas e polares de forma a permitir simulações em tubos de seção quadrada e circular. Além disso, as condições de contorno foram definidas de maneira a reproduzir um impulso (Dirac) aplicado inicialmente em algum ponto da fronteira externa, com o restante do contorno refletindo a energia acústica durante todo o tempo da simulação.

Discretização da Equação em Coordenadas Cartesianas

Fez-se a transformação do problema físico-matemático em um problema numérico e formulou-se rotinas de solução com base no Método de Diferenças Finitas (MDF), que apresenta algumas vantagens como, por exemplo, ser de fácil e rápida implementação.

Para simplificar a notação, não será utilizado o símbolo '*' daqui por diante nas equações adimensionalizadas. A equação (3.21) foi discretizada utilizando-se a formulação à esquerda no tempo e a formulação central no espaço. A derivada parcial de segunda ordem com relação ao tempo foi aproximada por diferenças à

esquerda de segunda ordem e as derivadas parciais com relação ao espaço foram aproximadas por diferenças centradas, também de segunda ordem, obtendo-se

$$\begin{aligned} & \frac{1}{\Delta t^2} (P_{i,j}^k - 2P_{i,j}^{k-1} + P_{i,j}^{k-2}) - \frac{c_{i,j}^2}{\Delta x^2} (P_{i+1,j}^k - 2P_{i,j}^k + P_{i-1,j}^k) - \\ & \frac{\lambda c_{i,j}^2}{\Delta y^2} (P_{i,j+1}^k - 2P_{i,j}^k + P_{i,j-1}^k) = 0. \end{aligned} \quad (3.29)$$

onde i e j indicam as coordenadas na malha, nas direções x e y , respectivamente, $\Delta t = 10^{-6}$ e

$$\Delta x = \frac{L_x}{N-1}, \quad (3.30)$$

$$\Delta y = \frac{L_y}{M-1}, \quad (3.31)$$

sendo que N e M são os números de pontos nos intervalos $[0, L_x]$ e $[0, L_y]$, nas direções x e y , respectivamente.

Reorganizando a equação (3.29) de modo a agrupar os termos em P , obtém-se

$$EP_{i,j}^k + AP_{i+1,j}^k + BP_{i-1,j}^k + CP_{i,j+1}^k + DP_{i,j-1}^k + FP_{i,j}^{k-1} + GP_{i,j}^{k-2} = 0, \quad (3.32)$$

onde

$$A = B = -\frac{c_{i,j}^2}{\Delta x^2} \quad (3.33)$$

$$C = D = -\frac{\lambda c_{i,j}^2}{\Delta y^2} \quad (3.34)$$

$$E = \frac{1}{\Delta t^2} + \frac{2c_{i,j}^2}{\Delta x^2} + \frac{2\lambda c_{i,j}^2}{\Delta y^2} \quad (3.35)$$

$$F = -\frac{2}{\Delta t^2} \quad (3.36)$$

$$G = \frac{1}{\Delta t^2} \quad (3.37)$$

A célula computacional correspondente é formada por cinco pontos: um central e dois em cada uma das coordenadas x e y (um anterior e um posterior ao ponto central), como pode ser visto na figura 3.6.

Figura 3.6: Célula computacional da geometria quadrada.

Através da transformação

$$P_{i,j} \rightarrow P_q, \quad (3.38)$$

com $q = i + (j - 1)N$, $q = 1, \dots, NM$, estabelece-se uma ordem para as P 's no vetor das incógnitas.

A equação (3.32) aplicada em cada um dos pontos da malha gera um sistema linear cujas incógnitas são os valores nodais das P 's. A dimensão desse sistema é NM .

Após a discretização, foram montadas todas as equações referentes a cada região de nós da seção transversal da tubulação quadrada, conforme as condições de Dirichlet e Neumann impostas no quadrado. Essas equações podem ser vistas nos casos abaixo:

- (i) Canto inferior esquerdo do quadrado (figura 3.7 (a)): neste caso, a equação que representa este nó ($i = 1, j = 1$) é

$$EP_{1,1}^k + (A + B)P_{2,1}^k + (C + D)P_{1,2}^k + FP_{1,1}^{k-1} + GP_{1,1}^{k-2} = 0. \quad (3.39)$$

- (ii) Borda inferior do quadrado (figura 3.7 (b)): para os pontos referentes à esta região, ou seja, para $i = 2, \dots, N - 1$ e $j = 1$, tem-se

$$EP_{i,1}^k + AP_{i+1,1}^k + BP_{i-1,1}^k + (C + D)P_{i,2}^k + FP_{i,1}^{k-1} + GP_{i,1}^{k-2} = 0. \quad (3.40)$$

Figura 3.7: Células computacionais relativas: (a) ao canto inferior esquerdo e (b) à borda inferior do quadrado.

- (iii) Canto inferior direito do quadrado (figura 3.8 (a)): este nó ($i = N, j = 1$) é representado pela equação

$$EP_{N,1}^k + (B + A)P_{N-1,1}^k + (C + D)P_{N,2}^k + FP_{N,1}^{k-1} + GP_{N,1}^{k-2} = 0. \quad (3.41)$$

- (iv) Borda lateral esquerda do quadrado (figura 3.8 (b)): para os pontos desta região, isto é, para $i = 1$ e $j = 2, \dots, M - 1$, tem-se

$$EP_{1,j}^k + (A + B)P_{2,j}^k + CP_{1,j+1}^k + DP_{1,j-1}^k + FP_{1,j}^{k-1} + GP_{1,j}^{k-2} = 0. \quad (3.42)$$

Figura 3.8: Células computacionais referentes: (a) ao canto inferior direito e (b) à borda lateral esquerda do quadrado.

- (v) Interior do quadrado (figura 3.9 (a)): nesta região, isto é, para $i = 2, \dots, N-1$ e $j = 2, \dots, M-1$, encontra-se

$$EP_{i,j}^k + AP_{i+1,j}^k + BP_{i-1,j}^k + CP_{i,j+1}^k + DP_{i,j-1}^k + FP_{i,j}^{k-1} + GP_{i,j}^{k-2} = 0. \quad (3.43)$$

- (vi) Borda lateral direita do quadrado (figura 3.9 (b)): para os pontos desta região, ou seja, para $i = N$ e $j = 2, \dots, M-1$, tem-se

$$EP_{N,j}^k + (B+A)P_{N-1,j}^k + CP_{N,j+1}^k + DP_{N,j-1}^k + FP_{N,j}^{k-1} + GP_{N,j}^{k-2} = 0. \quad (3.44)$$

Figura 3.9: Células computacionais referentes: (a) ao interior e (b) à borda lateral direita do quadrado.

- (vii) Canto superior esquerdo do quadrado (figura 3.10 (a)): este nó ($i = 1, j = M$) é representado pela equação

$$EP_{i,M}^k + (A + B)P_{2,M}^k + (D + C)P_{1,M-1}^k + FP_{1,M}^{k-1} + GP_{1,M}^{k-2} = 0. \quad (3.45)$$

(viii) Borda superior do quadrado (figura 3.10 (b)): para os pontos referentes à essa região, ou seja, para $i = 2, \dots, N - 1$ e $j = M$, tem-se

$$EP_{i,M}^k + AP_{i+1,M}^k + BP_{i-1,M}^k + (D + C)P_{i,M-1}^k + FP_{i,M}^{k-1} + GP_{i,M}^{k-2} = 0. \quad (3.46)$$

Figura 3.10: Células computacionais relativas: (a) ao canto superior esquerdo e (b) à borda superior do quadrado.

(ix) Canto superior direito do quadrado (figura 3.11): este nó ($i = N, j = M$) é representado pela equação

$$EP_{N,M}^k + (B + A)P_{N-1,M}^k + (D + C)P_{N,M-1}^k + FP_{N,M}^{k-1} + GP_{N,M}^{k-2} = 0. \quad (3.47)$$

Figura 3.11: Célula computacional concernente ao canto superior direito do quadrado.

Passando para o lado direito os termos referentes aos instantes $t = k - 1$ e $t = k - 2$, constrói-se o sistema linear

$$\tilde{A}x = \tilde{b} \quad (3.48)$$

cujas soluções produzem um vetor de soluções $x^k = (P_1^k, P_2^k, \dots, P_{NM}^k)$, no instante $t = k$.

A cada iteração no tempo, o vetor de soluções é atualizado, de modo a mostrar a evolução da onda acústica no tempo, ou seja,

$$x^{k-2} = x^{k-1} \quad (3.49)$$

$$x^{k-1} = x^k. \quad (3.50)$$

O sistema linear (3.48) possui certas características ilustradas no esquema abaixo (equação 3.51). Cada linha do sistema corresponde à equação (3.32) aplicada a um ponto nodal. Devido à escolha da célula computacional e da ordem para as incógnitas, a matriz $\tilde{A}$ possui 5 diagonais formadas pelas relações de A a E (equações 3.33, 3.34, 3.35), ou por uma combinação delas, no caso dos pontos de contorno.

$$\begin{pmatrix} E & A & & C & & \\ B & E & A & & C & \\ & B & E & & C & \\ & & & \dots & & \\ D & & & & & \\ & D & & & & \\ & & D & & \dots & \end{pmatrix} \begin{pmatrix} P_1^k \\ P_2^k \\ P_3^k \\ \vdots \\ \vdots \\ P_{NM}^k \end{pmatrix} = \begin{pmatrix} -FP_1^{k-1} - GP_1^{k-2} \\ -FP_2^{k-1} - GP_2^{k-2} \\ -FP_3^{k-1} - GP_3^{k-2} \\ \vdots \\ \vdots \\ -FP_{NM}^{k-1} - GP_{NM}^{k-2} \end{pmatrix} \quad (3.51)$$

Discretização da Equação em Coordenadas Polares

Novamente, fez-se a transformação do problema físico-matemático em um problema numérico, discretizando a equação (3.28) através do Método de Diferenças Finitas. Nos termos relacionados com o espaço, as derivadas parciais de primeira ordem foram aproximadas por diferenças centradas e as derivadas de segunda ordem por diferenças centradas de segunda ordem, e no termo relacionado com o tempo, utilizou-se a formulação à esquerda do método, obtendo-se

$$\begin{aligned} & \frac{1}{\Delta t^2} (P_{i,j}^k - 2P_{i,j}^{k-1} + P_{i,j}^{k-2}) - \frac{c_{i,j}^2}{\Delta r^2} (P_{i+1,j}^k - 2P_{i,j}^k + P_{i-1,j}^k) - \\ & \frac{c_{i,j}^2}{r_{i,j}} \frac{1}{2\Delta r} (P_{i+1,j}^k - P_{i-1,j}^k) - \frac{c_{i,j}^2}{r_{i,j}^2} \frac{1}{\Delta \theta^2} (P_{i,j+1}^k - 2P_{i,j}^k + P_{i,j-1}^k) = 0 \end{aligned} \quad (3.52)$$

onde i e j indicam as coordenadas na malha, nas direções r e θ , respectivamente, $\Delta t = 10^{-7}$ e

$$\Delta r = \frac{R}{N}, \quad (3.53)$$

$$\Delta \theta = \frac{2\pi}{M}, \quad (3.54)$$

sendo que N e M são o número de pontos nos intervalos $[0, R]$ e $[0, 2\pi]$, nas direções r e θ , respectivamente.

Reorganizando a equação (3.52) de modo a agrupar os termos em P , obtém-se

$$eP_{i,j}^k + aP_{i+1,j}^k + bP_{i-1,j}^k + cP_{i,j+1}^k + dP_{i,j-1}^k + fP_{i,j}^{k-1} + gP_{i,j}^{k-2} = 0, \quad (3.55)$$

onde

$$a = -\frac{c_{i,j}^2}{\Delta r^2} - \frac{c_{i,j}^2}{r_{i,j}} \frac{1}{2\Delta r} \quad (3.56)$$

$$b = -\frac{c_{i,j}^2}{\Delta r^2} + \frac{c_{i,j}^2}{r_{i,j}} \frac{1}{2\Delta r} \quad (3.57)$$

$$c = d = -\frac{c_{i,j}^2}{r_{i,j}^2} \frac{1}{\Delta \theta^2} \quad (3.58)$$

$$e = \frac{1}{\Delta t^2} + \frac{2c_{i,j}^2}{\Delta r^2} + \frac{2}{r_{i,j}^2} \frac{c_{i,j}^2}{\Delta \theta^2} \quad (3.59)$$

$$f = -\frac{2}{\Delta t^2} \quad (3.60)$$

$$g = \frac{1}{\Delta t^2} \quad (3.61)$$

Como no caso anterior, a célula computacional correspondente é formada por cinco pontos: um central e dois em cada uma das coordenadas r e θ (um anterior e um posterior ao ponto central), como pode ser visto na figura 3.12.

Figura 3.12: Célula computacional da geometria circular.

A ordem para as P 's no vetor das incógnitas é dada pela transformação

$$P_{i,j} \rightarrow P_q, \quad (3.62)$$

com

$$q = \begin{cases} i + (j - 1)N + 1, & \text{se } i \neq 0, \\ 1, & \text{se } i = 0, \end{cases} \quad (3.63)$$

para $i = 0, \dots, N$, $j = 1, \dots, M$, ou seja, $q = 1, \dots, NM + 1$.

A equação (3.22) apresenta um problema. Ela não é definida para $r = 0$. Existem pelo menos 3 modos diferentes de contornar essa situação, segundo Davis (1979), que são:

- (i) fazer com que nenhum nó da malha caia no ponto em que $r = 0$;
- (ii) usar a regra de L'Hôpital para eliminar a indefinição;
- (iii) utilizar coordenadas cartesianas no ponto central.

A primeira alternativa não é aconselhável no sentido de perder informações importantes ao problema. É sempre útil conhecer a solução no ponto central. Uma boa aproximação para tal ponto implica em uma malha muito fina ao seu redor, o que aumenta a chance de erros.

A segunda opção não é tão restritiva. Por exemplo, para o termo $\frac{1}{r} \frac{\partial P}{\partial r}$ que aparece na equação (3.22), a regra de L'Hôpital pode ser empregada da seguinte forma:

$$\lim_{r \rightarrow 0} \frac{1}{r} \frac{\partial P}{\partial r} = \lim_{r \rightarrow 0} \frac{\frac{\partial P}{\partial r}}{r} = \frac{\partial^2 P}{\partial r^2} \quad (3.64)$$

Fazendo o mesmo cálculo nos demais termos com r no denominador, é possível eliminar as indefinições e, portanto, obter a solução em $r = 0$.

Na terceira alternativa é necessário impor que o domínio seja dividido de tal forma que haja uma combinação entre as coordenadas polares e cartesianas, ou

seja, a malha deve ser gerada de forma que existam ângulos $\theta = 0, \theta = \frac{\pi}{2}, \theta = \pi$ e $\theta = \frac{3\pi}{2}$, para que coincida com a malha cartesiana nos eixos Ox e Oy .

As duas últimas opções por foram testadas por Kee (1974) * apud Davis, 1979, que concluiu que o uso de limites na equação diferencial resulta em uma aproximação menos precisa e mais difícil do que o uso de coordenadas cartesianas localmente.

Com base nestas conclusões, optou-se por utilizar a terceira alternativa neste trabalho, que tem por fundamento a discretização da equação (3.28) no ponto central, utilizando coordenadas cartesianas, ou seja, no modelo numérico do tubo circular, aplicou-se a equação (3.32) em $r = 0$ e a equação (3.55) nos demais pontos da malha, gerando um sistema linear cujas incógnitas são os valores nodais das P 's. A dimensão desse sistema é $NM + 1$.

No que se refere à aplicação dessa estratégia, é necessário um cuidado especial com relação ao particionamento do intervalo $[0, 2\pi]$, conforme mencionado anteriormente. O número de pontos nesse intervalo deve ser definido de tal modo que haja dois segmentos de reta contidos nos eixos cartesianos. Pode-se observar na figura 3.13 que a malha definida no domínio circular possui dois segmentos $\overline{AB}$ e $\overline{CD}$ ortogonais e cada um contido em um eixo cartesiano. No caso, $\overline{AB}$ contido no eixo x e $\overline{CD}$ no eixo y .

Depois da discretização, foram montadas todas as equações referentes a cada região de nós da seção transversal da tubulação circular, conforme as condições de Dirichlet e Neumann impostas no círculo. Essas equações podem ser vistas nos casos abaixo:

*KEE R. J., (1974). *A Numerical Study of Natural Convection inside Horizontal Cylinders with Asymmetric Boundary Conditions*, Ph.D. thesis, Univ. of California, Davis, Calif.

Figura 3.13: Malha em coordenadas polares.

- (i) Ponto central, ou seja, para $r = 0$ (figura 3.14): neste nó ($i = 0, j = 1$) foi aplicada a equação (3.32), com $\Delta x = \Delta y = \Delta r$, obtendo-se

$$EP_{i,j}^k + AP_{i+1,j}^k + BP_{i+1,j+\frac{2M}{4}}^k + CP_{i+1,j+\frac{M}{4}}^k + DP_{i+1,j+\frac{3M}{4}}^k + FP_{i,j}^{k-1} + GP_{i,j}^{k-2} = 0, \quad (3.65)$$

ou seja,

$$EP_{0,1}^k + AP_{1,1}^k + BP_{1,1+\frac{2M}{4}}^k + CP_{1,1+\frac{M}{4}}^k + DP_{1,1+\frac{3M}{4}}^k + FP_{0,1}^{k-1} + GP_{0,1}^{k-2} = 0. \quad (3.66)$$

Figura 3.14: Célula computacional relativa ao ponto central do círculo.

Nos demais casos, aplicou-se a equação (3.55).

- (ii) Pontos pertencentes ao semi-eixo positivo das abscissas que não têm intersecção com o círculo (figura 3.15): para os pontos desta região, onde $i = 1, \dots, N - 1$ e $j = 1$, tem-se

$$eP_{i,j}^k + aP_{i+1,j}^k + bP_{i-1,j}^k + cP_{i,j+1}^k + dP_{i,M}^k + fP_{i,j}^{k-1} + gP_{i,j}^{k-2} = 0, \quad (3.67)$$

isto é,

$$eP_{i,1}^k + aP_{i+1,1}^k + bP_{i-1,1}^k + cP_{i,2}^k + dP_{i,M}^k + fP_{i,1}^{k-1} + gP_{i,1}^{k-2} = 0. \quad (3.68)$$

Figura 3.15: Célula computacional relativa aos pontos pertencentes ao semi-eixo positivo das abscissas que não têm intersecção com o círculo.

- (iii) Ponto de intersecção do semi-eixo positivo das abscissas com o círculo (figura 3.16): este nó, $i = N, j = 1$, é representado pela equação

$$eP_{i,j}^k + (b + a)P_{i-1,j}^k + cP_{i,j+1}^k + dP_{i,M}^k + fP_{i,j}^{k-1} + gP_{i,j}^{k-2} = 0, \quad (3.69)$$

que implica em

$$eP_{N,1}^k + (b + a)P_{N-1,1}^k + cP_{N,2}^k + dP_{N,M}^k + fP_{N,1}^{k-1} + gP_{N,1}^{k-2} = 0. \quad (3.70)$$

Figura 3.16: Célula computacional relativa ao ponto de intersecção do semi-eixo positivo das abscissas com o círculo.

- (iv) Pontos do interior do círculo para $i = 1, \dots, N - 1$ e $j = 2, \dots, M - 1$ (figura 3.17): nesta região, tem-se

$$eP_{i,j}^k + aP_{i+1,j}^k + bP_{i-1,j}^k + cP_{i,j+1}^k + dP_{i,j-1}^k + fP_{i,j}^{k-1} + gP_{i,j}^{k-2} = 0. \quad (3.71)$$

Figura 3.17: Célula computacional relativa aos pontos do interior do círculo para $i = 1, \dots, N - 1$ e $j = 2, \dots, M - 1$.

- (v) Pontos do contorno para $j = 2, \dots, M - 1$ (figura 3.18): nestes pontos, onde $i = N$, tem-se

$$eP_{N,j}^k + (b + a)P_{N-1,j}^k + cP_{N,j+1}^k + dP_{N,j-1}^k + fP_{N,j}^{k-1} + gP_{N,j}^{k-2} = 0. \quad (3.72)$$

Figura 3.18: Célula computacional concernente aos pontos do contorno do círculo para $j = 2, \dots, M - 1$.

- (vi) Pontos do interior do círculo para $i = 1, \dots, N - 1$ e $j = M$, (figura 3.19): para os pontos desta região, tem-se

$$eP_{i,j}^k + aP_{i+1,j}^k + bP_{i-1,j}^k + cP_{i,1}^k + dP_{i,j-1}^k + fP_{i,j}^{k-1} + gP_{i,j}^{k-2} = 0, \quad (3.73)$$

ou seja,

$$eP_{i,M}^k + aP_{i+1,M}^k + bP_{i-1,M}^k + cP_{i,1}^k + dP_{i,M-1}^k + fP_{i,M}^{k-1} + gP_{i,M}^{k-2} = 0. \quad (3.74)$$

Figura 3.19: Célula computacional relativa aos pontos do interior do círculo para $i = 1, \dots, N - 1$ e $j = M$.

(vii) Ponto onde $i = N, j = M$ (figura 3.20): para este nó tem-se

$$eP_{i,j}^k + (b + a)P_{i-1,j}^k + cP_{i,1}^k + dP_{i,j-1}^k + fP_{i,j}^{k-1} + gP_{i,j}^{k-2} = 0, \quad (3.75)$$

isto é,

$$eP_{N,M}^k + (b + a)P_{N-1,M}^k + cP_{N,1}^k + dP_{N,M-1}^k + fP_{N,M}^{k-1} + gP_{N,M}^{k-2} = 0. \quad (3.76)$$

Figura 3.20: Célula computacional relativa ao ponto do círculo onde $i = N, j = M$.

Passando para o lado direito os termos referentes aos instantes $t = k - 1$ e $t = k - 2$, constrói-se o sistema linear

$$\widetilde{A}_1 x = \widetilde{b}_1 \quad (3.77)$$

cujas soluções produzem um vetor de soluções $x^k = (P_1^k, P_2^k, \dots, P_{NM}^k)$, no instante $t = k$.

Novamente, a cada iteração no tempo, o vetor de soluções é atualizado, produzindo a evolução da onda acústica no tempo, ou seja,

$$x^{k-2} = x^{k-1} \quad (3.78)$$

$$x^{k-1} = x^k. \quad (3.79)$$

Neste caso, o esquema do sistema linear é análogo ao do caso do tubo quadrado, ou seja, a matriz $\widetilde{A}_1$ tem as mesmas características da matriz $\widetilde{A}$.

3.2.3 Armazenamento

As matrizes $\widetilde{A}$ e $\widetilde{A}_1$ são matrizes esparsas, ou seja, apenas um pequeno número de seus elementos é diferente de zero. Há dois objetivos que devem ser apresentados em métodos de resolução de sistemas lineares com matrizes esparsas: economia de tempo e de espaço.

Baseado nisto, neste trabalho foi utilizada a idéia apresentada por Press *et al.* (1992), no livro *Numerical Recipes*, de um método de armazenamento indexado por linhas em que somente os elementos não nulos da matriz são estocados. Para isso, supondo uma matriz de dimensão $N \times N$, são utilizados dois vetores:

- (i) um vetor de elementos reais (com precisão simples ou dupla), responsável por armazenar os valores não nulos da matriz, de modo que nas N primeiras posições desse vetor estejam os elementos da diagonal da matriz (incluindo os elementos nulos), em ordem. A posição $N + 1$ é vazia, e da posição $N + 2$ em diante encontram-se os outros elementos não nulos da matriz, que estão fora da diagonal.
- (ii) um vetor de elementos inteiros, responsável pelo armazenamento dos índices dos valores não nulos, de maneira que a primeira posição deste vetor recebe o valor $N + 2$, e da segunda posição até a $N + 1$, o preenchimento é feito da seguinte forma: a posição i , $i = 2, \dots, N + 1$, recebe o valor da posição anterior somado com a quantidade de elementos não nulos fora da diagonal, que estão na mesma linha da matriz. Se em alguma linha da matriz só há elementos nulos fora da diagonal, então o valor da posição anterior é repetido. O valor da posição $N + 1$ indica a dimensão desses dois vetores. As demais posições

desse vetor, ou seja, da posição $N + 2$ em diante, armazenam o número da coluna da matriz onde os elementos não nulos (que não encontram-se na diagonal) estão localizados.

A principal vantagem desta técnica, além da otimização do uso da memória, é permitir que os dados armazenados sejam acessados de forma mais rápida, diminuindo consideravelmente o tempo gasto com a resolução do sistema linear. Quanto à montagem dos vetores de armazenamento, o algoritmo sugerido por Press *et al.* (1992) baseia-se na varredura de todos os elementos da matriz original para identificar aqueles diferentes de zero, demandando, portanto, um tempo excessivamente longo. Neste trabalho, a subrotina desenvolvida para o armazenamento dos elementos não nulos da matriz foi baseada na idéia de Press *et al.* (1992) e também no fato de se conhecer a estrutura da matriz. Assim, os dois vetores de armazenamento foram montados diretamente, reduzindo significativamente o tempo de execução dos cálculos.

3.2.4 Resolução do Sistema Linear

Os sistemas lineares (3.48) e (3.77) foram resolvidos pelo Método dos Gradientes Bi-Conjugados Pré-Condicionado (PBCG), que é uma generalização do Método dos Gradientes Conjugados, com a vantagem de resolver qualquer sistema linear desde que a matriz não seja singular. Na implementação foi utilizada uma rotina pronta do PBCG do pacote computacional *Numerical recipes in Fortran* (PRESS *et al.*, 1992).

Um dos motivos pelos quais adotou-se esse método para resolver os sistemas lineares do problema em questão reside no fato das matrizes dos coeficientes serem esparsas e, portanto, métodos iterativos são mais adequados por economia de memória. Outra razão está no fato das matrizes não serem definidas positivas^{*} ou simétricas e, assim, o Método dos Gradientes Conjugados não pode ser utilizado (PRESS *et al.*, 1992). Logo, o PBCG mostra ser o algoritmo mais indicado neste

^{*}Uma matriz A é definida positiva se $x^T A x > 0$ para todo $x \neq 0$.

caso, uma vez que independe do tipo de matriz. E também a velocidade é uma característica de grande relevância neste método.

Para a aplicação deste método, é necessário ter uma matriz pré-condicionada $\hat{A}$ de modo que $\hat{A}^{-1}\tilde{A} \cong I$, o que significa que $\hat{A}$ deve estar próxima de $\tilde{A}$. A idéia é resolver facilmente o sistema linear para alguma $\hat{A}$ perto de $\tilde{A}$, permitindo que o algoritmo convirja em poucas iterações. A matriz pré-condicionada foi escolhida como sendo a parte diagonal de $\tilde{A}$. Os sistemas auxiliares do algoritmo são calculados através das matrizes L e U (triangular inferior e superior), cujo arquivo é chamado a cada iteração do PBCG. Análogo para a matriz $\tilde{A}_1$.

Portanto, resolvendo os sistemas lineares (3.48) e (3.77), obtém-se o vetor de soluções $x^k = (P_1^k, P_2^k, \dots, P_{NM}^k)$, no instante $t = k$, para as geometrias quadrada e circular dos domínios, respectivamente.

Conceitos Básicos de Redes Neurais Artificiais

Este capítulo introduz os conceitos básicos de redes neurais artificiais e descreve o algoritmo de retropropagação do erro, utilizado para o treinamento das redes neurais implementadas neste trabalho.

4.1 O que é uma Rede Neural Artificial?

De acordo com Braga, Carvalho e Ludermir (2000), redes neurais artificiais são sistemas paralelos distribuídos compostos por unidades de processamento simples (neurônios), que calculam determinadas funções matemáticas (normalmente não-lineares). Tais unidades são dispostas em uma ou mais camadas e interligadas por um grande número de conexões, geralmente unidirecionais. Na maioria dos modelos estas conexões estão associadas a pesos, os quais armazenam o conhecimento representado no modelo e servem para ponderar a entrada recebida por cada neurônio da rede.

Em redes neurais, o procedimento usual na resolução de problemas passa inicialmente por uma fase de aprendizagem, em que um conjunto de exemplos é apresentado à rede, que extrai automaticamente as características necessárias para representar a informação fornecida. Posteriormente, estas características são utilizadas para gerar respostas para o problema.

A capacidade de aprender através de exemplos e de generalizar a informação

aprendida é, sem dúvida, o atrativo principal da solução de problemas através de RNAs. A generalização, que está associada à capacidade de a rede aprender através de um conjunto reduzido de exemplos e posteriormente dar respostas coerentes para dados não conhecidos, é uma demonstração de que a capacidade das RNAs vai muito além do que simplesmente mapear relações de entrada e saída.

4.2 Modelos Matemáticos de Neurônios

A descrição matemática do neurônio proposto por McCulloch e Pitts (nó MCP), ilustrado na figura 4.1, resultou em um modelo com p entradas $x_1, x_2, \dots, x_p$ e apenas uma saída y . As entradas são ponderadas por pesos sinápticos $w_1, w_2, \dots, w_p$, cujos valores podem ser positivos ou negativos (correspondendo à sinápses excitatórias ou inibitórias, respectivamente).

Figura 4.1: Neurônio de McCulloch e Pitts.

Neste modelo, a ativação do neurônio é obtida através da aplicação de uma função de ativação, que ativa ou não a saída, dependendo do valor da soma ponderada das suas entradas. Matematicamente, a saída do neurônio MCP é ativada quando

$$\sum_{j=1}^p x_j w_j \geq \theta, \quad (4.1)$$

sendo que p é o número de entradas do neurônio, w_j é o peso associado à entrada x_j e θ é o limiar (*threshold*) do neurônio.

Contudo, em sua definição original, o nó MCP foi proposto com pesos não-ajustáveis (fixos) (BRAGA, CARVALHO e LUDERMIR. 2000). Portanto, alguns outros modelos artificiais foram propostos com o objetivo de eliminar tal limitação. Estes modelos permitem a produção de uma saída qualquer e trabalham com diferentes funções de ativação.

Assim, foi definido um modelo de neurônio não-linear baseado no *perceptron*. Este neurônio, é atualmente o mais utilizado pelos pesquisadores da área de RNAs. A figura 4.2 mostra tal neurônio, onde são identificados três elementos básicos (HAYKIN, 1999):

Figura 4.2: Modelo de neurônio não-linear.

- (i) Um **conjunto de sinápses**, cada uma caracterizada por um peso próprio. Especificamente, um sinal x_j na entrada da sinapse j conectada ao neurônio k é multiplicado pelo peso sináptico w_{kj} . O primeiro índice do peso sináptico se refere ao neurônio em questão e o segundo, é referente ao sinal de entrada da sinapse à qual o peso está conectado;
- (ii) Um **somador** para somar os sinais de entrada, ponderados pelas respectivas

sinápses do neurônio;

- (iii) Uma **função de ativação** para restringir a amplitude da saída de um neurônio, também denominada função restritiva.

O modelo de neurônio da figura 4.2 também inclui um *bias* (*threshold*) aplicado externamente, representado por θ_k . O *bias* tem o efeito de aumentar ou diminuir a entrada da função de ativação, dependendo se ele é positivo ou negativo, respectivamente.

Em termos matemáticos, um neurônio k é descrito pelas seguintes equações

$$u_k = \sum_{j=1}^p w_{kj} x_j \quad (4.2)$$

$$v_k = u_k + \theta_k \quad (4.3)$$

$$y_k = \varphi(v_k), \quad (4.4)$$

sendo que $x_1, x_2, \dots, x_p$ são os sinais de entrada; $w_{k1}, w_{k2}, \dots, w_{kp}$ são os pesos do neurônio k ; u_k é a saída do somador devido aos sinais de entrada; v_k é o potencial de ativação do neurônio k ; θ_k é o *bias*; $\varphi(\cdot)$ é a função de ativação e y_k é o sinal de saída do neurônio.

Note que os pesos $w_{k1}, w_{k2}, \dots, w_{kp}$ e o *bias* θ_k são parâmetros escalares ajustáveis do neurônio. A idéia central das redes neurais é que tais parâmetros possam ser ajustados de modo que a rede apresente o comportamento desejado. Assim, pode-se treinar a rede para fazer um trabalho particular ajustando os pesos ou o *bias*, ou talvez a própria rede ajustará estes parâmetros para conseguir um desempenho desejado.

4.3 Funções de Ativação

A função de ativação, denotada por $\varphi(v)$, define a saída de um neurônio em termos do potencial de ativação v . De acordo com Haykin (1999), há três tipos básicos de funções de ativação, ilustrados na figura 4.3:

Figura 4.3: Exemplos de funções de ativação: (a) Degrau, (b) Linear por Partes e (c) Sigmóide.

(i) Função Degrau

$$\varphi(v) = \begin{cases} 1, & \text{se } v \geq 0 \\ 0, & \text{se } v < 0 \end{cases} \quad (4.5)$$

(ii) Função Linear por Partes

$$\varphi(v) = \begin{cases} 1, & \text{se } v \geq a \\ v, & \text{se } -a < v < a \\ -1, & \text{se } v \leq -a \end{cases} \quad (4.6)$$

(iii) Função Sigmóide

$$\varphi(v) = \frac{1}{1 + \exp(-v)} \quad (4.7)$$

4.4 Arquitetura de Redes Neurais

As redes neurais artificiais podem ser classificadas segundo três parâmetros que as compõem: arquitetura, tipo de neurônio e regra de aprendizagem.

A arquitetura de uma rede neural é formada pelo número de camadas, número de neurônios em cada camada, pelas ligações entre os neurônios e pelo fluxo de informação.

Usualmente as camadas de uma rede neural são classificadas em três grupos:

- (i) **Camada de Entrada** : onde os padrões são apresentados à rede;
- (ii) **Camadas Intermediárias ou Ocultas** : onde é feita a maior parte do processamento, através das conexões ponderadas, podem ser consideradas como extratoras de características;
- (iii) **Camada de Saída** : onde o resultado final é concluído e apresentado.

Segundo Haykin (1999), com base nas arquiteturas existentes, as redes neurais podem ser divididas em três classes principais: redes *feed-forward* (alimentadas adiante) com uma única camada (figura 4.4), redes *feed-forward* com múltiplas camadas (figura 4.5) e redes recorrentes (figura 4.6).

Figura 4.4: Rede neural *feed-forward* com uma única camada.

Figura 4.5: Rede neural *feed-forward* com múltiplas camadas.

Nas redes neurais *feed-forward* o fluxo de informação dos neurônios da camada de entrada até os neurônios da camada de saída segue estritamente adiante. O processamento dos dados pode ser estendido por várias camadas de neurônios, mas nenhuma conexão de realimentação está presente, isto é, as conexões de um neurônio são sempre para outros neurônios da camada posterior. A rede *perceptron* de múltiplas camadas (MLP) e a rede ADALINE são exemplos dessa forma de arquitetura.

As redes neurais recorrentes contêm conexões de realimentação, ou seja, as

Figura 4.6: Rede neural recorrente.

conexões de um neurônio podem estar ligadas a neurônios da mesma camada ou de camadas anteriores. Essas conexões de realimentação possuem como característica principal a criação de um comportamento temporal na rede, como uma memória de curto prazo. Um exemplo clássico de rede neural recorrente é a rede de Hopfield.

As redes neurais multi-camadas ainda podem ser subdivididas em três categorias: redes completamente conectadas, parcialmente conectadas e localmente conectadas. Nas redes totalmente conectadas todos os neurônios da camada l estão ligados a todos os neurônios da camada $l + 1$. Nas parcialmente conectadas, alguns neurônios da camada l estão ligados a alguns neurônios da camada $l + 1$. Já na rede localmente conectada, determinadas regiões (neurônios) de uma camada l estão ligadas a determinados neurônios da camada seguinte $l + 1$.

4.5 Tipos de Aprendizagem

Uma rede neural tem que ser configurada para que a aplicação de um conjunto de entradas produza uma saída desejada. Existem vários métodos para se ajustar os pesos das conexões entre os neurônios da rede. Uma maneira é ajustar os pesos das conexões explicitamente, a partir de conhecimento prévio sobre o problema a

ser resolvido. Outra maneira é 'treinar' a rede neural pela alimentação e ensino de padrões de entrada e mudança dos pesos das conexões de acordo com alguma regra de aprendizado.

Na literatura são encontrados três paradigmas de aprendizagem: aprendizagem supervisionada, aprendizagem não-supervisionada e aprendizagem por reforço. Vale lembrar que alguns autores não consideram a aprendizagem por reforço como um paradigma independente, mas sim como um modo de aprendizagem pertencente ao aprendizado não-supervisionado (HAYKIN, 1999), ou ainda como um caso particular do aprendizado supervisionado (BRAGA, CARVALHO e LUDERMIR 2000).

- (i) **Aprendizagem supervisionada** : é o método mais comum de aprendizado de redes neurais. A rede é treinada por um supervisor externo que fornece os padrões de entrada e a saída desejada para a rede. A saída gerada pela rede é comparada com a saída desejada, recebendo uma resposta do supervisor sobre o erro atual de sua saída. A cada padrão de entrada, compara-se a saída da rede com a saída desejada e efetua-se um ajuste nos pesos das conexões de forma a minimizar este erro;
- (ii) **Aprendizagem não-supervisionada** : como o próprio nome sugere, neste tipo de treinamento não há o supervisor externo para acompanhar o treinamento. Desta forma, a própria rede neural deve extrair do conjunto de entradas as informações necessárias para a realização do mapeamento entrada/saída;
- (iii) **Aprendizagem por reforço** : uma função, definida a priori, é utilizada para indicar se a saída gerada pela rede é boa ou ruim, gerando assim um processo de recompensa ou penalização para a rede neural. O ajuste nos pesos é realizado com base nesses valores de penalizações ou de recompensas obtidos a partir da função de avaliação.

4.6 *Perceptrons* de Múltiplas Camadas

As redes neurais *feed-forward* de múltiplas camadas consistem de um conjunto de nós sensoriais que constituem a camada de entrada, uma ou mais camadas intermediárias de neurônios e uma camada de saída. Para identificar a arquitetura deste tipo de RNA, será utilizada a seguinte notação:

$$R - S^1 - S^2 - S^3$$

onde o número R de entradas da rede é seguido pelo número de neurônios $S^i, i = 1, \dots, 3$ em cada uma das camadas posteriores. Por exemplo, na figura 4.7 está representada uma rede neural 5 - 5 - 4 - 3.

Figura 4.7: Rede neural 5 - 5 - 4 - 3.

Essas redes são chamadas de *perceptrons* de múltiplas camadas (MLP, *multilayer perceptrons*) e têm sido aplicadas com sucesso para solucionar diversos problemas difíceis, através de seu treinamento supervisionado com o algoritmo de retropropagação do erro, onde o sinal de entrada se propaga da entrada para a saída através da rede, camada por camada.

Basicamente, a aprendizagem pelo algoritmo de retropropagação de erro consiste de dois passos através das diferentes camadas da rede: um passo para a frente, a propagação, e um passo para trás, a retropropagação. No passo para

a frente, um vetor de entrada é aplicado aos nós sensoriais da rede e seu efeito se propaga através da rede, camada por camada. Finalmente, um conjunto de saídas é produzido como a resposta real da rede. Durante o passo de propagação, os pesos sinápticos da rede são todos fixos. Durante o passo para trás, os pesos sinápticos são todos ajustados de acordo com uma regra de correção de erro. Especificamente, a resposta real da rede é subtraída de uma resposta desejada para produzir um sinal de erro. Este sinal de erro é então propagado para trás através da rede, contra a direção das conexões sinápticas. Os pesos são ajustados com o objetivo de aproximar a resposta real da rede e a resposta desejada. O algoritmo de retropropagação do erro é também referido na literatura como algoritmo de retropropagação (*back-propagation*).

Na figura 4.8 está representada uma porção do *perceptron* de múltiplas camadas. Dois tipos de sinais são identificados nesta rede, de acordo com Parker (1987) * apud Haykin (1999):

Figura 4.8: Ilustração da propagação para frente dos sinais funcionais e da retropropagação dos sinais de erro em um MLP.

- (i) **Sinal Funcional** : é um sinal de entrada que recai no terminal de entrada da rede, propaga-se para a frente (neurônio por neurônio) emergindo no terminal de saída como um sinal de saída.
- (ii) **Sinal de Erro** : este sinal se origina em um neurônio de saída da rede e se propaga para trás (camada por camada) através da rede.

*PARKER, D. B., (1987). Optimal algorithms for adaptative networks: Second order back-propagation, second order direct propagation, and second order Hebbian learning, *IEEE 1st International Conference on Neural Networks*, v. 2, p. 593-600, San Diego, CA.

O principal problema de uma rede MLP esteve justamente em como realizar o ajuste dos pesos das camadas intermediárias da rede. Este problema esteve presente por toda a década de 70. Apenas em meados da década de 80, com o desenvolvimento do algoritmo de retropropagação do erro pelos pesquisadores Rumelhart, Hinton e Willians (1986), foi possível realizar o treinamento de uma rede MLP. O algoritmo de retropropagação do erro é descrito a seguir.

4.7 O Algoritmo de Retropropagação

Seja n a iteração (passo de tempo) onde o n -ésimo padrão de treinamento (exemplo) é apresentado à rede. O conjunto formado por todos os exemplos de treinamento é denominado conjunto de treinamento. Assim, o sinal de erro na saída do neurônio j , na iteração n , é definido por

$$e_j(n) = d_j(n) - y_j(n), \quad \text{o neurônio } j \text{ é um nó de saída} \quad (4.8)$$

sendo que $d_j(n)$ se refere à resposta desejada para o neurônio j , e $y_j(n)$ se refere ao sinal funcional que aparece na saída do neurônio j , na iteração n .

O valor do erro instantâneo para o neurônio j é dado por $\frac{1}{2}e_j^2(n)$, e o erro total correspondente à soma dos erros quadráticos instantâneos para todos os neurônios de saída é definido por

$$E(n) = \frac{1}{2} \sum_{j \in C} e_j^2(n), \quad (4.9)$$

onde o conjunto C inclui todos os neurônios da camada de saída da rede.

Considere que N represente o número total de exemplos contidos no conjunto de treinamento. O erro médio quadrático para todos os padrões é obtido somando-se os $E(n)$ para todos os N e então normalizando em relação ao tamanho do conjunto de treinamento, ou seja,

$$E_{med} = \frac{1}{N} \sum_{n=1}^N E(n), \quad (4.10)$$

O erro instantâneo $E(n)$, e conseqüentemente o erro médio E_{med} , é uma função de todos os parâmetros livres da rede, isto é, pesos sinápticos e *bias*. Para um dado conjunto de treinamento, E_{med} representa a função objetivo como a medida de desempenho de aprendizagem. A meta do processo de aprendizagem é ajustar os parâmetros livres da rede para minimizar E_{med} . Para fazer esta minimização, considera-se um método simples de treinamento, no qual os pesos são atualizados de padrão em padrão até formar uma época, isto é, uma apresentação completa do conjunto de treinamento inteiro está sendo processado. Os ajustes dos pesos são realizados de acordo com os respectivos erros calculados para cada padrão apresentado à rede.

Assim, considere o neurônio j sendo alimentado por um conjunto de sinais funcionais produzidos por uma camada de neurônios à sua esquerda. O potencial de ativação $v_j(n)$ do neurônio j é

$$v_j(n) = \sum_{i=0}^m w_{ji}(n) y_i(n), \quad (4.11)$$

onde m é o número total de entradas (excluindo o *bias*) aplicadas ao neurônio j , $w_{ji}(n)$ representa o peso sináptico conectando a saída do neurônio i à entrada do neurônio j , na iteração n e $y_i(n)$ o sinal funcional que aparece na saída do neurônio i , também na iteração n . O peso sináptico w_{j0} é igual ao *bias* aplicado ao neurônio j . A entrada correspondente ao *bias* é $y_0 = +1$.

Portanto, o sinal funcional $y_j(n)$ que aparece na saída do neurônio j na iteração n é dado por

$$y_j(n) = \varphi_j(v_j(n)). \quad (4.12)$$

O algoritmo de retropropagação aplica uma correção $\Delta w_{ji}(n)$ ao peso sináptico $w_{ji}(n)$, que é proporcional à derivada parcial $\frac{\partial E(n)}{\partial w_{ji}(n)}$. Aplicando a regra da cadeia a este gradiente tem-se

$$\frac{\partial E(n)}{\partial w_{ji}(n)} = \frac{\partial E(n)}{\partial e_j(n)} \frac{\partial e_j(n)}{\partial y_j(n)} \frac{\partial y_j(n)}{\partial v_j(n)} \frac{\partial v_j(n)}{\partial w_{ji}(n)} \quad (4.13)$$

A derivada parcial $\frac{\partial E(n)}{\partial w_{ji}(n)}$ determina a direção de busca no espaço de pesos, para o peso sináptico w_{ji} .

Diferenciando ambos os lados da equação (4.9) em relação a $e_j(n)$, obtém-se

$$\frac{\partial E(n)}{\partial e_j(n)} = e_j(n), \quad (4.14)$$

e diferenciando ambos os lados da equação (4.8) em relação a $y_j(n)$, tem-se

$$\frac{\partial e_j(n)}{\partial y_j(n)} = -1. \quad (4.15)$$

Em seguida, diferenciando a equação (4.12) em relação a $v_j(n)$, obtém-se

$$\frac{\partial y_j(n)}{\partial v_j(n)} = \varphi_j'(v_j(n)). \quad (4.16)$$

Finalmente, diferenciar a equação (4.11) em relação $w_{ji}(n)$ produz

$$\frac{\partial v_j(n)}{\partial w_{ji}(n)} = y_i(n). \quad (4.17)$$

O uso das equações (4.14) a (4.17) na equação (4.13) origina

$$\frac{\partial E(n)}{\partial w_{ji}(n)} = -e_j(n) \varphi_j'(v_j(n)) y_i(n). \quad (4.18)$$

A correção $\Delta w_{ji}(n)$ aplicada a $w_{ji}(n)$ é definida pela regra delta

$$\Delta w_{ji}(n) = -\eta \frac{\partial E(n)}{\partial w_{ji}(n)} \quad (4.19)$$

onde η é o parâmetro da taxa de aprendizado do algoritmo de retropropagação. O sinal negativo na equação (4.19) indica a descida do gradiente no espaço de pesos, ou seja, busca uma direção para a mudança de peso que reduza o valor de $E(n)$. Assim, utilizando a equação (4.18) em (4.19) tem-se

$$\Delta w_{ji}(n) = \eta \delta_j(n) y_i(n) \quad (4.20)$$

sendo que $\delta_j(n)$ é o gradiente local, definido por

$$\begin{aligned} \delta_j(n) &= -\frac{\partial E(n)}{\partial v_j(n)} \\ &= -\frac{\partial E(n)}{\partial e_j(n)} \frac{\partial e_j(n)}{\partial y_j(n)} \frac{\partial y_j(n)}{\partial v_j(n)} \\ &= e_j(n) \varphi'_j(v_j(n)) \end{aligned} \quad (4.21)$$

De acordo com a equação (4.21), o gradiente local $\delta_j(n)$ para o neurônio de saída j é igual ao produto do sinal de erro $e_j(n)$ correspondente para aquele neurônio pela derivada $\varphi'_j(v_j(n))$ da função de ativação associada. Este cálculo requer o conhecimento da derivada da função de ativação φ associada ao neurônio j . Para esta derivação existir, necessita-se que a função φ seja contínua. Em outras palavras, a diferenciabilidade é a única exigência que a função de ativação deve satisfazer.

Das equações (4.20) e (4.21) nota-se que um fator essencial envolvido no cálculo do ajuste de peso $\Delta w_{ji}(n)$ é o sinal de erro $e_j(n)$ na saída do neurônio j . Portanto, pode-se identificar dois casos distintos, dependendo da localização do neurônio j na rede.

4.7.1 O Neurônio j é um Nó da Camada de Saída

Quando o neurônio j está localizado na camada de saída da rede, ele é suprido com uma resposta desejada particular, fazendo com que o cálculo do sinal de erro associado seja direto. Assim, pode-se utilizar a equação (4.8) para calcular o sinal de erro $e_j(n)$ associado com este neurônio. Tendo-se determinado $e_j(n)$, calcula-se diretamente o gradiente local $\delta_j(n)$, através da equação (4.21).

4.7.2 O Neurônio j é um Nó em uma Camada Intermediária

Quando o neurônio j está localizado em uma camada intermediária da rede, não existe uma resposta desejada específica para aquele neurônio. O sinal de erro para um neurônio intermediário é determinado em termos dos sinais de erro de todos os neurônios aos quais o neurônio intermediário está diretamente conectado. De acordo com a equação (4.21), pode-se redefinir o gradiente local para o neurônio intermediário j como

$$\begin{aligned}\delta_j(n) &= -\frac{\partial E(n)}{\partial y_j(n)} \frac{\partial y_j(n)}{\partial v_j(n)} \\ &= -\frac{\partial E(n)}{\partial y_j(n)} \varphi'_j(v_j(n)), \quad \text{o neurônio } j \text{ é intermediário}\end{aligned}\tag{4.22}$$

onde na segunda linha aplicou-se a equação (4.16). Para calcular a derivada parcial $\frac{\partial E(n)}{\partial y_j(n)}$, seja

$$E(n) = \frac{1}{2} \sum_{k \in C} e_k^2(n), \quad \text{o neurônio } k \text{ é um nó de saída}\tag{4.23}$$

que é a equação (4.9) com o índice k utilizado no lugar do índice j , para não haver confusão, uma vez que neste caso, o índice j se refere ao neurônio da camada intermediária. Diferenciando a equação (4.23) em relação ao sinal funcional $y_j(n)$, obtém-se

$$\frac{\partial E(n)}{\partial y_j(n)} = \sum_k e_k \frac{\partial e_k(n)}{\partial y_j(n)} \quad (4.24)$$

Aplicando a regra da cadeia para a derivada parcial $\frac{\partial e_k(n)}{\partial y_j(n)}$, tem-se

$$\frac{\partial E(n)}{\partial y_j(n)} = \sum_k e_k(n) \frac{\partial e_k(n)}{\partial v_k(n)} \frac{\partial v_k(n)}{\partial y_j(n)} \quad (4.25)$$

Mas, como o neurônio k é um nó de saída, tem-se que

$$\begin{aligned} e_k(n) &= d_k(n) - y_k(n) \\ &= d_k(n) - \varphi_k(v_k(n)), \quad \text{o neurônio } k \text{ é um nó de saída} \end{aligned} \quad (4.26)$$

Assim,

$$\frac{\partial e_k(n)}{\partial v_k(n)} = -\varphi'_k(v_k(n)) \quad (4.27)$$

Para o neurônio k , o potencial de ativação é

$$v_k(n) = \sum_{j=0}^m w_{kj}(n) y_j(n), \quad (4.28)$$

onde m é o número de entradas (excluindo o *bias*) aplicadas ao neurônio k . Aqui, novamente, o peso sináptico $w_{k0}(n)$ é igual ao *bias* $\theta_k(n)$ aplicado ao neurônio k . A entrada correspondente ao *bias* está fixa no valor +1.

Diferenciar a equação (4.28) em relação a $y_j(n)$, produz

$$\frac{\partial v_k(n)}{\partial y_j(n)} = w_{kj}(n) \quad (4.29)$$

Utilizando as equações (4.27) e (4.29) em (4.25), obtém-se

$$\begin{aligned}
\frac{\partial E(n)}{\partial y_j(n)} &= - \sum_k e_k(n) \varphi'_k(v_k(n)) w_{kj}(n) \\
&= - \sum_k \delta_k(n) w_{kj}(n)
\end{aligned} \tag{4.30}$$

sendo que na segunda linha utilizou-se a definição de gradiente local dada na equação (4.21), com o índice j substituído por k .

Finalmente, utilizando a equação (4.30) em (4.22), obtém-se a fórmula de retropropagação para o gradiente local $\delta_j(n)$ como

$$\delta_j(n) = \varphi'_j(v_j(n)) \sum_k \delta_k(n) w_{kj}(n), \quad \text{o neurônio } j \text{ é intermediário.} \tag{4.31}$$

O fator $\varphi'_j(v_j(n))$ envolvido no cálculo do gradiente local $\delta_j(n)$ na equação (4.31) depende unicamente da função de ativação associada ao neurônio intermediário j . O fator restante envolvido neste cálculo, ou seja, o somatório em k depende de dois conjuntos de termos. O primeiro, os $\delta_k(n)$, requer conhecimento dos sinais de erro $e_k(n)$, para todos os neurônios que se encontram na camada imediatamente à direita do neurônio intermediário j e que estão diretamente conectados ao neurônio j . O segundo conjunto de termos, os $w_{kj}(n)$, consiste dos pesos sinápticos associados com estas conexões.

4.8 Treinamento por Lote

Em uma aplicação prática do algoritmo de retropropagação, o aprendizado resulta das muitas apresentações do conjunto de exemplos utilizados no treinamento para o *perceptron* de múltiplas camadas. O processo de aprendizagem é mantido de época em época até os pesos sinápticos e o *bias* se estabilizarem e o erro médio quadrado sobre todo o conjunto de treinamento convergir para um valor mínimo.

Conforme mencionado anteriormente, uma época é uma apresentação completa do conjunto de treinamento inteiro.

Para um determinado conjunto de treinamento, o aprendizado por retropropagação pode proceder de duas maneiras diferentes: treinamento por padrão ou por lote.

No modo de treinamento por padrão, também chamado de modo seqüencial, a atualização dos pesos é realizada após a apresentação de cada exemplo de treinamento à rede, já no modo de treinamento por lote, o ajuste dos pesos é efetuado após a apresentação de todos os exemplos de treinamento que constituem uma época.

Neste trabalho, foi utilizado o modo de treinamento por lote, por isso, a seguir encontra-se uma breve descrição desta forma de treinamento. Para uma época particular, define-se a função objetivo como o erro médio quadrado das equações (4.9) e (4.10), reproduzidos aqui na forma composta:

$$E_{med} = \frac{1}{2N} \sum_{n=1}^N \sum_{j \in C} e_j^2(n), \quad (4.32)$$

onde o sinal de erro $e_j(n)$ é relativo ao neurônio de saída j do exemplo de treinamento n , o qual é definido pela equação (4.8). O erro $e_j(n)$ é igual à diferença entre $d_j(n)$ e $y_j(n)$, que representam o j -ésimo elemento do vetor resposta desejada $\mathbf{d}(n)$ e o valor correspondente da saída da rede, respectivamente. Na equação (4.32), o somatório interno em relação a j é realizado sobre todos os neurônios da camada de saída de rede, enquanto que o somatório externo em relação a n é realizado sobre todo o conjunto de treinamento da época considerada. Para um parâmetro da taxa de aprendizagem η , o ajuste aplicado ao peso sináptico w_{ji} , conectando o neurônio i ao neurônio j , é definido pela regra delta

$$\Delta w_{ji} = -\eta \frac{\partial E_{med}}{\partial w_{ji}} = -\frac{\eta}{N} \sum_{n=1}^N e_j(n) \frac{\partial e_j(n)}{\partial w_{ji}} \quad (4.33)$$

Para calcular a derivada parcial $\frac{\partial e_j(n)}{\partial w_{ji}}$, procede-se da mesma forma como

anteriormente. De acordo com a equação (4.33), no modo por lote, o ajuste de peso Δw_{ji} é feito somente após o conjunto de treinamento inteiro ter sido apresentado à rede.

4.9 A Taxa de Aprendizagem

O algoritmo de retropropagação fornece uma "aproximação" para a trajetória no espaço de pesos calculada pelo método da descida mais íngreme (HAYKIN, 1999). Quanto menor for o parâmetro da taxa de aprendizagem η , menor serão as variações dos pesos sinápticos da rede, de uma iteração para a outra, e mais suave será a trajetória no espaço de pesos. Esta melhoria, no entanto, é obtida à custa de uma taxa de aprendizagem lenta. Por outro lado, se η for muito grande, para acelerar a taxa de aprendizagem, as grandes modificações nos pesos sinápticos resultantes podem tornar o processo de treinamento instável.

Um método simples de aumentar a taxa de aprendizagem, evitando entretanto o perigo de instabilidade, é modificar a regra delta da equação (4.20) incluindo um termo de momento (RUMELHART, HINTON e WILLIAMS, 1986), obtendo

$$\Delta w_{ji}(n) = \alpha \Delta w_{ji}(n-1) + \eta \delta_j(n) y_i(n) \quad (4.34)$$

onde α é usualmente um número positivo chamado constante de momento. A equação (4.34) é chamada de regra delta generalizada; ela inclui a regra delta da equação (4.20) como um caso especial, ou seja, com $\alpha = 0$.

4.10 Conjunto de Teste e Generalização

Um dos aspectos mais importantes de redes neurais é sua habilidade de generalização. Diz-se que uma rede generaliza bem quando o mapeamento de entrada/saída computado pela rede for correto (ou aproximadamente correto) para dados de teste não utilizados para o treinamento da rede.

Durante o treinamento, a rede neural aprende a partir dos exemplos incluídos no conjunto de treinamento. Depois de treinada, ou seja, depois de definido um conjunto de pesos e *bias* que minimiza o erro médio quadrático (equação 4.10), um conjunto de teste é apresentado à rede. Esse conjunto servirá para avaliar a capacidade de generalização da rede, ou seja, se ela fornece resultados aceitáveis para padrões não contidos no conjunto de treinamento. Isso evita que a rede "decore" os exemplos do conjunto de treinamento, ou seja, que ela forneça boa aproximação para os exemplos de treinamento mas não adquira as características básicas do problema em questão.

É importante observar que estas considerações são válidas quando o conjunto de teste é gerado a partir da mesma população usada para gerar o conjunto de treinamento. Para garantir que os dois conjuntos sejam representativos do mesmo fenômeno, eles foram gerados aleatoriamente a partir da base de dados formada através das simulações numéricas. Nesse trabalho, a divisão adotada foi de 80% dos exemplos para o conjunto de treinamento e 20% para o conjunto de teste.

A generalização é influenciada por três fatores:

- (i) o tamanho do conjunto de treinamento, e o quão representativo do ambiente de interesse ele é;
- (ii) a arquitetura da rede neural;
- (iii) a complexidade física do problema em questão.

Evidentemente, não se tem controle sobre o último fator. No contexto dos outros dois fatores, pode-se ver a questão da generalização sob duas perspectivas diferentes (HUSH e HORNE, 1993):

- (i) a arquitetura da rede é fixa (provavelmente de acordo com a complexidade física do problema relacionado), e a questão a ser resolvida é determinar o tamanho do conjunto de treinamento necessário para que ocorra uma boa generalização;
- (ii) o tamanho do conjunto de treinamento é fixo, e a questão de interesse é determinar a melhor arquitetura para alcançar uma boa generalização.

Estes dois pontos de vista são válidos em seus aspectos individuais.

4.11 Normalização dos Dados

O treinamento de uma rede neural artificial pode ser mais eficiente se determinadas etapas de pré-processamento dos dados de entrada e saída desejada forem executadas. Uma etapa bastante importante é a de normalização dos dados. Isto se deve ao fato de que as funções de ativação dos neurônios normalmente são limitadas inferior e superiormente. Ou seja, se não for realizada a normalização e os dados não estiverem entre os valores de máximo e mínimo da função, estes já entram saturados na rede, não ativando os neurônios e assim não sendo processados. Nesta seção, serão descritas algumas formas de normalização, sendo que a utilizada neste trabalho foi a divisão pelo máximo valor absoluto.

4.11.1 Normalização através da Divisão pelo Máximo Valor Absoluto

Como dito anteriormente, é muito importante a normalização dos dados de entrada e dos dados da saída desejada da rede neural, de modo que estes estejam sempre dentro de um intervalo especificado. Neste trabalho, foi utilizada a função `premnmx` do *Neural Networks Toolbox* do MATLAB 6.0, que faz com que os dados de entrada e saída da rede neural estejam no intervalo $[-1, 1]$.

As entradas e as saídas desejadas da rede são dados originalmente em matrizes. A normalização, neste caso, é realizada dividindo-se cada elemento das matrizes de entrada e de saída pelo máximo valor absoluto de cada linha das matrizes. Portanto, as entradas e saídas desejadas normalizadas possuem valores no intervalo $[-1, 1]$.

Assim, utilizando a função `premnmx` para normalizar as entradas e as saídas desejadas, a rede estará treinada para produzir saídas no intervalo $[-1, 1]$. Logo, as saídas da rede corresponderão às saídas desejadas normalizadas. Para converter as saídas da rede para as mesmas unidades usadas para as saídas originais, deve-se

utilizar a função `postmnmx`, também do *Neural Networks Toolbox* do MATLAB 6.0, obtendo-se, portanto, saídas não-normalizadas que serão comparadas às saídas desejadas, dadas originalmente à rede neural.

4.11.2 Normalização através da Média e do Desvio Padrão

Uma outra forma de pré-processamento para entradas e saídas desejadas da rede neural é normalizar a média e o desvio padrão do conjunto de treinamento. Este procedimento é executado pela função `prestd` do *Neural Networks Toolbox* do MATLAB 6.0, que normaliza as entradas e as saídas desejadas da rede neural de modo que tenham média igual a zero e desvio padrão igual a um.

Deste modo, a rede está treinada para produzir saídas com média igual a zero e desvio padrão igual a um, ou seja, as saídas da rede corresponderão às saídas desejadas normalizadas. Para converter as saídas da rede para as mesmas unidades usadas para as saídas originais, deve-se utilizar a função `poststd`, também do *Neural Networks Toolbox* do MATLAB 6.0, obtendo-se, portanto, saídas não-normalizadas que serão comparadas às saídas desejadas.

4.11.3 Análise das Componentes Principais

Em algumas situações, a dimensão do vetor da entrada é grande, mas os componentes dos vetores são altamente correlacionados, ou seja, são redundantes. Nesta situação, é útil reduzir a dimensão dos vetores da entrada (M1 *et al.* 2005). Um procedimento eficaz para executar esta operação é a análise da componente principal. Esta técnica tem três efeitos: ortogonalizar os componentes dos vetores de entrada (de modo que sejam não correlacionados um com o outro); obter as componentes ortogonais resultantes (componentes principais) de modo que aquelas com maior variância venham primeiramente; e dar uma indicação daquelas componentes que menos contribuem com a variação na série de dados. A função `prepca` do *Neural Networks Toolbox* do MATLAB 6.0 executa uma análise do com-

ponente principal.

Primeiramente os vetores da entrada são normalizados, usando a função `prestd`, de modo que tenham média igual a zero e desvio padrão igual a um. Este é um procedimento padrão ao usar-se componentes principais. A função `prepca` necessita de um argumento, por exemplo, 0.02. Isto significa que o `prepca` elimina aqueles componentes principais que contribuem menos de 2% à variação total na série de dados.

4.12 Modelo Neural

As redes utilizadas neste trabalho foram implementadas no *Neural Networks Toolbox* do MATLAB 6.0 em um computador AMD Athlon (TM) XP 2600+ com 2,08 GHz e 1,5 GB de RAM.

A finalidade das redes aqui implementadas foi estimar alguns parâmetros importantes dos escoamentos bifásicos óleo-água, como a fração volumétrica e a área interfacial. Implementou-se duas redes para determinar estes parâmetros: uma para mapear os dados das simulações numéricas em um tubo de seção quadrada e outra para mapear os dados dos testes em uma tubulação com geometria circular.

A arquitetura das redes foi definida a partir de estudos preliminares, da seguinte maneira: camada de entrada, duas camadas intermediárias e uma camada de saída. O número de neurônios das camadas intermediárias foi ajustado com o objetivo de capacitar as redes a aprender tarefas complexas pela extração progressiva de informações significantes dos padrões de entrada (HAYKIN, 1999). Conforme mencionado anteriormente, o número de neurônios na camada intermediária não é fixado previamente, ou seja, pode-se determinar um número qualquer de neurônios e posteriormente, de acordo com a necessidade da rede, pode-se aumentar ou diminuir a quantidade de neurônios na camada intermediária.

Para o mapeamento de entrada-saída dos dados referentes à geometria quadrada foi utilizada uma rede neural 24 - 16 - 12 - 2. Já para os dados concernentes à geometria circular, a arquitetura de rede neural utilizada foi 24 - 15 - 11 - 2. As

funções de ativação utilizadas nas duas redes são a tangente sigmoidal (figura 4.9) nas camadas intermediárias, dada por

$$\varphi(v) = \frac{2}{(1 + \exp(-2v))} - 1 \quad (4.35)$$

e a função linear (figura 4.10) na camada de saída, expressa por

$$\varphi(v) = v. \quad (4.36)$$

Cada neurônio da camada de saída é responsável por estimar um dos seguintes parâmetros do escoamento bifásico disperso: área interfacial e fração volumétrica.

Figura 4.9: Função de ativação utilizada nas camadas intermediárias das redes implementadas neste trabalho.

Figura 4.10: Função de ativação utilizada na camada de saída das redes implementadas neste trabalho.

O procedimento de treinamento constituiu em adquirir sinais de pressão acústica do escoamento de interesse para cada simulação numérica realizada e extrair algumas particularidades que caracterizam a complexidade do sinal em um nível suficiente para a inferência da área interfacial e da fração volumétrica para compor o conjunto de treinamento. Os dados utilizados no treinamento das redes são dispostos em duas matrizes: a matriz de entrada e a matriz de saída.

No caso da geometria quadrada, a matriz de entrada tem 24 linhas por 4000 colunas, onde 24 é o número de entradas, que significa 6 características extraídas de cada um dos 4 sinais obtidos em cada teste (a característica T foi descartada no treinamento da rede neural, conforme explicação no capítulo 5, e foram utilizadas somente as características T_1 , T_2 , T_3 , A_1 , A_2 e A_3), e 4000 é o número de exemplos utilizados no treinamento da rede neural. Cada exemplo corresponde a um valor conhecido e imposto de fração volumétrica e área interfacial. A matriz de saída tem 2 linhas por 4000 colunas, onde 2 é o número de saídas desejadas, sendo que na primeira linha estão os valores da fração volumétrica e na segunda encontram-se os valores da área interfacial de cada exemplo, e 4000 é o número de exemplos. Para a generalização, foram apresentados 1000 exemplos à rede, diferentes daqueles apresentados no treinamento. Análogo para a rede referente à geometria circular.

Resultados

O objetivo deste capítulo é apresentar os resultados das simulações numéricas realizadas a partir da ferramenta construída para o monitoramento acústico de escoamentos multifásicos. Também são expostos os resultados do treinamento e da generalização das redes neurais.

5.1 Resultados das Simulações Numéricas

As figuras 5.1 e 5.2 ilustram dois casos simulados, onde o contraste acústico ^{*} é de 1:1000 para ressaltar a propagação da energia acústica no interior do domínio. Esses são dados não reais, somente utilizados para uma melhor visualização dos resultados das simulações. O primeiro caso refere-se à geometria quadrada com uma inclusão (gota) igualmente quadrada localizada no centro do domínio. Pode-se observar o deslocamento das ondas acústicas para as posições laterais do domínio, devido à inclusão. A segunda simulação refere-se à geometria circular com uma inclusão segmental posicionada à esquerda do centro do domínio. Nesse caso, também pode-se observar deslocamento das ondas acústicas.

Nos demais testes numéricos, usados para obter os sinais de resposta, utilizou-se um contraste acústico real, com a velocidade de propagação acústica na água igual a $1,48 \frac{Km}{s}$ e no óleo igual a $1,7 \frac{Km}{s}$.

^{*}Razão entre as impedâncias de propagação acústica.

Figura 5.1: Simulação numérica da propagação da energia acústica em uma seção quadrada com uma inclusão localizada no centro do domínio.

Figura 5.2: Simulação numérica da propagação da energia acústica em uma seção circular com uma inclusão à esquerda do centro do domínio.

5.1.1 Simulações em Tubulação de Seção Quadrada

A análise numérica foi efetuada em uma seção transversal do tubo quadrado de largura e comprimento iguais a 1. Particionou-se o domínio conforme a figura 5.3, fixando um número de pontos $N = 30$ e $M = 30$, para as direções x e y , respectivamente. No problema direto, a distribuição de velocidade acústica é conhecida no interior do tubo, ou seja, a função velocidade é definida para cada ponto da malha.

Figura 5.3: Malha utilizada no domínio quadrado.

Nas figuras a seguir, encontram-se os sinais de pressão acústica obtidos nos 4 elementos de medida para uma simulação numérica onde a fração volumétrica é $\alpha = 0,098$ e a área interfacial é $\Gamma = 0,867$. Os sinais obtidos originalmente nas simulações foram afetados por baixas frequências devido ao algoritmo numérico (artefato numérico). Na figura 5.4 encontram-se tais sinais de resposta. Como consequência desse artefato, o sinal pode ser distorcido, dificultando a determinação de suas maiores amplitudes. Por isso, foi utilizado um filtro de Fourier atuando como passa-alta para eliminar os componentes de baixa frequência, pois estes não são originados da equação que modela o problema. A frequência de corte, neste caso, foi de 1000 Hz. Os sinais de resposta após a filtragem podem ser vistos na figura 5.5. As características T , T_1 , T_2 , T_3 , A_1 , A_2 e A_3 , citadas anteriormente, foram extraídas dos sinais filtrados.

Figura 5.4: Sinais acústicos obtidos nos testes da geometria quadrada.

Figura 5.5: Sinais acústicos filtrados obtidos nos testes da geometria quadrada.

Foram realizados 5000 testes numéricos simulando escoamento bifásico disperso em um tubo de seção quadrada, sendo 4000 testes numéricos utilizados no treinamento da rede neural e mais 1000 utilizados para testar a capacidade de generalização da rede neural. Para tanto, impôs-se que a fração volumétrica variasse de maneira uniforme entre 0 e 0,3 conforme mencionado anteriormente. A área interfacial resultante limitou-se entre 0 e 2,5 devido à malha utilizada nas simulações. A tabela 5.1 ilustra alguns destes casos, mostrando também as respectivas características extraídas dos sinais de resposta. Em seguida, fez-se uma análise dos histogramas destas características, mostrando que algumas delas são mais ricas em informação do que outras, sendo conveniente para o treinamento da rede excluir aquelas menos relevantes. Portanto, as características a serem consideradas doravante são as seguintes: T_1 , T_2 , T_3 , A_1 , A_2 e A_3 , por terem mostrado sofrer maior influência da distribuição interna de propagação acústica. A característica T foi descartada justamente por ser a característica que menos se alterou com a mudança da distribuição interna de propagação acústica. Abaixo são mostrados alguns destes histogramas a título de ilustração, sendo que na figura 5.6 estão os histogramas da amplitude do maior pico dos sinais de resposta obtidos nos elementos de medida 1, 2, 3 e 4, respectivamente. Já na figura 5.7, encontram-se os histogramas do tempo de chegada da onda acústica em cada um dos 4 elementos de medida.

Tabela 5.1: Características extraídas dos sinais de resposta para diferentes testes numéricos em uma tubulação quadrada.

Distribuição de gotas	Elemento de medida 1	Elemento de medida 2	Elemento de medida 3	Elemento de medida 4
$\alpha = 0,000$ $\Gamma = 0,000$	$T = 0,155$ $T_1 = 0,368$ $T_2 = 1,748$ $T_3 = 2,802$ $A_1 = 0,594$ $A_2 = 0,376$ $A_3 = 0,249$	$T = 0,525$ $T_1 = 0,794$ $T_2 = 1,283$ $T_3 = 3,181$ $A_1 = 0,585$ $A_2 = 0,351$ $A_3 = 0,247$	$T = 0,536$ $T_1 = 0,803$ $T_2 = 1,261$ $T_3 = 3,179$ $A_1 = 0,563$ $A_2 = 0,340$ $A_3 = 0,238$	$T = 0,173$ $T_1 = 0,392$ $T_2 = 1,745$ $T_3 = 2,792$ $A_1 = 0,585$ $A_2 = 0,368$ $A_3 = 0,257$
$\alpha = 0,027$ $\Gamma = 0,167$	$T = 0,144$ $T_1 = 0,343$ $T_2 = 0,455$ $T_3 = 1,743$ $A_1 = 0,576$ $A_2 = 0,376$ $A_3 = 0,332$	$T = 0,512$ $T_1 = 0,776$ $T_2 = 1,274$ $T_3 = 3,163$ $A_1 = 0,544$ $A_2 = 0,387$ $A_3 = 0,269$	$T = 0,533$ $T_1 = 0,800$ $T_2 = 1,248$ $T_3 = 3,155$ $A_1 = 0,628$ $A_2 = 0,319$ $A_3 = 0,209$	$T = 0,173$ $T_1 = 0,394$ $T_2 = 1,730$ $T_3 = 2,778$ $A_1 = 0,703$ $A_2 = 0,391$ $A_3 = 0,284$
$\alpha = 0,093$ $\Gamma = 0,530$	$T = 0,155$ $T_1 = 0,370$ $T_2 = 1,720$ $T_3 = 0,489$ $A_1 = 0,625$ $A_2 = 0,370$ $A_3 = 0,326$	$T = 0,524$ $T_1 = 0,791$ $T_2 = 1,255$ $T_3 = 3,794$ $A_1 = 0,654$ $A_2 = 0,279$ $A_3 = 0,209$	$T = 0,530$ $T_1 = 0,791$ $T_2 = 1,254$ $T_3 = 3,140$ $A_1 = 0,627$ $A_2 = 0,406$ $A_3 = 0,266$	$T = 0,162$ $T_1 = 0,372$ $T_2 = 1,274$ $T_3 = 1,434$ $A_1 = 0,477$ $A_2 = 0,383$ $A_3 = 0,305$
$\alpha = 0,187$ $\Gamma = 1,167$	$T = 0,154$ $T_1 = 0,359$ $T_2 = 0,474$ $T_3 = 0,561$ $A_1 = 0,885$ $A_2 = 0,606$ $A_3 = 0,440$	$T = 0,501$ $T_1 = 0,766$ $T_2 = 0,479$ $T_3 = 1,227$ $A_1 = 0,654$ $A_2 = 0,313$ $A_3 = 0,288$	$T = 0,519$ $T_1 = 0,777$ $T_2 = 0,478$ $T_3 = 0,366$ $A_1 = 0,779$ $A_2 = 0,322$ $A_3 = 0,292$	$T = 0,168$ $T_1 = 0,378$ $T_2 = 0,489$ $T_3 = 1,667$ $A_1 = 0,742$ $A_2 = 0,517$ $A_3 = 0,382$
$\alpha = 0,271$ $\Gamma = 1,667$	$T = 0,150$ $T_1 = 0,352$ $T_2 = 1,648$ $T_3 = 0,470$ $A_1 = 0,627$ $A_2 = 0,312$ $A_3 = 0,275$	$T = 0,492$ $T_1 = 0,746$ $T_2 = 1,213$ $T_3 = 3,630$ $A_1 = 0,454$ $A_2 = 0,220$ $A_3 = 0,213$	$T = 0,506$ $T_1 = 0,759$ $T_2 = 1,803$ $T_3 = 0,478$ $A_1 = 0,423$ $A_2 = 0,184$ $A_3 = 0,175$	$T = 0,150$ $T_1 = 0,433$ $T_2 = 0,343$ $T_3 = 0,497$ $A_1 = 0,333$ $A_2 = 0,296$ $A_3 = 0,294$

Figura 5.6: Histogramas da maior amplitude dos sinais de pressão acústica obtidos nos elementos (a) 1, (b) 2, (c) 3 e (d) 4.

Figura 5.7: Histogramas do tempo de chegada da onda acústica dos sinais de pressão acústica obtidos nos elementos (a) 1, (b) 2, (c) 3 e (d) 4.

5.1.2 Simulações em Tubulação de Seção Circular

A análise foi realizada em uma seção transversal de um tubo circular de raio $R = 1$, considerando o domínio particionado conforme a figura 5.8. Para as direções r e θ foram fixados um número de pontos $N = 28$ e $M = 36$, respectivamente, seguindo as restrições referentes ao particionamento no intervalo $[0, 2\pi]$ comentadas no capítulo (3). No problema direto, a distribuição de velocidade acústica é conhecida no interior do tubo, ou seja, a função velocidade é definida para cada ponto da malha.

Figura 5.8: Malha utilizada no domínio circular.

Nas figuras a seguir, encontram-se os sinais de pressão acústica obtidos nos 4 elementos de medida para uma simulação numérica onde a fração volumétrica é $\alpha = 0,212$ e a área interfacial é $\Gamma = 0,713$. Novamente, os sinais obtidos originalmente nas simulações numéricas foram afetados por baixas frequências. Por isso, foi utilizado um filtro de Fourier atuando como passa-alta para eliminar os componentes de baixa frequência. A frequência de corte, neste caso, foi de 3000 Hz. Na figura 5.9 estão os sinais de resposta originais e na figura 5.10 estão os sinais após a filtragem. Como no caso da geometria quadrada, as características T_1 , T_2 , T_3 , A_1 , A_2 , A_3 e T foram extraídas dos sinais filtrados.

Figura 5.9: Sinais acústicos obtidos nos testes da geometria circular.

Figura 5.10: Sinais acústicos filtrados obtidos nos testes da geometria circular.

Neste caso, também foram realizados 5000 testes numéricos simulando escoamento bifásico disperso em uma tubulação de seção circular, sendo 4000 testes numéricos utilizados no treinamento da rede neural e mais 1000 utilizados para testar a capacidade de generalização da rede neural. Novamente, a fração volumétrica variou de maneira uniforme entre 0 e 0,3 e a área interfacial resultante limitou-se entre 0 e 2,1 devido à malha utilizada nas simulações. A tabela 5.2 ilustra alguns destes casos, mostrando também as respectivas características dos sinais de medida mapeados pela rede neural.

No caso da geometria circular, o elemento de medida 1 foi responsável pela excitação e também por obter um dos sinais acústicos analisados em cada teste, diferentemente do que aconteceu no caso da geometria quadrada, onde o pulso de excitação tipo Dirac foi aplicado em um dos vértices do quadrado e assim, os elementos de medida ficaram responsáveis apenas por adquirir os sinais de resposta. Isso explica a diferença entre o sinal acústico obtido no elemento de medida 1 e os que foram obtidos nos demais elementos, como pode ser observado nas figuras 5.9 e 5.10.

Neste caso, também foi realizada uma análise dos histogramas das características extraídas dos sinais acústicos. Concluiu-se que as características a serem consideradas doravante são as mesmas consideradas no caso da geometria quadrada, ou seja, T_1 , T_2 , T_3 , A_1 , A_2 e A_3 . A característica T foi descartada por sofrer pouca influência da distribuição interna de velocidade de propagação acústica.

Tabela 5.2: Características extraídas dos sinais de resposta para diferentes testes numéricos em uma tubulação circular.

Distribuição de gotas	Elemento de medida 1	Elemento de medida 2	Elemento de medida 3	Elemento de medida 4
$\alpha = 0,000$ $\Gamma = 0,000$	$T_1 = 0,315$ $T_2 = 1,146$ $T_3 = 2,843$ $A_1 = 78,581$ $A_2 = 35,252$ $A_3 = 23,530$	$T_1 = 13,890$ $T_2 = 47,264$ $T_3 = 41,637$ $A_1 = 12,680$ $A_2 = 11,740$ $A_3 = 11,642$	$T_1 = 26,754$ $T_2 = 37,291$ $T_3 = 55,207$ $A_1 = 18,890$ $A_2 = 18,398$ $A_3 = 17,057$	$T_1 = 13,890$ $T_2 = 47,264$ $T_3 = 41,637$ $A_1 = 12,680$ $A_2 = 11,740$ $A_3 = 11,642$
$\alpha = 0,106$ $\Gamma = 0,688$	$T_1 = 0,315$ $T_2 = 1,146$ $T_3 = 3,674$ $A_1 = 79,182$ $A_2 = 35,936$ $A_3 = 23,791$	$T_1 = 57,524$ $T_2 = 35,052$ $T_3 = 13,678$ $A_1 = 14,951$ $A_2 = 11,531$ $A_3 = 11,498$	$T_1 = 26,721$ $T_2 = 55,058$ $T_3 = 14,301$ $A_1 = 18,699$ $A_2 = 16,135$ $A_3 = 16,135$	$T_1 = 41,527$ $T_2 = 13,893$ $T_3 = 47,234$ $A_1 = 14,848$ $A_2 = 12,715$ $A_3 = 11,977$
$\alpha = 0,155$ $\Gamma = 0,815$	$T_1 = 0,316$ $T_2 = 1,147$ $T_3 = 2,848$ $A_1 = 75,870$ $A_2 = 35,176$ $A_3 = 23,656$	$T_1 = 13,754$ $T_2 = 60,529$ $T_3 = 27,951$ $A_1 = 14,965$ $A_2 = 14,255$ $A_3 = 13,778$	$T_1 = 55,040$ $T_2 = 37,144$ $T_3 = 83,822$ $A_1 = 19,456$ $A_2 = 15,975$ $A_3 = 15,941$	$T_1 = 13,972$ $T_2 = 41,414$ $T_3 = 1,726$ $A_1 = 13,268$ $A_2 = 12,045$ $A_3 = 10,555$
$\alpha = 0,200$ $\Gamma = 0,942$	$T_1 = 0,319$ $T_2 = 1,149$ $T_3 = 2,842$ $A_1 = 70,733$ $A_2 = 35,834$ $A_3 = 23,092$	$T_1 = 41,328$ $T_2 = 83,240$ $T_3 = 27,995$ $A_1 = 15,743$ $A_2 = 14,081$ $A_3 = 13,009$	$T_1 = 26,639$ $T_2 = 37,160$ $T_3 = 83,957$ $A_1 = 19,035$ $A_2 = 18,430$ $A_3 = 17,932$	$T_1 = 86,142$ $T_2 = 41,341$ $T_3 = 51,778$ $A_1 = 13,959$ $A_2 = 13,628$ $A_3 = 13,483$
$\alpha = 0,200$ $\Gamma = 0,968$	$T_1 = 0,327$ $T_2 = 1,155$ $T_3 = 1,992$ $A_1 = 53,043$ $A_2 = 37,879$ $A_3 = 26,524$	$T_1 = 41,431$ $T_2 = 60,455$ $T_3 = 1,728$ $A_1 = 18,972$ $A_2 = 17,329$ $A_3 = 17,025$	$T_1 = 54,947$ $T_2 = 37,009$ $T_3 = 66,769$ $A_1 = 28,191$ $A_2 = 26,503$ $A_3 = 25,698$	$T_1 = 83,493$ $T_2 = 27,986$ $T_3 = 26,163$ $A_1 = 24,654$ $A_2 = 23,815$ $A_3 = 12,245$

5.2 Resultados das Redes Neurais Artificiais

Os resultados do treinamento das redes neurais mostraram que elas são capazes de reproduzir a relação entrada/saída dos dados do conjunto de treinamento, tanto para a fração volumétrica quanto para a área interfacial. Para avaliar a capacidade de generalização das RNAs, os dados característicos de testes numéricos com novas distribuições de velocidade local de propagação acústica foram apresentados às redes neurais, e pôde-se observar uma boa correlação entre os dados de entrada e de saída para os casos da fração volumétrica e da área interfacial, conforme pode ser visto nas subseções seguintes.

5.2.1 Geometria Quadrada

Conforme mencionado anteriormente, foram realizadas 5000 simulações numéricas para compor a base de dados da qual extraiu-se as características dos sinais fornecidas à RNA. O conjunto de treinamento foi composto por 4000 exemplos, onde cada exemplo representa um teste numérico realizado. Cada teste numérico possui uma configuração específica da distribuição de velocidade local de propagação acústica, uma vez que as gotas foram geradas aleatoriamente. Para cada configuração, ou seja, para cada exemplo apresentado à rede, obteve-se um valor para a fração volumétrica e um valor para a área interfacial, além das 6 características extraídas de cada um dos 4 sinais acústicos obtidos em cada teste. Assim, cada exemplo do conjunto de treinamento é formado por 24 entradas, que são as características dos sinais, e por duas saídas, que são a fração volumétrica e a área interfacial do escoamento bifásico.

Para avaliar a habilidade de generalização da RNA, os dados característicos de 1000 novos testes numéricos com novas distribuições de velocidade local de propagação acústica foram apresentados à rede neural.

No caso da geometria quadrada, a taxa de aprendizado utilizada na rede neural foi de 0,01. O tempo de treinamento foi de aproximadamente 9 horas e foram realizadas 100000 épocas durante o treinamento, uma vez que o erro do

treinamento da rede neural se estabilizou, conforme ilustra a figura 5.11. Apesar de o erro de treinamento da rede não ser tão pequeno (da ordem de 10^{-1}), ela conseguiu fornecer bons resultados para os exemplos de teste, que serão mostrados adiante.

Figura 5.11: Decaimento do erro durante o treinamento da rede para estimar a fração volumétrica e a área interfacial para os dados obtidos na geometria quadrada.

A seguir, as tabelas 5.3 e 5.4 exibem algumas respostas da rede neural para a fração volumétrica e para a área interfacial de alguns exemplos do conjunto de teste.

Tabela 5.3: Respostas da rede neural para a fração volumétrica em alguns testes realizados na geometria quadrada.

	valores reais	valores da RNA
teste 1	0,013	0,0164
teste 2	0,027	0,0262
teste 3	0,031	0,0292
teste 4	0,036	0,0376
teste 5	0,040	0,0457
teste 6	0,053	0,0524
teste 7	0,067	0,0643
teste 8	0,129	0,1280
teste 9	0,138	0,1409
teste 10	0,151	0,1490
teste 11	0,204	0,2093
teste 12	0,227	0,2263
teste 13	0,240	0,2403
teste 14	0,267	0,2776
teste 15	0,276	0,2778

Tabela 5.4: Resposta da rede neural para a área interfacial em alguns testes realizados na geometria quadrada.

	valores reais	valores da RNA
teste 1	0,267	0,2780
teste 2	0,300	0,2934
teste 3	0,333	0,3531
teste 4	0,367	0,3638
teste 5	0,400	0,4173
teste 6	0,600	0,5954
teste 7	0,767	0,7437
teste 8	0,833	0,8248
teste 9	1,000	1,0278
teste 10	1,067	1,0418
teste 11	1,133	1,1471
teste 12	1,533	1,5313
teste 13	1,733	1,7485
teste 14	1,800	1,7747
teste 15	2,033	1,9040

Os histogramas da figura 5.12 mostram que os valores da fração volumétrica e da área interfacial utilizados no treinamento da rede neural estão entre os intervalos $[0; 0,3]$ e $[0; 2,5]$, respectivamente. Pode-se observar que os valores de α e Γ dos exemplos do conjunto de teste também pertencem a estes intervalos, indicando que o conjunto de teste foi gerado a partir da mesma população usada para gerar o conjunto de treinamento e garantindo que os dois conjuntos sejam representativos do mesmo fenômeno. Nota-se ainda que o simulador numérico desenvolvido neste trabalho não gerou muitos valores maiores que 2,2 para a área interfacial, o que dificultou, mas não impediu o treinamento da rede neural. Isso ocorreu devido à malha utilizada nas simulações numéricas. Uma malha mais refinada poderia gerar um maior número de gotas, e portanto, aumentaria o valor da área interfacial.

Figura 5.12: Histogramas da fração volumétrica e da área interfacial calculadas para os exemplos de treinamento e de teste, respectivamente, na geometria quadrada.

A figura 5.13 apresenta os valores da fração volumétrica α calculados pela rede neural, referentes aos exemplos contidos no conjunto de teste. Os valores reais, obtidos nas simulações numéricas, são representados no eixo das abscissas, x , enquanto que os valores estimados pela rede neural são representados no eixo das ordenadas, y .

A reta $x = y$ (linha vermelha no gráfico), representa então o resultado ideal, ou seja, os valores estimados iguais aos valores reais. Os pontos azuis plotados no gráfico são os valores estimados através da rede neural. A equação para a reta representada no gráfico pela linha preta é $y = mx + b$, onde m é o grau de

Figura 5.13: Comparação entre os valores estimados pela rede neural e os valores reais para a fração volumétrica (geometria quadrada).

inclinação da reta e b é a intersecção da reta com o eixo y . As equações para calcular m e b podem ser encontradas em Young (1962) e Doebelin (1990) e são dadas por:

$$m = \frac{K \sum_{i=1}^K x_i y_i - (\sum_{i=1}^K x_i)(\sum_{i=1}^K y_i)}{K \sum_{i=1}^K x_i^2 - (\sum_{i=1}^K x_i)^2} \quad (5.1)$$

$$b = \frac{(\sum_{i=1}^K y_i)(\sum_{i=1}^K x_i^2) - (\sum_{i=1}^K x_i y_i)(\sum_{i=1}^K x_i)}{K \sum_{i=1}^K x_i^2 - (\sum_{i=1}^K x_i)^2} \quad (5.2)$$

onde K é o número total de pares (x_i, y_i) , em que x_i representa o valor real da fração volumétrica e y_i o valor estimado pela RNA para o teste i , para $i = 1, \dots, 1000$.

Neste caso, tem-se $m = 0,8936$ e $b = 0,0127$. Portanto, a reta (linha preta na figura 5.13)

$$y = 0,8936x + 0,0127 \quad (5.3)$$

é a que melhor representa a correlação entre os valores estimados pela RNA e os resultados reais para α . O coeficiente de correlação r é calculado por (YOUNG, 1962):

$$r = \frac{K \sum_{i=1}^K x_i y_i - \sum_{i=1}^K x_i \sum_{i=1}^K y_i}{[K \sum_{i=1}^K x_i^2 - (\sum_{i=1}^K x_i)^2]^{\frac{1}{2}} [K \sum_{i=1}^K y_i^2 - (\sum_{i=1}^K y_i)^2]^{\frac{1}{2}}} \quad (5.4)$$

Assim, $r = 1$ significa uma correlação perfeita e $r = 0$ significa que não há correlação entre os dados. Se há uma correlação imperfeita, espera-se um valor de r entre 0 e 1. Para os dados referentes à fração volumétrica, o coeficiente de correlação entre os valores foi $r = 0,93$, significando uma boa relação entre os dados de entrada e de saída fornecidos pela rede neural.

Para cada par (x_i, y_i) , pode-se definir o erro local da fração volumétrica e_i como

$$e_i = x_i - y_i \quad (5.5)$$

onde x_i e y_i representam o valor real e estimado pela RNA, respectivamente, da fração volumétrica para o i -ésimo teste numérico, para $i = 1, \dots, 1000$. O desvio padrão (medida do grau de dispersão dos valores da fração volumétrica em relação ao valor médio) de e_i é $\sigma_e = 0,032$.

Deste modo, é possível plotar as retas $y = mx + (b \pm 3\sigma_e)$, ou seja,

$$y = 0,8936x + 0,1093 \quad (5.6)$$

$$y = 0,8936x - 0,0839 \quad (5.7)$$

que formam um envelope em torno dos pontos (x_i, y_i) , como pode ser visto no gráfico da figura 5.14, mostrando o intervalo de confiança dos resultados obtidos pela RNA. Isso significa que para um determinado valor de α , a resposta da rede neural estará no intervalo $[\alpha - 3\sigma_e, \alpha + 3\sigma_e]$.

Figura 5.14: Intervalo de confiança dos resultados fornecidos pela RNA para os valores da fração volumétrica na geometria quadrada.

Também foram estimados os valores da área interfacial Γ . O resultado é mostrado no gráfico da figura 5.15. Assim como no caso da fração volumétrica, a melhor reta de regressão é dada por:

$$y = 0,9029x + 0,0599 \quad (5.8)$$

onde $m = 0,82$ e $b = 0,1907$ foram calculados pelas equações (5.1) e (5.2), respectivamente. O índice de correlação entre os valores estimados e os valores reais (valores das simulações numéricas) foi de 0,92 (equação 5.4), também indicando uma boa correlação entre os dados.

Figura 5.15: Comparação entre os valores estimados pela rede neural e os valores reais para a área interfacial (geometria quadrada).

Neste caso, utilizando o erro local definido na equação (5.5) para x_i e y_i representando os valores real e estimado pela RNA, respectivamente, da área interfacial para o i -ésimo teste numérico, para $i = 1, \dots, 1000$, pode-se calcular o desvio padrão do erro local, obtendo-se $\sigma_e = 0,22$. Assim, é possível plotar as retas $y = mx + (b \pm 3\sigma_e)$, ou seja,

$$y = 0,9029x + 0,7329 \quad (5.9)$$

$$y = 0,9029x - 0,6132 \quad (5.10)$$

que formam um envelope em torno dos pontos (x_i, y_i) , como pode ser visto no gráfico da figura 5.16, mostrando o intervalo de confiança dos resultados obtidos pela RNA para a área interfacial. Isso mostra que para um determinado valor de Γ , a resposta da rede neural estará no intervalo $[\Gamma - 3\sigma_e, \Gamma + 3\sigma_e]$.

Figura 5.16: Intervalo de confiança dos resultados fornecidos pela RNA para os valores da área interfacial na geometria quadrada.

Como pode ser observado, os resultados obtidos através da rede neural para os exemplos da geometria quadrada foram bastante satisfatórios, uma vez que ela foi capaz de mostrar a tendência da fração volumétrica e da área interfacial em quase todo o espaço de estudo.

5.2.2 Geometria Circular

No caso da geometria circular, também foram realizados 5000 testes, sendo que conjunto de treinamento foi composto por 4000 exemplos e o conjunto de teste continha 1000 exemplos, diferentes daqueles utilizados no conjunto de treinamento. A taxa de aprendizado utilizada na rede neural foi de 0,1. O tempo de treinamento foi de 15 horas aproximadamente. A rede foi treinada durante 200000 épocas, pois depois disto a variação do erro estabilizou conforme é mostrado na figura 5.17.

A seguir, as tabelas 5.5 e 5.6 exibem algumas respostas fornecidas pela rede neural para a fração volumétrica e para a área interfacial de alguns exemplos do conjunto de teste, para a geometria circular.

Figura 5.17: Decaimento do erro durante o treinamento da rede para estimar a fração volumétrica e a área interfacial para os dados obtidos na geometria circular.

Tabela 5.5: Resposta da rede neural para a fração volumétrica no caso da geometria circular.

	valores reais	valores da RNA
teste 1	0,013	0,0180
teste 2	0,032	0,0369
teste 3	0,042	0,0447
teste 4	0,110	0,1244
teste 5	0,143	0,1384
teste 6	0,169	0,1617
teste 7	0,175	0,1761
teste 8	0,182	0,1843
teste 9	0,201	0,1963
teste 10	0,205	0,2208
teste 11	0,214	0,2111
teste 12	0,234	0,2464
teste 13	0,260	0,2495
teste 14	0,276	0,2656
teste 15	0,283	0,2815

Tabela 5.6: Respostas da rede neural para a área interfacial no caso da geometria circular.

	valores reais	valores da RNA
teste 1	0,114	0,1260
teste 2	0,159	0,1595
teste 3	0,205	0,2011
teste 4	0,205	0,2065
teste 5	0,273	0,2973
teste 6	0,318	0,3178
teste 7	0,546	0,5850
teste 8	0,614	0,6211
teste 9	0,819	0,8100
teste 10	0,887	0,8909
teste 11	1,160	1,1212
teste 12	1,182	1,1628
teste 13	1,251	1,2778
teste 14	1,592	1,4516
teste 15	1,955	1,5032

Os histogramas da figura 5.18 mostram que os valores da fração volumétrica e da área interfacial utilizados no treinamento da rede neural estão entre os intervalos $[0; 0,3]$ e $[0; 2,1]$, respectivamente. Pode-se observar que os valores de α e Γ dos exemplos do conjunto de teste também pertencem a estes intervalos, indicando que o conjunto de teste foi gerado a partir da mesma população usada para gerar o conjunto de treinamento. Nota-se ainda que o simulador numérico desenvolvido neste trabalho não gerou muitos valores maiores que 1,5 para a área interfacial na geometria circular, dificultando o treinamento da rede neural. Isso ocorreu devido à malha utilizada nas simulações numéricas. Neste caso, uma malha mais refinada também poderia gerar um maior número de gotas, aumentando assim, o valor da área interfacial.

Figura 5.18: Histogramas da fração volumétrica e da área interfacial calculadas para os exemplos de treinamento e de teste, respectivamente, na geometria circular.

Na figura 5.19 está a comparação entre os valores estimados pela rede neural e os valores reais para a fração volumétrica α para os dados extraídos dos sinais obtidos na geometria circular, referentes aos exemplos contidos no conjunto de teste. Os valores reais, obtidos nas simulações numéricas, são representados no eixo das abscissas, x , enquanto que os valores estimados pela rede neural são representados no eixo das ordenadas, y . No gráfico, a reta $x = y$, representa o resultado ideal, isto é, os valores estimados iguais aos valores reais. Os pontos azuis são os valores estimados através da rede neural.

Figura 5.19: Comparação entre os valores estimados pela rede neural e os valores reais para a fração volumétrica (geometria circular).

Analogamente ao que foi mencionado no caso da geometria quadrada, a reta representada no gráfico pela linha preta é a que melhor representa a correlação entre os valores estimados pela RNA e os valores reais para α . E sua equação é:

$$y = 0,6054x + 0,0674 \quad (5.11)$$

sendo que $m = 0,6054$ e $b = 0,0674$ foram calculados através das equações (5.1) e (5.2), respectivamente. O coeficiente de correlação, neste caso, foi $r = 0,7$, calculado por (5.4). Neste caso, a relação entre os dados de entrada e de saída de rede neural não é tão boa quanto no caso da geometria quadrada. Mesmo assim, pode-se observar que os dados estão correlacionados e que a rede neural é capaz de fornecer bons resultados para a fração volumétrica.

Novamente, utilizando o erro local da fração volumétrica e_i como definido na equação (5.5), para cada par (x_i, y_i) , obtém-se $\sigma_e = 0,068$, que é o desvio padrão de e_i .

Deste modo, como pode ser visto no gráfico da figura 5.20, é possível plotar as retas $y = mx + (b \pm 3\sigma_e)$, ou seja,

Figura 5.20: Intervalo de confiança dos resultados fornecidos pela RNA para os valores da fração volumétrica na geometria circular.

$$y = 0,6054x + 0,2559 \quad (5.12)$$

$$y = 0,6054x - 0,1212 \quad (5.13)$$

que formam um envelope em torno dos pontos (x_i, y_i) , mostrando o intervalo de confiança dos resultados obtidos pela RNA. Isso significa que para um determinado valor de α , a resposta da rede neural estará no intervalo $[\alpha - 3\sigma_e, \alpha + 3\sigma_e]$.

A rede neural também estimou os valores da área interfacial Γ , na geometria circular. O resultado é mostrado no gráfico da figura 5.21. Neste caso, a melhor reta de regressão é dada por:

$$y = 0,5539x + 0,3688 \quad (5.14)$$

onde $m = 0,5539$ e $b = 0,3688$ também foram calculados pelas equações (5.1) e (5.2), respectivamente. O índice de correlação entre os valores estimados e reais foi de 0,66 (equação 5.4), indicando que há uma relação entre os dados, mas que eles não estão tão bem correlacionados como no caso da geometria quadrada.

Assim, utilizando o erro local definido na equação (5.5) para x_i e y_i representando os valores real e estimado pela RNA, respectivamente, da área interfacial para $i = 1, \dots, 1000$ testes numéricos, pode-se calcular o desvio padrão do erro local, obtendo-se $\sigma_e = 0,33$. Portanto, as retas $y = mx + (b \pm 3\sigma_e)$, ou seja,

$$y = 0,5539x + 1,3837 \quad (5.15)$$

$$y = 0,5539x - 0,6462 \quad (5.16)$$

formam um envelope em torno dos pontos (x_i, y_i) , ilustrado na figura 5.22, mostrando o intervalo de confiança dos resultados da RNA para a área interfacial.

Figura 5.21: Comparação entre os valores estimados pela rede neural e os valores reais para a área interfacial (geometria circular).

Figura 5.22: Intervalo de confiança dos resultados fornecidos pela RNA para os valores da área interfacial na geometria circular.

Isso significa que para um determinado valor de Γ , a resposta da rede neural estará no intervalo $[\Gamma - 3\sigma_e, \Gamma + 3\sigma_e]$.

Para o caso da geometria circular, pode-se concluir que a rede neural é capaz de fornecer boas respostas para a fração volumétrica e para a área interfacial, mas com os desvios padrões maiores do que os que foram calculados no caso da geometria quadrada.

Conclusões

Uma nova metodologia para a medida não intrusiva da fração volumétrica e da área interfacial de escoamentos bifásicos foi proposta neste trabalho, utilizando sensoramento acústico e resolvendo o problema inverso associado através do desenvolvimento e treinamento de redes neurais artificiais. O modelo matemático foi construído a partir das equações de propagação acústica. A distribuição geométrica das fases dentro do escoamento é mapeada pela velocidade local de propagação acústica, considerada na equação diferencial que governa o problema. Esta equação foi resolvida numericamente pelo método de diferenças finitas com as condições de contorno reproduzindo a estratégia de pulso/eco.

As redes neurais foram treinadas com dados oriundos de simulações numéricas. Para a realização de tais simulações, foi desenvolvida uma ferramenta numérica capaz de calcular a propagação acústica em meios heterogêneos. Um número significativo de distribuições (aleatórias) das velocidades de propagação acústica foi considerado na solução da equação diferencial para construir uma base de dados, da qual os parâmetros da rede (pesos e *bias*) podem ser ajustados. Mais especificamente, os modelos neurais foram construídos para mapear características extraídas de sinais obtidos de quatro sensores acústicos, localizados no contorno externo do domínio de sensoramento, estimando a fração volumétrica e a área interfacial correspondentes. Estas características correspondem às amplitudes e aos tempos de chegada dos três maiores picos da onda acústica.

Após concluída esta etapa, realizou-se a implementação das redes neurais artificiais e estudos preliminares visando obter arquiteturas otimizadas relativa-

mente ao problema tratado. Em seguida, fez-se o treinamento dos melhores modelos com exemplos extraídos da base de dados, buscando a determinação da fração volumétrica e da área interfacial, parâmetros importantes dos escoamentos bifásicos.

Os resultados das redes neurais foram bastante satisfatórios, principalmente na caso da geometria quadrada, uma vez que mostraram que elas são capazes de estimar os valores da fração volumétrica e da área interfacial para exemplos que não foram apresentados à elas no treinamento, e ainda, verificou-se que estes modelos são treináveis em um tempo computacional razoável e que os intervalos de confiança são compatíveis com futuras aplicações práticas.

Trabalhos futuros incluem estudos em um domínio tridimensional. Também podem ser analisados outros escoamentos, como o escoamento ar-água. Uma outra proposta de trabalho futuro é a implementação de outros tipos de redes neurais. Outra possibilidade, é refinar a malha nas simulações numéricas, afim de obter maiores valores para a área interfacial. Também há a possibilidade de obter sinais experimentais em laboratório para extrair as características da base de dados. Pode-se também, extrair outras características do sinais de resposta e estimar diferentes propriedades do escoamento, como densidade, viscosidade e outras.

Referências Bibliográficas

- [1] AL-SARKHI, A.; HANRATTY, T. J. (2002). Effect of pipe diameter on the drop size in a horizontal annular gas-liquid flow. *International Journal of Multiphase Flow*, v. 28, p. 1617-1629.
- [2] ALMEIDA, I. F. S.; CASTRO, M. S.; RODRIGUEZ, O. M. H. (2005). Non-Intrusive Capacitive Probe For Measurement Of Insitu Volumetric Fraction In Water-Oil And Water-Air Flows. *Proceedings of COBEM 2005 - 18th International Congress of Mechanical Engineering*. November 6-11, Ouro Preto, MG.
- [3] ALTMANN, J. (2001). Acoustic Weapons - A Prospective Assessment. *Science & Global Security*, v. 9, p. 165-234.
- [4] ANGELI, P.; HEWITT, G. F. (2000). Flow structure in horizontal oil-water flow. *International Journal of Multiphase Flow*, v. 26, p. 1117-1140.
- [5] BLITZ J. (1967). *Fundamentals of Ultrasonics*. 2nd ed. London. Butterworth & Co. (Publishers) Ltd.
- [6] BOND, L. J.; KEPLER, W. F.; FRANGOPOL, D. M. (2000). Improved assessment of mass concrete dams using acoustic travel time tomography. Part I - theory, *Construction and Building Materials*, v. 14, p. 133-146.
- [7] BRAGA, A. P.; CARVALHO, A. C. P. L. F.; LUDERMIR, T. B. (2000). *Redes Neurais Artificiais: Teoria e Aplicações*. LTC - Livros Técnicos e Científicos Editora S. A. Rio de Janeiro -RJ.

-
- [8] BUND, R. K., PANDIT, A. B. (2007). Sonocrystallization: Effect on lactose recovery and crystal habit, *Ultrasonics Sonochemistry*, v. 14, n. 2, p. 143-152, Fev.
- [9] BUTZ, P., TAUSCHER, B. (2002). Emerging technologies: chemical aspects. *Food Research International*, v. 35, n. (2/3), p. 279-284.
- [10] CRIVELARO, K. C. O.; SELEGHIM Jr. P., (2002). Identification of horizontal two-phase flow regimes through a neural network models. *Journal of the Brazilian Society of Mechanical Sciences*, v. 14, n. 1, p. 1-16.
- [11] CLOUTIER, G.; CHEN, D.; DURAND, L-G. (2001). Performance of time-frequency representation techniques to measure blood flow turbulence with pulsed-wave doppler ultrasound. *Ultrasound in Medicine & Biology*, v. 27, n. 4, p. 535-550.
- [12] DAIDZIC, N. E.; SCHMIDT, E.; HASAN, M. M.; ALTOBELLI, S. (2005). Gas-liquid phase distribution and void fraction measurements using MRI. *Nuclear Engineering and Design*, v. 235, n. 10-12, p. 1163-1178. May.
- [13] DAVIS, G. DE VAHL. (1979). A Note on a Mesh for use with Polar Coordinates, *Numerical Heat Transfer*, v.2, p. 261-266.
- [14] DEMUTH, H.; BEALE, M. (2000). *Neural Network Toolbox User's Guide*. The MathWorks, Inc.
- [15] DÍAZ, G. (2005). *Ecografia em cor*. Disponível em: < [http : //www.drgdiaz.com/eco/portugues/ecografia.shtml](http://www.drgdiaz.com/eco/portugues/ecografia.shtml) >. Acesso em : 22 nov 2005.
- [16] DOEBELIN, E. O. (1990). *Measurement Systems - Applications and Design*. 4th ed. McGraw-Hill, Inc. New-York.
- [17] 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*, v. 14, p. 183-192.

- [18] DONGJIAN, Z.; LIEJIN, G.; CHANGZHI, L.; XIMIN, Z. (2005). An experimental study on local interfacial area concentration using a double-sensor probe, *International Journal of Heat and Mass Transfer*, v. 48, n. 10, p. 1926-1935, May.
- [19] ESBENSEN K.H., HOPE B., LIED T.T., HALSTENSEN M., GRAVERMOEN T., SUNDBERG K. (1999). Acoustic chemometrics for fluid flow quantifications - II: A small constriction will go a long way, *Journal of Chemometrics*, v. 13, n.(3-4), p. 209-236. May-Aug.
- [20] EUH, D. J.; YUN, B. J.; SONG, C. H.; KWON, T. S.; CHUNG, M. K.; LEE, U. C. (2001). Development of the five-sensor conductivity probe method for the measurement of the interfacial area concentration, *Nuclear Engineering and Design*, v. 205, n. 1-2, p. 35-51, March.
- [21] FILLETTI, E. R. (2002). *Estudo sobre a Utilização da Tomografia Acústica para a Reconstrução de Campos Internos de Temperatura*. Dissertação (Mestrado), Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos - SP. 2002.
- [22] FUJII, M.; ZHANG, X. (2001). Noncontact measurement of internal temperature distribution in a solid material using ultrasonic computed tomography. *Experimental Thermal and Fluid Science*, v. 24, p. 107-116.
- [23] FUNAHASHI, K. (1989). On the Approximate Realization of Continuous Mappings by Neural Networks. *Neural Networks*, v. 2, p. 183-192.
- [24] GAO, H.; GU, H-Y; GUO L-J. (2003). Numerical study of stratified oil-water two-phase turbulent flow in a horizontal tube. *International Journal of Heat and Mass Transfer* v. 46, p. 749-754.
- [25] GISMONDI, R. C.; ALMEIDA, R. M. V. R. E INFANTOSI, A. F. C. (2002). Artificial neural networks for infant mortality modelling. *Computer Methods and Programs in Biomedicine*, v. 69, n. 3, p. 237-247. Nov.

-
- [26] HAGAN, M.T.; DEMUTH, H.B.; BEALE, M. (1996). *Neural network design*. Boston, PWS publishing company.
- [27] HAYKIN, S. (1999). *Neural Networks: a comprehensive foudation*. 2nd ed. New Jersey: Prentice Hall.
- [28] HENNING, B.; DAUR, P.; PRANGE, S.; DIERKS, K.; HAUPTMANN, P. (2000). In-line concentration measurement in complex liquids using ultrasonic sensors. *Ultrasonics*. v. 38, p. 799-803.
- [29] HILBERT, D. W.; OSTENDORF, B. (2001). The utility of artificial neural networks for modelling the distribution of vegetation in past, present and future climates. *Ecological Modelling*, v. 146, n. 1-3, p. 311-327. Dec.
- [30] HOPFIELD, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the national academy of science*, v. 79, p. 2554-2558, April.
- [31] HUANG, Z.; WANG, LI, H. (2003). Application of electrical capacitance tomography to the void fraction measurement of two-phase flow. *IEEE Transactions on Instrumentation and Measurement*, v. 52, n. 1, p. 7-12.
- [32] HUSH, D. R.; HORME, B. G. (1993). Progress in supervised neural networks: What's new since Lippman? *IEEE Signal Processing Magazine*, v. 10, p. 8-39.
- [33] INCE, R. (2004). Prediction of fracture parameters of concrete by Artificial Neural Networks. *Engineering Fracture Mechanics*, v. 71, n. 15, p. 2143-2159. Oct.
- [34] ISHII, M.; KIM, S. (2001). Micro four-sensor probe measurement of interfacial area transport for bubbly flow in round pipes, *Nuclear Engineering and Design*, v. 205, n. 1-2, p. 123-131, March.

- [35] JIN, N. D.; NIE, X. B.; WANG, J.; REN, Y. Y. (2003a). Flow pattern identification of oil-water two-phase flow based on kinematic wave theory. *Flow Measurement and Instrumentation*, v. 14, p. 177-182.
- [36] JIN, N. D.; NIE, X. B.; REN, Y. Y.; LIU, X. B. (2003b). Characterization of oil-water two-phase flow patterns based on nonlinear time series analysis. *Flow Measurement and Instrumentation*, v. 14, p. 169-175.
- [37] KARA, S.; KEMALOĞLU, S. E. GÜVEN, A. (2005). Detection of femoral artery occlusion from spectral density of Doppler signals using the artificial neural network. *Expert Systems with Applications*, v. 29, n. 4, p. 945-952. Nov.
- [38] KINSLER, L. E.; FREY, A. R. (1962). *Fundamentals of Acoustics*. 2nd ed. New York: John Wiley & Sons, Inc.
- [39] KIM, M. C.; KIM, S.; KIM, K. Y.; LEE, Y. J. (2001). Regularization Methods In Electrical Impedance Tomography Technique For The Two-Phase Flow Visualization. *International Communications In Heat And Mass Transfer*. v. 28, n. 6, p. 173-782.
- [40] KLEIN, F. L. (2004). *Análise Tempo-Frequência de Regimes de Escoamento Bifásico Gás-Líquido Intermitentes em tubo Horizontal*. Tese (Doutorado), Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos - SP.
- [41] KOHONEN, T. (2001). *Self-Organizing Maps* (3th Edition). Berlin. Springer-Verlag.
- [42] LEHNER, P.; WIRTH, K. E. (1999). Characterization of the flow pattern in a downer reactor. *Chemical engineering science*, v. 54, p. 5471-5483.
- [43] LEITE, C. C. (2005). Física Básica da Ultrasonografia. Apostila. Disponível em: <http://urutu.hcnet.usp.br/inrad/departamento/graduacao/mdr609.htm>. Acesso em: 16 Nov 2005.

- [44] LÉTANG, C.; PIAU, M.; VERDIER, C.; LEFEBVRE, L. (2001). Characterization of wheat-flour-water doughs: a new method using ultrasound, *Ultrasonics*, v. 39, p. 133-141.
- [45] LYNCH J. (1991). Non intrusive multiphase flow measurement techniques (bibliographic review), *Revue de l'Institut Francais du Petrole*, v. 46, n. 1, p. 59-88. Jan-Feb.
- [46] MA, Y.; ZHENG, Z.; XU, L.; LIU, X.; WU, Y. (2001). Application of electrical resistance tomography system to monitor gas/liquid two-phase flow in a horizontal pipe. *Flow Measurement and Instrumentation*, v. 12, p.259-265.
- [47] MALINEN, M.; HUTTUNEN, T.; KAPIO, J. P. (2002). Optimal control for the ultrasound induced heating of a tumor. In: *International Conference on Inverse Problems in Engineering*, 4., 26-31 May, Rio de Janeiro, 1 CD-ROM.
- [48] MI, X.; ZOU, Y.; WEI, W.; MA, K. (2005). Testing the generalization of artificial neural networks with cross-validation and independent-validation in modelling rice tillering dynamics. *Ecological Modelling*, v. 181, p.493-508.
- [49] PARLAK, A.; ISLAMOGU, Y.; YASAR, H.; EGRISOGUT, A. (2005). Application of artificial neural network to predict specific fuel consumption and exhaust temperature for a Diesel engine. *Applied Thermal Engineering*, In Press, Corrected Proof, Available online 18 November 2005.
- [50] PÉCORA, J. D.; GUERISOLI, D. M. Z. (2005). *Ultra-Som*. Disponível em: < [http : //www.forp.usp.br/restauradora/us01.htm](http://www.forp.usp.br/restauradora/us01.htm) >. Acesso em : 16 nov 2005.
- [51] PEREIRA, A. S. (1985). *Estudo da Velocidade e do Coeficiente de Atenuação do Ultra-Som em Tumores Ósseos pelo Método de pulso-Eco*. 100p. Dissertação (Mestrado), Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos - SP.

- [52] PINA FILHO, A. C., DUTRA, M. S., LIMA, A. S. (2006). Simulação de um robô bípede utilizando redes neurais para subida de uma rampa, *Anais do CONEM 2006 - IV Congresso Nacional de Engenharia Mecânica*. 22-25 Agosto, Recife-PE.
- [53] PRASSER, H. M.; MISAWA, M.; TISEANU, I. (2005). Comparison between wire-mesh sensor and ultra-fast X-ray tomograph for an air-water flow in a vertical pipe. *Flow Measurement and Instrumentation* v. 16, p. 73-83.
- [54] PRESS, W. H.; VETTERLING, W. T.; TEUKOLSKY, A. A.; FLANNERY, B. P. (1992). *Numerical Recipes in Fortran: the art of scientific computing*, 2nd ed. Cambridge (England); New York, NY, USA: Cambridge University Press.
- [55] PÜTTMER, A.; LUCKLUM, R.; HENNING, B.; HAUPTMANN, P. (1998). Improved ultrasonic density sensor with reduced diffraction influence, *Sensors and Actuators A*, v. 67, p. 8-12.
- [56] PÜTTMER, A.; HOPPE, N.; HENNING, B.; HAUPTMANN, P. (1999). Ultrasonic density sensor - analysis of errors due to thin layers of deposits on the sensor surface, *Sensors and Actuators*, v.76, p.122-126.
- [57] REVANKAR, S. T.; ISHII, M. (1993) Theory and measurement of local interfacial area using a four sensor probe in two-phase flow *International Journal of Heat and Mass Transfer*, v. 36, n. 12, p. 2997-3007, August.
- [58] ROSENBLATT, F. (1958). The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological review*, v. 65, n. 6, p. 386-408.
- [59] ROLNIK, V. P; SELEGHIM JR, P. (2002). On site calibration of a phase fraction meter by an inverse technique. *Revista Brasileira de Ciências Mecânicas*, v.24, n.4, p.1-14.
- [60] RUMELHART, D. E.; HINTON, G. E.; WILLIAMS, R. J. (1986). Learning representations of back-propagation errors. *Nature*, v. 323, p. 533-536.

- [61] SCHLABERG, H. I.; PODD, F. J. W.; HOYLE, B. S. (2000). Ultrasound process tomography system for hydrocyclones, *Ultrasonics*, v. 38, n. 1-8, p. 813-816. March.
- [62] TAN, M. J.; ISHII M. (1990). A method for measurement of local specific interfacial area *International Journal of Multiphase Flow*, v. 16, n. 2, p. 353-358, March-April.
- [63] TIKHONOV, A. N. (1963). Regularization of Incorrectly Posed Problems, *Soviet Math. Dokl.*, v. 4, n. 6, p. 1624-1627.
- [64] VEIGA, J. L .B. C.; CARVALHO, A. A.; SILVA, I. C.; REBELLO, J. M. A. (2005). The use of artificial neural networks in the classification of pulse-echo and TOFD ultra-sonic signals. *Journal of the Brazilian Society of Mechanical Sciences and Engineering*, v. 27, n. 4, p. 394-398. Oct-Dec.
- [65] WADE, G. (2000). Human uses of ultrasound: ancient and modern. *Ultrasonics*, v.38, p.1-5.
- [66] XU, L. J.; XU, L. A. (1998). Gas/liquid two-phase flow regime identification by ultrasonic tomography. *Flow Measurement and Instrumentation*, v. 8, p. 145-155, April.
- [67] XU *et al.* (1997). Application of ultrasonic tomography to monitoring gas/liquid flow. *Chemical Engineering Science*, v. 52, n. 13, p. 2171-2183.
- [68] YANG, A. S.; KUO, T. C.; LING, P. H. (2003). Application of neural networks to prediction of phase transport characteristics in high-pressure two-phase turbulent bubbly flows. *Nuclear Engineering and Design*, v. 223, p. 295-313.
- [69] YAN, H.; LIU, Y. H.; LIU, C. T. (2004). Identification of flow regimes using backpropagation networks trained on simulated data based on a capacitance tomography sensor. *Measurement Science and Technology*, v. 15, p. 432-436.

-
- [70] YOUNG, H. D. (1962). *Statistical Treatment of Experimental Data*. McGraw-Hill, New-York.

Índice Remissivo

- Área Interfacial Volumétrica, 25
- Época, 61
- Bias*, 52
- Algoritmo de Retropropagação do Erro, 60
- Amplitudes dos Sinais Acústicos, 24
- Análise do Componente Principal, 71
- Aprendizagem
 - Não-supervisionada, 57
 - por Reforço, 57
 - Supervisionada, 57
- Arquitetura de RNAS, 54
- Camada
 - de Entrada, 54
 - de Saída, 54
 - Intermediária ou Oculta, 54
- Coefficiente de Correlação, 93
- Condição
 - de Contorno, 22
 - Inicial, 22
- Configurações de Escoamento, 7
- Conjunto
 - de Sinápses, 51
 - de Teste, 69
 - de Treinamento, 60
- Contraste Acústico, 3
- Desvio Padrão, 93
- Elemento de Medida, 23
- Erro
 - Instantâneo, 60
 - Local, 93
 - Médio Quadrático, 60
 - Total, 60
- Fração Volumétrica, 25
- Função
 - de Ativação, 52, 53
 - Degrau, 54
 - Linear por Partes, 54
 - Sigmóide, 54
- Generalização, 68
- Gradiente Local, 63
- Matriz
 - de Entrada, 74
 - de Saída, 74
 - Definida Positiva, 47
 - Esparsa, 46
- Nó MCP, 16, 50

-
- Neurônio Não-linear, 51
- Normalização
- através da Divisão pelo Máximo Valor Absoluto, 70
 - através da Média e do Desvio Padrão, 71
 - dos Dados, 70
- Perceptrons de Múltiplas Camadas, 58
- Peso Sináptico, 51
- Potencial de Ativação, 61
- Propagação de Ondas Acústicas, 21
- Redes
- Feed-forward*, 55
 - Localmente Conectadas, 56
 - Neurais Artificiais, 15, 49
 - Parcialmente Conectadas, 56
 - Recorrentes, 55
 - Totalmente Conectadas, 56
- Regra
- Delta, 62
 - Delta Generalizada, 68
- Sensor Acústico, 22
- Sinais
- de Pressão Acústica, 23
 - de Resposta, 23
- Sinal
- de Erro, 59, 60
 - Funcional, 59, 61
- Somador, 51
- Taxa de Aprendizado, 63
- Tempo de Chegada
- da Onda Acústica, 24
 - das Amplitudes, 24
- Termo de Momento, 68
- Transferência Interfacial, 9
- Treinamento
- por Lote, 67
 - por Padrão, 67
- Ultra-som, 11