



Universidade do Estado do Rio de Janeiro
Instituto Politécnico

Rogério Campos de Oliveira

**Aplicação de máquinas de comitê de redes neurais artificiais na
solução de um problema inverso em transferência radiativa**

Nova Friburgo
2010

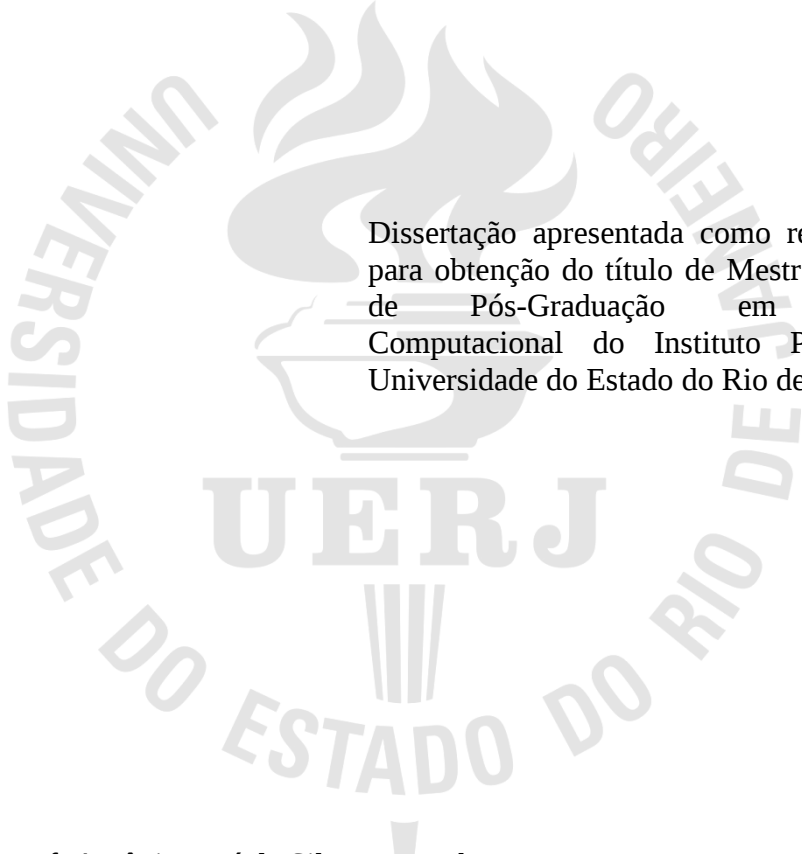
Livros Grátis

<http://www.livrosgratis.com.br>

Milhares de livros grátis para download.

Rogério Campos de Oliveira

**Aplicação de máquinas de comitê de redes neurais artificiais na
solução de um problema inverso em transferência radiativa**



Dissertação apresentada como requisito parcial
para obtenção do título de Mestre, ao Programa
de Pós-Graduação em Modelagem
Computacional do Instituto Politécnico, da
Universidade do Estado do Rio de Janeiro.

Orientador(es): Prof. Antônio José da Silva Neto, Ph.D.
Prof. Luiz Biondi Neto, D.Sc.

Nova Friburgo
2010

CATALOGAÇÃO NA FONTE
UERJ / REDE SIRIUS / BIBLIOTECA CTC/E

O48

Oliveira, Rogério Campos de.

Aplicação de máquinas de comitê de redes neurais artificiais na solução de um problema inverso em transferência radiativa / Rogério Campos de Oliveira.- 2010.

133 f.: il.

Orientadores: Antônio José da Silva Neto e Luiz Biondi Neto.

Dissertação (Mestrado) - Universidade do Estado do Rio de Janeiro, Instituto Politécnico.

1. Redes neurais (Computação) – Teses. 2. Radiação – Transferência – Teses. 3. Problemas inversos (Equações diferenciais) – Teses. I. Silva Neto, Antônio José da. II. Biondi Neto, Luiz. III. Universidade do Estado do Rio de Janeiro. Instituto Politécnico. IV. Título.

CDU 004.032.26

Autorizo, apenas para fins acadêmicos e científicos, a reprodução total ou parcial desta dissertação.

Assinatura

Data

Rogério Campos de Oliveira

**Aplicação de máquinas de comitê de redes neurais artificiais na
solução de um problema inverso em transferência radiativa**

Dissertação submetida ao corpo docente do Programa de Pós-Graduação em Modelagem Computacional do Instituto Politécnico, Campus Regional da UERJ em Nova Friburgo, como parte dos requisitos necessários à obtenção do grau de Mestre em Modelagem Computacional. Linha de Pesquisa: Matemática Aplicada e Computação Científica.

Aprovada em 26 de julho de 2010.

Banca examinadora:

Prof. Antônio José da Silva Neto, Ph.D.
IPRJ/UERJ

Prof. Luiz Biondi Neto, D.Sc.
UERJ.

Prof. Francisco Duarte Moura Neto, Ph.D.
IPRJ/UERJ

Prof. João Carlos Correia Baptista Soares de Mello, D.Sc.
UFF

Prof^a Rosana da Paz Ferreira Pinheiro, D.Sc.
UEZO

Nova Friburgo
2010

DEDICATÓRIA

À minha família, ao meu irmão Haroldo Campos de Oliveira (*in memorian*), à minha avó Maria Eulália de Campos (*in memorian*), a todas as pessoas de coragem e boa vontade, à natureza e à vida.

AGRADECIMENTOS

Primeiramente a Deus por ter me dado forças e coragem para seguir em frente nesta árdua caminhada.

Aos Professores Antônio José da Silva Neto e Luiz Biondi Neto, meus orientadores, pelos valiosos ensinamentos, amizade, confiança, motivação, incentivo e paciência na orientação deste trabalho.

À minha esposa e aos meus filhos pelo seu amor, compreensão e resignação diante do sacrifício a eles imposto.

Aos meus pais pelo seu amor, dignidade, exemplo, incentivo e apoio incondicional.

À minha saudosa avó Maria Eulália de Campos (*in memoriam*), que sempre me incentivou e acreditou em mim.

Aos meus irmãos e familiares pela amizade, apoio e incentivo.

À Dra. Maria Lina Vieira, Secretária de Tributação do RN, pela confiança e oportunidade.

Ao Professor Belchior de Oliveira Rocha, Reitor do IFRN, pela confiança e oportunidade.

Ao Professor Marcus Peigas Pacheco pela confiança, oportunidade e valiosos conselhos.

Aos Professores Joaquim Teixeira de Assis, Francisco Duarte Moura Neto e Júlio Alves Hermínio (IFRN) pela amizade, pelos valiosos ensinamentos e pelas cartas de apresentação.

A todos os professores do IPRJ dos quais fui aluno e que enriqueceram em muito a minha formação.

Ao meu amigo Antônio Merendaz do C. Neto, que tanto me ajudou durante todo o curso.

À D.Sc. Nancy Isabel Alvarez Acevedo pela amizade, presteza, simpatia e orientação neste trabalho.

Aos servidores do IPRJ sempre prestativos e zelosos em suas atribuições.

Aos colegas e amigos da SUFAC/SET e DIETIND/IFRN pelo apoio e incentivo.

Aos colegas e amigos de curso pelo espírito de luta, exemplo e incentivo.

Ao Governo do Estado do Rio Grande do Norte pelo apoio financeiro e oportunidade.

Ao Instituto Federal de Educação, Ciência e Tecnologia do Rio Grande do Norte pelo apoio financeiro e oportunidade.

À FAPERJ pelo apoio financeiro.

Aos amigos de todos os lugares pelas horas de lazer e descontração.

RESUMO

OLIVEIRA, Rogério Campos de. *Aplicação de máquinas de comitê de redes neurais artificiais na solução de um problema inverso em transferência radiativa*. 2010. 133 f. Dissertação (Mestrado em Modelagem Computacional) – Instituto Politécnico, Universidade do Estado do Rio de Janeiro, Nova Friburgo, 2010.

Este trabalho fundamenta-se no conceito de máquina de comitê de redes neurais artificiais e tem por objetivo resolver o problema inverso de transferência radiativa em um meio unidimensional, homogêneo, absorvedor e espalhador isotrópico. A máquina de comitê de redes neurais artificiais agrega e combina o conhecimento adquirido por um certo número de *especialistas* aqui representados, individualmente, por cada uma das redes neurais artificiais (RNA) que compõem a máquina de comitê de redes neurais artificiais. O objetivo é atingir um resultado final melhor do que o obtido por qualquer rede neural artificial separadamente, selecionando-se apenas àquelas redes neurais artificiais que apresentam os melhores resultados na fase de generalização descartando-se as demais, o que foi feito neste trabalho. Aqui são utilizados dois modelos estáticos de máquinas de comitê, usando a média aritmética de conjunto, que se diferenciam entre si apenas na composição do combinador de saída de cada máquina de comitê. São obtidas, usando-se máquinas de comitê de redes neurais artificiais, estimativas para os parâmetros de transferência radiativa, isto é, a espessura óptica do meio, o albedo de espalhamento simples e as refletividades difusas. Finalmente, os resultados obtidos com ambos os modelos de máquina de comitê são comparados entre si e com aqueles encontrados usando-se apenas redes neurais artificiais do tipo perceptrons de múltiplas camadas (MLP), isoladamente. Aqui essas redes neurais artificiais são denominadas *redes neurais especialistas*, mostrando que a técnica empregada traz melhorias de desempenho e resultados a um custo computacional relativamente baixo.

Palavras-chaves: Transferência radiativa; Problemas inversos; Máquina de comitê; Perceptron de Múltiplas Camadas; Redes neurais artificiais.

ABSTRACT

This work is based on the concept of neural networks committee machine and has the objective to solve the inverse radiative transfer problem in one-dimensional, homogeneous, absorbing and isotropic scattering media. The artificial neural networks committee machine adds and combines the knowledge acquired by an exact number of specialists which are represented, individually, by each one of the artificial neural networks (ANN) that composes the artificial neural network committee machine. The aim is to reach a final result better than the one obtained by any of the artificial neural network separately, selecting only those artificial neural networks that presents the best results during the generalization phase and discarding the others, what was done in this present work. Here are used two static models of committee machines, using the ensemble arithmetic average, that differ between themselves only by the composition of the output combinator by each one of the committee machine. Are obtained, using artificial neural networks committee machines, estimates for the radiative transfer parameters, that is, medium optical thickness, single scattering albedo and diffuse reflectivities. Finally, the results obtained with both models of committee machine are compared between themselves and with those found using artificial neural networks type multi-layer perceptrons (MLP), isolated. Here that artificial neural networks are named as *specialists neural networks*, showing that the technique employed brings performance and results improvements with relatively low computational cost.

Keywords: Radiative transfer; Inverse problems; Committee machine; Multi-layer-perceptron, Artificial neural networks.

LISTA DE FIGURAS

Figura 1 - Estrutura da máquina de comitê de redes neurais artificiais.....	37
Figura 2 - Meio participante unidimensional homogêneo com superfícies internas refletoras difusas e submetido à incidência de radiação externa de intensidades f_1 e f_2	39
Figura 3 - Representação esquemática dos dados experimentais e das condições de contorno do meio	43
Figura 4 - (a) domínio contínuo e (b) malha computacional.....	44
Figura 5 - Discretização do domínio angular, $-1 \leq \mu \leq 1$	44
Figura 6 - Discretização do domínio espacial para o meio homogêneo unidimensional, $0 \leq \tau \leq \tau_0$	44
Figura 7 - Esquema representativo da marcha para frente	47
Figura 8 - Esquema representativo da marcha para trás.....	48
Figura 9 - Representação esquemática do problema direto (a) e do problema inverso (b)	50
Figura10 - Representação esquemática de um meio participante homogêneo unidimensional e da radiação que emerge do mesmo com intensidade $\vec{Y}$, medida experimentalmente	51
Figura 11 - Rede Neural Biológica.....	56
Figura 12 - Neurônio artificial com uma única entrada.....	57
Figura 13 - Função de ativação tipo degrau, (a) sem bias e (b) com bias	58

Lista de figuras

Figura 14 - Função de ativação tipo linear, (a) sem bias e (b) com bias	58
Figura 15 - Função de ativação tipo log-sigmóide. (a) sem bias e (b) com bias	59
Figura 16 - Função de ativação tipo tan-sigmóide: (a) sem bias e (b) com bias	59
Figura 17 - Modelo do Neurônio Artificial de McCulloch e Pitts	60
Figura 18 - Estrutura de uma rede neural artificial.....	61
Figura 19 - Conexão entre dois neurônios artificiais.....	63
Figura 20 - Modelo do Perceptron.....	65
Figura 21 - Arquitetura do Perceptron de Múltiplas Camadas.....	67
Figura 22 - Fluxograma do algoritmo de retro-propagação (BP).....	69
Figura 23 - Treinamento com baixa taxa de aprendizado	70
Figura 24 - Treinamento com alta taxa de aprendizado	71
Figura 25 - Treinamento com alta taxa de aprendizado e com coeficiente de momento	72
Figura 26 - Diagrama em blocos de uma estrutura estática de um comitê usando o método da média aritmética.....	73
Figura 27 - Diagrama em blocos de uma estrutura estática de uma máquina de comitê usando o método do reforço.....	74
Figura 28 - Diagrama de blocos de uma estrutura dinâmica de uma máquina de comitê usando mistura de especialistas.....	75
Figura 29 - MLP com duas camadas ocultas para o problema inverso	76

Lista de figuras

Figura 30 - MLP com uma camada oculta para o problema inverso.....	80
Figura 31 - Modelo da máquina de comitê tipo estática onde o combinador efetua apenas a média aritmética das saídas das RNA especialistas.....	82
Figura 32 - Modelo da máquina de comitê tipo estática com MLP integrado ao combinador	82
Figura 33 - Modelo do MLP do combinador da máquina de comitê	84
Figura 34 - Curva de performance da RNA nº 10, com $N_{HL} = 2$; $N_H = 26$; $\omega = 0,5$ e $tol = 10^{-4}$	95
Figura 35 - Curva de performance da RNA nº 09, com $N_{HL} = 1$; $N_H = 24$; $\omega = 0,1$ e $tol = 10^{-3}$	99
Figura 36 - Curva de performance do MLP do combinador da máquina de comitê, para o caso $N_{HL} = 1$; $\omega = 0,1$; $tol = 10^{-3}$ e $M = 6$	100
Figura 37 - Curva de performance do MLP do combinador da máquina de comitê, para o caso $N_{HL} = 1$; $M = 6$; $\omega = 0,9$ e $tol = 10^{-3}$	102
Figura 38 - Curva de performance do MLP do combinador da máquina de comitê misto, com $N_{HL} = 1$ ou 2 ; $M = 6$; $\omega = 0,1$ e $tol = 10^{-3}$	103
Figura 39 - Curva de performance do MLP do combinador da máquina de comitê misto, com $N_{HL} = 1$ ou 2 ; $M = 6$; $\omega = 0,5$ e $tol = 10^{-3}$	104
Figura 40 - Curva de performance do MLP do combinador da máquina de comitê misto, com $N_{HL} = 1$ ou 2 ; $M = 6$; $\omega = 0,9$ e $tol = 10^{-3}$	104
Figura 41 - Curva de performance da RNA nº 03, com $N_{HL} = 1$; $N_H = 12$; $\omega = 0,1$, $0,5$ ou $0,9$ e $tol = 10^{-4}$	122

Lista de figuras

Figura 42 - Curva de performance da RNA nº 10, com $N_{HL} = 2$; $N_H = 26$; $\omega = 0,1$, $0,5$ ou $0,9$ e $tol = 10^{-3}$	125
--	-----

LISTA DE TABELAS

Tabela 1 - Pontos de colocação, μ_i , e pesos correspondentes, a_i , para a quadratura Gauss – Legendre com $M = 20$	46
Tabela 2 - Comparativo entre os erros relativos das estimativas obtidas para $\vec{Z} = \{\tau_0, \omega, \rho_1, \rho_2\}$ com redes neurais e máquinas de comitês para todos os casos testes com $\omega = 0,1; 0,5$ ou $0,9$ $tol = 10^{-3}$ e/ou 10^{-4}	91 - 93
Tabela 3 - Estimativas obtidas para $N_{HL} = 2; \omega = 0,5$ e $tol = 10^{-4}$	93
Tabela 4 - Erros relativos percentuais por rede neural especialista para $N_{HL} = 2; \omega = 0,5$ e $tol = 10^{-4}$	94
Tabela 5 - Estimativas obtidas para $N_{HL} = 1; \omega = 0,1$ e $tol = 10^{-3}$	97
Tabela 6 - Erros relativos percentuais por rede neural especialista para $N_{HL} = 1; \omega = 0,1$ e $tol = 10^{-3}$	98
Tabela 7 - Estimativas obtidas para $N_{HL} = 1; \omega = 0,9$ e $tol = 10^{-3}$	101
Tabela 8 - Erros relativos percentuais por rede neural especialista para $N_{HL} = 1; \omega = 0,9$ e $tol = 10^{-3}$	101
Tabela 9 - Estimativas obtidas para $N_{HL} = 1; \omega = 0,5$ e $tol = 10^{-3}$	119
Tabela 10 - Erros relativos percentuais por rede neural especialista para $N_{HL} = 1; \omega = 0,5$ e $tol = 10^{-3}$	119
Tabela 11 - Estimativas obtidas para $N_{HL} = 1; \omega = 0,1$ e $tol = 10^{-4}$	120

Lista de tabelas

Tabela 12 - Erros relativos percentuais por rede neural especialista para $N_{HL} = 1$; $\omega = 0,1$ e $tol = 10^{-4}$	120
Tabela 13 - Estimativas obtidas para $N_{HL} = 1$; $\omega = 0,5$ e $tol = 10^{-4}$	120
Tabela 14 - Erros relativos percentuais por rede neural especialista para $N_{HL} = 1$; $\omega = 0,5$ e $tol = 10^{-4}$	121
Tabela 15 - Estimativas obtidas para $N_{HL} = 1$; $\omega = 0,9$ e $tol = 10^{-4}$	121
Tabela 16 - Erros relativos percentuais por rede neural especialista para $N_{HL} = 1$; $\omega = 0,9$ e $tol = 10^{-4}$	121
Tabela 17 - Estimativas obtidas para $N_{HL} = 2$; $\omega = 0,1$ e $tol = 10^{-3}$	122
Tabela 18 - Erros relativos percentuais por rede neural especialista para $N_{HL} = 2$; $\omega = 0,1$ e $tol = 10^{-3}$	123
Tabela 19 - Estimativas obtidas para $N_{HL} = 2$; $\omega = 0,5$ e $tol = 10^{-3}$	123
Tabela 20 - Erros relativos percentuais por rede neural especialista para $N_{HL} = 2$; $\omega = 0,1$ e $tol = 10^{-4}$	123
Tabela 21 - Estimativas obtidas para $N_{HL} = 2$; $\omega = 0,9$ e $tol = 10^{-3}$	124
Tabela 22 - Erros relativos percentuais por rede neural especialista para $N_{HL} = 2$; $\omega = 0,9$ e $tol = 10^{-3}$	124
Tabela 23 - Estimativas obtidas para $N_{HL} = 2$; $\omega = 0,1$ e $tol = 10^{-4}$	124
Tabela 24 - Erros relativos percentuais por rede neural especialista para $N_{HL} = 2$; $\omega = 0,1$ e $tol = 10^{-4}$	125
Tabela 25 - Estimativas obtidas para $N_{HL} = 2$; $\omega = 0,9$ e $tol = 10^{-4}$	126

Lista de tabelas

Tabela 26 - Erros relativos percentuais por rede neural especialista para $N_{HL} = 2$; $\omega = 0,9$ e $tol = 10^{-4}$	126
Tabela 27 - Resultados da máquina de comitê para $M = 10$ e $M = 6$, com $N_{HL} = 1$; $\omega = 0,1$ e $tol = 10^{-3}$	127
Tabela 28 - Resultados da máquina de comitê para $M = 10$ e $M = 6$, com $N_{HL} = 1$; $\omega = 0,5$ e $tol = 10^{-3}$ (MLP)	127
Tabela 29 - Resultados da máquina de comitê para $M = 10$ e $M = 6$, com $N_{HL} = 1$; $\omega = 0,5$ e $tol = 10^{-3}$	127
Tabela 30 - Resultados da máquina de comitê para $M = 10$ e $M = 6$, com $N_{HL} = 1$; $\omega = 0,1$ e $tol = 10^{-4}$	127
Tabela 31 - Resultados da máquina de comitê para $M = 10$ e $M = 6$, com $N_{HL} = 1$; $\omega = 0,5$ e $tol = 10^{-4}$	128
Tabela 32 - Resultados da máquina de comitê para $M = 10$ e $M = 6$, com $N_{HL} = 1$; $\omega = 0,9$ e $tol = 10^{-4}$	128
Tabela 33 - Resultados da máquina de comitê para $M = 10$ e $M = 6$, com $N_{HL} = 2$; $\omega = 0,1$ e $tol = 10^{-3}$	128
Tabela 34 - Resultados da máquina de comitê para $M = 10$ e $M = 6$, com $N_{HL} = 2$; $\omega = 0,5$ e $tol = 10^{-3}$	128
Tabela 35 - Resultados da máquina de comitê para $M = 10$ e $M = 6$, com $N_{HL} = 2$; $\omega = 0,9$ e $tol = 10^{-3}$	129
Tabela 36 - Resultados da máquina de comitê para $M = 10$ e $M = 6$, com $N_{HL} = 2$; $\omega = 0,1$ e $tol = 10^{-4}$	129

Lista de tabelas

Tabela 37 - Resultados da máquina de comitê para $M = 10$ e $M = 6$, com $N_{HL} = 2$; $\omega = 0,5$ e $tol = 10^{-4}$	129
Tabela 38 - Resultados da máquina de comitê para $M = 10$ e $M = 6$, com $N_{HL} = 2$; $\omega = 0,9$ e $tol = 10^{-4}$	129
Tabela 39 - Resultados da máquina de comitê para $M = 10$ e $M = 6$, com $N_{HL} = 1$; $\omega = 0,1$ e $tol = 10^{-3}$ (MLP).....	130
Tabela 40 - Resultados da máquina de comitê para $M = 10$ e $M = 6$, com $N_{HL} = 1$; $\omega = 0,5$ e $tol = 10^{-3}$ (MLP).....	130
Tabela 41 - Resultados da máquina de comitê para $M = 10$ e $M = 6$, com $N_{HL} = 1$; $\omega = 0,9$ e $tol = 10^{-3}$ (MLP).....	130
Tabela 42 - Resultados da máquina de comitê para $M = 10$ e $M = 6$, com $N_{HL} = 1$; $\omega = 0,1$ e $tol = 10^{-4}$ (MLP).....	130
Tabela 43 - Resultados da máquina de comitê para $M = 10$ e $M = 6$, com $N_{HL} = 1$; $\omega = 0,5$ e $tol = 10^{-4}$ (MLP).....	131
Tabela 44 - Resultados da máquina de comitê para $M = 10$ e $M = 6$, com $N_{HL} = 1$; $\omega = 0,9$ e $tol = 10^{-4}$ (MLP).....	131
Tabela 45 - Resultados da máquina de comitê para $M = 10$ e $M = 6$, com $N_{HL} = 2$; $\omega = 0,1$ e $tol = 10^{-3}$ (MLP).....	131
Tabela 46 - Resultados da máquina de comitê para $M = 10$ e $M = 6$, com $N_{HL} = 2$; $\omega = 0,5$ e $tol = 10^{-3}$ (MLP).....	131
Tabela 47 - Resultados da máquina de comitê para $M = 10$ e $M = 6$, com $N_{HL} = 2$; $\omega = 0,9$ e $tol = 10^{-3}$ (MLP).....	132

Lista de tabelas

Tabela 48 - Resultados da máquina de comitê para $M = 10$ e $M = 6$, com $N_{HL} = 2$; $\omega = 0,1$ e $tol = 10^{-4}$ (MLP).....	132
Tabela 49 - Resultados da máquina de comitê para $M = 10$ e $M = 6$, com $N_{HL} = 2$; $\omega = 0,5$ e $tol = 10^{-4}$ (MLP).....	132
Tabela 50 - Resultados da máquina de comitê para $M = 10$ e $M = 6$, com $N_{HL} = 2$; $\omega = 0,9$ e $tol = 10^{-4}$ (MLP).....	132
Tabela 51 - Resultados da máquina de comitê misto para $M = 6$, com $N_{HL} = 1$ e/ou 2; $\omega = 0,1$ e $tol = 10^{-3}$ e/ou 10^{-4}	133
Tabela 52 - Resultados da máquina de comitê misto para $M = 6$, com $N_{HL} = 1$ e/ou 2; $\omega = 0,5$ e $tol = 10^{-3}$ e/ou 10^{-4}	133
Tabela 53 - Resultados da máquina de comitê misto para $M = 6$, com $N_{HL} = 1$ e/ou 2; $\omega = 0,9$ e $tol = 10^{-3}$ e/ou 10^{-4}	133

LISTA DE ABREVIATURAS E SIGLAS

ACO	Ant Colony Optimization (otimização por colônia de formigas)
ADALINE	Adaptative linear neuron ou adaptative linear element
ART	Adaptative resonance theory
BP	Backpropagation
COMBINADOR	Combinador da máquina de comitê de redes neurais artificiais
CONE	Construct neural network ensembles
FORTTRAN 77	Fórmula Translation (compilador em linguagem científica de programação)
GA	Genetic Algorithm (algoritmo genético)
GEO	Generalized extremal optimization (otimização extrema generalizada)
GLP	Gás liquefeito de petróleo
IA	Inteligência artificial
IC	Inteligência computacional
IOP	Inherent optical properties
LDA	Linear discriminant analysis
LIDAR	Light detection and ranging
LM	Levenberg-Maraquardt
MATLAB	Matrix Laboratory (software para operações com matrizes)
MC	Método de Monte Carlo
MEM	Mixture expert model
MLP	Multi-layer Perceptron
PD	Problema direto
PI	Problema inverso
PSO	Particle swarm optimization (otimização por enxames de partículas)
purelin	Função de ativação tipo linear do Neural Network Toolbox do MATLAB2008®
RNA	Rede neural artificial
SA	Simulated annealing (recozimento simulado)

Lista de abreviaturas e siglas

SLP	Single layer perceptron
SM	Método das ordenadas discretas
SVM	Support vector machine
tansig	Função de ativação tipo tangente hiperbólica do Neural Network Toolbox do MATLAB2008®
trainlm	Função de treinamento com base na otimização de LM, do Neural Network Toolbox do MATLAB2008®

LISTA DE SÍMBOLOS

a_m^q	Combinação linear dos valores de entrada Y_l^q com $l = 1, 2, \dots, N$ ponderadas pelos pesos sinápticos $w_{m,l}^{(1)q}$ e $w_{m,0}^{(1)q}$ que conectam o neurônio m da primeira camada oculta (1), com os N nós da camada de entrada, onde $m = 1, 2, \dots, N_{H1}$, $q = 0, \dots, Q-1$, do MLP das Figuras 29 e 30.
b_k^q	Combinação linear dos valores das respostas p_m^q fornecidas pela primeira camada oculta (1), ponderadas pelos pesos sinápticos $w_{k,m}^{(2)q}$ e $w_{k,0}^{(2)q}$ que conectam o neurônio m da primeira camada oculta (1) com o neurônio k da segunda camada oculta (2), do MLP da Figura 29, onde $k = 1, 2, \dots, N_{H2}$ e $m = 1, 2, \dots, N_{H1}$.
E	Erro médio quadrático considerando os J neurônios da camada de saída do MLP da Figura 21 ao final de uma época.
E_n^q	Sinal de erro na estimativa do parâmetro de ordem n , $n = 1, 2, \dots, N_u$ no padrão q , obtido pelo confronto entre o valor alvo $Z_{n,alvo}$ do parâmetro n e o seu valor estimado $t_n^q \equiv Z_{n,i}^q$ (para MLP das Figuras 29 e 30).
E_j^q	Sinal de erro na estimativa do parâmetro radiativo de ordem j , $j = 1, 2, \dots, N_s$ no padrão q , obtido pelo confronto entre o valor alvo $Z_{j,alvo}$ do parâmetro e o seu valor estimado $\dot{Z}_j^q$ (Figura 33).
e	Limitante superior para o número de épocas de treinamento das RNA especialistas e do MLP da Figura 33.
$f(.)$	Função de ativação da primeira camada oculta do MLP nas Figuras 29 e 30.
$f_1(\mu)$	Radiação externa incidente sobre o meio em $x = 0$.
$f_2(\mu)$	Radiação externa incidente sobre o meio em $x = L$.
f_n	Coefficientes da expansão dos polinômios de Legendre.
$f'(a_m^q)$	Derivada da função de ativação f da primeira camada oculta (1) em relação à combinação linear a_m^q entre cada neurônio m da primeira camada oculta (1) e a entrada Y_l^q de cada nó l da camada de entrada na iteração q (Figuras 29 e 30).
$g(.)$	Função de ativação da segunda camada oculta (2) do MLP da Figura 29.

$g'(b_k^q)$	Derivada da função de ativação g da segunda camada oculta (2) em relação à combinação linear b_k^q entre cada neurônio k da segunda camada oculta (2) e a resposta p_m^q de cada neurônio m da primeira camada oculta (1) na iteração (Figura 29).
g_i	Função em x não linear dada pela equação 30, da rede de passagem da máquina de comitê da Figura 28, $i = 1, \dots, N$.
$h(.)$	Função de ativação das camadas de saída (3) e (2) dos MLP das Figuras 29 e 30, respectivamente.
h^*	Camada oculta do MLP da Figura 33.
$h'(s_n^q)$	Derivada da função de ativação h da camada de saída em relação à combinação linear s_n^q entre cada neurônio n da camada de saída (3) e a resposta v_k^q de cada neurônio k da segunda camada oculta (2) na iteração q (Figura 29). Para o MLP da Figura 30, é a derivada da função de ativação h da camada de saída (2) em relação à combinação linear s_n^q entre cada neurônio n da camada de saída (2) e a resposta p_m^q de cada neurônio m da camada oculta (1) na iteração q .
I	Intensidade da radiação que se propaga no meio.
I_{exp_l}	Intensidade de radiação experimental simulada (o mesmo que Y_l).
$I_{calc_l}(\vec{Z})$	Intensidade de radiação calculada na solução do problema direto de transferência radiativa.
$I(x, \mu)$	Valores calculados da intensidade de radiação em função de x e de μ .
$I^+(0, \mu)$	Condição de contorno, intensidade de radiação que entra em $\tau = 0$.
$I^+(\tau_0, \mu)$	Dado experimental, intensidade de radiação que sai em $\tau = \tau_0$.
$I^-(\tau_0, \mu)$	Condição de contorno, intensidade de radiação que entra em $\tau = \tau_0$.
$I^-(0, \mu)$	Dado experimental, intensidade de radiação que sai em $\tau = 0$.
I_m^i	Discretização da intensidade de radiação nos domínios espacial e angular.
I_m^1	Aproximação da condição de contorno em $\tau = 0$.
$I_m^{N_x}$	Aproximação da condição de contorno em $\tau = \tau_0$.
I_n^i	Discretização da intensidade de radiação devido ao espalhamento interno nos domínios espacial e angular.
$(I_m^i)^p$	Estimativa para a intensidade de radiação na iteração p .

Lista de símbolos

$(I_m^i)^{p+1}$	Estimativa para a intensidade de radiação na iteração $p + 1$.
I_{mef}^1	Intensidade efetiva de radiação que emerge do meio em $\tau = 0$.
$I_{mef}^{N_x}$	Intensidade efetiva de radiação que emerge do meio em $\tau = \tau_0$.
k_a	Coefficiente de absorção.
L	Espessura física do meio medida ao longo do eixo $-x$.
N_{HL}	Número de camadas ocultas das RNA especialistas tipo MLP de acordo com as Figuras 29 e 30, para efeito de análise de resultados na Tabela 2 (p 90 - 92).
N_H	Número de neurônios por camada oculta do MLP de acordo com as Figuras 29 e 30.
N_{H1} e N_{H2}	Número de neurônios na primeira e segunda camadas ocultas do MLP de acordo com a Figura 29.
N_h	Número de neurônios na camada oculta do MLP (Figura 33) do combinador de saída da máquina de comitê com modelo de acordo com o mostrado na Figura 32.
N_s	Número de neurônios da camada de saída s do MLP da Figura 33.
N_x	Número de pontos no domínio espacial.
NET_i	Representa o produto do vetor peso sináptico pelo vetor de entrada do neurônio especialista.
NET_j	Representa o produto do vetor de entrada pelo vetor de pesos sinápticos da rede de passagem da máquina de comitê tipo dinâmica (Figura 28) na equação 30.
O	Vetor das saídas do MLP da Figura 21.
O_j	Saída calculada para o neurônio j , $j = 1, 2, \dots, N_o$ na Figura 21.
p	Número de iterações das marchas para frente e para trás na equação (19).
p_m^q	Resposta do neurônio m , $m = 1, 2, \dots, N_{HL}$, $q = 0, \dots, Q-1$, da primeira camada oculta (1) do MLP das Figuras 29 e 30, de acordo com a equação (33).
$p(\mu, \mu')$	Função de fase.
$P_n(\mu)$	Polinômios de Legendre em μ .
$P_n(\mu')$	Polinômios de Legendre em μ' .
q	Ordem do padrão de treinamento.
Q	Limite superior do número de padrões de treinamento q .

Lista de símbolos

q^i	Discretização espacial do espalhamento interno isotrópico usando quadratura de Gauss-Legendre.
r_l	Número pseudo-randômico no intervalo $[-1,1]$.
RNA_i	Rede neural artificial especialista de ordem i , $i = 1, \dots, M$.
s	Camada de saída do MLP da Figura 33.
$S(x, \mu)$	Fonte interna de radiação.
S_M	Método das Ordenadas Discretas.
s_n^q	Combinação linear dos valores das respostas v_k^q fornecidas pela segunda camada oculta (2), ponderadas pelos pesos sinápticos $w_{n,k}^{(3)q}$ e $w_{n,0}^{(3)q}$ que conectam o neurônio k da segunda camada oculta (2) com o neurônio n da camada de saída, do MLP da Figura 29 e equação (36), onde $k = 1, 2, \dots, N_{H2}$ e $n = 1, 2, \dots, N_u$. Para Figura 30 é a combinação linear dos valores das respostas p_m^q fornecidas pela camada oculta (1), ponderadas pelos pesos sinápticos $w_{n,m}^{(2)q}$ e $w_{n,0}^{(2)q}$ que conectam o neurônio m da camada oculta (1) ao neurônio n da camada de saída, do MLP da Figura 30 de acordo com a equação (43d), onde $m = 1, 2, \dots, N_{H1}$ e $n = 1, 2, \dots, N_u$.
T	Vetor alvo (<i>target</i>) do MLP da Figura 21.
T_j	Saída desejada para o neurônio j , $j = 1, 2, \dots, N_o$ na Figura 21.
t_n^q	Resposta do neurônio n , $n = 1, 2, \dots, N_u$, $q = 0, \dots, Q-1$, da camada de saída (3) e (2) do MLP da Figura 29 e 30, de acordo com as equações (38) e (42), respectivamente.
$\vec{t}_i^q$	Vetor das respostas t_n^q , onde $n = 1, 2, \dots, N_u$, $q = 0, \dots, Q-1$ e $i = 1, 2, \dots, M$, para os MLP das Figuras 29 e 30.
$\vec{t}_i^{q=0}$	Estimativa inicial para o vetor $\vec{t}_i^q$.
u	Neurônio da camada oculta do MLP (ver Figura 33) do combinador da máquina de comitê de acordo com o modelo da Figura 32.
v_k^q	Resposta do neurônio k , $k = 1, 2, \dots, N_{H2}$, $q = 0, \dots, Q-1$, da segunda camada oculta (2) do MLP da Figura 29, de acordo com a equação (35).
$w_{m,l}^{(1)}$	Peso sináptico que conecta o nó l de entrada, onde $l = 1, 2, \dots, N$, ao neurônio m da primeira camada oculta (1), com N_{H1} neurônios, do MLP das Figuras 29 e 30.

Lista de símbolos

$w_{m,0}^{(1)}$	Bias do neurônio m da primeira camada oculta (1) do MLP das Figuras 29 e 30, onde $m = 1, 2, \dots, N_{H1}$.
$\Delta w_{m,l}^{(1)}$	Variação do peso sináptico que conecta o nó l de entrada com o neurônio m da primeira camada oculta (1) (Figuras 29 e 30).
$w_{k,m}^{(2)}$	Peso sináptico que conecta o neurônio m da primeira camada oculta (1) ao neurônio k da segunda camada oculta (2) do MLP da Figura 29, onde $k = 1, 2, \dots, N_{H2}$, $m = 1, 2, \dots, N_{H1}$.
$w_{k,0}^{(2)}$	Bias do neurônio k da segunda camada oculta (2) do MLP da Figura 29, onde $k = 1, 2, \dots, N_{H2}$.
$\Delta w_{k,m}^{(2)}$	Variação do peso sináptico que conecta o neurônio m da primeira camada oculta (1) ao neurônio k da segunda camada oculta (2) (Figura 29).
$w_{n,k}^{(3)}$	Peso sináptico que conecta o neurônio k da segunda camada oculta (2) ao neurônio n da camada de saída (3) do MLP da Figura 29, onde $n = 1, 2, \dots, N_u$, $k = 1, 2, \dots, N_{H2}$.
$w_{n,0}^{(3)}$	Bias do neurônio n da camada de saída (3) do MLP da Figura 29, onde $n = 1, 2, \dots, N_u$.
$\Delta w_{n,k}^{(3)}$	Variação do peso sináptico que conecta o neurônio k da segunda camada oculta (2) ao neurônio n da camada de saída (3) (Figura 29).
$w_{u,n}^{(h^*)}$	Peso sináptico que conecta o nó n de entrada do MLP da Figura 33 ao neurônio u da camada oculta h^* .
$w_{u,0}^{(h^*)}$	Bias do neurônio u da camada oculta h^* do MLP da Figura 33, onde $u = 1, 2, \dots, N_h$.
$w_{j,u}^{(s)}$	Peso sináptico que conecta o neurônio u da camada oculta h^* do MLP da Figura 33 ao neurônio j da camada de saída s .
$\Delta w_{j,u}^{(s)}$	Valor atual do peso sináptico que conecta o neurônio u da camada oculta h^* ao neurônio j da camada de saída s (Figura 33).
$w_{j,0}^{(s)}$	Bias do neurônio j da camada de saída s do MLP da Figura 33, onde $s = 1, 2, \dots, N_s$.

Lista de símbolos

$w_{n,m}^{(2)}$	Peso sináptico que conecta o neurônio m da camada oculta (1) do MLP da Figura 30 ao neurônio n da camada de saída (2), onde $m = 1,2, \dots, N_{H1}$, $n = 1,2, \dots, N_u$.
$w_{n,0}^{(2)}$	Bias do neurônio n da camada de saída (2) do MLP da Figura 30, onde $n = 1,2, \dots, N_u$.
$\Delta w_{n,m}^{(2)}$	Valor atual do peso sináptico que conecta o neurônio m da camada oculta (1) ao neurônio n da camada de saída (2) (Figura 30).
x	Coordenada espacial medida ao longo do eixo $-x$.
$\mathbf{X}$	Vetor das entradas nas Figuras 21, 26 e 28.
$\mathbf{X}_j$	Vetores de entradas da máquina de comitê da Figura 27, $j = 1, \dots, N$.
$\mathbf{Y}$	Saída global das máquinas de comitê das Figuras 26, 27 e 28.
Y_l	Número de dados experimentais obtidos de acordo com a equação (25), com $l = 1, 2, \dots, N$.
$\vec{Y}^q$	Vetor das excitações de entradas Y_l^q das RNA $_i$, $i = 1, \dots, M$, $l = 1, 2, \dots, N$ no padrão q , $q = 0, \dots, Q-1$.
$\vec{Y}$	Vetor padrão de dados experimentais que não participaram da fase de treinamento.
$\vec{Z}$	Vetor dos parâmetros radiativos, incógnitas do problema inverso de transferência radiativa, dado pela equação (24) e saídas da máquina de comitê de acordo com o modelo representado na Figura 1.
$\vec{Z}_{exato}$	Vetor cujos componentes são os valores exatos dos parâmetros radiativos, usados para obter os dados experimentais simulados Y_l para a solução do problema inverso em transferência radiativa.
$\vec{Z}_{alvo}^q$	Vetor alvo, cujos elementos são os valores padrões para os parâmetros radiativos, usados na fase de treinamento e gerados randomicamente.
Z_{nalvo}^q	Componentes do vetor $\vec{Z}_{alvo}^q$, dos padrões de saídas usadas na fase de treinamento das RNA $_i$, $i = 1,2, \dots, M$, $n = 1,2, \dots, N_u$ e $q = 0, \dots, Q-1$.
Z_{jalvo}^q	Componentes do vetor $\vec{Z}_{alvo}^q$, dos padrões de saídas usadas na fase de treinamento do MLP da Figura 33, $j = 1,2, \dots, N_s$ e $q = 0, \dots, Q-1$.
$\vec{Z}_i^q$	Vetor de saída da rede neural especialista de ordem i , $i = 1, 2, \dots, M$, e padrão de ordem q , $q = 0, \dots, Q-1$, que é idêntico ao vetor $\vec{t}_i^q$ (Figuras 29 e 30).

Lista de símbolos

$\vec{Z}_i$	Vetor de saída da rede neural especialista de ordem i , $i = 1, 2, \dots, M$, na fase de generalização.
$Z_{n,i}^q$	Componente do vetor $\vec{Z}_i^q$, de saída da RNA $_i$, $i = 1, 2, \dots, M$, $n = 1, 2, \dots, N_u$ e $q = 0, \dots, Q-1$, que assumem os valores das estimativas t_n^q (ver Figuras 29 e 30).
$Z_{n,i}$	Componente do vetor $\vec{Z}_i$, de saída da RNA $_i$, $i = 1, 2, \dots, M$, $n = 1, 2, \dots, N_u$ na fase de generalização.
$\vec{Z}_i^{q=0}$	Estimativa inicial para o vetor $\vec{Z}_i^q$, na fase de treinamento.
$\vec{Z}^q$	Vetor média aritmética dos vetores $\vec{Z}_i^q$ de saída das redes neurais especialistas (Figuras 29 e 30), até a ordem i , $i = 1, 2, \dots, M$, e padrão q , $q = 0, \dots, Q-1$.
$\vec{Z}$	Vetor média aritmética dos vetores $\vec{Z}_i$ de saída das redes neurais especialistas (Figuras 29 e 30), até a ordem i , $i = 1, 2, \dots, M$, na fase de generalização.
$\bar{Z}_n^q$	Componente do vetor média aritmética das saídas das RNA $_i$ especialistas, $i = 1, 2, \dots, M$, $n = 1, 2, \dots, N_u$ (Figuras 29 e 30).
$\bar{Z}_n$	Componente do vetor média aritmética das saídas das RNA $_i$ especialistas, $i = 1, 2, \dots, M$, $n = 1, 2, \dots, N_u$ (Figuras 29 e 30), na fase de generalização.
z_u^q	Combinação linear dos valores de entrada $\bar{Z}_n^q$, com $n = 1, 2, \dots, N_u$ ponderadas pelos pesos sinápticos $w_{u,n}^{(h^*)q}$ e $w_{u,0}^{(h^*)q}$ que conectam o neurônio u da camada oculta h^* com os N_u nós da camada de entrada, onde $u = 1, 2, \dots, N_h$ e $q = 0, 1, \dots, Q-1$, no MLP da Figura 33.
$\hat{z}_u^q$	Resposta do neurônio u , $u = 1, 2, \dots, N_h$, $q = 0, \dots, Q-1$, da camada oculta h^* do MLP da Figura 33, de acordo com a equação (46).
$\hat{z}_j^q$	Combinação linear dos valores das respostas $\hat{z}_u^q$ fornecidas pela camada oculta h^* , ponderadas pelos pesos sinápticos $w_{j,u}^{(s)q}$ e $w_{j,0}^{(s)q}$ que conectam os neurônios da camada oculta h^* com os neurônios da camada de saída s , do MLP da Figura 33, onde $j = 1, 2, \dots, N_s$.
$\dot{z}_j^q$	Resposta do neurônio j , $j = 1, 2, \dots, N_s$, $q = 0, \dots, Q-1$, da camada de saída s do MLP da Figura 33, de acordo com a equação (48).

Lista de símbolos

$\vec{Z}^q$	Vetor de saída do MLP (Figura 33) do combinador de saída da máquina de comitê de acordo com o modelo da Figura 32, cujos componentes são as respostas $\dot{Z}_j^q$ do neurônio j , $j = 1, 2, \dots, N_s$, $q = 0, \dots, Q-1$, da camada de saída s do MLP da Figura 33.
α_n	Pesos da quadratura de Gauss-Legendre correspondentes aos pontos de colocação, μ_n , $n = 1, 2, \dots, M$.
β	Coeficiente de extinção total.
$\gamma(\cdot)$	Função de ativação da camada oculta h^* do MLP da Figura 33.
$\gamma'(z_u^q)$	Derivada da função de ativação γ da camada oculta h^* em relação à combinação linear z_u^q entre os valores de entrada $\bar{Z}_n^q$, com $n = 1, 2, \dots, N_u$ e cada neurônio u da camada oculta h^* na iteração q (Figura 33).
$\delta_n^{(3)q}$	Gradiente local ou sinal de retro-propagação do erro partindo do neurônio n da camada de saída (3) do MLP, para o neurônio k da segunda camada oculta (2), na iteração anterior q (Figura 29).
$\delta_n^{(2)q}$	Gradiente local ou sinal de retro-propagação do erro que parte do neurônio n da camada de saída (2) para o neurônio m da camada oculta (1), na iteração anterior q (Figura 30).
$\delta_k^{(2)q}$	Gradiente local ou sinal de retro-propagação do erro que parte do neurônio k da segunda camada oculta (2) para o neurônio m da primeira camada oculta (1), na iteração anterior q (Figura 29).
$\delta_m^{(1)q}$	Gradiente local ou sinal de retro-propagação do erro que parte do neurônio m da primeira camada oculta (1) para o nó l da camada de entrada, na iteração anterior q (Figuras 29 e 30).
$\delta_j^{(s)q}$	Gradiente local ou sinal de retro-propagação do erro que parte do neurônio j da camada de saída s para o neurônio u da camada oculta h^* , na iteração anterior q (Figura 33).
$\Delta\tau$	Incremento da variável óptica na malha computacional.
ε	Tolerância requerida na solução do problema direto.
η	Taxa de aprendizado.
θ	Ângulo entre o eixo-x positivo e a direção do feixe de radiação.
θ'	Ângulo entre o eixo-x positivo e a direção do feixe de radiação refletida.
μ	Cosseno do ângulo polar θ .

Lista de símbolos

μ'	Cosseno do ângulo polar θ' .
μ_m	Variável angular discretizada, $m = 1, \dots, M$.
ρ_1	Refletividade difusa na superfície interna em $x = 0$.
ρ_2	Refletividade difusa na superfície interna em $x = L$.
σ	Desvio padrão dos erros de medição.
σ_s	Coeficiente de espalhamento.
τ	Variável óptica.
τ_0	Espessura óptica do meio.
$\bar{\tau}_0$	Valor máximo escolhido para a τ_0 (limite superior artificial).
τ_i	Variável óptica discretizada, $i = 1, \dots, N_x$.
χ	Função de ativação da camada de saída s do MLP da Figura 33.
$\chi'(\hat{Z}_j^q)$	Derivada da função de ativação χ da camada de saída s em relação à combinação linear $\hat{Z}_j^q$ entre cada neurônio j da camada de saída e cada neurônio u da camada oculta h^* na iteração q (Figura 33).
ω	Albedo de espalhamento simples.

SUMÁRIO

1	INTRODUÇÃO	29
1.1	O tema e sua relevância	29
1.2	O problema e sua solução	36
1.3	Estrutura da dissertação	38
2	PROBLEMAS DIRETO E INVERSO EM TRANSFERÊNCIA RADIATIVA	39
2.1	Problema direto em transferência radiativa	39
2.1.1	<u>Formulação matemática do problema direto em transferência radiativa</u>	39
2.1.2	<u>Solução numérica do problema direto em transferência radiativa</u>	42
2.2	Problema inverso em transferência radiativa	49
2.2.1	<u>Formulação e solução do problema inverso em transferência radiativa</u>	51
3	FUNDAMENTOS SOBRE REDES NEURAIS E MÁQUINAS DE COMITÊ E SUA APLICAÇÃO À SOLUÇÃO DO PROBLEMA INVERSO EM TRANSFERÊNCIA RADIATIVA	53
3.1	Histórico sobre redes neurais e máquinas de comitê de redes neurais artificiais	53
3.2	Fundamentação teórica sobre redes neurais artificiais	55
3.2.1	<u>O neurônio biológico</u>	55
3.2.2	<u>O neurônio artificial</u>	56
3.2.3	<u>Tipos de funções de ativação</u>	58
3.2.3.1	Função degrau (threshold function)	58
3.2.3.2	Função linear	58
3.2.3.3	Função log-sigmóide	59
3.2.3.4	Função tan-sigmóide	59
3.2.4	<u>O neurônio artificial de entrada vetorial</u>	60
3.2.5	<u>A rede neural artificial</u>	60
3.2.6	<u>Paradigmas de treinamento de redes neurais artificiais</u>	62
3.2.7	<u>Perceptron</u>	65
3.2.8	<u>Perceptron de múltiplas camadas</u>	66
3.3	Fundamentação teórica sobre máquinas de comitê de redes neurais artificiais	72
3.3.1	<u>Classificação das máquinas de comitê de redes neurais artificiais</u>	73
3.3.1.1	Estruturas Estáticas	73

3.3.1.2 Estruturas Dinâmicas	74
3.4 Modelagem e formulação matemática para a solução do problema inverso em transferência radiativa usando redes neurais artificiais.....	76
3.5 Modelagem e formulação matemática para a solução do problema inverso em transferência radiativa usando máquinas de comitê de redes neurais artificiais.....	81
3.5.1 <u>Modelagem</u>	81
3.5.2 <u>Formulação</u>	83
4 RESULTADOS E DISCUSSÃO.....	87
4.1 Metodologia aplicada.....	87
4.2 Comparativo dos resultados obtidos com as RNA especialistas e com as máquinas de comitê	91
5 CONCLUSÃO	106
REFERÊNCIAS BIBLIOGRÁFICAS.....	108
APÊNDICE A – Resultados dos demais testes efetuados	118

1 INTRODUÇÃO

1.1 O tema e sua relevância

A solução de problemas diretos e inversos associados à transferência radiativa em meios participantes, i.e. meios absorvedores, espalhadores e emissores, tem atraído a atenção de muitos pesquisadores devido à sua relevante aplicação em várias áreas de elevado interesse, científico e tecnológico (SILVA NETO; BECCENERI, 2009; SILVA NETO; MOURA NETO, 2001). Spurr et al. (2007), Verhoef e Bach (2003), Hanan (2001), Fause, Isaka e Guillemet (2001) apresentaram soluções em trabalhos de pesquisa em avaliação de risco de incêndios, modelos climáticos ambientais regionais e globais e, sensoriamento remoto da Terra. Kim e Charette (2007), Montero, Roberty e Silva Neto (2004) propuseram soluções de problemas em reconstrução de imagens, eletricidade e magnetismo e tomografia.

Maffione, Voss, e Honey (1993), Maffiione e Dana, (1997) e Tao, Sanchez e McCormick (1990) realizaram trabalhos em problemas inversos que foram feitos com o propósito de calcular as propriedades óticas inerentes a um meio (inherent optical properties – IOP's), no caso a água, através de fontes luminosas artificiais introduzidas no meio.

Gordon e Zhang, (1995), Zhang e Gordon, (1997) e Casttrall (2001) pesquisaram problema análogo ao problema inverso em ótica atmosférica que consiste na estimação de propriedades, por exemplo: função de espalhamento, albedo de espalhamento simples através de medidas de radiância do céu.

Barichello, Siewert e Garcia (1997) estimaram a intensidade da radiação incidente em um meio unidimensional, com dependência polar, a partir da medida da densidade da radiação (momento zero) em algumas posições do interior do meio.

Fukshansky-Kazarinova et al. (1998) estimaram o coeficiente de extinção total e a função de fase de espalhamento anisotrópico usando medidas experimentais do fluxo radiativo no interior do meio.

Sedano (2007) estimou simultaneamente a temperatura do gás e a emissividade em um forno utilizando a análise inversa em radiação térmica.

França et al. (2002) utilizaram problemas inversos no projeto de sistemas térmicos com dominância em transferência radiativa. França e Howell (2006) apresentaram uma metodologia para solucionar problemas inversos transientes de radiação confinada em

processos de aquecimento que requerem controle refinado de temperatura. França et al. (2009) aplicaram problemas inversos em projetos de iluminação em um recinto retangular tridimensional.

Knupp, Silva Neto e Sacco (2009), Galski, et al. (2009), Sousa et al. (2007), Zhou et al. (2002) e Oliveira et al. (2008) propuseram soluções em trabalhos de óptica hidrológica e de estimativa de propriedades radiativas.

Em algumas aplicações pode ser mais interessante usar medidas experimentais dos momentos da intensidade da radiação, sendo até mesmo possível em alguns casos obter medidas no interior do meio, e não somente no contorno do mesmo, e.g. aplicações em oceanografia McCormick (1994, 1996, 2000), Leathers e McCormick (1997), Sundman, Sanchez e McCormick (1990), McCormick e Kaskas (2000).

Montero, Roberty e Silva Neto (2001) estimaram o coeficiente de absorção em um meio heterogêneo bidimensional, absorvedor puro, visando aplicações em tomografia computadorizada, empregando uma partição de domínio consistente com feixes divergentes de radiação.

Li e Özisik (1993) estimaram o termo fonte e a refletividade em uma superfície de contorno a partir da medida experimental da intensidade da radiação que sai do meio para diferentes ângulos polares. Silva Neto e Özisik (1993) obtiveram estimativas para a condutividade térmica e propriedades radiativas em um problema de transferência de calor por condução e radiação térmica usando medidas experimentais sintéticas da temperatura no interior do meio e das intensidades da radiação em função do ângulo polar nas superfícies externas do mesmo. Silva Neto e Özisik (1995a) estimaram o albedo de espalhamento simples, a espessura óptica e os coeficientes da expansão da função de fase de espalhamento anisotrópico em polinômios de Legendre, também usando medidas da intensidade da radiação emergente com dependência no ângulo polar. Li e Yang (1997) resolveram este mesmo problema, com o mesmo tipo de dado experimental, empregando um algoritmo genético, enquanto Silva Neto e Özisik (1995b) usaram o método de Levenberg-Marquardt. Pinheiro (2004) usou em sua tese de doutorado o método de Levenberg-Marquardt, o funcional de Tikhonov e a função objetivo construída com as distâncias de Bregman e momentos da q -discrepância para solucionar o problema inverso em transferência radiativa.

Kauati, Silva Neto e Roberty (1999; 2001), Silva Neto e Roberty (1998a; 1998b e 1998c) estimaram o coeficiente de extinção total em um meio unidimensional homogêneo e espalhador isotrópico usando a técnica fonte-detector com medidas experimentais da intensidade da radiação que sai do meio. Kauati, Silva Neto e Roberty (2000) estenderam a

aplicação da técnica fonte-detector para meios espalhadores anisotrópicos.

Soeiro, Carvalho e Silva Neto (2000) usaram o método do recozimento simulado (*Simulated Annealing*), para obter estimativas das propriedades térmicas de polímeros. Silva Neto e Soeiro (2000) usaram os métodos de otimização global, LM e SA, para solucionar o problema inverso de transferência radiativa. Silva Neto e Soeiro (2002) estimaram a função de fase de um espalhamento anisotrópico com uma combinação dos métodos baseados no gradiente e estocásticos de otimização global. Silva Neto e Soeiro (2003) usaram métodos híbridos para obter a solução de formulações implícitas de problemas inversos em transferência de calor.

Flores et al. (2001) obtiveram a solução do problema inverso de transferência radiativa usando programação matemática e métodos híbridos.

Formulações implícitas de problemas inversos para estimativas de parâmetros de transferência radiativa, na qual uma função custo é minimizada, têm sido largamente empregadas para a solução de tais problemas.

Soeiro e Silva Neto (2006), Silva Neto (2002) e Pinheiro (2004) na maioria dos casos têm utilizado métodos baseados no gradiente para a minimização da função custo. Esses métodos são de convergência local, e muitas vezes ficam confinados em mínimos locais, no entanto, apresentam a vantagem de terem uma alta eficiência computacional.

Silva Neto e Soeiro (2002, 2003), Galski et al. (2009), Sousa et al. (2007) têm utilizado outra aproximação para a minimização da função custo que envolve métodos de otimização global, com um crescente interesse em direção aos métodos estocásticos, em particular as heurísticas bio-inspiradas. Com a implementação desses métodos consegue-se uma aproximação muito precisa do mínimo global, mas a um custo computacional muito elevado.

Silva Neto e Soeiro (2005), Soeiro, Soares e Silva Neto (2004) em vários trabalhos têm empregado uma combinação de métodos baseados no gradiente com os métodos de otimização estocásticos para a solução de problemas inversos. Nessas aproximações, os métodos estocásticos têm sido usados para obter estimativas para os métodos baseados no gradiente. É uma tentativa de aliar a eficiência computacional dos métodos baseados no gradiente com a garantia de convergência de uma cuidadosa implementação dos métodos de otimização estocásticos.

Muitas abordagens têm sido desenvolvidas para a obtenção da solução dos problemas diretos e inversos em transferência radiativa. Hansen e Travis (1974) e Lenoble (1977) revisaram métodos para a solução do problema direto em transferência radiativa, produzindo

excelentes artigos, e McCormick (1992) realizou o mesmo com relação aos problemas inversos.

Recentemente métodos estocásticos de Monte Carlo para a solução do problema direto têm despertado interesse crescente entre os pesquisadores, entre eles, Maurense, Viêlmo e França (2007), Chen e Liou (2006), Battaglia e Mantovani (2005) e Postylyakov, (2004), bem como variantes do Método das Ordenadas Discretas por Çayan e Selçuk (2007) e Chalhoub (2003, 2005), que foi inicialmente proposto por Wick (1943) e Chandrasekhar (1944, 1950).

Comparações entre os vários métodos para a solução de problemas diretos e inversos em transferência radiativa têm sido realizadas por alguns pesquisadores: Jensen et al. (2007) Bulgarelli, e Doyle (2004), Chalhoub et al. (2003), Chalhoub, Campos Velho e Silva Neto (2007), com o intuito de obter maior rapidez e precisão na solução de tais problemas, mais particularmente no que diz respeito a aliar convergência com baixo custo computacional.

Redes neurais artificiais (RNA) têm sido utilizadas para a solução de problemas inversos em transferência radiativa e condução de calor. Bokar (1999) usou RNA para estimar parâmetros de transferência radiativa e Boillereaux, Cadet e Le Bail (2003) usaram RNA para obter propriedades térmicas de amostras de alimentos com o objetivo de manter a segurança e qualidade dos alimentos.

Recentemente, numa tentativa de obter uma hibridização entre os métodos para a solução do problema inverso em transferência radiativa que sejam mais rápidos e eficientes, Soeiro et al. (2004) e Soeiro, Soares e Silva Neto (2004) têm empregado RNA para obter estimativas iniciais para o método de otimização de Levenberg-Marquardt (LM). Soeiro e Silva Neto (2006) usaram RNA para obter estimativas iniciais para a solução do problema inverso de transferência de calor para um meio com duas camadas. Chalhoub, Silva Neto e Soeiro (2007) têm obtido soluções pelo método de Monte Carlo (MC), com diferentes níveis de precisão, para serem utilizadas no treinamento de RNA, tornando-as bastante robustas para produzir estimativas para as incógnitas, mesmo para os casos teste com dados experimentais corrompidos por ruídos.

No entanto, o uso de RNA para a solução de tarefas computacionalmente complexas pode levar a resultados pouco exatos, devido principalmente ao confinamento em mínimos locais sobre a superfície de erro. Uma solução possível para tal problema é aplicar o princípio cartesiano de se dividir um problema complexo em vários problemas mais simples, distribuindo tarefas entre um determinado número de especialistas, que são RNA do tipo Perceptron de Múltiplas Camadas (MLP), que dividirão o espaço das entradas em um conjunto de subespaços.

Haykin (2001) afirma que uma combinação de especialistas constitui o que se chama máquina de comitê. Basicamente, a máquina de comitê combina todas as soluções individuais dos especialistas em uma solução única e global, que supostamente é superior àquela alcançada por qualquer um dos especialistas individualmente.

As pesquisas com máquinas de comitê tiveram início com Nilsson (1965) que foi quem idealizou a primeira máquina de comitê, que era constituída por uma estrutura formada por uma camada de perceptrons elementares seguida de um perceptron de votação na segunda camada. A partir daí, devido à sua praticidade, eficácia na generalização do aprendizado e baixo custo computacional extra, as máquinas de comitê têm sido utilizadas de forma crescente durante a última década em diversos trabalhos de vários pesquisadores em amplas áreas de interesse científico e tecnológico.

Magoulas e Dounias (2007) escreveram sobre o uso da Inteligência Computacional na medicina, mostrando o contínuo crescimento do interesse nos métodos de inteligência artificial a partir de 2001, mais notadamente de 2005 para cá, onde a técnica de máquina de comitê ocupa o quarto lugar das metodologias de IC utilizadas, com 273 publicações até 2006, onde 35% das pesquisas se concentram em diagnósticos médicos. Dentre esses trabalhos de pesquisa realizados muitos se concentram em classificadores múltiplos, com o intuito de melhorar a performance de classificação. Aqui citaremos alguns deles, em várias áreas e de forma geral.

Silva (2005) em sua dissertação de mestrado propôs a utilização de uma máquina estática de comitê formada por redes neurais artificiais tipo MLP, como uma nova estratégia para classificar nódulos mamográficos, que podem ser identificados por radiologistas por meio de exames de mamografia como benignos ou malignos, usando processamento de imagens para classificar as características dos nódulos de acordo com a forma, a textura e a nitidez da borda.

Cho e Won (2007) utilizaram uma máquina de comitê composta por um conjunto de redes neurais artificiais e um combinador Bayesiano na saída, aliados a métodos estocásticos sobre subconjuntos de genes para classificar vários tipos de cânceres. Minku e Ludermir (2008) apresentaram o método CONE (Construct Neural Network Ensembles) que explicitamente particiona o espaço de entrada através de um método de agrupamento que assegura a diferentes redes neurais se especializarem em diferentes regiões do espaço de entrada, trabalhando de acordo com a máxima dividir e conquistar, com o objetivo de aumentar a exatidão dos resultados.

Carcenac (2007) apresentou uma arquitetura versátil e modular de redes neurais artificiais integradas para reconhecimento em super resolução de imagens de faces humanas de baixa resolução inicial. Cada módulo era um MLP, treinado com o método de Levenberg-Marquardt. Juntos esses módulos de construção genérica, foram utilizados para obter resultados finais de imagens de faces humanas em super resolução a partir da combinação de imagens de saídas de baixa resolução com imagens de saídas de alta resolução, obtidas individualmente por cada bloco genérico.

Maglogiannis, Zafiroopoulos e Anagnostopoulos (2007) propuseram o uso de um SVM (Support Vector Machine) para prognósticos e diagnósticos de câncer de mama, comparando os resultados com os obtidos usando classificadores Bayesianos e com os obtidos usando redes neurais artificiais.

Silva, Hernandez, e Rangayan (2008) utilizaram uma máquina de comitê, cujos componentes eram redes neurais artificiais do tipo MLP, para classificar massas achadas em mamografias, como benignas ou malignas, comparando os resultados com os obtidos por MLP (Multi-Layer Perceptrons), SLP (Single Layer Perceptron) e por LDA (Linear Discriminant Analysis) os testes foram realizados utilizando a distribuição student-t.

Silva e Thomé (2003) abordaram o desenvolvimento de um protótipo que utiliza um time de redes neurais artificiais para o reconhecimento de caracteres manuscritos. O processo de desenvolvimento envolve também a avaliação de alguns métodos de extração de características. Na análise dos resultados o desempenho alcançado pelo protótipo é comparado ao alcançado pelo ser humano.

Puma-Villanueva e Zuben (2008) utilizaram uma máquina de comitê de redes neurais artificiais para predição de séries temporais. Fernandes et al, (2004) usaram uma máquina de comitê de redes neurais artificiais para classificar imagens obtidas por LIDAR, para detecção remota de incêndios florestais.

Castro, Braga e Andrade (2005) aplicaram um modelo ensemble de redes neurais artificiais para previsão de séries temporais não-estacionárias.

Goodband, Haas e Mills (2006) apresentaram um novo algoritmo, MEM – mixture expert model, que produz uma arquitetura aproximada ótima para mapeamento contínuo, na aplicação de terapia com radiação modulada à base de espalhamento de fótons. O algoritmo utiliza o fuzzy – C means clustering algorithm para a partição dos dados antes de iniciar o treinamento. Uma redução do tamanho dos conjuntos de treinamento permite a utilização do algoritmo de Levenberg-Marquardt, resultando numa redução do erro de validação e no tempo de treinamento em relação ao que uma simples RNA alcança.

Hornburg (2007) utilizou uma máquina de comitê de redes neurais artificiais para filtragem de e-mails classificando-os como spams ou não-spams, comparando os resultados obtidos com outros métodos.

Paula et al. (2005) propuseram um sistema de autenticação baseado no padrão de digitação humana, usando uma combinação de RNA e Support Vector Machine (SVM).

Hirayama et al. (2004) usaram uma máquina de comitê composta por dez RNA, em paralelo, para reconhecimento da qualidade de vapor de álcool combustível e do poder calorífico do gás combustível GLP em um nariz eletrônico, obtendo 97% de acerto.

Coelho (2006) propôs em seu trabalho de dissertação de mestrado, o uso de uma metodologia de criação de ensembles, para problemas de classificação de padrões através de voto majoritário em fronteiras de decisão, onde os componentes são redes neurais artificiais tipo MLP. Coelho afirma que para que um ensemble apresente um desempenho superior aos dos componentes individualmente, é preciso que os componentes tenham bom desempenho individual, o que não ocorre sempre. Portanto, devem-se escolher os componentes que apresentem os melhores desempenhos individuais para compor o ensemble (Esta metodologia é válida também para as máquinas de comitê de RNA).

Magalhães (2007) propôs em seu trabalho de dissertação de mestrado, a elaboração de um estudo em arquiteturas de agrupamentos de redes neurais artificiais (máquinas de comitê e redes modulares) fazendo uso de outros paradigmas e algoritmos de aprendizagem de máquina em um ambiente de alto poder de processamento (cluster). As pesquisas com máquinas de comitê de redes neurais artificiais para a solução de problemas inversos em transferência radiativa é algo ainda bastante incipiente. Destacam-se entre essas pesquisas os seguintes trabalhos realizados:

1. Gil et al. (2008) utilizaram uma máquina de comitê de redes neuro-fuzzy combinada com redes neurais artificiais em cascata, para a solução de problemas inversos em transferência radiativa.
2. Oliveira et al. (2008) propuseram uma máquina de comitê de redes neurais artificiais do tipo estática para solucionar o problema inverso de transferência radiativa, em que o combinador de saída da máquina de comitê consiste de uma rede neural tipo MLP desmembrada em quatro sub-redes neurais simplificadas, cada qual usada para obter estimativas para cada parâmetro radiativo isoladamente.

1.2 O problema e sua solução

O objetivo geral deste trabalho é a solução de um problema inverso em transferência radiativa, ou seja, a determinação das propriedades radiativas: a espessura óptica τ_0 , o albedo de espalhamento simples ω , e as refletividades difusas ρ_1 e ρ_2 , que representam o processo físico do modelo, i.e., um meio participante, cinza, homogêneo e unidimensional, a partir do conhecimento de dados simulados computacionalmente, $\vec{Y}$, obtidos adicionando-se um ruído pseudo-randômico às intensidades de radiação calculadas obtidas a partir da solução do problema direto em transferência radiativa.

Para a solução do problema inverso em transferência radiativa são usados dois modelos de máquinas de comitê de redes neurais artificiais do tipo estática que utilizam a média aritmética de conjunto (ensemble), conforme a configuração mostrada na Figura 1.

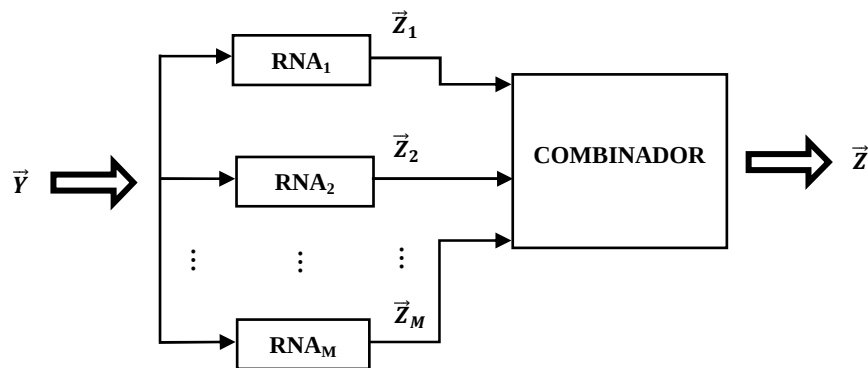


Figura 1: Estrutura da máquina de comitê de redes neurais artificiais.

O primeiro modelo de máquina de comitê usada para solucionar o problema inverso objeto deste trabalho consiste de uma máquina do tipo estática clássica, que efetua a média aritmética do conjunto das saídas das redes neurais especialistas que a compõem produzindo então, a saída global da máquina de comitê que será uma estimativa para o vetor dos parâmetros radiativos, solução do problema inverso em questão.

O segundo modelo de máquina de comitê usado é uma proposta inovadora deste trabalho de dissertação, que consiste de uma variação do modelo da máquina estática anterior adicionando-se ao combinador de saída desta uma rede neural do tipo MLP, que utiliza o vetor média aritmética das saídas das redes neurais especialistas como excitação de entrada, sendo então treinada através de um processo contínuo de mapeamento entrada-saída

supervisionado, produzindo após a fase de treinamento, estimativas para os parâmetros radiativos que são a solução do problema inverso objeto deste trabalho.

A motivação principal para a utilização de máquinas de comitê de redes neurais artificiais na solução do problema inverso em transferência radiativa deve-se ao fato de que o seu uso nas diversas áreas do conhecimento científico e tecnológico tem comprovado a ocorrência de todos aqueles benefícios decorrentes do uso de redes neurais artificiais, como por exemplo, não-linearidade, mapeamento de entrada-saída, adaptabilidade, robustez, plasticidade, etc; além de apresentarem, de forma geral, um resultado global na fase de generalização melhor do que aqueles obtidos por cada uma das RNA isoladamente, incorrendo apenas em um pequeno acréscimo computacional extra .

O objetivo específico deste trabalho está na comparação dos resultados obtidos com as máquinas de comitê com aqueles obtidos com as RNA isoladamente, comprovando que os primeiros são em geral melhores do que os segundos, além de mostrar que a máquina de comitê com um MLP integrado ao seu combinador, uma proposta inovadora deste trabalho de dissertação, apresenta resultados tão bons ou melhores do que os obtidos com o modelo clássico de máquina de comitê do tipo estática, que efetua apenas a média aritmética das estimativas obtidas pelas RNA especialistas em seu combinador.

As RNA especialistas usadas para compor ambos os modelos de máquina de comitê são do tipo MLP e apresentam quantidade de camadas ocultas variando de uma a duas.

A obtenção de estimativas para os parâmetros radiativos usando as redes neurais especialistas isoladamente e a obtenção de estimativas para os parâmetros radiativos usando máquinas de comitê são buscadas para uma posterior comparação dos resultados obtidos quanto à exatidão, precisão e do custo computacional envolvido da seguinte forma:

1. Comparação entre os resultados obtidos por meio de cada uma das RNA especialistas isoladamente com aqueles obtidos por meio de cada um dos dois modelos de máquina de comitê usados; e
2. Comparação dos resultados obtidos com ambos os modelos de máquina de comitê entre si, inclusive antes e depois da seleção das melhores RNA especialistas que as constituem.

Os padrões usados no treinamento das RNA especialistas e da rede neural do combinador da máquina de comitê, de modelo inovador aqui proposto, foram gerados através de repetidas soluções da versão linear da equação de Boltzmann, que é usada para modelar o problema direto de transferência radiativa (ÖZISIK, 1973).

1.3 Estrutura da dissertação

Esta dissertação divide-se em cinco capítulos e um apêndice. No capítulo 1 tem-se a introdução onde se discute a relevância do tema e suas várias aplicações, além de uma breve revisão bibliográfica, a motivação para este trabalho e dos objetivos geral e específico do mesmo.

No capítulo 2 são apresentadas as formulações e soluções do problema direto e do problema inverso em transferência radiativa.

No capítulo 3 é descrito um breve histórico e uma resumida fundamentação teórica sobre redes neurais artificiais e máquinas de comitê de redes neurais artificiais, além da modelagem e formulação matemática de redes neurais artificiais e máquinas de comitê de redes neurais artificiais para a solução do problema inverso em transferência radiativa.

No capítulo 4 são apresentados os melhores resultados e os maiores progressos obtidos, no que diz respeito à precisão e exatidão, além do desempenho das RNA especialistas e dos modelos de máquinas de comitês aplicados para a sua obtenção, por meio de tabelas e curvas de performance, geradas a partir do Neural Network Toolbox do software MATLAB[®]. A metodologia aplicada é descrita e uma discussão sobre os resultados, comparando os obtidos com as RNA isoladamente, com aqueles obtidos com as máquinas de comitês e com as máquinas de comitês entre si.

O capítulo 5 contém a conclusão do trabalho, onde é descrita e analisada a metodologia aplicada para a realização do mesmo e uma discussão quanto a sua eficácia e dos objetivos atingidos, além de sugerir a aplicação de outros métodos e linhas de trabalhos com o objetivo de resolver o problema inverso em transferência radiativa.

Finalmente, o apêndice A, onde são mostrados por meio de tabelas e curvas de performance, geradas a partir do Neural Network Toolbox do software MATLAB[®], todos os testes realizados, incluindo aqueles que não foram considerados entre os melhores resultados e, que portanto, não foram incluídos no capítulo 4.

2 PROBLEMAS DIRETO E INVERSO EM TRANSFERÊNCIA RADIATIVA

2.1 Problema direto em transferência radiativa

O problema direto em transferência radiativa em meios participantes, i.e. meios absorvedores, emissores e espalhadores isotrópicos, homogêneos ou heterogêneos, consiste no cálculo dos valores da intensidade de radiação, em todo domínio espacial e angular, considerando-se conhecidas a geometria (a espessura ótica do meio), as propriedades radiativas do meio (a espessura ótica, o albedo de espalhamento simples e as refletividades difusas), o termo fonte e as condições de contorno, i.e., as radiações externas incidentes nas superfícies laterais do meio (SILVA NETO; BECCENERI, 2009, p.25 - 27).

2.1.1 Formulação matemática do problema direto em transferência radiativa

Considerando-se o problema de transferência radiativa em um meio participante, (absorvedor, emissor e espalhador), unidimensional, homogêneo, (os coeficientes de absorção e espalhamento independem da posição), espalhador isotrópico, com simetria azimutal, cinza (as propriedades radiativas do meio independem do comprimento de onda da radiação), com superfícies refletoras difusas e sujeitas a incidência de radiação externa com intensidades f_1 e f_2 . A situação física aqui considerada está representada esquematicamente na Figura 2, onde L é a espessura física do meio e τ_0 é a espessura ótica do meio, caracterizando o comprimento de penetração da radiação no meio.

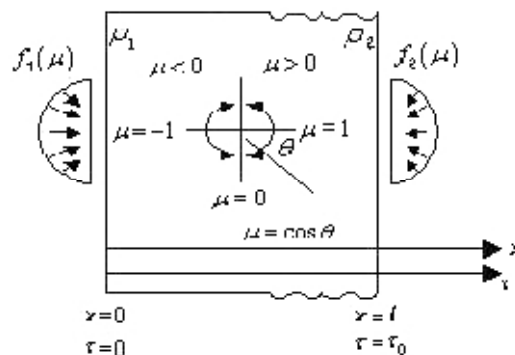


Figura 2: Meio participante, cinza, homogêneo, unidimensional com superfícies internas refletoras difusas e submetido à incidência de radiação externa de intensidades f_1 e f_2 (SILVA NETO; BECCENERI, 2009, p.18).

Problema direto em transferência radiativa

A formulação matemática do problema direto de transferência radiativa no meio representado na Figura 2, considerada para regime permanente, representando a interação entre a radiação e o meio, é dada pela versão linear da equação de Boltzmann, conhecida também como equação de transporte (ÖZISIK, 1973),

$$\mu \frac{\partial I(x, \mu)}{\partial x} + \beta I(x, \mu) = \frac{\sigma_s}{2} \int_{-1}^1 p(\mu, \mu') I(x, \mu') d\mu' + S(x, \mu);$$

$$0 < x < L, -1 \leq \mu \leq 1 \quad (1a)$$

$$I(0, \mu) = f_1(\mu) + 2\rho_1 \int_0^1 I(0, -\mu') \mu' d\mu'; \mu > 0 \quad (1b)$$

$$I(L, \mu) = f_2(\mu) + 2\rho_2 \int_0^1 I(L, \mu') \mu' d\mu'; \mu < 0 \quad (1c)$$

onde I representa a intensidade da radiação, x é a coordenada espacial, μ o cosseno do ângulo polar θ (ângulo entre o eixo- x positivo e a direção do feixe de radiação espalhada), σ_s o coeficiente de espalhamento, $S(x, \mu)$ uma fonte interna de radiação, e ρ_1 e ρ_2 as refletividades difusas nas superfícies internas do meio, respectivamente em $x = 0$ e $x = L$, $p(\mu, \mu')$ é a função de fase, e β é o coeficiente de extinção total.

$$\beta = k_a + \sigma_s \quad (2)$$

onde k_a é o coeficiente de absorção.

Aqui é utilizado o sistema de coordenadas esféricas, considerando simetria azimutal.

A função de fase $p(\mu, \mu')$ representa a densidade de probabilidade da radiação incidente na direção θ' , onde $\mu' = \cos\theta'$, ser espalhada na direção θ , onde $\mu = \cos\theta$, e aqui é escrita como uma expansão em polinômios de Legendre (ÖZISIK, 1973), de acordo com a equação (3)

$$p(\mu, \mu') \equiv \sum_{n=0}^M (2n+1) f_n P_n(\mu) P_n(\mu') \quad (3)$$

Problema direto em transferência radiativa

com $f_0 = 1$, sendo f_n os coeficientes da expansão. A função de fase representa a distribuição angular da radiação espalhada e independe do ângulo azimutal. Como neste trabalho considera-se o caso de espalhamento isotrópico, i.e., $M = 0$, da equação (3) resulta

$$p(\mu, \mu') = P_0(\mu)P_0(\mu') = 1 \quad (4)$$

Quando a geometria, as propriedades radiativas e as condições de contorno são conhecidas, o problema de transferência radiativa formulado com as equações (1) pode ser resolvido fornecendo os valores calculados da intensidade de radiação $I(x, \mu)$, em todo o domínio espacial, $0 \leq x \leq L$, e angular, $-1 \leq \mu \leq 1$. Este é o problema direto de transferência radiativa em um meio homogêneo unidimensional.

Em problemas referentes à transferência radiativa utiliza-se a variável ótica, τ , definida como

$$d\tau = \beta dx \Rightarrow \tau = \int_0^x \beta dx' \quad (5)$$

Considerando a integração em todo o domínio espacial obtém-se a espessura ótica do meio,

$$\tau_0 = \int_0^L \beta dx' \quad (6)$$

que para um meio homogêneo corresponde a

$$\tau_0 = \beta L \quad (7)$$

Dividindo a equação (1a) pelo coeficiente de extinção total β definido na equação (2) e empregando as definições da variável ótica τ e da espessura ótica τ_0 , nas equações (5) e (6), respectivamente, e considerando a emissão do termo fonte $S(x, y)$ como desprezível em relação a radiação externa incidente sobre o meio, obtém-se a equação do problema de transferência radiativa na sua forma adimensional (ÖZISIK, 1973),

$$\mu \frac{\partial I(\tau, \mu)}{\partial \tau} + I(\tau, \mu) = \frac{\omega}{2} \int_{-1}^1 I(\tau, \mu') d\mu';$$

$$0 < \tau < \tau_0, -1 \leq \mu \leq 1 \quad (8a)$$

$$I(0, \mu) = f_1(\mu) + 2\rho_1 \int_0^1 I(0, -\mu') \mu' d\mu'; \mu > 0 \quad (8b)$$

$$I(\tau_0, \mu) = f_2(\mu) + 2\rho_2 \int_0^1 I(\tau_0, \mu') \mu' d\mu'; \mu < 0 \quad (8c)$$

onde ω é o albedo de espalhamento simples, que corresponde à razão entre o coeficiente de espalhamento e o coeficiente de extinção total

$$\omega = \frac{\sigma_s}{\beta} = \frac{\sigma_s}{k_a + \sigma_s} \quad (9)$$

Os demais símbolos nas equações (8) já foram definidos anteriormente.

2.1.2 Solução numérica do problema direto em transferência radiativa

A abordagem utilizada aqui para a solução do problema direto em transferência radiativa é o emprego do Método das Ordenadas Discretas (S_M), de Chandrasekhar (1950) e o Método das Diferenças Finitas (SMITH, 1985).

Os efeitos devidos a uma possível diferença dos índices de refração do meio em estudo com relação à sua vizinhança não são levados em consideração.

Supondo conhecidas as propriedades materiais do meio e as condições de contorno impostas, o problema descrito pelas equações (1) pode ser resolvido, fornecendo os valores das intensidades de radiação calculadas em todos os pontos do domínio espacial e angular (PINHEIRO, 2004, p.22).

As intensidades de radiações calculadas que emergem do meio e as condições de contorno são representados esquematicamente na Figura 3.

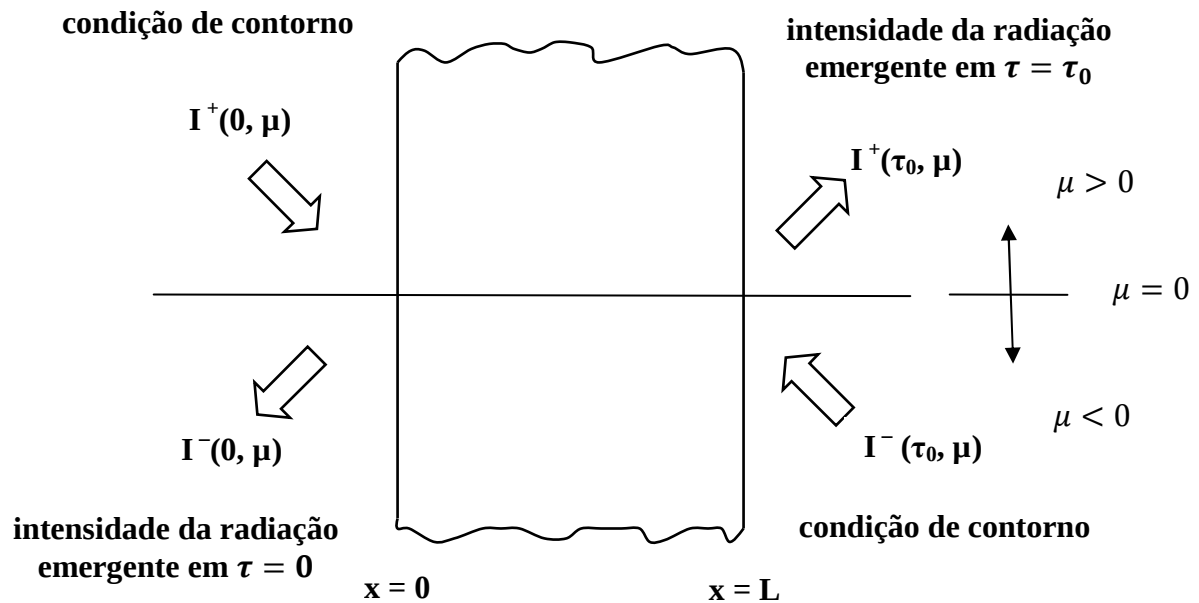


Figura 3: Representação esquemática das intensidades de radiação calculadas e das condições de contorno do meio (PINHEIRO, 2004, p.21).

A solução de uma equação íntegro-diferencial em uma região implica na obtenção dos valores para a variável dependente em cada ponto da região. Computacionalmente, apenas pode-se lidar com uma região contínua se for determinada uma fórmula analítica para a solução do problema. O computador pode, então, ser utilizado para calcular a solução em qualquer ponto desejado da região, com o uso dessa fórmula. Todavia, no caso de técnicas numéricas de solução, não é possível tratar a região estudada como contínua, já que o método numérico obtém a solução em pontos discretos (τ, μ) . Contudo, nada impede que sejam escolhidos alguns pontos dentro da região e apenas neles calcular a solução do problema (PINHEIRO, 2004).

Com base no parágrafo anterior, a Figura 4 mostra o domínio contínuo e a respectiva discretização que é usada na obtenção da solução do problema considerado, assim como as fronteiras do meio onde são impostas as condições de contorno.

Na Figura 5 é representada a discretização do domínio angular. Na Figura 6 é representada a discretização do domínio espacial para o meio participante homogêneo

unidimensional.

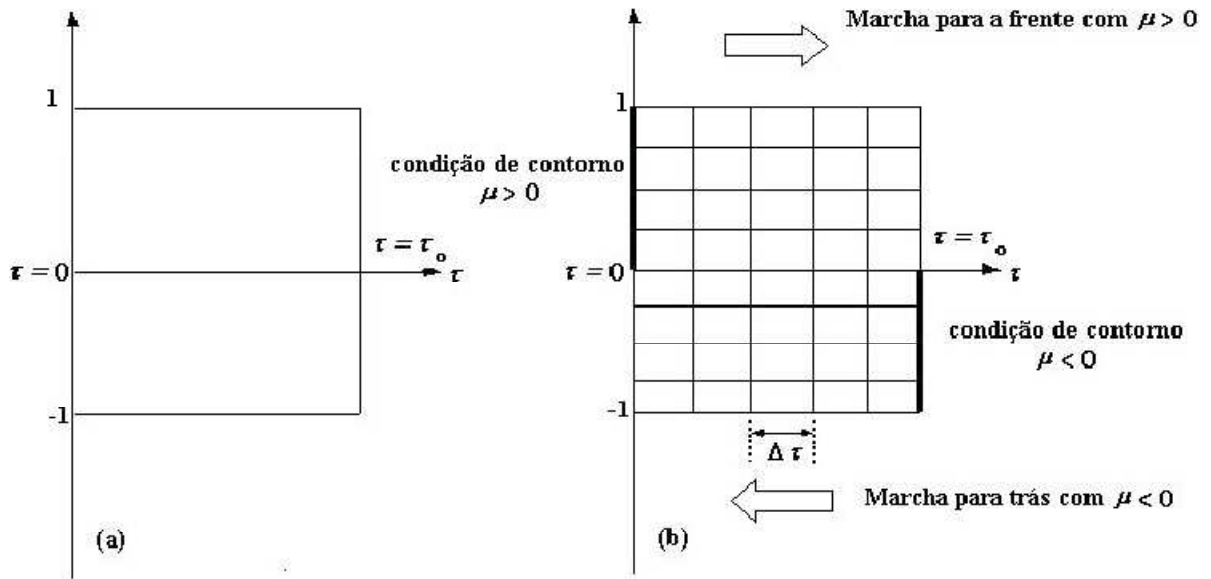


Figura 4: (a) domínio contínuo e (b) malha computacional (PINHEIRO, 2004 p.24).

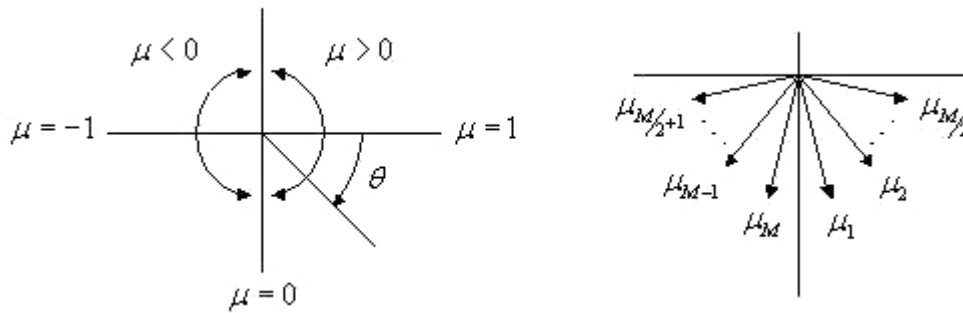


Figura 5: Discretização do domínio angular, $-1 \leq \mu \leq 1$ (SILVA NETO; BECCENERI, 2009, Cap. 3, p. 24).

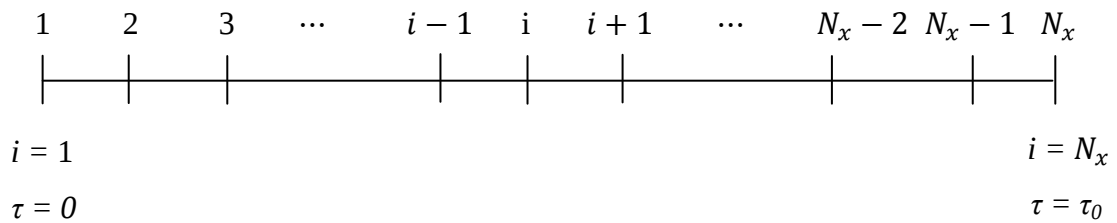


Figura 6: Discretização do domínio espacial para o meio homogêneo unidimensional, $0 \leq \tau \leq \tau_0$.

Problema direto em transferência radiativa

Considerando as discretizações dos domínios angular e espacial representadas respectivamente nas Figuras 5 e 6, obtém-se as discretizações para as intensidades de radiação para N_x pontos do domínio espacial e para M pontos do domínio angular

$$I_m^i = I(\tau_i, \mu_m) \quad (10)$$

com

$$\tau_i = (i - 1)\Delta\tau, i = 1, 2, \dots, N_x \text{ e } m = 1, 2, \dots, M \quad (11)$$

onde

$$\Delta\tau = \frac{\tau_0}{N_x - 1} \quad (12)$$

A solução da equação linearizada de Boltzmann, juntamente com as condições de contorno, representados pelo problema de valor de contorno dado pelas equações (8) é feita através de uma discretização angular empregando o método das ordenadas discretas com quadratura Gauss-Legendre (CHANDRASEKHAR, 1950), (ABRAMOWITZ; STEGUN, 1965) de ordem M , i.e. S_M , e uma aproximação por diferenças finitas para a discretização do domínio espacial.

O termo integral no lado direito da equação (8a) é substituído então, pela quadratura de Gauss-Legendre (CHANDRASEKHAR, 1950), (ABRAMOWITZ; STEGUN, 1965)

$$q^i = \frac{\omega}{2} \int_{-1}^1 I(\tau, \mu') d\mu' \cong \frac{\omega}{2} \sum_{n=1}^M a_n I_n^i \quad (13)$$

onde, a_n , $n = 1, 2, \dots, M$, são os pesos da quadratura correspondentes aos pontos de colocação, μ_n .

Os valores dos pesos e dos pontos de colocação da quadratura de Gauss-Legendre aqui utilizados, estão na Tabela 1, com $M = 20$.

Tabela 1: Pontos de colocação, μ_i , e pesos correspondentes, a_i , para a quadratura Gauss-Legendre com $M = 20$ (ABRAMOWITZ; STEGUN, 1965).

μ	a
$\mu_1 = -\mu_{20} = 0,0765265211$	$0,1527533871$
$\mu_2 = -\mu_{19} = 0,2277858511$	$0,1491729864$
$\mu_3 = -\mu_{18} = 0,3737060887$	$0,1420961093$
$\mu_4 = -\mu_{17} = 0,5108670019$	$0,1316886384$
$\mu_5 = -\mu_{16} = 0,6360536807$	$0,1181945319$
$\mu_6 = -\mu_{15} = 0,7463319064$	$0,1019301198$
$\mu_7 = -\mu_{14} = 0,8391169718$	$0,0832767415$
$\mu_8 = -\mu_{13} = 0,9122344282$	$0,0626720483$
$\mu_9 = -\mu_{12} = 0,9639719272$	$0,0406014298$
$\mu_{10} = -\mu_{11} = 0,9931285991$	$0,0176140071$

Na Tabela 1 observa-se que

$$\mu_{M-i+1} = -\mu_i, \text{ com } \mu_i > 0, i = 1, 2, \dots, \frac{M}{2}, \quad (14)$$

e

$$a_{M-i+1} = a_i, i = 1, 2, \dots, \frac{M}{2} \quad (15)$$

Considerando aproximações por diferenças finitas avançadas e atrasadas para o primeiro termo do lado esquerdo da equação (8a) obtém-se

$$\left. \frac{\partial I(\tau, \mu)}{\partial \tau} \right|_{(\tau_i, \mu_m)} \cong \frac{I_m^{i+1} - I_m^i}{\Delta \tau}, \quad (16a)$$

e

$$\left. \frac{\partial I(\tau, \mu)}{\partial \tau} \right|_{(\tau_i, \mu_m)} \cong \frac{I_m^i - I_m^{i-1}}{\Delta \tau} \quad (16b)$$

Das equações (8), (13) e (16) obtém-se então

$$I_m^{i+1} = \left(1 - \frac{\Delta\tau}{\mu_m}\right) I_m^i + \frac{\Delta\tau}{\mu_m} q^i \quad (17a)$$

$$I_m^{i-1} = \left(1 + \frac{\Delta\tau}{\mu_m}\right) I_m^i - \frac{\Delta\tau}{\mu_m} q^i \quad (17b)$$

A equação (8b) que representa a condição de contorno da superfície em $\tau = 0$, é aproximada por

$$I_m^1 = f_1(\mu_m) + 2\rho_1 \sum_{n=\frac{M}{2}+1}^M a_n |\mu_n| I_n^1, \text{ onde } m = 1, 2, \dots, \frac{M}{2} \quad (18a)$$

A equação (8c) que representa a condição de contorno da superfície em $\tau = \tau_0$ é aproximada por

$$I_m^{N_x} = f_2(\mu_m) + 2\rho_2 \sum_{n=1}^{M/2} a_n \mu_n I_n^{N_x}, \text{ onde } m = \frac{M}{2} + 1, \frac{M}{2} + 2, \dots, M \quad (18b)$$

A partir da equação (18a) é realizada uma marcha para frente, de acordo com a Figura 7, sendo calculados os valores de I_m^{i+1} com a equação (17a), para $m = 1, 2, \dots, \frac{M}{2}$, e $i = 1, 2, \dots, N_x - 1$.



Figura 7: Esquema representativo da marcha para frente.

A partir da equação (18b) é realizada então uma marcha para trás, de acordo com a Figura 8, sendo calculados os valores de I_m^{i-1} com a equação (17b), para $m = \frac{M}{2} + 1, \frac{M}{2} + 2, \dots, M$, e $i = N_x, N_x - 1, \dots, 2$.

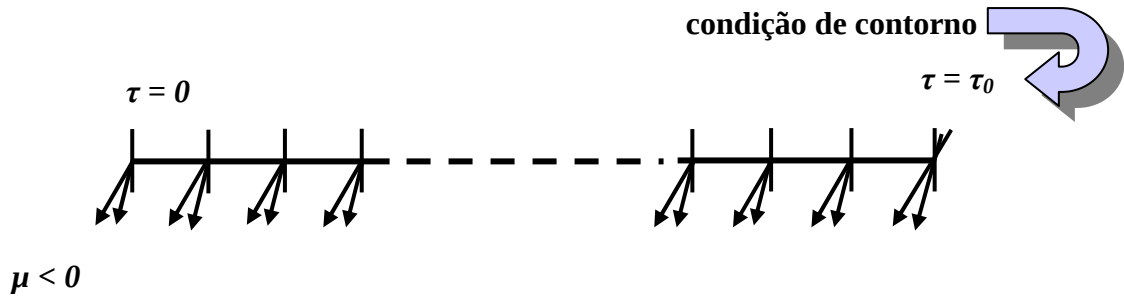


Figura 8: Esquema representativo da marcha para trás.

As marchas para frente e para trás obedecem a um processo iterativo, que segue até que um critério de parada seja satisfeito, como exemplo tem-se a convergência dos valores da intensidade de radiação, que pode ser verificada através do erro relativo dado pela equação (19)

$$\left| \frac{(I_m^i)^{p+1} - (I_m^i)^p}{(I_m^i)^p} \right| < \varepsilon \quad (19)$$

onde ε é uma tolerância requerida e p um número de iterações estabelecido a priori. Neste trabalho o critério escolhido foi o número de iterações p .

Ao final do processo iterativo das marchas para frente e para trás, as intensidades de radiação que emergem efetivamente do meio em $\tau = 0$ e $\tau = \tau_0$ são dadas por

$$I_{m_{ef}}^1 = (1 - \rho_1) I_m^1, \quad \text{para } m = \frac{M}{2} + 1, \frac{M}{2} + 2, \dots, M \quad (20a)$$

$$I_{m_{ef}}^{N_x} = (1 - \rho_2) I_m^{N_x}, \quad \text{para } m = 1, 2, \dots, \frac{M}{2} \quad (20b)$$

O albedo de espalhamento simples ω e as refletividades difusas ρ_j , com $j = 1$ ou 2 , são grandezas cujos valores estão no intervalo $[0,1]$. A espessura ótica τ_0 é uma grandeza positiva, não possuindo limite superior. Porém, para a solução dos problemas inversos em transferência radiativa usando métodos iterativos, é necessário definir uma região de busca para os valores

das incógnitas, sendo então escolhido um valor máximo $\bar{\tau}_0$ usualmente entre 2 e 10, exceto quando indicado de outra forma. Aqui escolhemos $\bar{\tau}_0 = 1,0$. Pode-se então resumir,

$$0 \leq \omega \leq 1 \quad (21)$$

$$0 \leq \rho_j \leq 1, \quad j = 1 \text{ ou } 2 \quad (22)$$

$$0 \leq \tau_0 \leq \bar{\tau}_0 \quad (23)$$

onde $\bar{\tau}_0$ é um limite superior artificial escolhido de acordo com o problema que se deseja resolver. Analogamente, este procedimento usado para a espessura ótica é adotado para as intensidades das fontes externas, f_1 e f_2 , bem como para os coeficientes de absorção e de espalhamento, k_a e σ_s (SILVA NETO ; BECCENERI, 2009, p.23).

2.2 Problema inverso em transferência radiativa

Antes de se definir o que é um problema inverso em transferência radiativa deve-se considerar o problema direto descrito na subseção 2.1, representado esquematicamente na Figura 2 para um meio homogêneo unidimensional.

A radiação externa incidente nas superfícies de contorno entra no meio e interage com o mesmo. Para este processo, que envolve os fenômenos de absorção e de espalhamento, dispõe-se de um modelo que é descrito matematicamente pela versão linear da equação de Boltzmann. Para o caso aqui considerado, i.e. um meio homogêneo unidimensional, a formulação matemática do problema é descrita pelas equações (1) ou (8).

Quando a geometria, as propriedades radiativas e as condições de contorno são conhecidas, o problema de transferência radiativa pode ser resolvido fornecendo os valores calculados da intensidade de radiação. Este é o problema direto (PD).

De acordo com a representação esquemática na Figura 9(a), a partir do conhecimento dos fenômenos envolvidos em um determinado processo, aqui sendo considerado o fenômeno de transferência radiativa, e do modelo matemático usado para descrevê-lo, incluindo a geometria do meio sob análise, e de uma causa conhecida, por exemplo as intensidades de radiação incidentes nas superfícies de contorno do meio participante, pode ser calculado o efeito, ou seja a intensidade de radiação calculada em qualquer ponto do domínio espacial e

angular do meio. Portanto, a incógnita do problema direto de transferência radiativa, é a intensidade de radiação calculada.

Considerando-se agora que a condição de contorno, i.e. a intensidade de uma ou das duas fontes externas de radiação, e/ou as propriedades radiativas do meio, i.e. os coeficientes de absorção e de espalhamento e as refletividades difusas, não são conhecidos. Porém, estão disponíveis valores experimentais da intensidade da radiação, $\vec{Y}$, medidos externamente ao meio conforme representado na Figura 10.

É possível então determinar as incógnitas do problema a partir destes dados experimentais, ou seja, é possível determinar a causa, i.e. as intensidades da radiação externa incidente, e/ou o que se desconhece do processo que está sendo investigado, i.e. às propriedades radiativas, a partir dos valores medidos experimentalmente da radiação que emerge do meio. Este é o Problema Inverso (PI), representado esquematicamente na Figura 9(b).

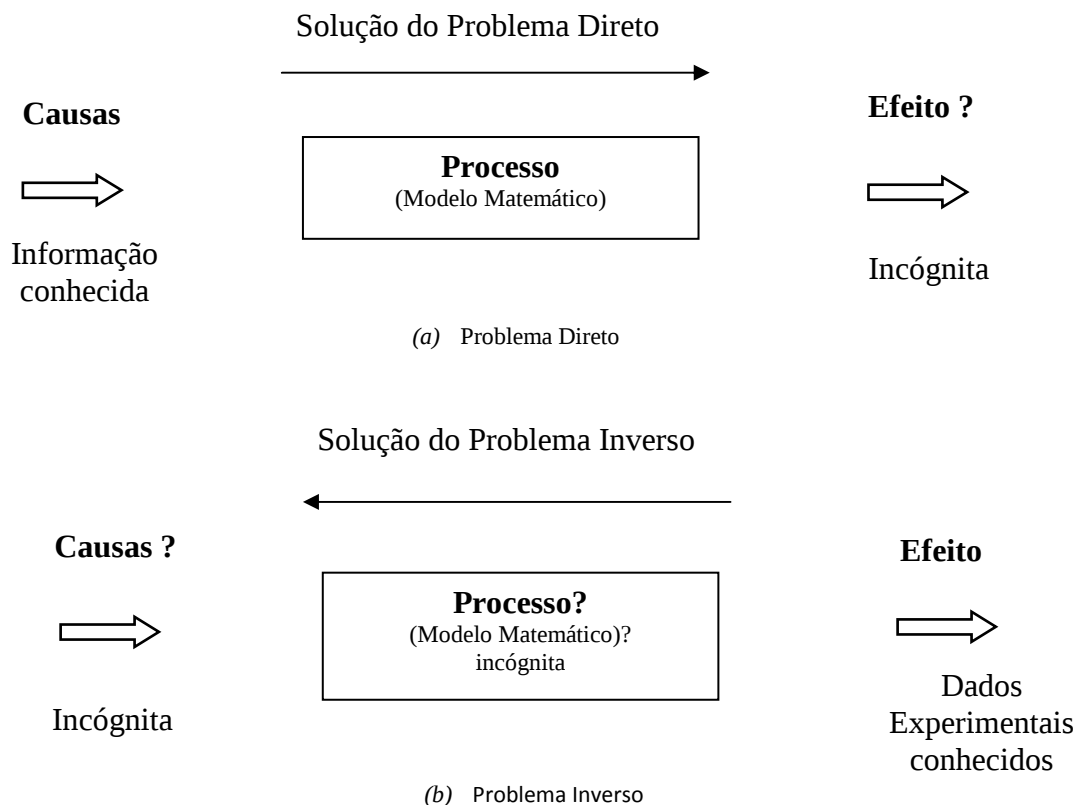


Figura 9: Representação esquemática do problema direto (a) e do problema inverso (b) [Fonte adaptada de: SILVA NETO, 2007].

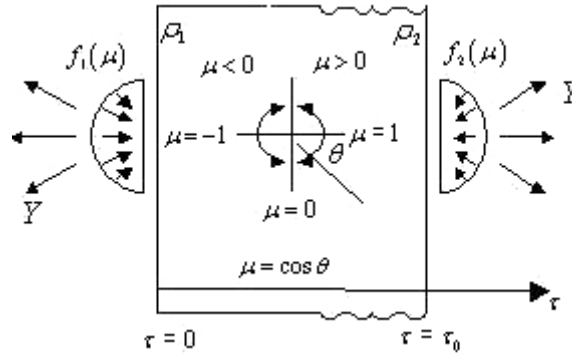
2.2.1 Formulação e solução do problema inverso em transferência radiativa

Figura 10: Representação esquemática de um meio participante homogêneo unidimensional e da radiação que emerge do mesmo com intensidade $\bar{Y}$ e que é medida experimentalmente (SILVA NETO; CAMPOS VELHO, 2009, p. 35).

Neste trabalho, como já mencionado anteriormente no capítulo 1, o objetivo é a solução do problema inverso em transferência radiativa para a determinação das propriedades radiativas do meio homogêneo unidimensional representado na Figura 2, ou seja, a obtenção do vetor das incógnitas $\vec{Z}$, dado pela equação (24)

$$\vec{Z} = \{\tau_0, \omega, \rho_1, \rho_2\}^T \quad (24)$$

cujos elementos são determinados utilizando os valores das intensidades experimentais $\bar{Y} = \{Y_l\}^T$, $l = 1, 2, \dots, N$. O superescrito T é indicativo da transposta.

Como dados experimentais reais não estão disponíveis, é construído um conjunto de dados computacionais simulados com

$$Y_l = I_{exp_l} = I_{calc_l}(\vec{Z}_{exato}) + \sigma r_l \quad (25)$$

onde I_{calc_l} representa os valores calculados da intensidade de radiação usando os valores exatos dos parâmetros, $\vec{Z}_{exato}$, que em uma aplicação real não estão disponíveis e cujas estimativas se quer determinar com a solução do problema inverso, σ representa o desvio

Problema inverso em transferência radiativa

padrão dos erros dos dados experimentais, e r_l representa um número pseudo-aleatório
gerado no intervalo $[-1,1]$.

3 FUNDAMENTOS SOBRE REDES NEURAIS E MÁQUINAS DE COMITÊ E SUA APLICAÇÃO À SOLUÇÃO DO PROBLEMA INVERSO EM TRANSFERÊNCIA RADIATIVA

Antes de ser apresentada a solução do problema inverso em transferência radiativa aplicando redes neurais e máquinas de comitê, será dissertado um breve histórico e fundamentação teórica sobre redes neurais e máquinas de comitê.

3.1 Histórico sobre redes neurais e máquinas de comitê

As redes neurais artificiais e as máquinas de comitê de redes neurais artificiais estão ambas inseridas na linha conexonista do estudo da Inteligência Artificial (IA), linha essa que visa à modelagem da inteligência humana imitando a estrutura neuro-física e a funcionalidade do cérebro, de seus neurônios, e de suas interligações (BIONDI NETO et. al., 2009).

O cérebro humano é uma estrutura biológica que processa informações de forma não linear e paralela. A proposta de se criar uma estrutura artificial imitando a estrutura neuro-física e a funcionalidade do cérebro humano, modelando assim, a inteligência humana, foi formalizada quando McCulloch e Pitts (1943) propuseram um primeiro modelo matemático simples para um neurônio, que modelava o funcionamento de operações da lógica binária, através de um sistema de inequações lineares de primeira ordem para representar o conhecimento relacionado. A especificação automática dos parâmetros de um neurônio ou de um conjunto de neurônios (a aprendizagem) começou a ser estudada por Hebb (1949) que propôs uma formulação explícita para a primeira regra de aprendizagem para atualização sináptica entre células neuronais com funcionamento descrito pelo modelo de McCulloch e Pitts. Rosenblatt (1958), propôs o primeiro modelo de rede neural artificial, que consistia de um conjunto de neurônios dispostos em uma única camada que foi denominado de Perceptron. Minsky e Papert (1969), analisaram o modelo proposto por Rosenblatt, demonstrando matematicamente que o Perceptron solucionava apenas problemas que apresentassem superfícies linearmente separáveis entre as classes, impedindo desta forma que a rede mapeasse problemas simples não-lineares, como por exemplo, a solução do problema do OR exclusivo (HAYKIN, 2001).

Este fato provocou uma falta de motivação e interesse pela pesquisa com redes neurais artificiais por um longo período de tempo, chegando mesmo à paralisação. Foi quando Werbos (1974), propôs um novo algoritmo que possibilitava o treinamento de redes neurais

com múltiplas camadas, introduzindo o conceito da retro-propagação do erro, superando dessa forma os obstáculos levantados por Minsky e Papert . Mesmo, assim, à época, não foi dada a atenção e relevância devidas à descoberta de Werbos tendo tido pouca repercussão e divulgação na comunidade científica. Foi então, que pesquisas realizadas por Rumelhart, Hinton e William (1986), resultaram na divulgação das regras de treinamento inovadoras idealizadas por Werbos, reavivando o interesse e dando novo impulso às pesquisas com redes neurais artificiais. A partir daí, o algoritmo da retro-propagação do erro de Werbos tem sido amplamente divulgado e utilizado (BIONDI NETO, et. al., 2009).

Podem-se citar, no contexto, trabalhos importantes de outros pesquisadores na linha connexionista, que influíram no desenvolvimento das redes neurais artificiais e máquinas de comitê de redes neurais, dentre eles os seguintes: Widrow e Hoff (1960) que propuseram uma poderosa regra de aprendizado, produzindo o algoritmo do mínimo quadrado médio, baseado na regra delta que é fundamentada no método do gradiente descendente; Cowan (1989) criou um neurônio artificial de disparo suave, baseado na função logística; Grossberg (1988) criou o modelo aditivo do neurônio envolvendo equações diferenciais não lineares; Malsburg (1973) demonstrou o mapeamento auto-organizável; Little e Shaw (1975) propuseram um modelo probabilístico de um neurônio; Willshaw e Malsburg (1976) publicaram o primeiro trabalho sobre mapas auto-organizáveis; Kohonen (1974) iniciou pesquisa envolvendo memórias associativas; Carpenter e Grossberg (1988) desenvolveram a teoria da ressonância adaptativa (Adaptative Resonance Theory, ART); Kohonen desenvolveu o algoritmo não supervisionado com aprendizagem competitiva, que é um modelo de mapa auto-organizável (1982); Hopfield (1982) introduziu a arquitetura de redes recorrentes, trabalhando também na área de memórias, associativas; Broomhead e Lowe (1988) descreveram os procedimentos para utilização de redes com funções de base radial (RBF); Vapnik (1982) desenvolveu a Máquina de Vetor de Suporte. Nilsson (1965) introduziu o conceito de máquina de comitê de redes neurais ao idealizar uma estrutura formada por uma camada de perceptrons elementares seguida de um perceptron de votação na segunda camada.

O recente e vertiginoso desenvolvimento de novos microprocessadores de elevado poder de processamento e baixo custo, aliando a alta capacidade de armazenamento com alta velocidade, somando a isso o desenvolvimento de novos algoritmos específicos e muito eficientes, inicialmente baseados nas pesquisas de Werbos (redes de múltiplas camadas), possibilitou o aprendizado de problemas complexos de elevado nível de não linearidade, des-

de que fossem rigorosamente mapeados. Dessa forma, possibilitou um novo impulso às pesquisas com RNA e máquinas de comitê de redes neurais artificiais, com forte dependência computacional (hardware e software) (WASSERMAN, 1989).

3.2 Fundamentação teórica sobre redes neurais artificiais

A finalidade do uso de redes neurais artificiais é imitar a estrutura e o processamento paralelo do cérebro humano. Portanto, para se compreender a lógica do funcionamento das RNA é de fundamental importância o conhecimento de alguns conceitos básicos do funcionamento do cérebro humano e de seus componentes essenciais, os neurônios. A formação das conexões entre os neurônios e o entendimento de como ocorre a aquisição do conhecimento (aprendizagem) pelos mesmos, e assim poder formalizá-lo através de um modelo matemático é de fundamental importância para o entendimento e desenvolvimento da neurocomputação (JUNIOR; COSTA, 2007, p. 9 -14).

3.2.1 O neurônio biológico

O neurônio (do grego: célula nervosa) é a célula responsável pelo processamento do sinal cerebral, sendo considerada a célula mais complexa do organismo humano no que se refere a sua estrutura e funcionalidade. Além das funções inerentes a qualquer célula do organismo, ela possui a capacidade de processar e transmitir sinais contendo informações (BIONDI et. al., 2009, p 84).

O cérebro humano possui cerca de cem bilhões de neurônios. Cada neurônio forma entre 10.000 a 100.000 sinapses com outros neurônios, totalizando aproximadamente 10^{16} sinapses de capacidade máxima de memória.

Os neurônios operam sempre em grupos formando redes neurais ou redes neuronais biológicas. Através de reações elétricas e bioquímicas permitem que o cérebro execute, de modo paralelo, uma série de funções particulares ou tarefas de interesse, inclusive de aprendizado.

A Figura 11 representa de forma bastante simplificada a conexão entre dois neurônios de uma rede biológica.

Um neurônio biológico qualquer apresenta a seguinte estrutura: um corpo celular, também chamado de soma, que é responsável pela captação e combinação das informações

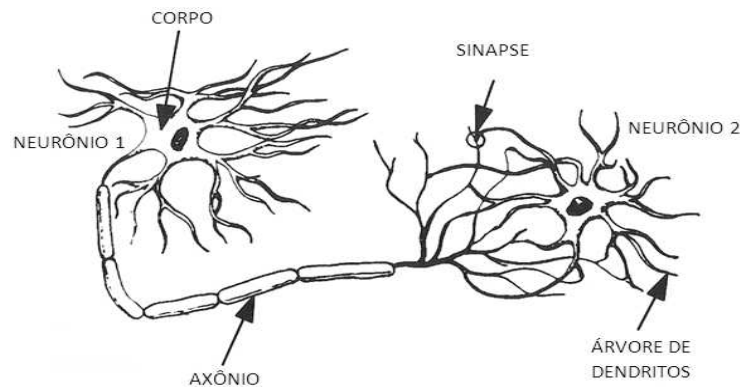


Figura 11: Rede Neural Biológica (BIONDI NETO, et. al., 2009).

oriundas de outros neurônios; uma árvore de dendritos, cuja função é receber os estímulos provenientes de outros neurônios; e um axônio, que é um tubo filamentar limitado pela membrana da célula, cuja finalidade é transmitir pulsos elétricos regenerativos em todo o percurso do sistema nervoso, onde na parte terminal forma um telodendro propiciando um grande número de pontos de contato com outras células nervosas (KOVACS; JOHNSON; SAX, 1979, p. 281-293).

Com a finalidade de aumentar a velocidade de condução dos sinais elétricos dos impulsos nervosos no interior do axônio e por questões de isolamento, uma camada chamada bainha de mielina recobre o axônio (LENT, 2001). A sinapse não exerce apenas a função estrutural de transferência de sinal entre dois neurônios. Nela os sinais podem ser amplificados ou atenuados de forma parcial ou total, denotando a sinapse uma característica de um local de decisão no sistema nervoso (KOVACS, 1991).

A árvore de dendritos, funciona de forma geral como um captador de estímulos oriundos de neurônios adjacentes, que então são conduzidos para o corpo celular do neurônio, onde são processados gerando um pulso na base do axônio. Esse sinal é transmitido pelo axônio ao neurônio alvo através da árvore de dendritos desse, tendo como ponto de conexão a sinapse do neurônio emissor (BIONDI NETO et. al., 2009).

3.2.2 O neurônio artificial

Um neurônio artificial ou elemento processador é uma unidade lógico-matemática que modela um neurônio biológico. Similarmente ao neurônio natural, recebe um ou mais sinais

de entrada e responde com um único sinal de saída, que pode ser transmitido para um ou mais neurônios adjacentes posteriores através de ligações sinápticas artificiais. As intensidades dessas ligações sinápticas são representadas pelos pesos sinápticos e bias que ponderam o sinal de entrada amplificando ou atenuando o seu valor.

A Figura 12 representa os elementos que compõem um neurônio artificial com uma única entrada. À esquerda da Figura 12, temos um neurônio de uma única entrada sem bias. O sinal de entrada p , um escalar, é transmitido através de uma conexão que multiplica sua excitação de entrada p pelo peso escalar w para formar o produto ponderado wp , também um escalar. O produto ponderado $n = wp$, chamado de função de propagação ou transfer function net input

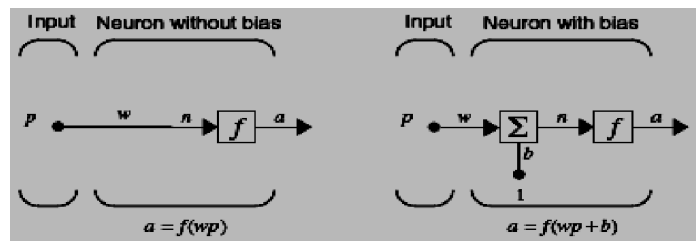


Figura 12: Neurônio artificial com uma única entrada [Neural Network Toolbox, MATLAB®].

(entrada líquida da função de ativação), é então, o único sinal de entrada ou argumento da função de ativação (ou função de transferência) f , que responde com a saída $a = f(n) = f(wp)$. À direita da Figura 12, tem-se um neurônio com bias, um escalar que funciona como um peso que independe do sinal de entrada p , e tem valor de entrada fixado em 1. O bias b é um limiar fixo para a função de ativação f que dispara quando a soma ponderada $wp + b$ é maior ou igual a zero, produzindo a saída $a = f(wp + b)$. O bias possibilita que um neurônio apresente saída não nula ainda que todas as suas entradas sejam nulas. O peso w e o bias b são os parâmetros escalares ajustáveis do neurônio, cujos valores são atualizados durante o processo de treinamento ou aprendizado. O peso w e o bias b constituem a memória do neurônio, após a fase de treinamento ou aprendizado.

A função de ativação exerce uma função similar ao do axônio do neurônio biológico, que dispara o sinal de saída do neurônio emissor para o neurônio alvo, ao se atingir um determinado limiar. Na sequência são descritas algumas das principais funções de ativação comumente empregadas no processamento neural artificial.

3.2.3 Tipos de funções de ativação

Há várias possibilidades para a função de ativação f (HAYKIN, 2001, p.37, 195). Aqui são mostradas as principais. São as seguintes:

3.2.3.1 Função degrau (*threshold function*)

A função degrau representada na Figura 13, limita a saída do neurônio a 0, se a função de propagação n for menor do que 0, ou limita a saída do neurônio a 1, se o valor da função de propagação n for maior ou igual a 0. É usada com Perceptrons para classificação de decisões.

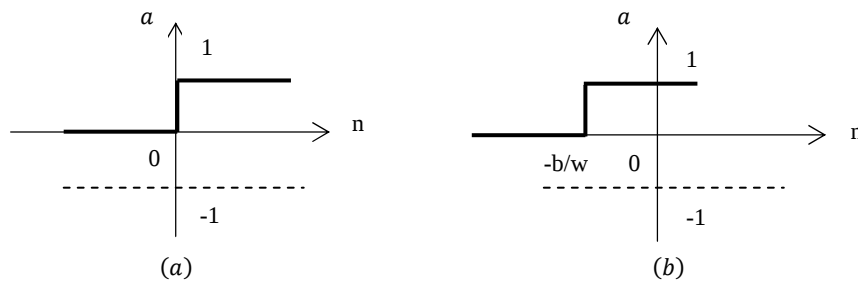


Figura 13: Função de ativação tipo degrau, (a) sem bias e (b) com bias.

3.2.3.2 Função linear

A função linear representada na Figura 14 é usada com aproximadores lineares em filtros.

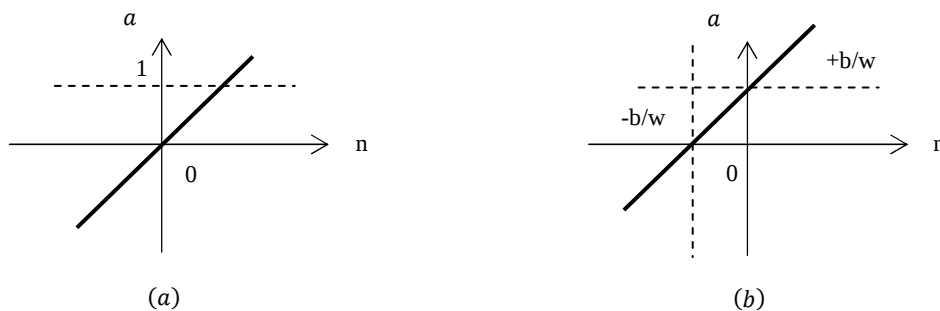


Figura 14: Função de ativação tipo linear, (a) sem bias e (b) com bias.

3.2.3.3 Função log-sigmóide

A função log-sigmóide, representada na Figura 15, é uma função estritamente crescente que exibe um balanceamento adequado entre comportamento linear e não-linear (HAYKIN, 2001, p. 40). Ela é comumente usada em redes neurais com retro-propagação do erro, principalmente devido à sua diferenciabilidade. A entrada da função sigmóide pode assumir qualquer valor entre menos infinito e mais infinito, que então, restringe a sua saída aos valores entre 0 e 1.

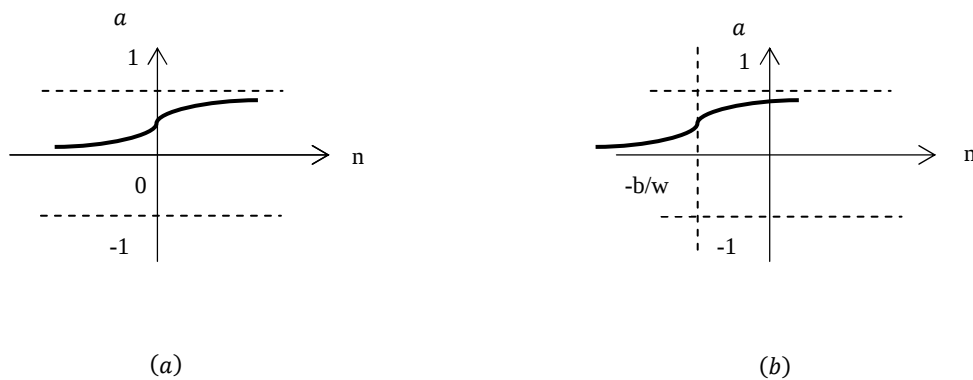


Figura 15: Função de ativação tipo log-sigmóide, (a) sem bias e (b) com bias.

3.2.3.4 Função tan-sigmóide

A função de ativação tan-sigmóide, representada na Figura 16, é derivada da função tangente hiperbólica e tem a mesma aplicação da função log-sigmóide, derivada da função logaritmo. Ambas são estritamente crescentes no intervalo de menos infinito a mais infinito. A resposta da função restringe-se ao intervalo $[-1, 1]$.

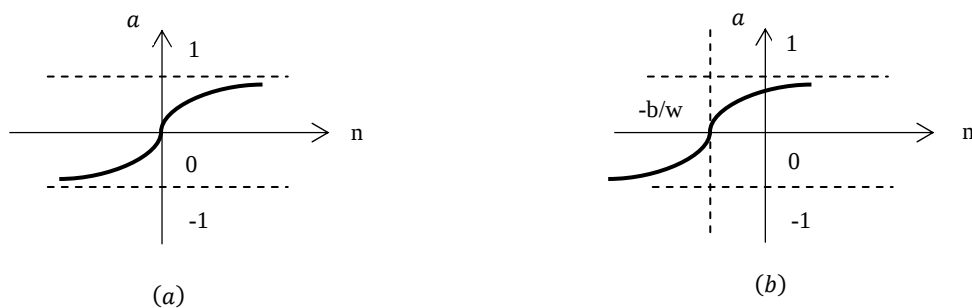


Figura 16: Função de ativação tan-sigmóide, (a) sem bias e (b) com bias.

3.2.4 O neurônio artificial de entrada vetorial

A Figura 17 mostra o modelo simplificado do neurônio artificial proposto por McCulloch e Pitts (1943).

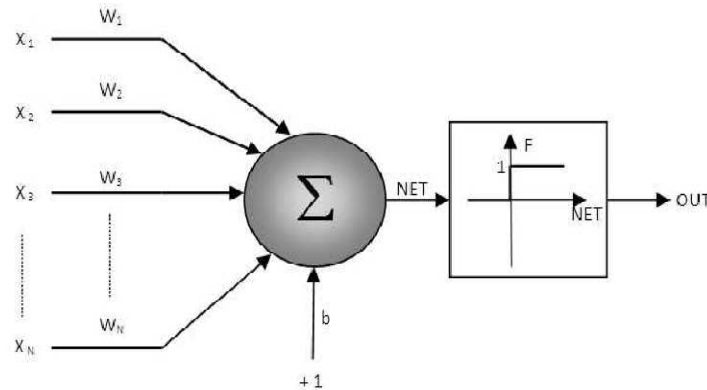


Figura 17: Modelo do Neurônio Artificial (BIONDI NETO, et. al., 2009).

McCulloch e Pitts (1943), propuseram um modelo matemático para um neurônio biológico onde o corpo celular é representado por um somatório (Σ), o axônio por uma função de ativação (F), geralmente com limiar que pode ser uma das funções descritas no subitem 3.2.3, entre outras possibilidades, e que restringe a amplitude do sinal de saída OUT do neurônio. As sinapses são representadas por pesos (W) emulando as sinapses que ponderam as entradas do modelo, podendo ser positivos ou negativos em função das características das sinapses serem excitatórias ou inibitórias (HAYKIN, 2001).

A função NET representa a soma ponderada das entradas x_i , $i = 1, \dots, N$, pelos respectivos pesos sinápticos ($NET = \sum W.x + b$), onde b é o bias. Caso o neurônio apresente um limiar fixo (representado pelo bias b), esse dispara quando $NET \geq 0$. O valor de b pode ser positivo ou negativo, aumentando ou diminuindo a entrada líquida da função de ativação F e pode ser representado diretamente como mais um peso associado a uma entrada fixada em “1”. Assim, para $NET \geq 0$, $OUT = F(NET) = 1$, caso contrário $OUT = 0$.

3.2.5 A rede neural artificial

Redes neurais artificiais são estruturas maciçamente paralelas e distribuídas, formadas por neurônios artificiais densamente interconectados através de pesos sinápticos, que imitam

o processamento do cérebro humano, com capacidade para adquirir, armazenar e utilizar o conhecimento experimental, podendo ser implementadas em software e hardware.

Devido à sua similaridade com a estrutura do cérebro humano, as redes neurais artificiais apresentam características semelhantes ao comportamento biológico do mesmo, que são:

- Procura paralela e endereçamento pelo conteúdo;
- Capacidade de aprendizado;
- Capacidade de adaptação;
- Capacidade de associação;
- Capacidade de generalização;
- Capacidade de abstração; e
- Robustez e degradação gradual.

A Figura 18 mostra a estrutura de uma rede neural artificial com múltiplas camadas, onde se distinguem claramente as diferenças entre as camadas de entrada, formada por nós de transferência que têm unicamente a função de distribuir os sinais de entrada às demais camadas intermediárias ou ocultas e daí para a camada de saída da RNA, que são todas compostas integralmente por neurônios artificiais e, portanto, com funcionalidades neurais.

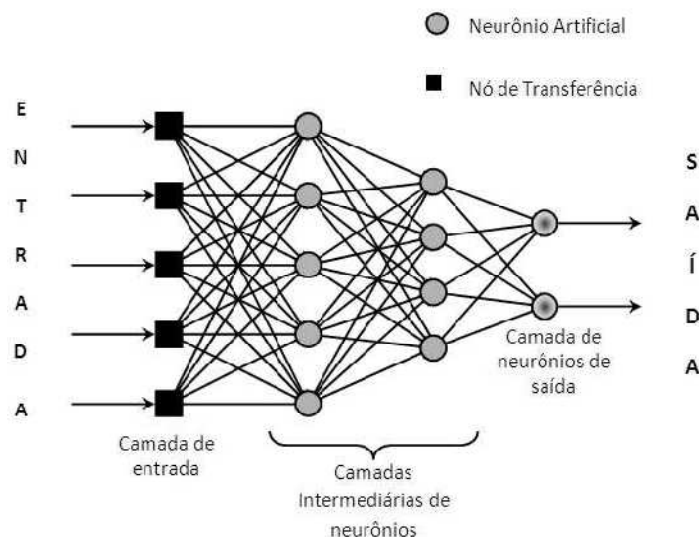


Figura 18: Estrutura de uma rede neural artificial (BIONDI NETO, et. al., 2009).

O processamento com RNA divide-se em duas fases distintas: fase de treinamento ou de aprendizado e fase de execução.

A fase de treinamento é realizada, com a utilização de algoritmos específicos, e é quando as RNA “aprendem” um certo conhecimento pela apresentação sucessiva de exemplos chamados de “padrões de treinamento”. Durante esta fase informações relevantes são extraídas e distribuídas por toda rede neural a partir das informações contidas nos padrões de treinamento que ficam mapeados sobre toda a rede.

A fase de treinamento é um processo iterativo e contínuo durante o qual os pesos e bias das conexões sinápticas entre os neurônios são ajustados e atualizados de forma que ao final desta etapa o conhecimento adquirido fica armazenado na forma de uma matriz de pesos W e bias b . Como na fase de treinamento são utilizados algoritmos que envolvem a minimização do erro, e que em certas circunstâncias podem ter convergência lenta, paralisada ou até mesmo oscilar indefinidamente, essa fase geralmente é a mais demorada.

A fase de execução consiste na utilização do conhecimento adquirido pela rede neural na sua fase de treinamento ou aprendizado, representado pelas informações contidas na matriz de pesos W e bias b , que constituem a memória da rede, para classificar padrões de entradas com um certo grau de ruído, e que não participaram do treinamento da rede. Se a rede neural conseguir classificar corretamente essas entradas de acordo com os padrões estabelecidos na fase de treinamento, podemos então afirmar que a rede aprendeu ou que está treinada (BIONDI NETO et. al., 2009).

Em todo o processo de construção de uma rede neural para resolver um determinado problema de classificação de padrões, possivelmente as partes mais difíceis e penosas de serem trabalhadas sejam a da construção de sua arquitetura, a da definição dos parâmetros a serem usados e o método a ser aplicado. Isto se deve principalmente ao aspecto estocástico do processo. Portanto, não existe uma forma eficaz e determinística para se construir uma rede neural artificial, isto vai depender das características do problema a ser solucionado.

Existem diversas classificações e técnicas de treinamento de redes neurais bem desenvolvidas na literatura especializada (HAYKIN, 2001). No subitem 3.2.6 serão descritos os dois principais paradigmas de treinamento de RNA.

3.2.6 Paradigmas de treinamento de redes neurais artificiais

Os dois principais paradigmas do treinamento de redes neurais artificiais são o aprendizado supervisionado e o aprendizado não supervisionado.

O aprendizado supervisionado, que será o aplicado a este trabalho, caracteriza-se pela existência de um “supervisor”. No início do treinamento a rede neural não tem nenhum conhecimento do problema a ser solucionado. Pares de treinamento que constituem os “padrões de treinamento” estimulam a rede fazendo com que ela tome conhecimento do ambiente a ser representado. A cada padrão de entrada apresentado à rede, a saída resultante é confrontada com um alvo ou saída desejada, caracterizando assim um processo de supervisão. Essa confrontação resulta na geração e medição de um erro, que é a diferença entre o valor da saída obtida e o valor da saída desejada. Se o valor do erro estiver acima da tolerância requerida, um sinal de erro é retro-propagado da camada de saída para as camadas intermediárias anteriores até a camada de entrada da rede neural, ajustando todos os pesos e bias. Esse processo é iterativo e contínuo e prossegue até que se atinja um valor mínimo requerido para o erro. Quando isso ocorre a rede é considerada como treinada e apta para a fase de execução.

Vários tipos de RNA utilizam o treinamento supervisionado, dentre eles o Perceptron de Múltiplas Camadas (MLP), que será estudado no subitem 3.2.8, e que é o tipo de rede neural que tem sido utilizada na solução do problema inverso em transferência radiativa inclusive neste trabalho. Como exemplo tem-se o trabalho de Soeiro et. al. (2004a).

O outro paradigma é o do aprendizado não supervisionado, que como o próprio nome indica, caracteriza-se pela inexistência de um “supervisor”. Hebb (1949), em sua declaração, dá uma idéia bem clara do que é o aprendizado não supervisionado. A declaração afirma o seguinte:

Quando um axônio de uma célula A está próximo o suficiente para excitar uma célula B e repetida ou persistentemente participa desta ativação, um processo de crescimento ou de mudança metabólica ocorre em uma ou mais células, de tal forma que a eficiência de A em ativar B seja aumentada.

O esquema da Figura 19 ilustra de forma bem clara a declaração de Hebb (1949).

Sejam dois neurônios (i) e (j), sendo (i) o neurônio fonte, com saída OUT_i e (j) o neurônio de destino com saída OUT_j . Seja W_{ij} o peso sináptico entre o neurônio fonte e o neurônio de destino.

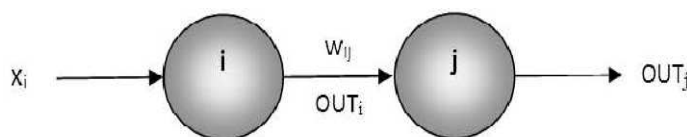


Figura 19: Conexão entre dois neurônios artificiais (BIONDI NETO, et. al., 2009).

De acordo com a declaração de Hebb (1949), se os neurônios (i) e (j) na Figura 19 trabalharem em sincronia, a conexão entre os dois, representada pelo peso sináptico W_{ij} , se fortalecerá de forma proporcional ao produto da saída do neurônio (i), OUT_i , pela saída do neurônio (j), OUT_j . Por outro lado se os neurônios (i) e (j) trabalharem de forma assíncrona a conexão W_{ij} enfraquecerá podendo mesmo deixar de existir. Assim, considerando a declaração de Hebb, pode-se estabelecer uma equação que modele a variação da intensidade da conexão sináptica entre os neurônios (i) e (j), como $\Delta W_{ij}(n) = \eta \cdot OUT_i \cdot OUT_j$, que, por conseguinte, representa a variação do peso sináptico entre ambos os neurônios. O parâmetro η representa uma constante positiva denominada taxa de aprendizado e (n) representa a iteração atual. Com base na regra de Hebb (1949), pode-se escrever a equação (26), que permite a atualização dos pesos sinápticos a cada iteração. (HAYKIN, 2001; ROGERS; McCLELLAND, 2004).

$$W_{ij}(n + 1) = W_{ij}(n) + \Delta W_{ij}(n) = W_{ij}(n) + \eta \cdot OUT_i \cdot OUT_j \quad (26)$$

Acrescentando ao que foi descrito, a rede, no aprendizado não supervisionado, desenvolve suas próprias representações de associação. A cada novo padrão apresentado, a rede verifica a sua similaridade com os padrões apresentados anteriormente, de acordo com um critério estabelecido. Caso o padrão apresente certa semelhança com alguns dos padrões anteriores, este é então, representado de tal forma que a rede o insere na mesma classe destes padrões de natureza semelhante. Caso contrário, o padrão é representado dando origem a uma nova classe de informação, que ainda não existe.

Na fase de execução ou generalização, as matrizes de pesos e bias geradas na fase de treinamento estão disponíveis, por conseguinte, não há mais nenhum processo iterativo a ser realizado nesta fase. Com o objetivo de recuperar a informação, na fase de execução a RNA calcula uma saída para um determinado padrão teste de entrada que não tenha participado do treinamento. Neste processo, ocorre apenas a multiplicação da matriz de pesos W e bias b pelo padrão de teste x_t , que não fez parte do treinamento. O resultado é então aplicado diretamente na função de transferência F , obtendo-se a resposta $OUT = F(W \cdot x_t + b)$, sendo, portanto, um processo bastante rápido.

3.2.7 Perceptron

O Perceptron é a RNA de estrutura mais simples que existe usada com sucesso na classificação de padrões. Proposto por Rosenblatt (1958), é o primeiro modelo bem sucedido de aprendizado supervisionado. O neurônio artificial usado no Perceptron, segue o modelo proposto por McCulloch e Pitts (1943), e sua topologia é composta por uma camada de entrada e uma camada de saída contendo um ou mais neurônios artificiais, sem nenhuma camada intermediária ou escondida (JUNIOR; COSTA, 2007, p. 27-29).

A Figura 20 representa a topologia de um Perceptron com um neurônio j de “ n ” entradas X_i ($i = 1, \dots, n$), ponderadas pelos pesos W_{ij} , usando uma função de ativação do tipo degrau simétrico (bipolar) com maior poder de discriminação e com saída OUT_j . A “supervisão” é feita por T_j (resposta desejada) e o sinal de erro correspondente vale $T_j - OUT_j$. O ajuste dos pesos é feito segundo a fórmula $\Delta W_{ij} = \eta \cdot (T_j - OUT_j) \cdot X_i$

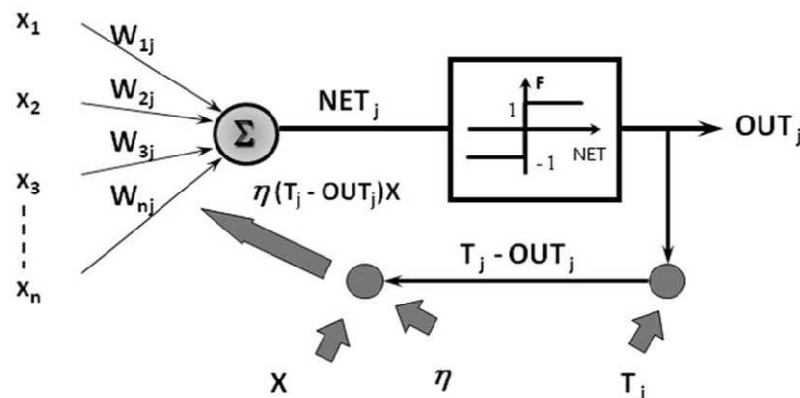


Figura 20: Modelo do Perceptron (BIONDI NETO et. al., 2009).

O algoritmo de treinamento do Perceptron, conhecido como regra delta, consiste dos seguintes passos:

1. Inicializar os pesos sinápticos W_{ij} com valores aleatórios;
2. Aplicar um padrão de entrada X_i e calcular a saída OUT_j ;
3. Se a saída $OUT_j = T_j$ e existir um novo padrão a ser apresentado, então voltar ao passo 2, senão ir para o passo 7;
4. Se a saída $OUT_j \neq T_j$ calcular a variação dos pesos $\Delta W_{ij}(n) = \eta \cdot (T_j - OUT_j) \cdot X_i$;

5. Atualizar os pesos sinápticos $W_{ij}(n+1) = W_{ij}(n) + \Delta W_{ij}(n)$;
6. Se existir um novo padrão a ser apresentado voltar para o passo 2;
7. FIM

O elemento que usa a regra delta é chamado de ADALINE (Adaptative Linear Neuron ou Adaptative Linear Element) (JUNIOR; COSTA, 2007, p.29).

A taxa de aprendizado η mede o passo a ser adotado pela rede em busca da convergência, i.e., em busca do mínimo global. Se a taxa η for pequena a convergência se torna muito lenta, podendo até paralisar o treinamento devido ao confinamento num mínimo local, ao contrário se a taxa η for muito elevada, o treinamento pode oscilar (BIONDI NETO, et. al., 2009).

A proposta da criação do Perceptron por Rosenblatt (1958) é a classificação de padrões. Como mencionando anteriormente no item 3.1, Minsky e Papert (1969) demonstraram que o Perceptron não podia ser usado como solução para qualquer tipo de problema de classificação, como se pensava inicialmente. Ficou provado (HAYKIN, 2001) que o Perceptron não resolvia problemas que apresentassem superfícies linearmente não separáveis entre as classes, como o caso do OR exclusivo. A solução apresentada para a solução do problema foi à construção de redes neurais semelhantes ao Perceptron, porém com múltiplas camadas escondidas, denominadas Perceptron de Múltiplas Camadas. No entanto, naquela época, ainda não existia um algoritmo de aprendizado capaz de treinar redes neurais com mais de uma camada de neurônios artificiais, causando grande pessimismo entre os pesquisadores com o futuro da área de RNA (HAYKIN, 2001). Com a criação do algoritmo da retro-propagação do erro por Werbos (1974) o problema ficou completamente resolvido.

3.2.8 Perceptron de múltiplas camadas

O Perceptron de Múltiplas Camadas (MLP) é uma generalização da rede Perceptron proposto por Rosenblatt (1958), onde cada neurônio da camada antecedente se conecta com todos os neurônios artificiais da camada conseqüente, e assim como esta, é treinada de forma supervisionada, usando o algoritmo da retropropagação do erro fundamentado na regra delta generalizada, que ajusta os pesos sinápticos através de estimativas do vetor gradiente do erro em relação aos pesos sinápticos (vide equação 28).

O MLP é uma rede cuja topologia constitui-se de uma camada sensorial ou de entrada, que possui tantos nós de entrada quantos forem os sinais de entrada, que dependem

intrinsecamente do problema, uma ou mais camadas intermediárias ou ocultas de neurônios e uma camada de saída com um número de neurônios idêntico ao número de sinais de saída, que são tantos quantos forem necessários à modelagem do problema. O sinal de entrada se propaga para frente (*feedforward*) através das camadas ocultas, até a camada de saída. Portanto, o MLP é uma rede alimentada para frente.

O número de camadas ocultas e a quantidade de neurônios artificiais de cada uma dessas camadas dependem da complexidade do problema que se deseja solucionar. A Figura 21 mostra a arquitetura de um MLP, onde N_i , N_h e N_o correspondem respectivamente, ao número de elementos dos vetores da camada de entrada, da camada oculta e da camada de saída.

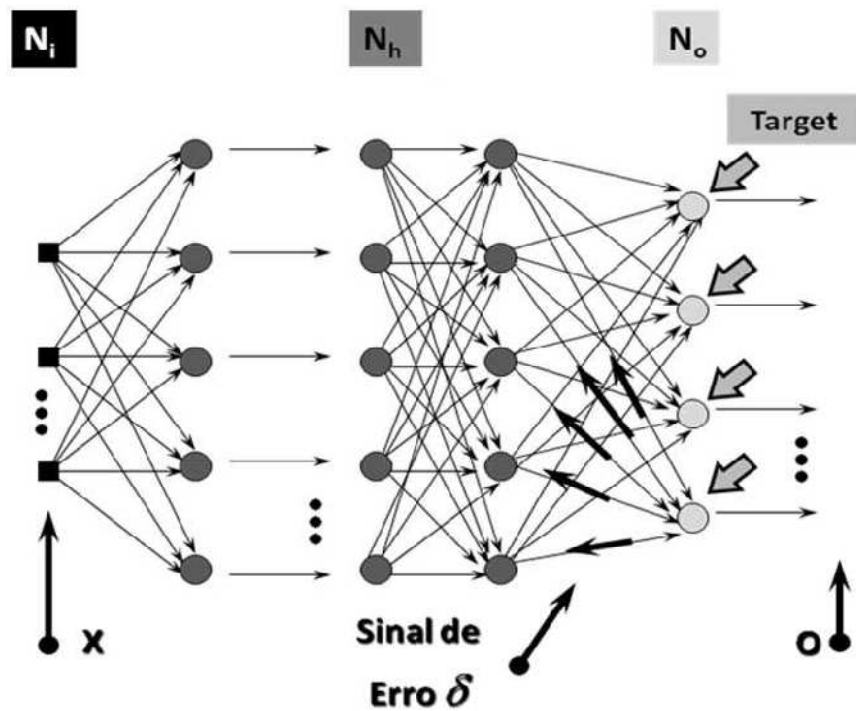


Figura 21: Arquitetura do Perceptron de Múltiplas Camadas (BIONDI NETO, et. al., 2009).

Os vetores X , O e T (*Target*) correspondem respectivamente aos vetores dos sinais de entrada, dos sinais de saída da RNA, e das saídas desejadas impostas pelo “supervisor”.

O treinamento do MLP, usando o algoritmo da retropropagação do erro (ver Figura 22), “Back-propagation”, BP, é realizado em duas etapas: a primeira é chamada de propagação e a segunda de retropropagação (HAYKIN, 2001). Durante a primeira etapa do

treinamento o vetor dos sinais de entrada $\mathbf{X} = [X_1, X_2, \dots, X_{N_i}]^T$ é aplicado a cada neurônio da primeira camada oculta da rede cada qual calculando a sua respectiva saída, que é então transmitida para frente como sinal de entrada para todos os neurônios da camada oculta subsequente, que repetem o mesmo procedimento de cálculo e transmissão, e assim por diante até a camada de saída da rede quando o valor da saída obtida pela rede $\mathbf{O} = [O_1, O_2, \dots, O_{N_o}]^T$ é comparado com o valor da saída desejada $\mathbf{T}$, *Target*, obtendo-se o sinal de erro $e_j = T_j - O_j$ de cada neurônio j desta camada para cada padrão apresentado.

O erro médio quadrático, considerando todos os N_o neurônios da camada de saída J e após a apresentação de todos os Q padrões de treinamento completando assim uma época, é calculado e definido por

$$E = \frac{1}{Q} \sum_{q=1}^Q E^q = \frac{1}{Q} \sum_{q=1}^Q \frac{1}{2} \sum_{j=1}^{N_o} (e_j)^2 = \frac{1}{Q} \sum_{q=1}^Q \frac{1}{2} \sum_{j=1}^{N_o} (T_j - O_j)^2 \quad (27)$$

onde O_j e T_j são respectivamente a saída calculada e a saída desejada para o neurônio j , onde $j = 1, 2, \dots, N_o$, e E^q é o erro total instantâneo no padrão q (HAYKIN, 2001).

A segunda etapa ocorre em sentido contrário. Há uma retropropagação do sinal de erro calculando-se δ (i.e., o gradiente local) de cada neurônio, atualizando os pesos e bias das conexões a partir da camada de saída J , propagando-se recursivamente pelas camadas ocultas até a camada de entrada, pelo uso da regra delta generalizada representada pela equação (28)

$$\Delta W_{j,i} = -\eta \cdot \frac{\partial E}{\partial W_{j,i}} \quad (28)$$

onde η é a taxa de aprendizado e i indica indexação referente à camada anterior I .

Verifica-se na equação (28) que o algoritmo BP, cujo fluxograma é representado na Figura 22, computa as derivadas parciais de uma função de erro E , com relação ao vetor peso $\mathbf{W}$ de certo vetor de entrada $\mathbf{X}$. O algoritmo BP, apesar de eficiente em várias aplicações pode apresentar alguns problemas como tempo de convergência longo, paralisia da rede e mínimo local (WASSERMAN, 1993).

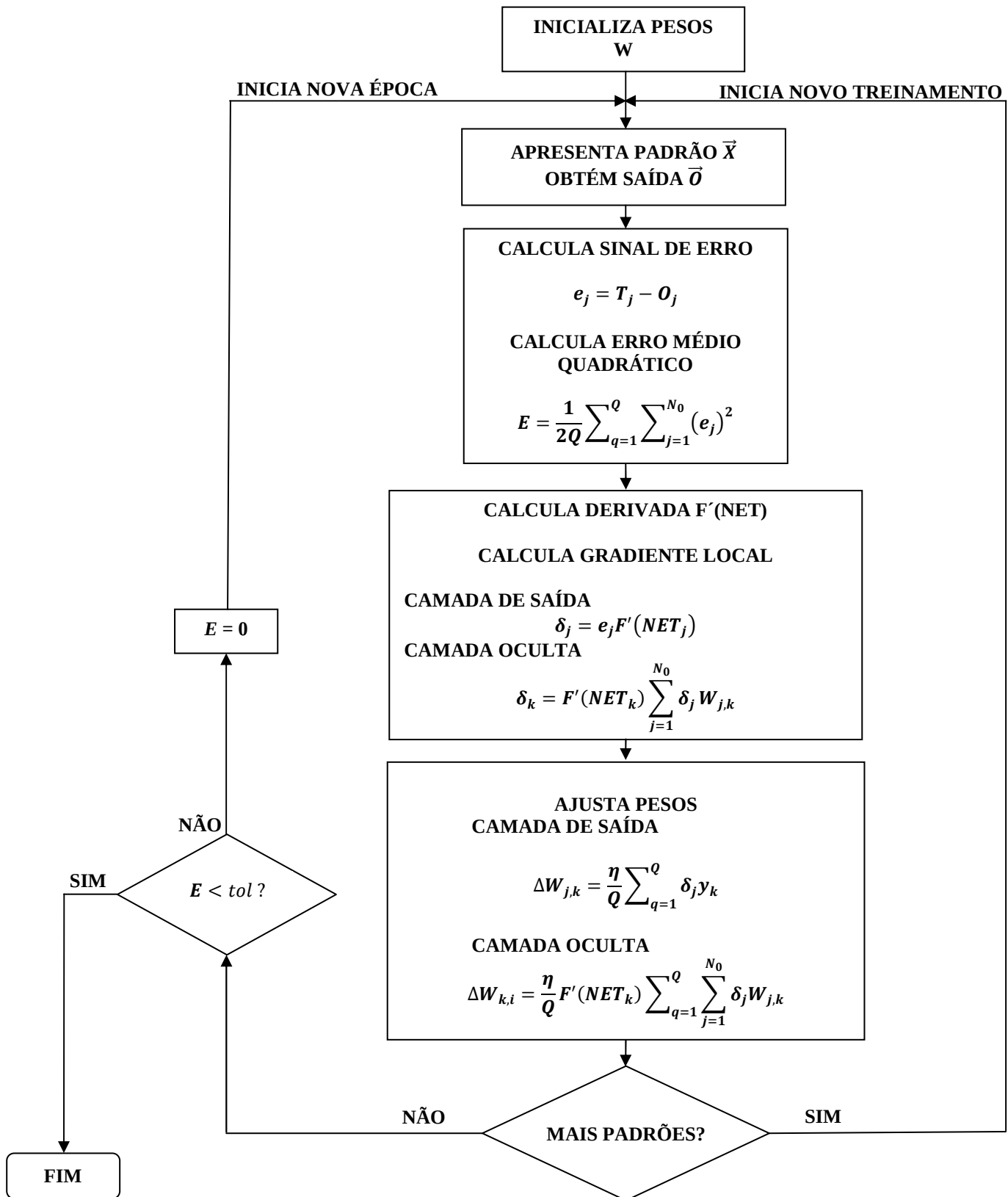


Figura 22: Fluxograma do algoritmo de retropropagação segundo o modo por lote de aprendizagem.

A taxa de aprendizado η é que responde pela velocidade com que ocorre a busca no espaço dos pesos, em direção aos valores que resultam em um erro global mínimo. Se a taxa de aprendizado não for adequadamente escolhida, podem ocorrer problemas durante o treinamento da rede. Quanto menor o valor de η mais suave será a curva de treinamento, porém o aprendizado se tornará muito lento ocorrendo um ajuste infinitesimal dos pesos, podendo inclusive ocorrer uma paralisia da rede em pontos onde a derivada da função de ativação seja nula ou infinitamente pequena.

A Figura 23 ilustra um treinamento onde a taxa de aprendizado é muito baixa, com uma variação muito pequena de $\Delta W_{j,i}$ de uma iteração para outra, provocando uma suave trajetória no espaço dos pesos, aumentando dessa forma, as chances de a rede ficar presa em um mínimo local, ou, na melhor das hipóteses, o treinamento ficar muito lento. Em contrapartida, se for usado um valor muito alto para a taxa de aprendizado, ocorrerão modificações muito intensas nos pesos, devido a uma alta variação de $\Delta W_{j,i}$, provocando oscilações intensas na curva de treinamento, causando instabilidade na rede. A Figura 24 ilustra um treinamento onde a taxa de aprendizado é muito alta, com uma variação intensa de $\Delta W_{j,i}$ de uma iteração para outra.

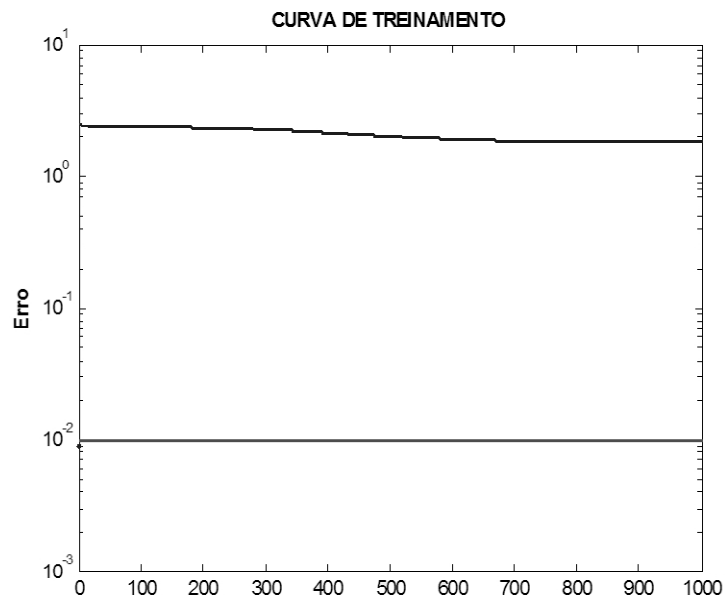


Figura 23: Treinamento com baixa taxa de aprendizado (BIONDI NETO, et. al., 2009).

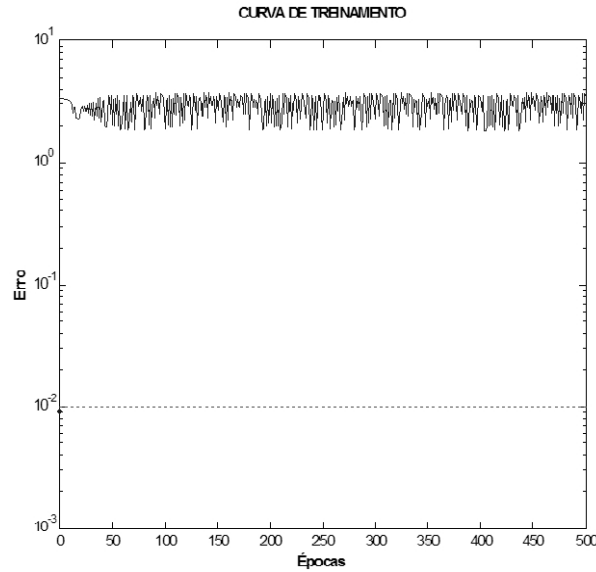


Figura 24: Treinamento com alta taxa de aprendizado (BIONDI NETO, et. al., 2009).

O problema da ocorrência de oscilações na curva de treinamento da RNA devido a utilização de altas taxas de aprendizado, pode ser bastante amenizado através do uso de Coeficiente de Momento na regra delta, de acordo com a equação (29). O uso do coeficiente de momento permite aumentar η sem comprometer a estabilidade da rede. A técnica consiste em amortecer a amplitude de $\Delta W_{j,i}(q)$ a cada iteração após a troca de sinais da derivada $\partial E(q)/\partial W_{j,i}(q)$, ou seja, após a passagem pelo valor $\Delta W_{j,i}(q)$ que resulta em um $E(q)$ mínimo. A mecânica do amortecimento ocorre quando na iteração em que a parcela $-\eta \cdot (\partial E(q)/\partial W_{j,i}(q))$ muda de sinal, o termo $\alpha \cdot \eta \cdot \partial E(q-1)/\partial W_{j,i}(q-1)$, ainda preserva o seu antigo sinal e, desta forma, $\Delta W_{j,i}(q)$ é amortecido, gerando um efeito estabilizador e, assim, evitando a oscilação no processo de busca no espaço dos pesos (JUNIOR; COSTA, 2007, p. 58).

A Figura 25, mostra um exemplo de treinamento com o uso do coeficiente de momento, α que é positivo e geralmente menor do que 0,9.

$$\Delta W_{j,i}(q) = -\eta \cdot \frac{\partial E(q)}{\partial W_{j,i}(q)} + \alpha \cdot \Delta W_{j,i}(q-1) \quad (29)$$

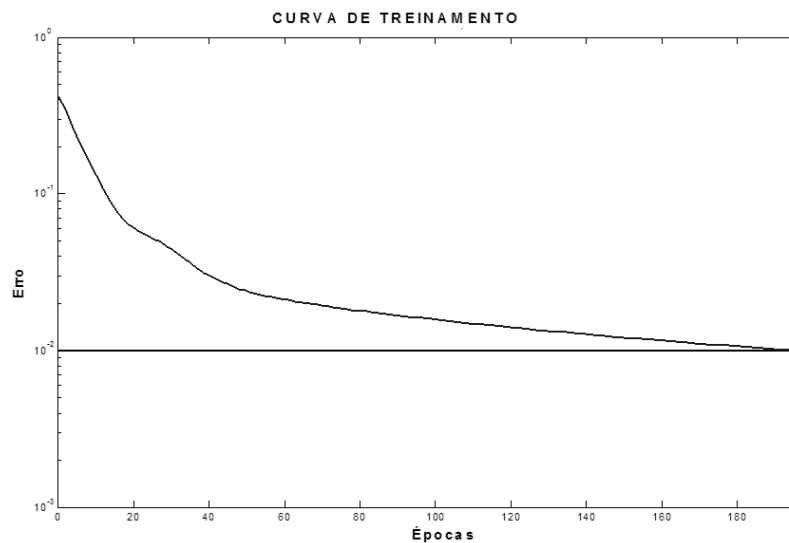


Figura 25: Treinamento com alta taxa de aprendizado e com coeficiente de momento (BIONDI NETO, et. al., 2009).

3.3 Fundamentação teórica sobre máquinas de comitê de redes neurais artificiais

As máquinas de comitê de redes neurais artificiais são técnicas usadas para resolver problemas encontrados no treinamento supervisionado, especialmente no tocante a generalização. Fundamentam-se no método cartesiano de dividir problemas complexos em diversos problemas mais simples resolvendo-os computacionalmente.

As máquinas de comitê consistem de estruturas de processamento de dados formadas por conjuntos de redes neurais artificiais em paralelo, chamadas de especialistas ou classificadores, cada qual realizando a mesma tarefa de forma independente no conjunto do problema. O resultado isolado obtido por cada especialista é combinado com os resultados isolados obtidos por cada uma das demais especialistas para se obter o resultado final da máquina de comitê. Resultado este que em geral, é melhor e mais confiável do que aquele obtido por cada especialista individualmente (HAYKIN, 2001; DUDA; HART; STORK, 2001; SILVA, 2005). Com o objetivo de obter um bom conjunto é preciso melhorar a capacidade de generalização das RNA especialistas (ZHOU; WU; TANG, 2002).

3.3.1 Classificação das máquinas de comitê de redes neurais artificiais

As máquinas de comitê de redes neurais, podem ser classificadas em duas grandes categorias (HAYKIN, 2001), a saber :

3.3.1.1 Estruturas Estáticas:

Nessa classe de máquinas de comitê de redes neurais, as saídas das RNA especialistas, são combinadas sem a participação do sinal de entrada. Essa classe usa principalmente dois métodos de tratamento do sinal de saída das RNA: a média aritmética de conjunto e o método do reforço. No método da média aritmética de conjunto as entradas X das RNA especialistas são rigorosamente as mesmas e a saída global Y é dada pela combinação linear das saídas Y_i , $i = 1, \dots, N$, das RNA especialistas que são treinadas individualmente. É de se esperar que as RNA especialistas treinadas de forma diferente e individual, converjam para diferentes mínimos locais na superfície de erro, e que, o desempenho global seja aumentado combinando-se as saídas de alguma forma. São permitidas variações nas condições iniciais usadas no treinamento e modificações nos parâmetros e algoritmos de minimização do erro. A Figura 26, mostra um diagrama em blocos de uma máquina de comitê de redes neurais artificiais do tipo estática (que é o tipo adotado neste trabalho de dissertação, com pequenas alterações)

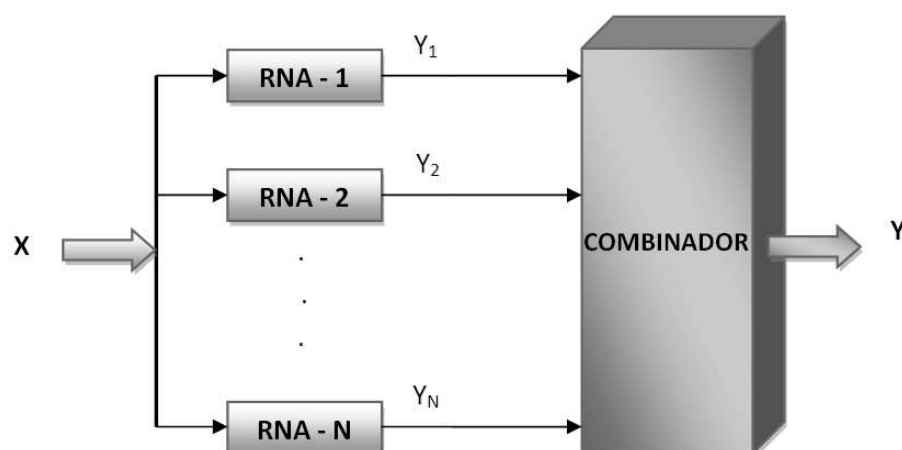


Figura 26: Diagrama em blocos de uma estrutura estática de um comitê usando o método da média aritmética (BIONDI NETO, et. al., 2009).

No método de reforço, representado pela Figura 27, o conjunto de dados de treinamento das RNA especialistas apresentam distribuições distintas, isto é diferentes sinais de entrada. Nesse caso o objetivo é melhorar qualquer algoritmo de treinamento (HAYKIN, 2001).

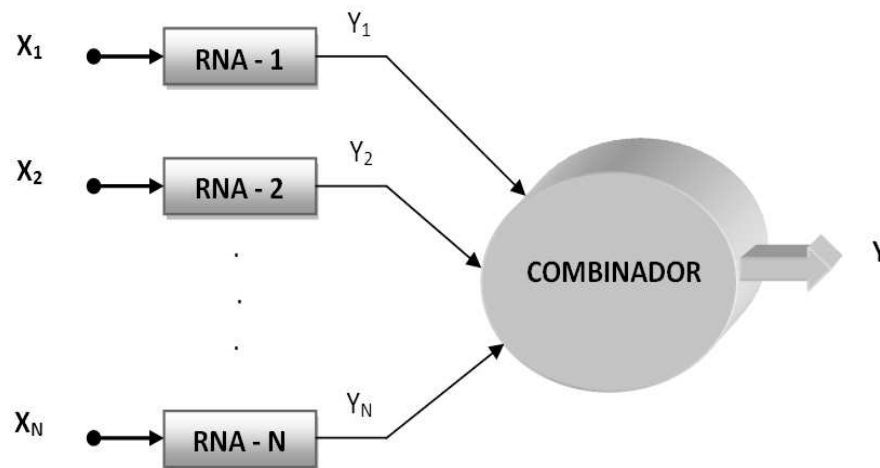


Figura 27: Diagrama em blocos de uma estrutura estática de uma máquina de comitê usando o método do reforço (BIONDI NETO et. al., 2009).

3.3.1.2 Estruturas Dinâmicas:

Nesta categoria de máquinas de comitê, cujo método mais conhecido é a mistura de especialistas, as saídas das RNA especialistas são ponderadas por uma estrutura combinadora não linear, que geralmente é outra RNA denominada de rede de passagem, que funciona como um mediador entre as RNA especialistas. A rede de passagem apresenta apenas uma camada de neurônios, um neurônio para cada RNA especialista. A função de ativação desses neurônios é uma generalização da função logística para múltiplas entradas definida pela equação (30), e representa uma generalização diferenciável da operação “o vencedor leva tudo” de escolha do valor máximo e conhecida como *softmax* (HAYKIN, 2001). O termo NET_j representa o produto do vetor peso sináptico pelo vetor de entrada do neurônio especialista. O termo NET_i representa o produto do vetor de entrada pelo vetor de pesos sinápticos da rede de passagem. As redes neurais especialistas e a rede de passagem são treinadas simultaneamente

$$g_i = \frac{\exp(NE_{T_i})}{\sum_{j=1}^N \exp(NE_{T_j})}, i = 1, 2, \dots, N \quad (30)$$

A saída global Y do modelo é dada pela equação (31), onde Y_i representa a saída da RNA especialista e g_i a função em x não linear da rede de passagem (HAYKIN, 2001; DUDA; HART ; STORK, 2001).

$$Y = \sum_{j=1}^N g_i Y_i \quad (31)$$

Se a saída da rede de passagem for $g_i = 0$ ($i = 1, 2, \dots, N$) indica que a RNA especialista correspondente não tem influência sobre a saída global, enquanto se $g_i = 1$ ($i = 1, 2, \dots, N$) indica controle total da RNA especialista sobre a saída global. Assim as saídas são combinadas de forma mais inteligente permitindo que as melhores RNA especialistas contribuam para a saída global de um dado padrão. A Figura 28 representa o diagrama em blocos de uma estrutura dinâmica de uma máquina de comitê usando mistura de especialistas.

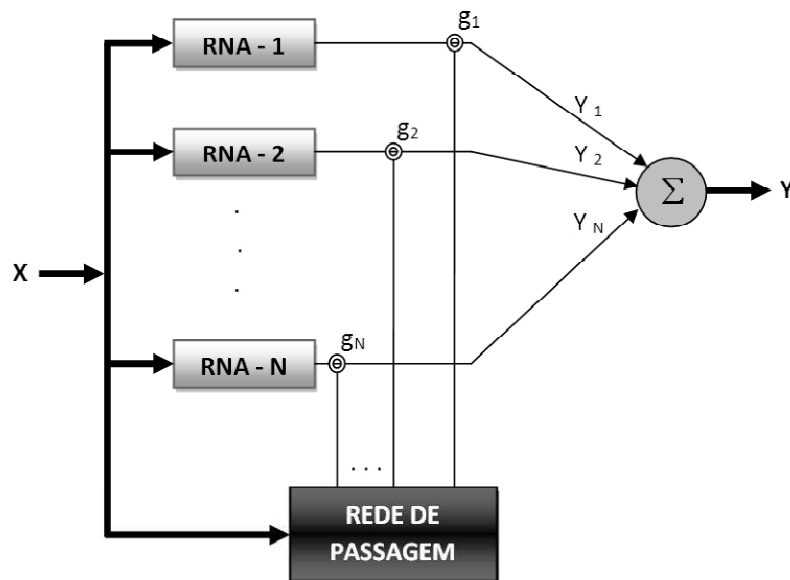


Figura 28: Diagrama de blocos de uma estrutura dinâmica de uma máquina de comitê usando mistura de especialistas (BIONDI NETO et al., 2009).

3.4 Modelagem e formulação matemática para a solução do problema inverso em transferência radiativa usando redes neurais artificiais

Primeiramente será desenvolvida uma modelagem e uma formulação matemática considerando que as redes neurais especialistas que compõem as máquinas de comitê tenham duas camadas ocultas. Essa modelagem e formulação englobam àquelas para redes neurais especialistas com apenas uma camada oculta, cuja modelagem e formulação serão mostradas em seguida.

A Figura 29 representa um MLP, que modela cada uma das redes neurais especialistas, que compõem as máquinas de comitê aqui aplicadas para a solução do problema inverso em transferência radiativa.

O MLP da Figura 29 é constituído de uma camada de entrada, duas camadas ocultas com N_{H1} e N_{H2} neurônios respectivamente, e uma camada de saída com N_u neurônios (aqui $N_u = 4$), sendo um neurônio para cada parâmetro radiativo.

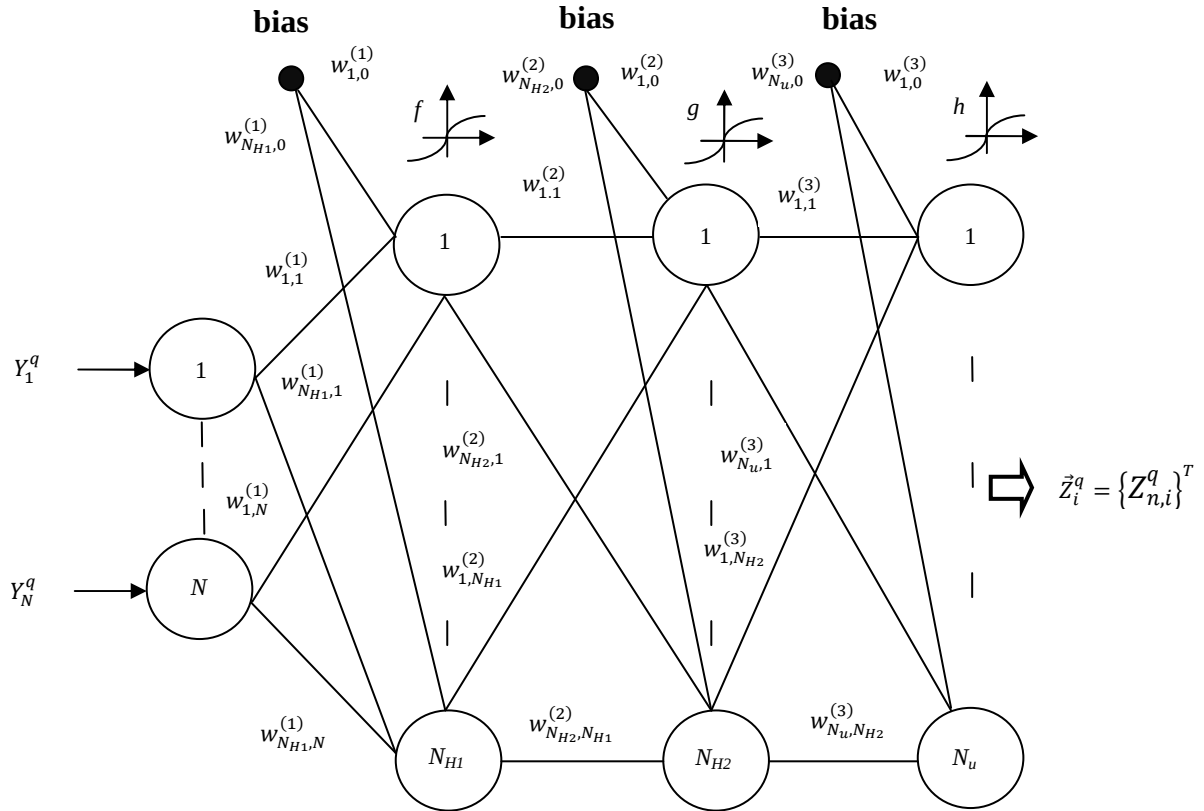


Figura 29: MLP com duas camadas ocultas para o problema inverso.

Nas redes neurais especialistas representadas pelo modelo da Figura 29 cada neurônio m , com $m = 1, 2, \dots, N_{H1}$, na primeira camada oculta realiza uma combinação linear dos valores de entrada Y_l^q , com $l = 1, 2, \dots, N$, obtidos aleatoriamente, e fornecidas pela camada de entrada

$$a_m^q = \sum_{l=1}^N w_{m,l}^{(1)} Y_l^q + w_{m,0}^{(1)} \quad (32)$$

onde $w_{m,l}^{(1)}$, $m = 1, 2, \dots, N_{H1}$, $l = 1, 2, \dots, N$, são os pesos sinápticos que conectam os nós das entradas com os neurônios da primeira camada oculta, N é o número de nós na camada de entrada, N_{H1} é o número de neurônios da primeira camada oculta, e o índice q representa a ordem do padrão de treinamento que está entrando na RNA, $q = 0, \dots, Q-1$, onde aqui $Q = 10.000$.

A soma ponderada a_m^q dada pela equação (32) é vista como uma excitação para o neurônio m da primeira camada oculta, que fornece a resposta

$$p_m^q = f(a_m^q), m = 1, 2, \dots, N_{H1} \quad (33)$$

onde $f(.)$ é uma função de ativação. Várias escolhas para a função $f(.)$ são possíveis (HAYKIN, 2001).

Cada neurônio k , com $k = 1, 2, \dots, N_{H2}$, na segunda camada oculta realiza uma combinação linear dos valores das respostas p_m^q fornecidas pela primeira camada oculta

$$b_k^q = \sum_{m=1}^{N_{H1}} w_{k,m}^{(2)} p_m^q + w_{k,0}^{(2)}, k = 1, 2, \dots, N_{H2} \quad (34)$$

onde $w_{k,m}^{(2)}$, $k = 1, 2, \dots, N_{H2}$, $m = 1, 2, \dots, N_{H1}$, são os pesos sinápticos que conectam os neurônios da primeira camada oculta com os neurônios da segunda camada oculta.

A soma ponderada b_k^q dada pela equação (34) é vista como uma excitação para o neurônio k da segunda camada oculta, que fornece a resposta

$$v_k^q = g(b_k^q), k = 1, 2, \dots, N_{H2} \quad (35)$$

onde $g(\cdot)$ é uma função de ativação. Várias escolhas para a função $g(\cdot)$ são possíveis (HAYKIN, 2001).

Cada neurônio n , $n = 1, 2, \dots, N_u$ da camada de saída, realiza uma combinação linear das respostas v_k^q , $k = 1, 2, \dots, N_{H2}$, dos neurônios da segunda camada oculta

$$s_n^q = \sum_{k=1}^{N_{H2}} w_{n,k}^{(3)} v_k^q + w_{n,0}^{(3)}, \quad n = 1, 2, \dots, N_u \quad (36)$$

onde $w_{n,k}^{(3)}$, $n = 1, 2, \dots, N_u$, $k = 1, 2, \dots, N_{H2}$, são os pesos sinápticos entre os neurônios da segunda camada oculta e os neurônios da camada de saída, e N_u é o número de neurônios na camada de saída, que coincide com o número de incógnitas do problema inverso. Aqui temos $N_u = 4$, conforme a equação (24), no capítulo 2.

A soma ponderada s_n^q dada pela equação (36) é vista como uma excitação para o neurônio n da camada de saída, que produz uma resposta

$$t_n^q = h(s_n^q), \quad n = 1, 2, \dots, N_u, \quad (37)$$

onde $h(\cdot)$ é uma função de excitação. Várias escolhas para a função $h(\cdot)$ são possíveis (HAYKIN, 2001).

Combinando as equações (32) à (37), para duas camadas ocultas, obtemos

$$t_n^q = h \left(\sum_{k=1}^{N_{H2}} w_{n,k}^{(3)} g \left(\sum_{m=1}^{N_{H1}} w_{k,m}^{(2)} f \left(\sum_{l=1}^N w_{m,l}^{(1)} Y_l^q + w_{m,0}^{(1)} \right) + w_{k,0}^{(2)} \right) + w_{n,0}^{(3)} \right), \quad n = 1, 2, \dots, N_u, \quad (38)$$

Considerando as entradas Y_l^q , $l = 1, 2, \dots, N$, $q = 0, \dots, Q-1$, observamos na equação (38) que t_n^q , $n = 1, 2, \dots, N_u$, componentes dos vetores $\vec{t}_i^q$, são as estimativas, durante o treinamento, para os parâmetros radiativos, $Z_{n,i}^q$, com $n = 1, 2, \dots, N_u$, que são as componentes do vetor de saídas $\vec{Z}_i^q$ de cada rede neural especialista, RNA_i , $i = 1, 2, \dots, M$, aqui inicialmente $M = 10$ e finalmente $M = 6$. Mas antes que possamos usar a equação (38) devemos determinar os pesos sinápticos $w^{(1)}$, $w^{(2)}$ e $w^{(3)}$.

A determinação dos pesos $w^{(1)}$, $w^{(2)}$ e $w^{(3)}$ é feita pela apresentação de um conjunto de padrões: entradas conhecidas $\vec{Y}^q = \{Y_l^q\}^T$ e saídas desejadas $\vec{Z}_{alvo}^q$, gerados

randomicamente. Calculam-se então, os pesos que produzam o melhor ajuste entre o valores $\vec{Z}_i^q$ estimados e os vetores das saídas desejadas, $\vec{Z}_{alvo}^q$. As entradas Y_l^q usadas no treinamento supervisionado das RNA foram geradas calculando-se os valores $I_{calcl}(\vec{Z}_{alvo}^q)$ a partir do vetor de parâmetros radiativos $\vec{Z}_{alvo}^q$ gerados aleatoriamente com a função DRNUNF() do MSIMSL do software FORTRAN 77, que gera números aleatórios no intervalo (0,1).

Para a determinação de $w^{(1)}$, $w^{(2)}$ e $w^{(3)}$ é usado o algoritmo de retropropagação (vide Figura 22). Parte-se de uma estimativa inicial para os pesos $w^{(1)}$, $w^{(2)}$ e $w^{(3)}$, e após a apresentação de todos os padrões de treinamento e os vetores das entradas $\vec{Y}^q = \{Y_l^q\}^T$ terem sido propagados à frente através da rede produzindo estimativas sucessivas $\vec{t}_i^q \equiv \vec{Z}_i^q$ para as saídas (parâmetros radiativos), que comparadas com os vetores das saídas desejadas $\vec{Z}_{alvo}^q$ produzem os sinais de erros,

$$E_n^q = Z_{nalvo}^q - Z_{n,i}^q, \quad n = 1, 2, \dots, N_u \text{ e } i = 1, 2, \dots, M \quad (39)$$

os pesos são então ajustados usando a informação fornecida pelo erro da saída

$$\Delta w_{n,k}^{(3)} = \frac{\eta}{Q} \sum_{q=1}^Q \delta_n^{(3)q} v_k^q \quad (40a)$$

$$\Delta w_{k,m}^{(2)} = \frac{\eta}{Q} \sum_{q=1}^Q \delta_k^{(2)q} p_m^q \quad (40b)$$

$$\Delta w_{m,l}^{(1)} = \frac{\eta}{Q} \sum_{q=1}^Q \delta_m^{(1)q} Y_l^q \quad (40c)$$

onde

$$\delta_n^{(3)q} = E_n^q h'(s_n^q) \quad (41a)$$

$$\delta_k^{(2)q} = g'(b_k^q) \sum_{n=1}^{N_u} \delta_n^{(3)q} w_{n,k}^{(3)} \quad (41b)$$

$$\delta_m^{(1)q} = f'(a_m^q) \sum_{k=1}^{N_{H2}} \delta_k^{(2)q} w_{k,m}^{(2)} \quad (41c)$$

são gradientes locais de retropropagação dos sinais de erro desde a camada de saída até a camada de entrada, possibilitando o ajuste dos pesos e bias desde a camada de saída até a primeira camada oculta da rede neural.

A letra grega η representa a taxa de aprendizado, que pode assumir diferentes valores entre 0 e 1. Aqui $\eta = 0,8$, e h' , g' e f' representam as derivadas das funções de ativação.

Considerando agora que seja eliminada a segunda camada oculta com N_{H2} neurônios, do MLP representado na Figura 29, permanecendo apenas a primeira camada oculta com N_{H1} neurônios e a camada de saída com N_u (aqui $N_u = 4$) neurônios, temos então o MLP representado pela Figura 30.

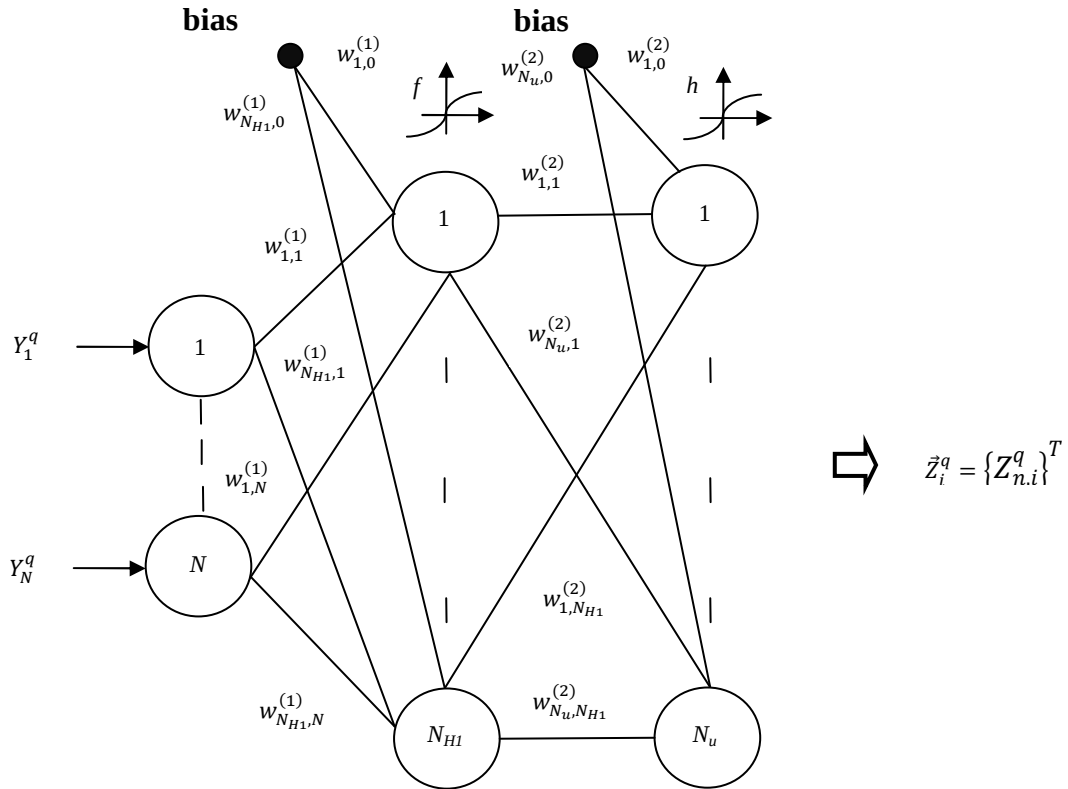


Figura 30: MLP com uma camada oculta para o problema inverso.

Os componentes t_n^q , com $n = 1, 2, \dots, N_u$, dos vetores $\vec{t}_i^q$, os quais são os valores calculados e atribuídos às estimativas, $Z_{n,i}^q$, com $n = 1, 2, \dots, N_u$, dos parâmetros radiativos que são as componentes do vetor de saídas $\vec{Z}_i^q$ de cada rede neural especialista, RNA_i , $i = 1, 2, \dots, M$, terão a seguinte formulação

$$t_n^q = h \left(\sum_{m=1}^{N_{H1}} w_{n,m}^{(2)} f \left(\sum_{l=1}^N w_{m,l}^{(1)} Y_l^q + w_{m,0}^{(1)} \right) + w_{n,0}^{(2)} \right), n = 1, 2, \dots, N_u, \quad (42)$$

O cálculo do sinal de erro é obtido usando as equações (39) e (43a) à (43e)

$$\Delta w_{n,m}^{(2)} = \frac{\eta}{Q} \sum_{q=1}^Q \delta_n^{(2)q} p_m^q \quad (43a)$$

$$\Delta w_{m,l}^{(1)} = \frac{\eta}{Q} \sum_{q=1}^Q \delta_m^{(1)q} Y_l^q \quad (43b)$$

onde

$$\delta_n^{(2)q} = E_n^q h'(s_n^q) \quad (43c)$$

com

$$s_n^q = \sum_{m=1}^{N_{H1}} w_{n,m}^{(2)} p_m^q + w_{n,0}^{(2)}, n = 1, 2, \dots, N_u \quad (43d)$$

$$\delta_m^{(1)q} = f'(a_m^q) \sum_{n=1}^{N_u} \delta_n^{(2)q} w_{n,m}^{(2)} \quad (43e)$$

onde as equações (43c) e (43e) são os gradientes locais, de retropropagação dos sinais de erro desde a camada de saída até a camada de entrada, possibilitando o ajuste dos pesos e bias da camada de saída até a primeira camada oculta da rede neural.

3.5 Modelagem e formulação matemática para a solução do problema inverso em transferência radiativa usando máquinas de comitê de redes neurais artificiais

A solução do problema inverso em transferência radiativa utilizando máquinas de comitê é apresentada neste item, tendo como modelo básico uma máquina de comitê do tipo estática que utiliza a média aritmética de conjunto de acordo com a configuração mostrada na Figura 1, contida no capítulo 1 desta dissertação.

3.5.1 Modelagem

Quanto à modelagem têm-se duas configurações que são aqui consideradas: a primeira é o modelo clássico de máquina de comitê do tipo estática com a média aritmética de conjunto, representado na Figura 31, onde o combinador combina as saídas resultantes da

generalização das RNA especialistas através de sua média aritmética (combinação linear), que será então a saída final da máquina de comitê; a segunda configuração, representada na Figura 32, integra um MLP ao combinador da máquina de comitê, que agrega as saídas das RNA especialistas através da média aritmética das mesmas que será então a excitação da camada de entrada do MLP que integra o combinador, que por sua vez fornecerá o vetor de saídas da máquina de comitê que será confrontado com o vetor alvo na fase de treinamento para cada padrão q apresentado.

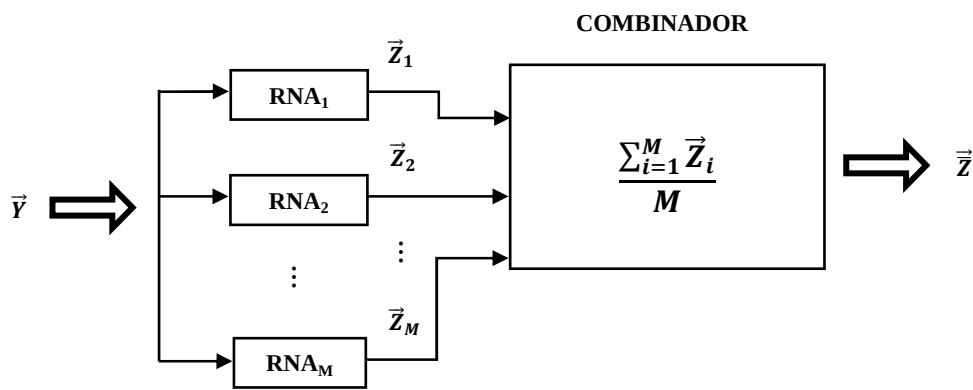


Figura 31: Modelo da máquina de comitê tipo estática onde o combinador efetua apenas a média aritmética das saídas das RNA especialistas.

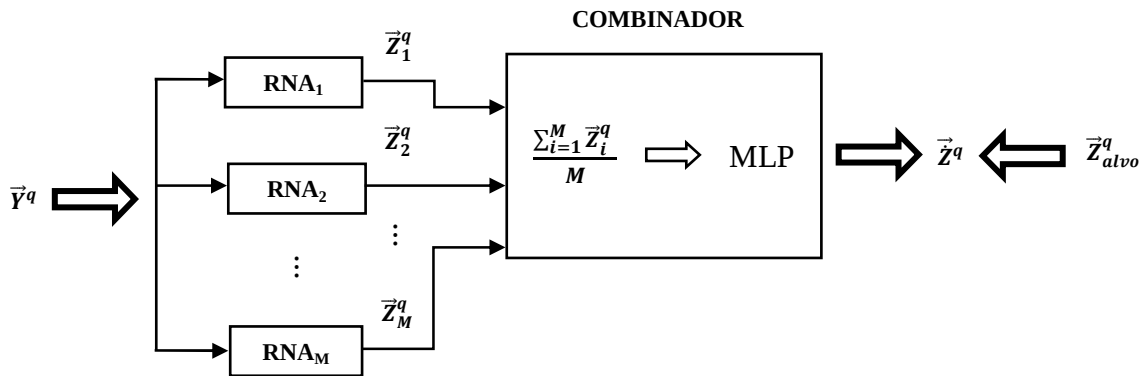


Figura 32: Modelo da máquina de comitê tipo estática com MLP integrado ao combinador.

Com o objetivo de solucionar o problema inverso de transferência radiativa os dois modelos são usados, de forma separada e independente, sendo cada qual composto por M redes neurais especialistas do tipo Perceptron de Múltiplas Camadas (MLP), tendo cada especialista uma ou duas camadas ocultas e uma camada de saída com quatro neurônios, um

para cada parâmetro radiativo. Aqui inicialmente $M = 10$, e após a seleção das melhores redes neurais especialistas quanto à generalização, $M = 6$ (para $\omega = 0,1, 0,5$ ou $0,9$). As M redes neurais especialistas, para cada modelo de máquina, apresentam o mesmo número de camadas ocultas com igual número de neurônios (porém diferentes de uma RNA especialista para outra), mesmas funções de treinamento e de ativação, e mesma taxa de aprendizado ($\eta = 0,8$). O MLP que integra o combinador de saída da máquina de comitê é constituído por uma camada oculta com N_h neurônios, aqui $N_h = 12$ (em todas as simulações realizadas) e uma camada de saída com quatro neurônios, um para cada parâmetro radiativo.

3.5.2 Formulação

Para uma máquina de comitê de acordo com o modelo representado na Figura 31, i.e., efetuando apenas a média aritmética das saídas das RNA especialistas no combinador, na fase de generalização, temos que os vetores $\vec{Z}_i$ (que são as saídas das redes neurais especialistas RNA_i , $i = 1, 2, \dots, M$), cujas componentes são $Z_{n,i}$, $n = 1, 2, \dots, N_u$ (estimativas para as componentes do vetor $\vec{Z}$ da equação (24)), são combinados linearmente por meio de uma média aritmética, de acordo com a equação (44) e, então o vetor $\vec{Z}$, obtido através da equação (44), será a saída da máquina de comitê representada de acordo com o modelo da Figura 31.

$$\vec{Z} = \sum_{i=1}^M \frac{\vec{Z}_i}{M} \quad (44)$$

O vetor $\vec{Z}$, é um vetor coluna cujos elementos $\bar{Z}_n$, $n = 1, 2, \dots, N_u$ (aqui $N_u = 4$) é cada um deles a média aritmética dos correspondentes componentes $Z_{n,i}$ dos vetores $\vec{Z}_i$, $i = 1, 2, \dots, M$, das estimativas dos parâmetros radiativos obtidos pelas redes neurais especialistas na fase de generalização.

Para uma máquina de comitê de acordo com o modelo representado pela Figura 32, um MLP, conforme representado na Figura 33, é parte integrante do combinador do modelo proposto para a máquina de comitê.

Na camada oculta h^* do MLP de saída do combinador da máquina de comitê de redes neurais representado na Figura 33, o neurônio u , com $u = 1, 2, \dots, N_h$, realiza uma combinação linear, através da equação (45), dos valores dos resultados $\bar{Z}_n^q$, $n = 1, 2, \dots, N_u$, fornecidos pela

média aritmética das saídas das redes neurais especialistas RNA_i , $i = 1, 2, \dots, M$, obtida através da equação (44) (considerada agora para a fase de treinamento a cada padrão q apresentado)

$$z_u^q = \sum_{n=1}^{N_u} w_{u,n}^{(h^*)} \bar{z}_n^q + w_{u,0}^{(h^*)} \quad (45)$$

onde $w_{u,n}^{(h^*)}$, $u = 1, 2, \dots, N_h$, $n = 1, 2, \dots, N_u$, $q = 0, 1, 2, \dots, Q-1$, (aqui $Q = 10.000$, número de padrões de treinamento que é igual a uma época) são os pesos sinápticos entre o neurônio u da camada oculta h^* do MLP de saída do combinador da máquina de comitê de redes neurais e os nós de suas entradas, que são os componentes $\bar{z}_n^q$, $n = 1, 2, \dots, N_u$ do vetor $\vec{\bar{z}}^q$, média aritmética das saídas das especialistas para cada padrão q apresentado, obtido através da equação (44), considerada agora para a fase de treinamento a cada padrão q apresentado.

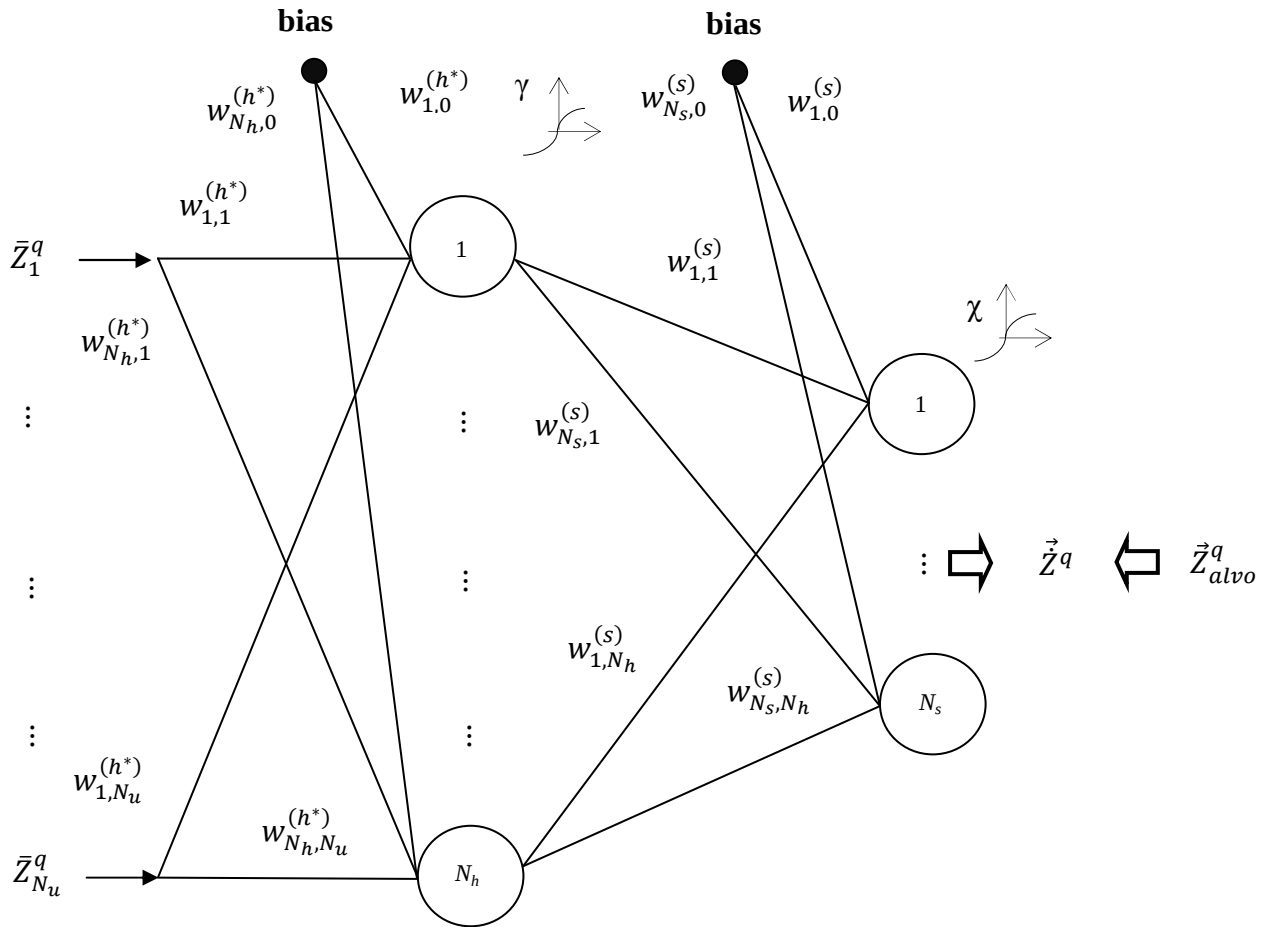


Figura 33: Modelo do MLP do combinador da máquina de comitê.

A soma ponderada z_u^q dada pela equação (45) é vista como uma excitação para o neurônio u , que fornece a resposta

$$\hat{z}_u^q = \gamma(z_u^q), u = 1, 2, \dots, N_h \quad (46)$$

onde $\gamma(\cdot)$ é uma função de ativação. Várias escolhas para a função $\gamma(\cdot)$ são possíveis (HAYKIN, 2001).

Na camada de saída s do MLP de saída do combinador da máquina de comitê de redes neurais na Figura 33, cada neurônio j , com $j = 1, 2, \dots, N_s$, aqui $N_s = 4$, realiza uma combinação linear com os valores das respostas $\hat{z}_u^q$ fornecidas pela camada oculta h^*

$$\hat{Z}_j^q = \sum_{u=1}^{N_h} w_{j,u}^{(s)} \hat{z}_u^q + w_{j,0}^{(s)} \quad (47)$$

onde $w_{j,u}^{(s)}$, $j = 1, 2, \dots, N_s$, $u = 1, 2, \dots, N_h$, são os pesos sinápticos entre os neurônios da camada oculta e a camada de saída.

A soma ponderada $\hat{Z}_j^q$ dada pela equação (47) é vista como uma excitação para o neurônio j , com $j = 1, 2, \dots, N_s$, da camada de saída do MLP de saída do combinador da máquina de comitê de redes neurais, que produz a resposta

$$\dot{Z}_j^q = \chi(\hat{Z}_j^q), j = 1, 2, \dots, N_s \quad (48)$$

onde $\chi(\cdot)$ é uma função ativação. Várias escolhas para a função $\chi(\cdot)$ são possíveis (HAYKIN, 2001).

As respostas $\dot{Z}_j^q$, $j = 1, 2, \dots, N_s$, aqui $N_s = 4$, são as estimativas para os parâmetros radiativos, obtidas pela máquina de comitê de redes neurais, e que são os componentes do vetor de saída $\vec{\dot{Z}}^q$, que é comparado com vetor alvo $\vec{\dot{Z}}_{alvo}^q$ a cada iteração q , $q = 0, 1, 2, \dots, Q-1$.

Para a determinação dos pesos $w_{u,n}^{(h^*)}$ e $w_{j,u}^{(s)}$ o procedimento é o mesmo que o utilizado para a obtenção dos pesos $w^{(1)}$, $w^{(2)}$ e $w^{(3)}$, das RNAs especialistas, sendo que aqui as entradas são a média aritmética das saídas das redes neurais especialistas, $\bar{Z}_n^q$, $n = 1, 2, \dots, N_u$, aqui $N_u = 4$.

Os sinais de erros são dados por

$$E_j^q = Z_{jalvo}^q - \dot{Z}_j^q, \quad j = 1, 2, \dots, N_s \quad (49)$$

Onde $Z_{jalvo}^q, j = 1, 2, \dots, N_s$ são os componentes do vetor alvo $\vec{Z}_{alvo}^q$.

Os pesos são então ajustados usando a informação fornecida pelo erro da saída (HAYKIN, 2001),

$$\Delta w_{j,u}^{(s)} = \frac{\eta}{Q} \sum_{q=1}^Q \delta_j^{(s)q} \hat{Z}_u^q \quad (50a)$$

$$\Delta w_{u,n}^{(h^*)} = \frac{\eta}{Q} \sum_{q=1}^Q \delta_u^{(h^*)q} \bar{Z}_n^q \quad (50b)$$

onde

$$\delta_j^{(s)q} = E_j^q \chi'(\hat{Z}_j^q) \quad ..(51a)$$

$$\delta_u^{(h^*)q} = \gamma'(z_u^q) \sum_{j=1}^{N_s} \delta_j^{(s)q} w_{j,u}^{(s)} \quad (51b)$$

são os gradientes locais de retropropagação dos sinais de erro desde a camada de saída até a camada de entrada do MLP do combinador de saída da máquina de comitê, possibilitando assim o ajuste dos pesos e bias.

A taxa de aprendizado η , pode assumir diferentes valores entre 0 e 1. Aqui $\eta = 0,8$ e χ' e γ' representam as derivadas das funções de ativação.

4 RESULTADOS E DISCUSSÃO

4.1 Metodologia aplicada

Os resultados aqui mostrados foram obtidos a partir dos testes de generalização realizados de forma independente para todas as RNA especialistas e para os modelos de máquinas de comitê representados nas Figuras 31 e 32, do capítulo 3, ambos utilizando as mesmas saídas $\vec{Z} = \{\tau_0, \omega, \rho_1, \rho_2\}$, obtidas pelas RNA especialistas como entradas para o combinador de cada um dos modelos de máquinas de comitê aplicados.

Os resultados obtidos nos testes de generalização realizados são mostrados através de tabelas geradas a partir de seis configurações distintas:

1. Os resultados obtidos com uma máquina de comitê do tipo estática cujo combinador realiza a média aritmética das saídas das RNA especialistas (vide Figura 31), todas do tipo MLP, cada qual com uma camada oculta (vide Figura 30);
2. Os resultados obtidos com uma máquina de comitê do tipo estática (vide Figura 32) tendo integrado ao seu combinador uma rede neural do tipo MLP com uma camada oculta (vide Figura 33) cujas excitações de entrada são a média aritmética das saídas das RNA especialistas, todas do tipo MLP, cada qual com uma camada oculta (vide Figura 30);
3. Os resultados obtidos com uma máquina de comitê do tipo estática cujo combinador realiza a média aritmética das saídas das RNA especialistas (vide Figura 31), todas do tipo MLP cada qual com duas camadas ocultas (vide Figura 29);
4. Os resultados obtidos com uma máquina de comitê do tipo estática (vide Figura 32) tendo integrado ao seu combinador uma rede neural do tipo MLP com uma camada oculta (vide Figura 33), cujas excitações de entrada são a média aritmética das saídas das RNA especialistas, todas do tipo MLP, cada qual com duas camadas ocultas (vide Figura 29);
5. Os resultados obtidos com uma máquina de comitê do tipo estática cujo combinador realiza a média aritmética das saídas das RNA especialistas (vide Figura 31), todas do tipo MLP, cada qual com uma ou duas camadas ocultas (vide Figuras 30 e 29); e

6. Os resultados obtidos com uma máquina de comitê do tipo estática (vide Figura 32) tendo integrado ao seu combinador uma rede neural do tipo MLP com uma camada oculta (Figura 33) cujas excitações de entrada são a média aritmética das saídas das RNA especialistas, todas do tipo MLP, cada qual com uma ou duas camadas ocultas (vide Figuras 30 e 29).

Estes resultados são comparados entre si e com aqueles obtidos com as redes neurais especialistas isoladamente, através da comparação dos erros relativos percentuais calculados de acordo com a equação (52)

$$ER = \left| \frac{\vec{Z}_{exato} - \vec{Z}_{estimado}}{\vec{Z}_{exato}} \right| \times 100 \% \quad (52)$$

Para todos os testes de generalização realizados foram utilizados os seguintes valores para as variáveis e parâmetros envolvidos:

- Valores exatos dos parâmetros radiativos $\vec{Z}_{exato} = \{\tau_0, \omega, \rho_1, \rho_2\}$, tem-se: espessura ótica $\tau_0 = 1,0$; albedo de espalhamento simples $\omega = 0,1$; 0,5 ou 0,9 (baixo, médio ou alto espalhamento); refletividades difusas $\rho_1 = 0,5$ e $\rho_2 = 0,5$;
- Valores das radiações externas incidentes: $f_1 = 1,0$ e $f_2 = 0,0$;
- Na obtenção dos dados experimentais $Y_l = I_{exp_l} = I_{calc_l}(\vec{Z}_{exato}) + \sigma r_l$, tem-se: $r_l \in [-1, 1]$ um número aleatório com distribuição uniforme e $\sigma = 0,002$ o desvio padrão do ruído nos dados experimentais;
- Taxa de aprendizado: $\eta = 0,8$;
- Número de neurônios nas camadas ocultas das RNA especialistas: $N_{H1} = N_{H2} = 8, 10, 12, 14, 16, 18, 20, 22, 24$ ou 26 (idem para RNA especialistas com apenas uma camada oculta);
- Número de neurônios na camada oculta da rede neural do tipo MLP representado na Figura 33, que integra o combinador da máquina de comitê com modelo de acordo com o mostrado na Figura 32: $N_h = 12$, para todos os testes realizados. Foram testados $N_h = 4, 5, 6, 7, 8, 9, 10, 11$ ou 12, porém, entre todos os valores, o que apresentou melhores resultados na generalização foi $N_h = 12$;

- Número de neurônios na camada de saída de cada RNA especialista e do MLP que integra o combinador de saída da máquina de comitê de modelo conforme a Figura 32: $N_u = 4$;
- Limitante superior para o número de épocas $e = 3000$; e
- Limitante superior estabelecido para o erro médio quadrático E : $tol = 10^{-3}$ ou 10^{-4} .

No treinamento e testes de generalização de todas as redes neurais artificiais usadas, foram utilizados programas implementados com o software MATLAB®, e para a obtenção dos dados experimentais simulados Y_l usados nos testes de generalização, e na geração randômica dos padrões de treinamento $\{\vec{Y}^q, \vec{Z}_{alvo}^q\}$, foram utilizados programas e subprogramas em FORTRAN 77. A máquina utilizada para processar todos os programas foi um processador Intel Core 2 Duo processor T5500 1,66 GHz, 667 MHz FSB, 2MB L2 cache, com 160 GB HDD e 4GB DDR2.

Todas as redes neurais especialistas e a rede neural que integra o combinador (vide Figura 33) da máquina de comitê, descrita no modelo da Figura 32, foram treinados usando o algoritmo da retropropagação do erro, o *Backpropagation*, baseado no método do gradiente decrescente, que minimiza o erro médio quadrático E a cada época completada, de acordo com a otimização de Levenberg-Marquardt, usando o Neural Network Toolbox do software MATLAB®. Foram usadas funções de ativação (ou de transferência ou restritiva) tipo tangente hiperbólica, para as camadas ocultas e a linear para a camada de saída.

Foram treinadas e depois testadas quanto à generalização, de forma independente, RNA especialistas com uma ou duas camadas ocultas, de acordo com os modelos de MLP mostrados nas Figuras 30 e 29, respectivamente. Todas as RNA especialistas e o MLP (vide Figura 33) do combinador da máquina de comitê, de acordo com o modelo mostrado na Figura 32, apresentam funções de treinamento (integrantes do Neural Network Toolbox do software MATLAB®), funções de ativação e taxas de aprendizado idênticas.

Para cada valor do albedo de espalhamento simples, $\omega = 0,1; 0,5$ ou $0,9$, e para cada valor do limitante superior para o erro médio quadrático E , $tol = 10^{-3}$ ou 10^{-4} , foram construídas, na ordem abaixo descrita, três tabelas cujos conteúdos são os seguintes:

1. Tabela contendo para cada uma das RNA especialistas: o número de ordem na máquina de comitê, RNA_i , $i = 1, 2, \dots, M$, o número de camadas ocultas N_{HL} , o número de neurônios por camada oculta N_H , o tempo de treinamento, o número de épocas completadas durante a fase de

treinamento, e as estimativas obtidas para cada parâmetro radiativo na fase de generalização;

2. Tabela contendo os erros relativos percentuais das estimativas de cada parâmetro radiativo: $ER_{\tau_0}, ER_{\omega}, ER_{\rho_1}$ e ER_{ρ_2} , para cada rede neural especialista, com grifo para as seis com os melhores resultados;
3. Tabela contendo as estimativas obtidas para os parâmetros radiativos e os seus respectivos erros relativos percentuais, para cada modelo de máquina de comitê usada, com resultados anteriores ($M = 10$) e posteriores à seleção das seis ($M = 6$) melhores RNA especialistas. Na Tabela 2 do item 4.2, estão condensados os resultados considerando todos os casos.

Serão mostrados, por meio de tabelas no item 4.2, os melhores resultados obtidos na fase de generalização para as estimativas do vetor solução $\vec{Z} = \{\tau_0, \omega, \rho_1, \rho_2\}$, conseguidas com as RNA especialistas (com uma ou duas camadas ocultas) individualmente e com as máquinas de comitês com modelos de acordo com as Figuras 31 e 32, com $M = 10$ ou $M = 6$, e $tol = 10^{-3}$ ou 10^{-4} , além dos casos em que as RNA especialistas usadas isoladamente apresentaram os piores resultados na fase de generalização e cujos progressos foram bastante relevantes após essas RNA especialistas comporem máquinas de comitê.

As figuras mostradas são dos gráficos das *curvas de performance*, obtidas durante os testes de validação (que utilizam 20 % dos dados de entrada), que foram geradas durante a fase de treinamento usando o Neural Network Toolbox do software MATLAB[®], para a rede neural especialista (com uma ou duas camadas ocultas), que apresentou o menor número de épocas realizadas no seu treinamento, considerando $\omega = 0,1; 0,5$ ou $0,9$ e $tol = 10^{-3}$ ou 10^{-4} , e para a rede neural do combinador de saída da máquina de comitê, de acordo com o modelo mostrado na Figura 32, com $M = 10$ ou $M = 6$.

Os resultados da generalização e treinamento dos demais casos serão mostrados através de tabelas e gráficos, no Apêndice A, da seguinte forma e ordem:

1. Primeiro para as redes neurais especialistas isoladamente com uma ou duas camadas ocultas para valores do albedo de espalhamento simples $\omega = 0,1; 0,5$ ou $0,9$ e $tol = 10^{-3}$ ou 10^{-4} .
2. Segundo para as máquinas de comitê efetuando apenas a média aritmética no combinador e constituídas por RNA especialistas com uma e/ou duas camadas ocultas para valores do albedo de espalhamento simples $\omega = 0,1; 0,5$ ou $0,9$ e $tol = 10^{-3}$ ou 10^{-4} .

3. Por último, e da mesma forma, para máquinas de comitê que combinam a média aritmética mais uma rede neural tipo MLP (vide Figura 33) no combinador de saída.

4.2. Comparativo dos resultados obtidos com as RNA especialistas e com as máquinas de comitê.

Na Tabela 2 é mostrado um comparativo geral entre os resultados obtidos, com base nos erros relativos das estimativas obtidas para os parâmetros radiativos, considerando RNA especialistas com uma ou duas camadas ocultas, $\omega = 0,1; 0,5$ ou $0,9$; $tol = 10^{-3}$ ou 10^{-4} e modelos de máquinas de comitê conforme representados nas Figuras 31 e 32, com $M = 10$ ou $M = 6$. São mostrados, na primeira coluna da Tabela 2, as médias dos erros relativos das estimativas obtidas para $\vec{Z} = \{\tau_0, \omega, \rho_1, \rho_2\}$, com as RNA especialistas.

Tabela 2: Comparativo entre os erros relativos das estimativas obtidas para $\vec{Z} = \{\tau_0, \omega, \rho_1, \rho_2\}$ com redes neurais e máquinas de comitê para todos os casos testes com $\omega = 0,1; 0,5$ ou $0,9$; $tol = 10^{-3}$ e/ou 10^{-4} .

$\overline{ER} (\%)$ (RNA)	N_{HL}	tol	ω	$ER (\%)$ MA M=10	$ER (\%)$ MA M=6	$ER (\%)$ MLP M=10	$ER (\%)$ MLP M=6
$\tau_0 = 1,18$	01	10^{-3}	0,1	0,01	0,83	0,14	0,03
$\omega = 12,46$	01	10^{-3}	0,1	1,2	2,6	0,9	0,4
$\rho_1 = 3,28$	01	10^{-3}	0,1	1,68	0,06	0,06	0,06
$\rho_2 = 4,77$	01	10^{-3}	0,1	4,78	4,3	1,86	1,16
$\tau_0 = 2,08$	01	10^{-3}	0,5	2,08	2,42	0,97	0,86
$\omega = 2,86$	01	10^{-3}	0,5	1,68	0,66	0,16	0,06
$\rho_1 = 4,31$	01	10^{-3}	0,5	3,06	2,14	1,32	0,52
$\rho_2 = 2,07$	01	10^{-3}	0,5	2,08	2,1	0,44	0,04
$\tau_0 = 11,58$	01	10^{-3}	0,9	11,58	9,31	6,04	4,95
$\omega = 1,69$	01	10^{-3}	0,9	0,28	0,44	0,27	0,53
$\rho_1 = 6,66$	01	10^{-3}	0,9	2,56	2,0	0,38	1,64
$\rho_2 = 4,53$	01	10^{-3}	0,9	2,44	0,04	1,22	1,12
$\tau_0 = 1,14$	02	10^{-3}	0,1	0,49	0,23	0,43	0,09
$\omega = 15,95$	02	10^{-3}	0,1	14,8	4,8	3,6	2,2
$\rho_1 = 3,69$	02	10^{-3}	0,1	2,28	3,44	2,52	1,56
$\rho_2 = 4,07$	02	10^{-3}	0,1	3,96	3,96	3,46	1,74

Tabela 2 (Cont.): Comparativo entre os erros relativos das estimativas obtidas para $\vec{Z} = \{\tau_0, \omega, \rho_1, \rho_2\}$ com redes neurais e máquinas de comitê para todos os casos testes com $\omega = 0,1; 0,5$ ou $0,9$; $tol = 10^{-3}$ e/ou 10^{-4} .

$\overline{ER}$ (%) (RNA)	N_{HL}	tol	ω	ER (%) MA M=10	ER (%) MA M=6	ER (%) MLP M=10	ER (%) MLP M=6
$\tau_0 = 2,07$	02	10^{-3}	0,5	1,57	1,48	0,18	0,06
$\omega = 2,86$	02	10^{-3}	0,5	0,34	0,48	0,60	0,1
$\rho_1 = 3,93$	02	10^{-3}	0,5	2,96	0,26	0,68	0,22
$\rho_2 = 1,96$	02	10^{-3}	0,5	1,66	1,52	0,88	0,08
$\tau_0 = 7,15$	02	10^{-3}	0,9	7,10	6,89	3,37	3,9
$\omega = 1,92$	02	10^{-3}	0,9	0,66	1,61	0,23	0,86
$\rho_1 = 4,54$	02	10^{-3}	0,9	2,32	2,28	0,54	0,06
$\rho_2 = 1,81$	02	10^{-3}	0,9	0,38	0,3	0,7	0,62
$\tau_0 = 1,56$	01	10^{-4}	0,1	0,39	0,35	0,1	0,01
$\omega = 15,73$	01	10^{-4}	0,1	1,9	0,8	1,1	0,8
$\rho_1 = 3,25$	01	10^{-4}	0,1	1,9	2,74	2,76	0,38
$\rho_2 = 3,96$	01	10^{-4}	0,1	3,96	4,78	1,68	1,66
$\tau_0 = 2,19$	01	10^{-4}	0,5	2,19	1,22	0,42	0,28
$\omega = 3,28$	01	10^{-4}	0,5	0,14	1,54	0,6	0,16
$\rho_1 = 3,93$	01	10^{-4}	0,5	2,34	0,28	0,9	0,22
$\rho_2 = 2,14$	01	10^{-4}	0,5	2,14	1,44	0,62	0,04
$\tau_0 = 11,60$	01	10^{-4}	0,9	11,60	2,46	6,87	2,3
$\omega = 1,82$	01	10^{-4}	0,9	0,06	0,48	1,59	1,03
$\rho_1 = 5,25$	01	10^{-4}	0,9	2,18	0,48	1,34	0,88
$\rho_2 = 3,49$	01	10^{-4}	0,9	3,3	0,26	1,7	0,5
$\tau_0 = 1,63$	02	10^{-4}	0,1	1,13	0,84	1,17	0,75
$\omega = 8,95$	02	10^{-4}	0,1	8,8	4,9	7,5	4,8
$\rho_1 = 4,42$	02	10^{-4}	0,1	4,42	4,78	4,12	0,14
$\rho_2 = 3,87$	02	10^{-4}	0,1	3,86	4,66	3,16	0,36
$\tau_0 = 1,33$	02	10^{-4}	0,5	1,33	1,17	1,17	0,75
$\omega = 1,59$	02	10^{-4}	0,5	1,38	1,42	0,52	0,16
$\rho_1 = 1,9$	02	10^{-4}	0,5	1,06	0,32	0,56	0,12
$\rho_2 = 1,82$	02	10^{-4}	0,5	1,84	1,04	0,6	0,3
$\tau_0 = 3,53$	02	10^{-4}	0,9	3,53	2,08	3,31	1,43
$\omega = 1,03$	02	10^{-4}	0,9	0,94	0,49	0,98	0,03
$\rho_1 = 2,1$	02	10^{-4}	0,9	0,90	0,16	0,14	0,06
$\rho_2 = 1,43$	02	10^{-4}	0,9	0,24	0,1	0,32	0,02

Tabela 2 (Cont.): Comparativo entre os erros relativos das estimativas obtidas para $\vec{Z} = \{\tau_0, \omega, \rho_1, \rho_2\}$ com redes neurais e máquinas de comitê para todos os casos testes com $\omega = 0,1; 0,5$ ou $0,9$; $tol = 10^{-3}$ e/ou 10^{-4} .

$\overline{ER} (%)$ (RNA)	N_{HL}	tol	ω	$ER (%)$ MA M=10	$ER (%)$ MA M=6	$ER (%)$ MLP M=10	$ER (%)$ MLP M=6
$\tau_0 = 0,65$	CM	10^{-3}	0,1	-	0,5		0,25
$\omega = 2,23$	CM	10^{-3}	0,1	-	2,0		0,9
$\rho_1 = 1,85$	CM	10^{-3}	0,1	-	1,8		0,26
$\rho_2 = 2,29$	CM	10^{-3}	0,1	-	3,34		2,24
$\tau_0 = 0,49$	CM	10^{-3}	0,5	-	0,41		0,18
$\omega = 1,0$	CM	10^{-3}	0,5	-	0,64		0,9
$\rho_1 = 0,63$	CM	10^{-3}	0,5	-	0,92		0,84
$\rho_2 = 0,6$	CM	10^{-3}	0,5	-	1,04		0,6
$\tau_0 = 1,25$	CM	10^{-3}	0,9	-	2,08		1,14
$\omega = 0,36$	CM	10^{-3}	0,9	-	0,49		1,29
$\rho_1 = 0,93$	CM	10^{-3}	0,9	-	0,16		0,16
$\rho_2 = 0,73$	CM	10^{-3}	0,9	-	0,1		0,12

Nota: Aqui CM significa Máquina de Comitê Misto.

Verifica-se na Tabela 2 que os melhores resultados obtidos para as estimativas de $\vec{Z}$ com as RNA especialistas, foram aquelas usando $\omega = 0,5$ (médio espalhamento). Isto se deve, possivelmente, ao equilíbrio entre a absorção e o espalhamento da radiação que se propaga através do meio. Nas Tabelas 3 e 4 são mostrados os melhores resultados obtidos para $\omega = 0,5$, nos testes de generalização, pelas RNA especialistas.

Tabela 3: Estimativas obtidas para o caso $N_{HL} = 2$; $\omega = 0,5$ e $tol = 10^{-4}$.

RNA	N_H	tempo	épocas	τ_0	ω	ρ_1	ρ_2
1	8	0:05:48	154	0,9627	0,4796	0,4729	0,5066
2	10	0:08:08	172	0,9861	0,5024	0,5029	0,5318
3	12	01:05:55	757	0,9987	0,4898	0,5113	0,5085
4	14	00:27:52	299	0,9988	0,5029	0,4837	0,5095
5	16	00:17:11	147	0,9900	0,4886	0,4967	0,5128
6	18	00:17:58	112	0,9830	0,4891	0,5047	0,5116
7	20	00:20:34	98	0,9927	0,4924	0,4954	0,5007
8	22	00:17:32	68	0,9935	0,4927	0,4941	0,5058
9	24	00:19:56	61	0,9789	0,4973	0,4830	0,5018
10	26	00:19:48	51	0,9829	0,4961	0,5020	0,5029

Tabela 4: Erros relativos percentuais por rede neural especialista para o caso $N_{HL} = 2$; $\omega = 0,5$ e $tol = 10^{-4}$.

RNA	$ER\tau_0$ (%)	$ER\omega$ (%)	$ER\rho_1$ (%)	$ER\rho_2$ (%)
1	3,73	4,08	5,42	1,32
2	1,39	0,48	0,58	6,36
3	0,13	2,04	2,26	1,7
4	0,12	0,58	3,26	1,9
5	1,0	2,28	0,66	2,56
6	1,7	2,18	0,94	2,32
7	0,73	1,52	0,92	0,14
8	0,65	1,46	1,18	1,16
9	2,11	0,54	3,4	0,36
10	1,71	0,78	0,4	0,4
ER	1,33	1,59	1,9	1,82

Nota: As RNA destacadas em negrito são as seis com os melhores resultados.

Observa-se na Tabela 4, que os resultados mais precisos foram os obtidos pelas RNA 3, 5, 6, 7, 8 e 10, todas com erros relativos menores ou iguais a 2,56%. A rede neural especialista que apresentou o menor tempo de processamento foi a RNA 1 com 5 minutos e 48 segundos e a que apresentou o menor número de épocas foi a RNA 10, com 51 épocas, que consideramos como a de melhor desempenho na fase de treinamento, tendo em vista que precisou de um menor número de épocas para o seu aprendizado.

Na Figura 34 é mostrada a curva de performance da RNA 10 acima mencionada.

Observa-se na Figura 34, que o processo de treinamento é interrompido com 51 épocas, ponto onde a curva de validação atinge o ponto ótimo de treinamento, quando o gradiente do erro médio quadrático E em relação ao peso sináptico W atinge um valor muito próximo de zero (ponto de sela), passando de negativo para positivo, caracterizando a minimização da função erro médio quadrático E em torno da tolerância requerida de 10^{-4} .

Ainda na Tabela 2, verifica-se que as RNA especialistas que apresentaram os piores resultados foram aquelas para os casos extremos $\omega = 0,1$ e $\omega = 0,9$ (exceto para o caso $\omega = 0,9$ com duas camadas ocultas e $tol = 10^{-4}$, que apresentou bons resultados juntamente com aqueles obtidos para $\omega = 0,5$), ou seja, para baixo espalhamento e alta absorção, e alto espalhamento e baixa absorção, respectivamente. Ambos apresentaram médias dos erros relativos $\overline{ER}_\omega$, $\overline{ER}_{\tau_0}$ e $\overline{ER}_{\rho_1}$ maiores que 4,0 % e menores que 16%. Sendo que $8,95 \% \leq \overline{ER}_\omega \leq 15,95 \%$ ocorreu para o caso $\omega = 0,1$, o que já era esperado devido a alta

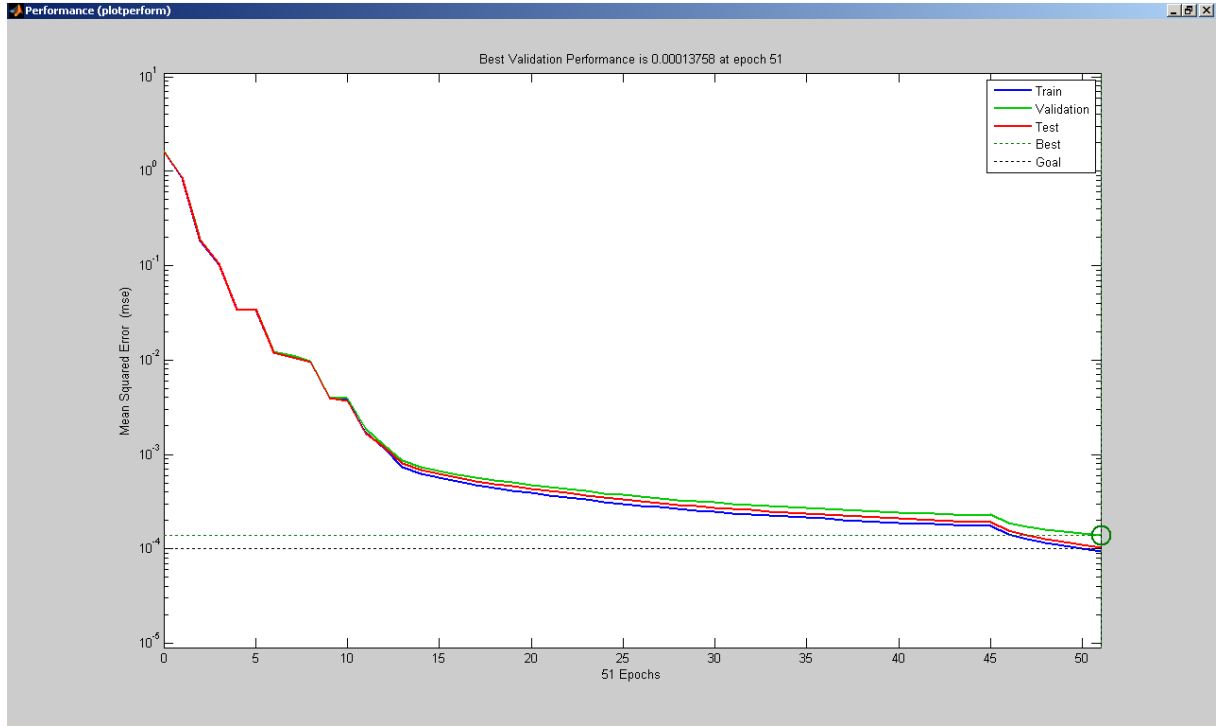


Figura 34: Curva de performance da RNA nº 10 com $N_{HL} = 2$, $N_H = 26$, $\omega = 0,5$ e $tol = 10^{-4}$.

absorção e baixo espalhamento imposto pelo problema físico, e $7,15\% \leq \overline{ER}_{\tau_0} \leq 11,60\%$ e $4,54\% \leq \overline{ER}_{\rho_1} \leq 6,6\%$ ocorreram para o caso $\omega = 0,9$.

Observa-se na Tabela 02 que as estimativas para $\vec{Z}$, obtidas usando as máquinas de comitê tanto com o modelo da Figura 31 quanto com o modelo da Figura 32, foram no geral melhores do que as médias daquelas obtidas com as RNA especialistas isoladamente. Notadamente as estimativas obtidas pelas máquinas de comitê, com ambos os modelos após a seleção das seis melhores RNA especialistas na fase de generalização, podem ser consideradas quase na totalidade boas e melhores do que as médias daquelas obtidas com as RNA especialistas isoladamente. Ainda no geral, verifica-se na Tabela 2 que os resultados conseguidos usando o modelo de máquina de comitê da Figura 32, ou seja, com um MLP integrado ao combinador, foram melhores do que os resultados obtidos com o modelo da Figura 31, tanto para $M = 10$ como para $M = 6$.

Na Tabela 2, verifica-se que os maiores progressos em termos de melhorias de resultados das estimativas para $\vec{Z}$, obtidos pelas máquinas de comitê em relação às médias das estimativas obtidas pelas RNA especialistas isoladamente, foram para os seguintes casos:

1. Para o caso $\omega = 0,1$; $N_{HL} = 1$ e $tol = 10^{-3}$, a média obtida dos erros relativos percentuais com as RNA especialistas para o parâmetro

radiativo ω foi $\overline{ER}_\omega = 12,46$ %, e com uma máquina de comitê, de acordo com o modelo da Figura 32 e $M = 6$, o erro relativo percentual obtido foi $ER_\omega = 0,4$ %, representando um progresso em termos de melhora na capacidade de generalização de 3.115 %.

2. Para o caso $\omega = 0,9$; $N_{HL} = 1$ e $tol = 10^{-3}$, as médias obtidas dos erros relativos percentuais com as RNA especialistas para os parâmetros radiativos τ_0 e ρ_1 , foram respectivamente $\overline{ER}_{\tau_0} = 11,58$ % e $\overline{ER}_{\rho_1} = 6,66$ %, e com uma máquina de comitê, de acordo com o modelo da Figura 32, os erros relativos percentuais obtidos foram $ER_{\tau_0} = 4,95$ %, com $M = 6$ e $ER_{\rho_1} = 0,38$ % com $M = 10$, representando um progresso em termos de melhora na capacidade de generalização de 233,94 % e 1.752,63 %, respectivamente.
3. Para o caso $\omega = 0,1$; $N_{HL} = 2$ e $tol = 10^{-3}$, a média obtida dos erros relativos percentuais com as RNA especialistas para o parâmetro radiativo ω foi $\overline{ER}_\omega = 15,95$ % e com uma máquina de comitê, de acordo com o modelo da Figura 32 e $M = 6$, o erro relativo percentual obtido foi $ER_\omega = 2,2$ %, representando um progresso em termos de melhora na capacidade de generalização de 725 %.
4. Para o caso $\omega = 0,9$; $N_{HL} = 2$ e $tol = 10^{-3}$, a média obtida dos erros relativos percentuais com as RNA especialistas para o parâmetro radiativo τ_0 foi $\overline{ER}_{\tau_0} = 7,15$ % e com uma máquina de comitê, de acordo com o modelo da Figura 32 e $M = 6$, o erro relativo percentual obtido foi $ER_{\tau_0} = 3,9$ %, representando um progresso em termos de melhora na capacidade de generalização de 183,33 % ;
5. Para o caso $\omega = 0,1$; $N_{HL} = 1$ e $tol = 10^{-4}$, a média obtida dos erros relativos percentuais com as RNA especialistas para o parâmetro radiativo ω foi $\overline{ER}_\omega = 15,73$ % e com uma máquina de comitê, de acordo com o modelo da Figura 32 e $M = 6$, o erro relativo percentual obtido foi $ER_\omega = 0,8$ %, representando um progresso em termos de melhora na capacidade de generalização de 1.966,25 % ;
6. Para o caso $\omega = 0,9$; $N_{HL} = 1$ e $tol = 10^{-4}$, as médias obtidas dos erros relativos percentuais com as RNA especialistas para os parâmetros radiativos τ_0 e ρ_1 , foram $\overline{ER}_{\tau_0} = 11,60$ % e $\overline{ER}_{\rho_1} = 5,25$ %

respectivamente, e com uma máquina de comitê, de acordo com o modelo da Figura 32, com $M = 6$, o erro relativo percentual obtido foi $ER_{\tau_0} = 2,3 \%$, para o parâmetro radiativo τ_0 , e com uma máquina de comitê, de acordo com o modelo da Figura 31 e $M = 6$, o erro relativo percentual obtido foi $ER_{\rho_1} = 0,48 \%$ para o parâmetro radiativo ρ_1 , representando um progresso em termos de melhora na capacidade de generalização de 504,35 % e 1.093,75 %, respectivamente.

7. Para o caso $\omega = 0,1$; $N_{HL} = 2$ e $tol = 10^{-4}$, a média obtida dos erros relativos percentuais com as RNA especialistas para o parâmetro radiativo ω foi $\overline{ER}_{\omega} = 8,95 \%$ e com uma máquina de comitê, de acordo com o modelo da Figura 32 e $M = 6$, o erro relativo percentual obtido foi $ER_{\omega} = 4,8 \%$, representando um progresso em termos de melhora na capacidade de generalização de 186,46 %;

Serão mostrados de forma mais detalhada, por meio de tabelas, os resultados das generalizações das RNA especialistas, considerando apenas os casos 1 e 2, relatados no parágrafo anterior, que foram os que apresentaram os maiores progressos nos resultados de generalização após a utilização de máquinas de comitê, dentre os sete casos supra citados. Para os demais casos os resultados serão mostrados no Apêndice A.

Inicialmente nas Tabelas 5 e 6 são mostrados os resultados obtidos nos testes de generalização com as RNA especialistas para o caso 1: $\omega = 0,1$; $N_{HL} = 1$ e $tol = 10^{-3}$.

Tabela 5: Estimativas obtidas para o caso $N_{HL} = 1$; $\omega = 0,1$ e $tol = 10^{-3}$.

RNA	N_H	tempo	épocas	τ_0	ω	ρ_1	ρ_2
1	8	00:22:00	464	0,9508	0,1124	0,5540	0,5304
2	10	02:57:30	3000	0,9964	0,1124	0,5075	0,5294
3	12	00:03:36	90	1,0050	0,1181	0,5061	0,5203
4	14	00:34:11	307	1,0235	0,1103	0,4732	0,5296
5	16	00:30:26	186	1,0060	0,0853	0,5200	0,5126
6	18	00:15:18	70	1,000	0,0819	0,5111	0,5326
7	20	00:26:12	103	1,0178	0,0848	0,4871	0,5160
8	22	00:19:41	72	1,0053	0,0979	0,4997	0,5312
9	24	00:04:56	51	1,0008	0,0940	0,5142	0,5099
10	26	00:17:47	53	0,9929	0,1153	0,5113	0,5265

Tabela 6: Erros relativos percentuais por rede neural especialista para $N_{HL} = 1$; $\omega = 0,1$ e $tol = 10^{-3}$.

RNA	ER_{τ_0} (%)	ER_{ω} (%)	ER_{ρ_1} (%)	ER_{ρ_2} (%)
1	4,92	12,4	10,8	6,08
2	0,36	12,4	1,5	5,88
3	0,5	18,1	1,22	4,06
4	2,35	10,3	5,36	5,92
5	0,6	14,7	4	2,52
6	0	18,1	2,22	6,52
7	1,78	15,2	2,58	3,2
8	0,53	2,1	0,06	6,24
9	0,08	6,0	2,84	1,98
10	0,71	15,3	2,26	5,3
ER	1,18	12,46	3,28	4,77

Nota: As RNA destacadas em negrito são as seis com os melhores resultados.

Pode-se verificar nas Tabelas 2 e 6 que o erro relativo ER_{ω} resultante do cálculo da estimativa para o parâmetro radiativo ω passou de 18,1 %, obtido com as RNA 3 e 6, para 0,4 %, obtido com a máquina de comitê de acordo com o modelo da Figura 32 com $M = 6$ e $tol = 10^{-3}$, representando um progresso em termos de melhora na capacidade de generalização de 4.525 %. Para o parâmetro radiativo τ_0 o erro relativo ER_{τ_0} passou de 4,92 %, obtido com a RNA 1, para 0,01 %, obtido com a máquina de comitê, de acordo com o modelo da Figura 31, com $M = 10$ e $tol = 10^{-3}$, representando um progresso em termos de melhora na capacidade de generalização de 49.200 %. Para o parâmetro radiativo ρ_1 o erro relativo ER_{ρ_1} passou de 10,8 %, obtido com RNA 1, para 0,06 %, obtido tanto com a máquina de comitê de acordo com o modelo da Figura 31, com $M = 6$ e $tol = 10^{-3}$, como com a máquina de comitê de acordo com o modelo da Figura 32, com $M = 6$ e $tol = 10^{-3}$, representando um progresso em termos de melhora na capacidade de generalização de 18.000 %. Para o parâmetro radiativo ρ_2 o erro relativo ER_{ρ_2} passou de 6,52 %, obtido com RNA 6, para 1,16 %, obtido com a máquina de comitê de acordo com o modelo da Figura 32, com $M = 6$ e $tol = 10^{-3}$, representando um progresso em termos de melhora na capacidade de generalização de 562,07 % (vide as estimativas obtidas com as máquinas de comitê nas Tabelas 27 e 39 no Apêndice A) .

Na Tabela 5 verifica-se que a rede neural especialista que apresentou o menor tempo de processamento foi a RNA 3 com 3 minutos e 36 segundos e a que apresentou o menor número de épocas foi a RNA 9, com 51 épocas, que consideramos como a de melhor

desempenho na fase de treinamento, tendo em vista que precisou de um menor número de épocas para o seu aprendizado.

Na Figura 35 é mostrada a curva de performance gerada na fase de treinamento ou aprendizado da RNA 9.

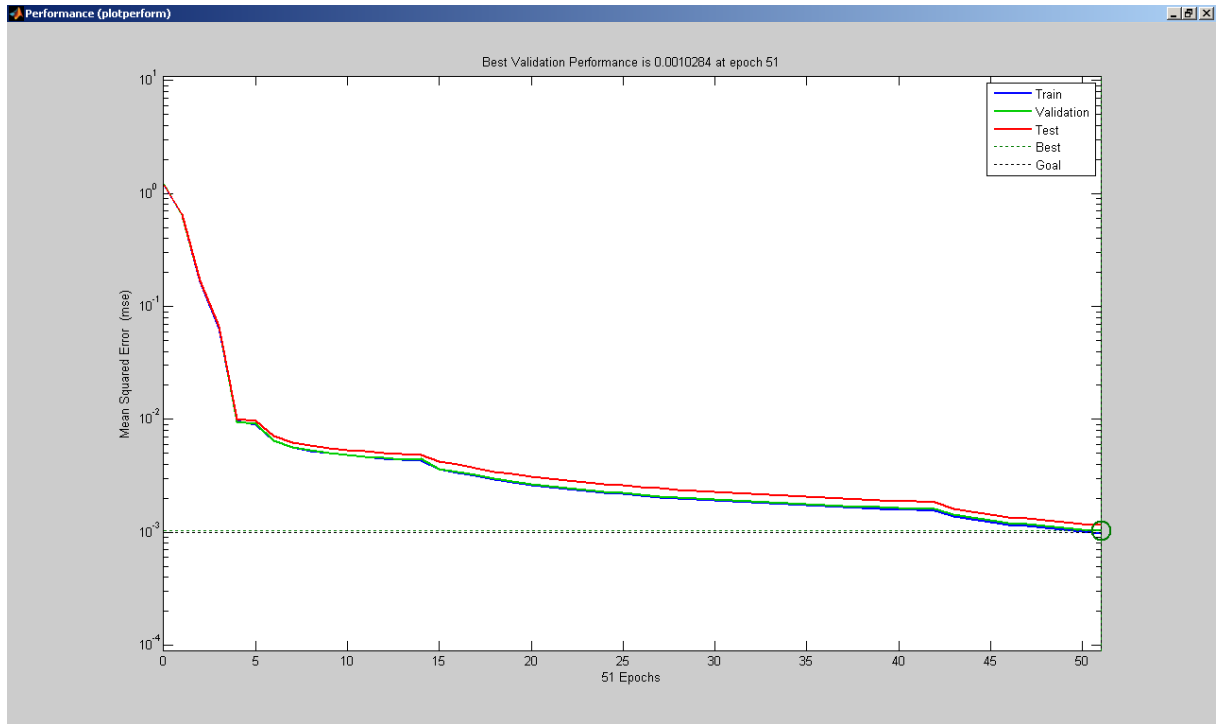


Figura 35: Curva de performance da RNA nº 9 com $N_{HL} = 1$, $N_H = 24$, $\omega = 0,1$ e $tol = 10^{-3}$.

Analogamente ao relatado para o caso mostrado na Figura 34 anterior, observa-se na Figura 35, que o processo de treinamento é interrompido com 51 épocas, ponto onde a curva na cor verde, que representa o teste de validação, atinge o ponto ótimo de treinamento, que ocorre quando o gradiente do erro médio quadrático E em relação ao peso sináptico W atinge um valor muito próximo de zero (ponto de sela), passando de negativo para positivo, caracterizando a minimização da função erro médio quadrático E em torno da tolerância requerida de 10^{-3} .

Na Figura 36 é mostrada a curva de performance para o MLP do combinador da máquina de comitê, de acordo com o modelo da Figura 32, com $M = 6$ e $tol = 10^{-3}$, mostrando que o erro médio quadrático E atingiu em apenas 13 épocas a tolerância requerida de 10^{-3} , na fase de treinamento, acarretando os bons resultados obtidos na fase de generalização, após a seleção das melhores RNA especialistas, com um pequeno esforço computacional extra.

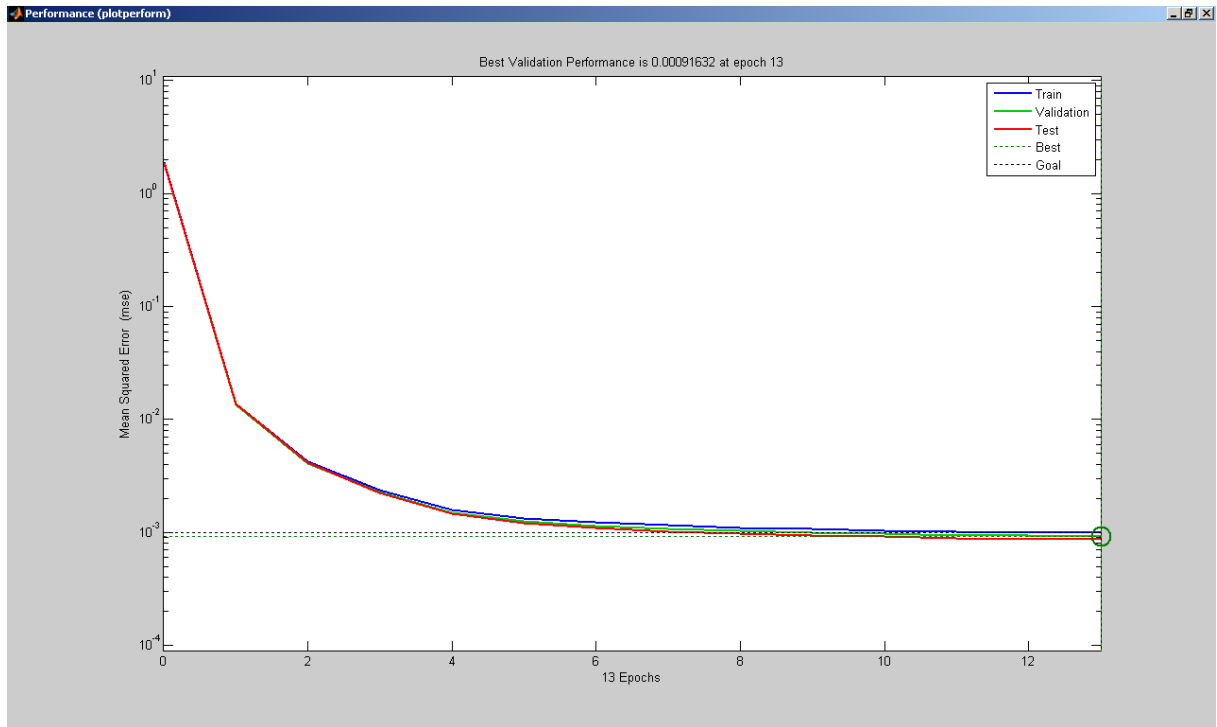


Figura 36: Curva de performance do MLP do combinador da máquina de comitê, para o caso $N_{HL} = 1$; $\omega = 0,1$; $tol = 10^{-3}$ e $M = 6$.

Nas Tabelas 7 e 8 são mostrados os resultados obtidos nos testes de generalização com as RNA especialistas para o caso 2: $\omega = 0,9$; $N_{HL} = 1$ e $tol = 10^{-3}$.

Pode-se verificar nas Tabelas 2 e 8 que o erro relativo ER_{τ_0} resultante do cálculo da estimativa para o parâmetro radiativo τ_0 passou de 18,1 %, obtido com a RNA 1, para 4,95 %, obtido com a máquina de comitê de acordo com o modelo da Figura 32, com $M = 6$ e $tol = 10^{-3}$, representando um progresso em termos de melhora na capacidade de generalização de 365,66%. Para o parâmetro radiativo ω o erro relativo ER_{ω} passou de 2,74 %, obtido com a RNA 8, para 0,27 %, obtido com a máquina de comitê, de acordo com o modelo da Figura 32, com $M = 10$ e $tol = 10^{-3}$, representando um progresso em termos de melhora na capacidade de generalização de 1.014,81 %. Para o parâmetro radiativo ρ_1 o erro relativo ER_{ρ_1} passou de 16,08 %, obtido com a RNA 1, para 0,38 %, obtido com a máquina de comitê de acordo com o modelo da Figura 32, com $M = 10$ e $tol = 10^{-3}$, representando um progresso em termos de melhora na capacidade de generalização de 4.231,58 %. Para o parâmetro radiativo ρ_2 o erro relativo ER_{ρ_2} passou de 9,14 %, obtido com a RNA 3, para 0,04 %, obtido com a máquina de comitê, de acordo com o modelo da Figura 31, com $M = 6$ e $tol = 10^{-3}$, representando um progresso em termos de melhora na capacidade de generalização de

Tabela 7: Estimativas obtidas para o caso $N_{HL} = 1$; $\omega = 0,9$ e $tol = 10^{-3}$.

RNA	N_H	tempo	épocas	τ_0	ω	ρ_1	ρ_2
1	8	00:22:00	464	0,8187	0,9199	0,4196	0,5153
2	10	02:57:30	3000	0,8556	0,9134	0,5576	0,5317
3	12	00:03:36	90	0,8405	0,8878	0,4360	0,5457
4	14	00:34:11	307	0,9112	0,9116	0,5137	0,5262
5	16	00:30:26	186	0,9103	0,8928	0,4807	0,5251
6	18	00:15:18	70	0,8805	0,8760	0,4733	0,4925
7	20	00:26:12	103	0,8861	0,8798	0,5186	0,5301
8	22	00:19:41	72	0,8752	0,9247	0,5112	0,4799
9	24	00:04:56	51	0,9278	0,9020	0,4597	0,4878
10	26	00:17:47	53	0,9366	0,9167	0,5013	0,4875

Tabela 8: Erros relativos percentuais por rede neural especialista para $N_{HL} = 1$; $\omega = 0,9$ e $tol = 10^{-3}$.

RNA	$ER\tau_0$ (%)	$ER\omega$ (%)	$ER\rho_1$ (%)	$ER\rho_2$ (%)
1	18,1	2,21	16,08	3,06
2	14,4	1,49	11,52	6,34
3	16,0	1,36	12,8	9,14
4	8,88	1,29	2,74	5,24
5	8,97	0,8	3,86	5,02
6	12,0	2,67	5,34	1,5
7	11,4	2,24	3,72	6,02
8	12,5	2,74	2,24	4,02
9	7,22	0,22	8,06	2,44
10	6,34	1,86	0,26	2,5
ER	11,58	1,69	6,66	4,53

Nota: As RNA destacadas em negrito são as seis com os melhores resultados.

22.850% (vide as estimativas obtidas com as máquinas de comitê nas Tabelas 29 e 41 no Apêndice A).

Verifica-se também nas Tabelas 5 e 7, que os resultados obtidos para o tempo de processamento e para o número de épocas realizadas, são idênticos para os casos $\omega = 0,1$; $N_{HL} = 1$ e $tol = 10^{-3}$ e $\omega = 0,9$; $N_{HL} = 1$ e $tol = 10^{-3}$. Isto se deve ao fato de que houve um único treinamento das RNA especialistas para o caso $N_{HL} = 1$ e $tol = 10^{-3}$, considerando os valores para o albedo de espalhamento, $\omega = 0,1$; 0,5 ou 0,9, havendo diferenças apenas quanto aos resultados obtidos na fase de generalização, para cada RNA especialista e para cada valor de ω . Todavia, o treinamento do MLP do combinador da máquina de comitê, de acordo com a Figura 32 e após a seleção das seis RNA especialistas com os melhores resultados na generalização, $M = 6$, foi realizado de forma independente para cada valor de ω . Portanto, a

RNA especialista com melhor desempenho para o caso $\omega = 0,9$; $N_{HL} = 1$ e $tol = 10^{-3}$ é exatamente a mesma que para o caso $\omega = 0,1$; $N_{HL} = 1$ e $tol = 10^{-3}$, e a sua curva de performance está mostrada na Figura 35.

Na Figura 37 é mostrada a curva de performance gerada na fase de treinamento do MLP da máquina de comitê, para $\omega = 0,9$; $N_{HL} = 1$ e $tol = 10^{-3}$, com $M = 6$.

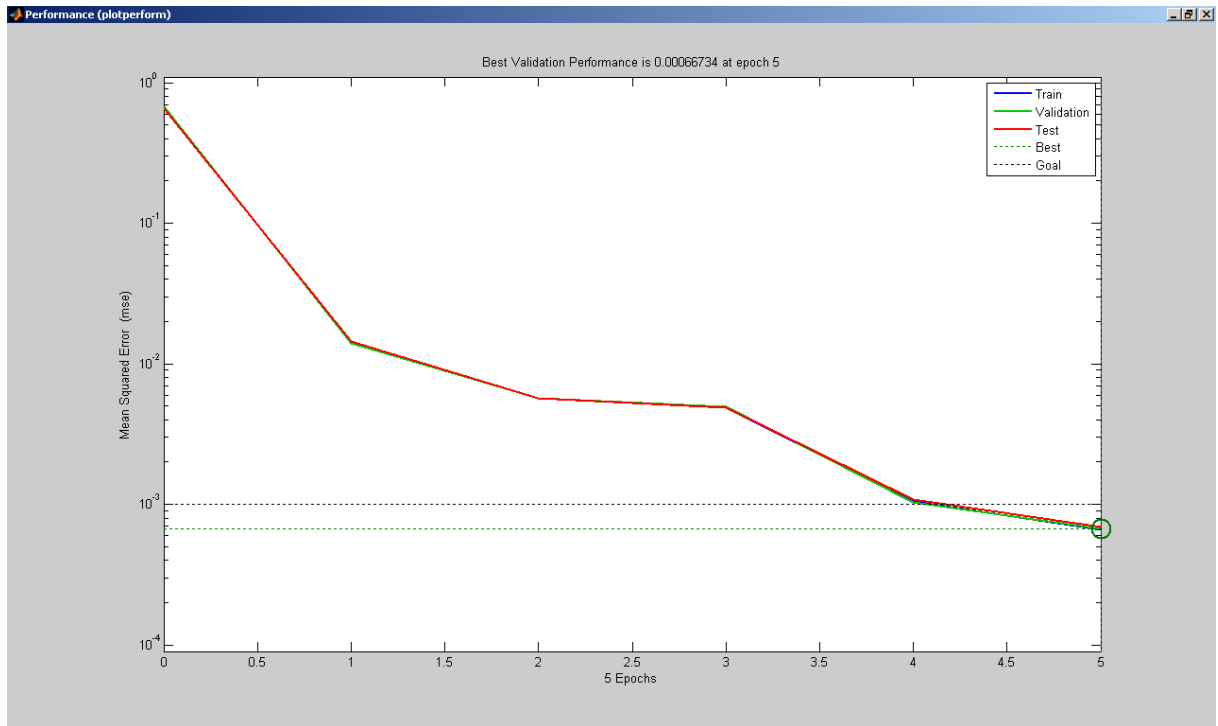


Figura 37: Curva de performance do MLP do combinador da máquina de comitê, para o caso $N_{HL} = 1$; $\omega = 0,9$; $tol = 10^{-3}$ e $M = 6$.

Verifica-se na Figura 37, que o número necessário de épocas para o treinamento do MLP do combinador da máquina de comitê para o caso $N_{HL} = 1$; $\omega = 0,9$; $tol = 10^{-3}$ e $M = 6$, foi de apenas cinco, mostrando uma convergência bem mais rápida do que a da RNA especialista com melhor desempenho, que foi de 51 épocas.

Foram realizados também testes com máquinas de comitê compostas por RNA especialistas com uma e/ou duas camadas ocultas, aqui designadas de máquinas de comitê misto, considerando-se apenas as seis RNA especialistas ($M = 6$) com os melhores resultados dentre todos os testes realizados para cada valor do albedo de espalhamento ($\omega = 0,1$; 0,5 ou 0,9) e para cada valor da tolerância requerida ($tol = 10^{-3}$ ou 10^{-4}). Estes resultados são mostrados na Tabela 2, e são identificados pela sigla CM. Verifica-se que os resultados obtidos pelas máquinas de comitê misto foram bons para todos os casos estudados, ficando os

erros relativos referentes ao cálculo das estimativas para cada parâmetro radiativo abaixo de 5,0 %.

Nas Figuras 38, 39 e 40, são mostradas as curvas de performance obtidas com as máquinas de comitê misto formados por RNA especialistas com uma e/ou duas camadas ocultas, para $\omega = 0,1$; 0,5 e 0,9 respectivamente.

Na Figura 38, pode-se observar que o resultado do treinamento da máquina de comitê misto para o caso $\omega = 0,1$ e $tol = 10^{-3}$ com $M = 6$ foi bom, tendo em vista a rápida convergência do erro médio quadrático E para a tolerância requerida, inclusive atingindo um

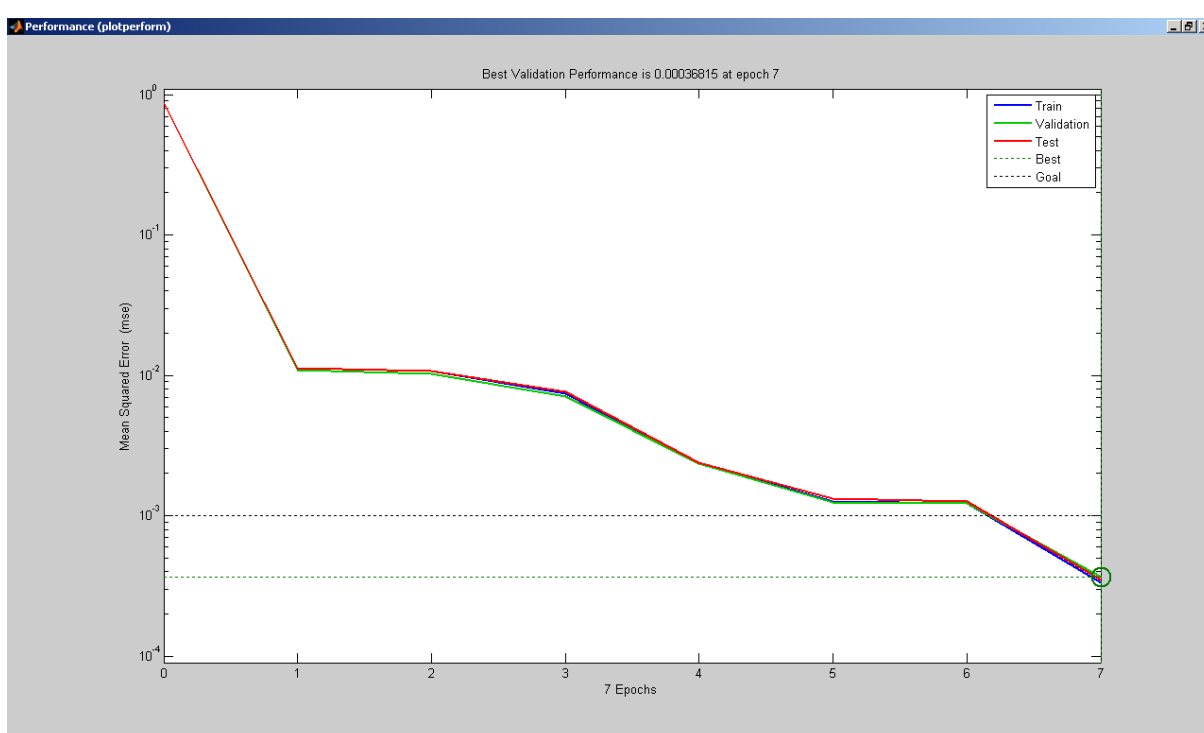


Figura 38: Curva de performance do MLP do combinador da máquina de comitê misto, com $N_{HL} = 1$ ou 2; $M = 6$; $\omega = 0,1$ e $tol = 10^{-3}$.

valor menor que 10^{-3} com apenas 7 épocas completadas, o que justifica os baixos valores para os erros relativos obtidos nas estimativas dos parâmetros radiativos com esta máquina de comitê.

De modo análogo podemos comprovar através das Figuras 39 e 40 os bons desempenhos das máquinas de comitê misto na avaliação dos parâmetros radiativos para $\omega = 0,5$ e 0,9 respectivamente.

Nas Figuras 39 e 40, observa-se que o erro médio quadrático E convergiu para valor abaixo de 10^{-3} com apenas 5 e 6 épocas respectivamente, o que acarretou nos bons resultados

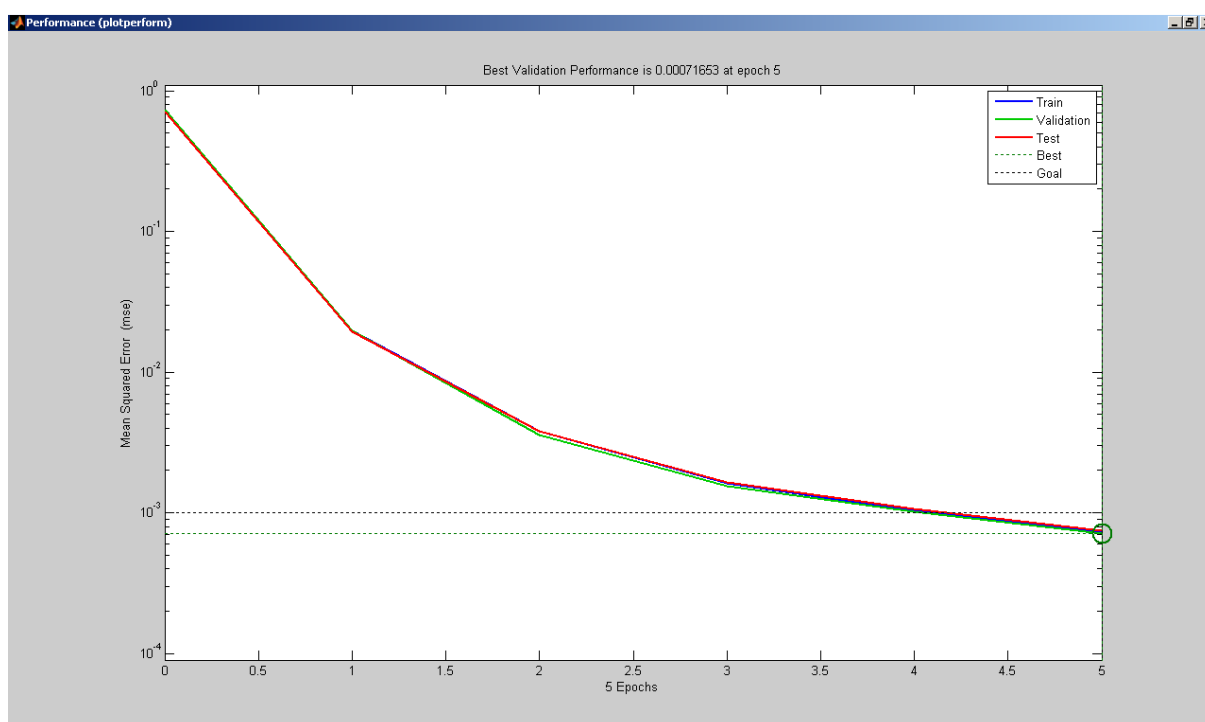


Figura 39: Curva de performance do MLP do combinador da máquina de comitê misto, com $N_{HL} = 1$ ou 2; $M = 6$; $\omega = 0,5$ e $tol=10^{-3}$.

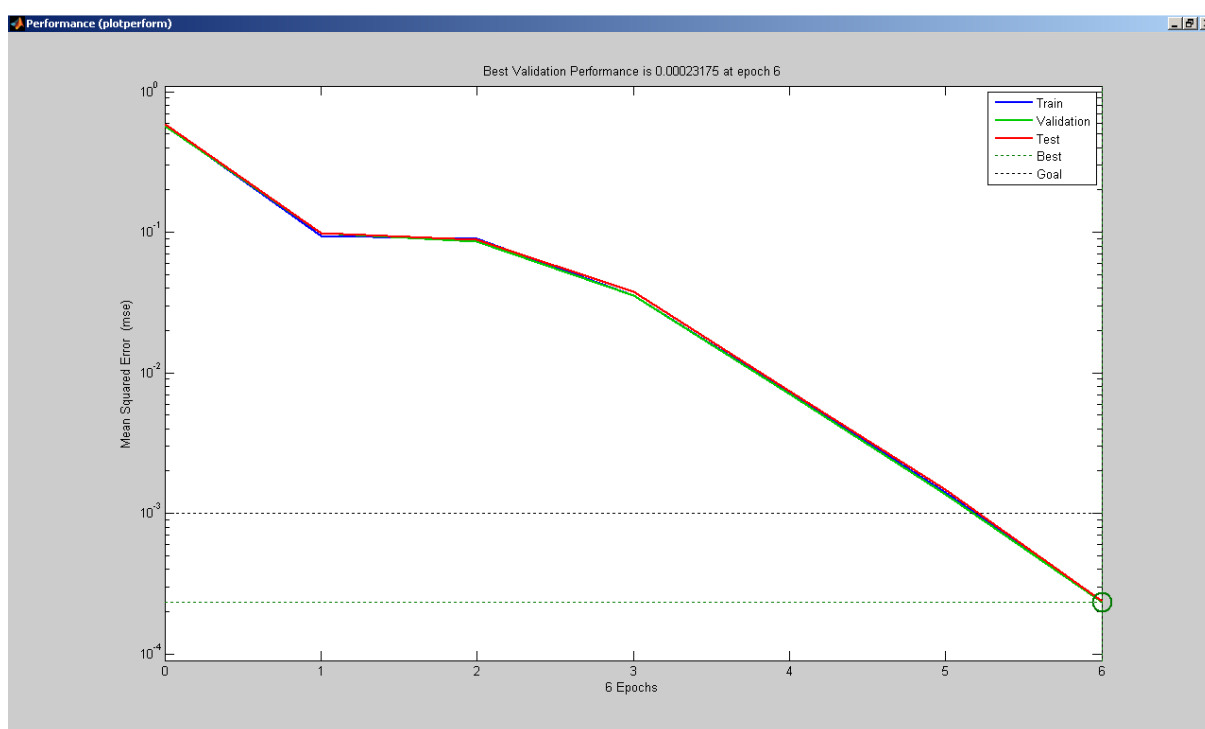


Figura 40: Curva de performance do MLP do combinador da máquina de comitê misto, com $N_{HL} = 1$ ou 2; $M = 6$; $\omega = 0,9$ e $tol=10^{-3}$.

obtidos para as estimativas dos parâmetros radiativos como se pode constatar na Tabela 2.

Fundamentando-se nos resultados apresentados na Tabela 2 e nas demais tabelas, além das curvas de performance - que utilizam 60% dos dados de entrada fornecidos para a finalidade de treinamento das RNA especialistas e do MLP das máquinas de comitê, 20% desses dados para testes de generalização e os outros 20% restantes para os testes de validação - geradas na fase de treinamento das RNA especialistas e da RNA do combinador de saída da máquina de comitê (conforme o modelo da Figura 32), todos mostrados neste capítulo e no Apêndice A, pode-se verificar que os resultados obtidos com ambos os modelos de máquina de comitê para todos os casos aqui tratados, foram, em geral, melhores do que os obtidos com as RNA especialistas isoladamente. Pode-se ainda constatar, que no geral, os melhores resultados obtidos foram aqueles através do uso da máquina de comitê de acordo com o modelo da Figura 32 - que é uma proposta inovadora deste trabalho de dissertação - que apresentou na maioria dos casos resultados significativamente superiores aos obtidos com as RNA especialistas isoladamente e daqueles obtidos com a máquina de comitê do tipo estático clássica, que efetua apenas a média aritmética de conjunto no seu combinador de saída (conforme o modelo da Figura 31), notadamente para os casos com $M = 6$, ou seja, após a seleção das seis melhores RNA especialistas na fase de generalização para compor as máquinas de comitê, atendendo aos objetivos específicos deste trabalho.

Verifica-se ainda, que os resultados obtidos com as máquinas de comitê para o caso $M = 6$, mesmo compostas por RNA especialistas com apenas uma camada oculta - que apresentaram resultados ruins para valores de $\omega = 0,1$ ou $0,9$ - foram em geral muito bons e melhores do que os obtidos pelas RNA especialistas isoladamente tendo apresentado melhoras significativas. Portanto, conclui-se que para a solução do problema inverso em transferência radiativa usando máquinas de comitê, a utilização de RNA especialistas com apenas uma camada oculta, $N_{HL} = 1$, e $tol = 10^{-3}$, é o suficiente.

5 CONCLUSÃO

Neste trabalho de dissertação, foram empregadas redes neurais artificiais do tipo MLP, aqui chamadas de RNA especialistas, e máquinas de comitê de redes neurais artificiais, para a solução do problema inverso em transferência radiativa, que consistiu na obtenção do vetor dos parâmetros radiativos $\vec{Z} = \{\tau_0, \omega, \rho_1, \rho_2\}$ a partir dos dados experimentais simulados Y_l , minimizando o erro médio quadrático E , a cada época completada, decorrente do confronto entre as saídas obtidas e as saídas desejadas, usando o algoritmo de retro-propagação do erro para reajustar os pesos e bias das camadas ocultas das RNA especialistas, pesos estes que constituem a “memória” das RNA especialistas e, por conseguinte, das máquinas de comitê por elas compostas, caracterizando assim o processo de aprendizagem e a sua resultante capacidade de generalização e classificação de padrões.

Os resultados aqui obtidos e descritos no capítulo 4 e no Apêndice A, demonstram que os objetivos, tanto do ponto de vista geral quanto do ponto de vista específico deste trabalho foram alcançados, ou seja, a solução do problema inverso em transferência radiativa usando RNA e máquinas de comitê de RNA, e a comprovação de que as máquinas de comitê, no geral, produzem estimativas na fase de generalização tão boas ou melhores do que as RNA especialistas usadas isoladamente, com um pequeno custo computacional extra.

Cabe aqui destacar que o modelo de máquina de comitê representado na Figura 32, cuja configuração difere da máquina de comitê clássica do tipo estática com média aritmética de conjunto (vide Figura 31), pelo fato de agregar ao combinador de saída da máquina de comitê uma rede neural do tipo MLP e cujo vetor de entrada é a média aritmética das saídas das redes neurais especialistas, não consta, nas referências bibliográficas pesquisadas, ter sido testado antes, e que em parte das simulações realizadas neste trabalho apresentou resultados tão bons ou melhores do que os obtidos com o modelo clássico, demonstrando ter sido uma experiência positiva e que vale ser mais explorada em trabalhos posteriores (vide Tabela 2).

Comprova-se através da análise dos valores dos erros relativos percentuais devido às estimativas obtidas para cada um dos parâmetros radiativos pela aplicação de máquinas de comitê, contidos na Tabela 2 do item 4.2, que o critério adotado de se selecionar as redes neurais especialistas com os melhores resultados na fase de generalização ($M = 6$), para compor a média aritmética das saídas das redes neurais especialistas no combinador da máquina de comitê, para ambos os modelos de máquina (vide Figuras 31 e 32), efetuando então novos testes de generalização, acarretou de modo geral, na obtenção de resultados mais precisos para as estimativas dos parâmetros radiativos do que aqueles obtidos sem efetuar a

Conclusão

seleção das RNA especialistas com os melhores resultados nos testes de generalização ($M = 10$).

Finalmente, deixamos algumas sugestões para trabalhos futuros aplicando máquinas de comitê de redes neurais artificiais isoladamente ou em conjunto com outros métodos na linha de IA.

Para trabalhos futuros, sugere-se:

- A aplicação de uma máquina de comitê de redes neurais artificiais do tipo estática usando o método de reforço com diferentes sinais de entrada para estimar os parâmetros radiativos;
- A aplicação de uma máquina de comitê de redes neurais artificiais do tipo dinâmica usando mistura de especialistas e uma rede neural de passagem como estrutura combinadora não-linear para estimar os parâmetros radiativos;
- A hibridização de máquinas de comitê de redes neurais artificiais tipo estática com métodos de otimização global, como LM, SA, GA, PSO, GEO, Monte Carlo, etc. para estimar os parâmetros radiativos;
- A hibridização de máquinas de comitê de redes neurais artificiais tipo dinâmicas com métodos de otimização global, como LM, SA, GA, PSO, GEO, Monte Carlo, etc. para estimar os parâmetros radiativos;
- Resolver o problema inverso em transferência radiativa usando valores exatos para os parâmetros radiativos diferentes dos aqui utilizados, como por exemplo, $\tau_0 \geq 2,0$.
- Para valores de $\tau_0 \geq 2,0$, possivelmente serão obtidos valores piores para as estimativas dos parâmetros radiativos, pois uma maior espessura ótica acarretará intensidades de radiações calculadas mais fracas e, por conseguinte dados computacionais simulados mais fracos.
- Sugere-se, então à aplicação da análise de sensibilidade para melhorar as estimativas dos parâmetros radiativos quando se usar valores exatos de τ_0 maior ou igual a 2,0.

REFERÊNCIAS BIBLIOGRÁFICAS

ABRAMOWITZ, M.; STEGUN, I. A. (Ed.). *Handbook of mathematical functions, with formulas, graphs and mathematical tables*. New York: Dover, 1965. 1046 p.

BARICHELO, L. B.; SIEWERT, C. E.; GARCIA, R. D. M. On inverse boundary condition problems in radiative transfer. *J. of Quant. Spectrosc. and Radiat. Transfer*, v. 57, n. 3, p.405-410, 1997.

BATTAGLIA, A.; MANTOVANI, S. Forward Monte Carlo computations of fully polarized microwave radiation in non-isotropic media. *J. Quant. Spectrosc. Radiat. Transfer*, v.95, p. 285-308, 2005.

BIONDI NETO, L. et al. *Tópicos especiais em problemas inversos, redes neurais e lógica nebulosa..* SILVA NETO, A. J; BECCENERI, J. C. (Ed.). *Técnicas de Inteligência Computacional Inspiradas na Natureza: Aplicações em Problemas Inversos em Transferência Radiativa* Cap. 7 p.83 – 102, SBMAC, 2009.

BOILLEREAUX, L.; CADET, C.; LE BAIL. A thermal properties estimation during thawing via real-time neural network learning. *Journal of Food Engineering*, v. 57, n.1, p. 17-23, mar. 2003.

BOKAR, J. C. The estimation of spatially varying albedo and optical thickness in a radiating slab using artificial neural networks. *International Communications in Heat and Mass Transfer*, vol. 26, n.3, p. 359-367, 1999.

BROOMHEAD, D. S.; LOWE, D. G. *Radial basis functions: multivariable interpolation and adaptive networks*. Malvern: Royal Signals Radar Est., 1988. (RSRE memo 4148).

BULGARELLI, B.; DOYLE J. P. Comparison between numerical models for radiative transfer simulation in the atmosphere-ocean system. *J. Quant. Spectrosc. Radiat. Transfer*, v.86, p. 315-334, 2004.

ÇAYAN, F. N.; SELÇUK, N. The method of lines solution of discrete ordinates method for non-grey media. *J. Quant. Spectrosc. Radiat. Transfer*, v.104, p. 228-237, 2007.

OLIVEIRA, R. C. et al. Aplicação da técnica de máquina de comitê de redes neurais artificiais para a solução de problemas inversos em transferência radiativa. In: ENCONTRO DE MODELAGEM COMPUTACIONAL, 11., 2008, Volta Redonda. [Anais...]. Volta Redonda: UERJ; UFF, 2008.

CARCENAC, M. A modular neural network for super-resolution of human faces. *Appl. Intell.*, DOI 10.1007/s10489-007-0109-4, 2007.

CARPENTER, G. A.; GROSSBERG, S. The art of adaptive pattern recognition by a selforganizing neural network. *Computer*, v.21, n.3, p. 77-88, 1988.

CASTTRALL, C. *Retrieval of the columnar aerosol phase function and single-scattering albedo from sky radiance over the ocean: measurement of African dust*. 2001. 75 f. Tese (Doutorado) – Universidade do Sul da Flórida, Flórida, 2001.

Referências bibliográficas

CASTRO, C. L. ; BRAGA, A. P.; ANDRADE, A. V. Aplicação de um modelo ensemble de redes neurais artificiais para previsão de séries temporais não estacionárias. In: CONGRESSO DA SOCIEDADE BRASILEIRA DE COMPUTAÇÃO, 25., 2005, São Leopoldo do Sul. [Anais...]. São Leopoldo do Sul: UNISINOS, 2005.

CHALHOUB, E. S.; CAMPOS VELHO, H. F. de. Multispectral reconstruction of bioluminescence term in natural waters. *Appl. Num. Math.*, v. 47, p. 365-376, 2003.

_____. Discrete-ordinates solution for radiative transfer problems. *J. Quant. Spectrosc. Radiat. Transfer*, v. 76, p.193-206, 2003.

_____. Discrete-ordinates solution for uncoupled multi-wavelength radiative transfer problems. *J. Quant. Spectrosc. Radiat. Transfer*, v. 92, p. 335-349, 2005.

_____ et al. A comparison of radiances generated by selected methods of solving the radiative-transfer equation. *Transport Theor. Stat. Phys.*, v. 32, p. 473-503, 2003.

_____; CAMPOS VELHO, H. F.; SILVA NETO, A. J. A comparison of the one-dimensional radiative transfer problem solutions obtained with the monte carlo method and three variations of the discrete ordinates method. In: INTERNATIONAL CONGRESS OF MECHANICAL ENGINEERING – COBEM, 19., 2007, Brasília. *Proceedings...* Brasília: Associação Brasileira de Engenharia e Ciências Mecânicas, 2007.

_____; SILVA NETO, A. J.; SOEIRO, F. J. C. P. Estimation of optical thickness and single scattering albedo with artificial neural network and a Monte Carlo method. In: INTERNATIONAL CONGRESS OF MECHANICAL ENGINEERING – COBEM, 19., 2007, Brasília. *Proceedings...* Brasília: Associação Brasileira de Engenharia e Ciências Mecânicas, 2007.

CHANDRASEKHAR, S. On the radiative equilibrium of a stellar atmosphere II. *Astrophys. J.*, v.100, p. 76-86, 1944.

_____. *Radiative transfer*. London: Oxford University Press, 1950.

CHEN, Y.; LIOU, K. N. A Monte Carlo method for 3D thermal infrared radiative transfer. *J. Quant. Spectrosc. Radiat. Transfer*, v.101, p. 166-178, 2006.

CHO, Sung-Bae; WON, Hong-Hee. Cancer classification using ensemble of neural networks with multiple significant gene subsets. *Appl. Intell.*, v. 26, n. 3, p. 243–250, jun. 2007.

COELHO, G. P. *Geração, seleção e combinação de componentes para ensembles de redes neurais aplicadas a problemas de classificação*. 2006. 137 f. Dissertação (Mestrado em Engenharia de Computação) – Faculdade de Engenharia Elétrica de Computação, Universidade Estadual de Campinas, Campinas, 2006.

COWAN, J. D.; DAVID, S. H. *The artificial intelligence debate: false starts, real foundations*. Cambridge: MIT Press, 1989.

DUDA, R.; HART P. E.; STORK, D. G. *Pattern classification*. 2nd ed. New York: Wiley-Interscience, 2000. 654 p.

Referências bibliográficas

- FAUSE, T.; ISAKA, H.; GUILLEMET, B. Neural network retrieval of cloud parameters of inhomogeneous and fractional clouds. *Remote Sens. Environ.*, v. 77, p. 123-128, 2001.
- FERNANDES, A. M. et al. Development of a neural network committee machines for automatic forest fire detection using lidar. *Pattern Recognition*, v. 37, n.10, p. 2039-2047, oct. 2004.
- FLORES, E. O. et al. Solução de um problema inverso em transferência radiativa com programação matemática e métodos de otimização global. In: CONGRESSO BRASILEIRO DE ENGENHARIA MECÂNICA, 16., 2001, Uberlândia. *Anais...* Uberlândia: Associação Brasileira de Engenharia e Ciências Mecânicas, 2001.
- FRANÇA, Francis H. R. et al. Inverse design of thermal systems with dominant radiative transfer. *Advances in Heat Transfer*, v. 36, p. 1-110, 2002.
- FRANÇA, Francis H. R.; HOWELL, John R. Transient inverse design of radiative enclosures for thermal processing of materials. *Inverse Problems in Science and Engineering*, v. 14, n. 4, p. 423-436, Jun. 2006.
- FRANÇA, Francis H. R. et al. Application of inverse analysis to illumination design. *Inverse Problems in Science and Engineering*, v. 17, n. 6, p. 737-753, sept. 2009.
- FUKSHANSKY-KAZARINOVA, et al. Solution of the inverse problem of radiative transfer on the basis of measured internal fluxes. *Journal of Quantitative Spectroscopy and Radiative Transfer*, v. 59, n. 1/2, p. 77– 89, 1998.
- GALSKI, R. L. et al. Application of a GEO+SA hybrid optimization algorithm to the solution of an inverse radiative transfer problem. *Inverse Problems in Science and Engineering, Special Issue: Selected papers of IPDO 2007*, v. 17, n. 3, p. 321-334, 2009.
- GIL, M. C. C. et al. Aplicação de um método de comitês de redes neuro-fuzzy combinado com redes neurais artificiais em cascata para a solução de problemas inversos em transferência radiativa. In: ENCONTRO DE MODELAGEM COMPUTACIONAL, 11., 2008, Volta Redonda. [Anais...]. Volta Redonda: UERJ ; UFF, 2008.
- GOODBAND, J. H.; HAAS, O. C. L.; MILLS, J. A. A mixture of experts committee machine to design compensators for intensity modulated radiation therapy. *Pattern Recognition*, v. 39, n. 9, Elsevier B.V., Convetry, UK, p. 1704 – 1714, sep. 2006.
- GORDON, H. R.; ZHANG, T. Columnar aerosol properties over oceans by combining surface and aircraft measurements: simulations, *Applied Optics* v.34, p. 5552–5555. 1995.
- GROSSBERG, S. *Neural networks and natural intelligence*. Cambridge: MIT Press, 1988. 637 p.
- HANAN, N. P. Enhanced two-layer radiative transfer scheme for a land surface model with a discontinuous upper canopy. *Agr. Forest Meteorol.*, v. 109, p. 265-281, 2001.
- HANSEN, J. E.; TRAVIS, L. D. Light scattering in planetary atmospheres. *Space Sci. Rev.*, v. 16, p. 527-610, 1974.

Referências bibliográficas

- HAYKIN, S. *Redes neurais: princípios e prática*. 2. ed. Porto Alegre: Bookman, 2001. 900 p. Tradução de Paulo Martins Engel.
- HEBB, D. O. *The organizations of behavior*. New York: J.Wiley, 1949. 335 p.
- HOPFIELD, J. J. Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences, USA*, p. 2554-2558, 1982.
- HORNBURG, J. E. *Máquinas de comitê aplicadas à filtragem de spam*. 2007. 7 f. Monografia (Bacharelado em Ciências da Computação) – Universidade Federal de Santa Catarina, Florianópolis, 2007.
- HIRAYAMA, V. et al. *Classificador de qualidade de álcool combustível e poder calorífico de gás GLP*. Biblioteca Digital de Teses e Dissertações da USP. Disponível em: <http://www.teses.usp.br/teses/disponiveis/3/3140/tde-27022005-164712/>. Acesso em: 8 jun. 2004.
- JENSEN, K. A. et al. On various modeling approaches to radiative heat transfer in pool fires. *Combust. Flame*, v. 148, p. 263-279, 2007.
- JUNIOR, Ludwing. O.; COSTA, E. Montgomery. Meira. *Redes neurais: fundamentos e aplicações com programas em C*. Rio de Janeiro: Ciência Moderna, 2007. 125 p.
- KAUATI, A. T.; ROBERTY, N. C.; SILVA NETO, A. J. A source detector methodology for the construction and solution of the one-dimensional inverse transport equation. In: INTERNATIONAL CONFERENCE ON INVERSE PROBLEMS IN ENGINEERING: THEORY AND PRACTICE, 3., 1999, Port Ludlow. *Proceedings...* Port Ludlow, 1999.
- KAUATI, A. T.; ROBERTY, N. C.; SILVA NETO, A. J. Uso da técnica fonte-detector em tomografia com espalhamento anisotrópico. In: CONGRESSO BRASILEIRO DE ENGENHARIA BIOMÉDICA, 17., 2000, Florianópolis. *Anais...* Florianópolis, 2000.
- _____; SILVA NETO, A. J.; ROBERTY, N. C. A source-detector methodology for the construction and solution of the one-dimensional inverse transport equation. *Inverse Prob. Eng.*, v. 9, p. 45-66, 2001.
- KIM, H. K.; CHARETTE, A. A sensitivity function-based conjugate gradient method for optical tomography with the frequency domain equation of radiative transfer. *J. Quant. Spectrosc. Radiat. Transfer*, v. 104, p. 24-39, 2007.
- KNUPP, D. C.; SILVA NETO, A. J.; SACCO, W. F. Radiative properties estimation with the particle collision algorithm based on a sensitivity analysis. *High Temperatures.-High Pressures*, v.38, n.2, p. 137-151, 2009.
- KOHONEN, T. An adaptive associative memory principle. *IEEE Transactions on Computers*, v. C-24, n. 4, p. 444-445, 1974.
- KOHONEN, T. Self-organized formation of topologically correct feature maps. *Biological Cybernetics*, v.43, p. 59-69, 1982.

Referências bibliográficas

KOVACS, Z. L. *Redes neurais artificiais: fundamentos e aplicações*. São Paulo: Editora Acadêmica, 1996. 164 p.

_____; JOHNSON, T. L.; SAX, D. S. Estimation of the distribution of conduction velocities in peripheral nerves. *Comput. Biol. Med.*, v. 9, p. 281-293, 1979.

LEATHERS, R. A.; McCORMICK, N. J. Absorption and scattering coefficient estimation with asymptotic apparent optical properties. In: ACKLESON, S.G.; FROUIN, R. (Ed.). *Ocean Optics XIII* - SPIE - The International Society for Optical Engineering, v. 2963, p. 21-25, 1997.

LENOBLE, J. (Ed.). *Standard procedures to compute atmospheric radiative transfer in a scattering atmosphere*. Boulder: National Center for Atmospheric Research, 1977. 135 p.

LENT, R. *Cem bilhões de neurônios: conceitos fundamentais de neurociência*. São Paulo: Atheneu, 2002. 698 p.

LI, H. Y.; ÖZISIK, M. N. Inverse radiation problem for simultaneous estimation of temperature profile and surface reflectivity. *Journal of Thermophysics and Heat Transfer* v.7, n.1, p. 88 – 93, 1993.

LI, H. Y.; YANG, C. Y. A genetic algorithm for inverse radiation problems. *Int. J. Heat MassTransfer*, v. 40, n. 7, p. 1545–1549, 1997.

LITTLE, W.; SHAW, M. L. G. A statistical theory of short and long term memory. *Behavioral, Biology*, v. 14, p. 115-133, 1975.

MAFFIONE, R. A.; DANA, D. R. Instruments and methods for measuring the backscattering coefficient of ocean waters: applied optics. *Applied Optics*, v. 36, n. 2, p. 6057–6067, 1997.

MAFFIONE, R. A.; VOSS, K. J.; HONEY, R. C. Measurement of the spectral absorption coefficient in the ocean with an isotropic point source. *Applied Optics*, v. 32, p. 3273–3279, 1993.

MAGALHÃES, R. M. *Processamento paralelo no projeto de máquinas de comitê*. 2007. 98 f. Dissertação (Mestrado em Engenharia Elétrica) –Programa de Pós –Graduação em Engenharia Elétrica, Universidade Federal do Rio Grande do Norte, Natal, 2007.

MAGOULAS, G.; DOUNIAS, G. Computational intelligence in medicine and biology. *Appl. Intell.*, v. 27, p. 189–192, 2007.

MAGLOGIANNIS, I.; ZAFIROPOULOS, E.; ANAGNOSTOPOULOS, I. An intelligent system for automated breast câncer diagnosis and prognosis using SVM based classifiers. *Appl. Intell.*, DOI 10.1007/s10489-007-0073-z, 2007.

MALSBURG, C. von der. Self-organization of orientation selective cells in the striate cortex. *Kybernetik*, v. 14, 1973.

Referências bibliográficas

- MAURENTE, A.; VIELMO, H. A.; FRANÇA, F. H. R. A Monte Carlo implementation to solve radiation heat transfer in non-uniform media with spectrally dependent properties. *J. Quant. Spectrosc. Radiat. Transfer*, v. 108, p. 295-307, 2007.
- McCORMICK, N. J. Inverse radiative transfer problems: a review. *Nucl. Sci. Eng.*, v. 112, p. 185-198, 1992.
- McCORMICK, N. J. Ocean source estimation using irradiance measurements at only one depth. In: JAFFE, J. S. (Ed.). *Ocean Optics XII SPIE -- The International Society for Optical Engineering*, v. 2258, p. 711-722, 1994.
- McCORMICK, N. J. Analytical transport theory applications in optical oceanography. *Annals Nuclear Energy*, v. 23, 1996, p. 381-395, 1996.
- McCORMICK, N. J. *Inverse problems: methods and applications*. In: CONGRESSO NACIONAL DEMATEMÁTICA APLICADA E COMPUTACIONAL, 23., 2000, Santos. *Proceedings...* Santos: Sociedade Brasileira de Matemática Aplicada e Computacional, 2000.
- _____; KASKAS, A. Isotropic spherical source analysis for ocean optics. *Applied Optics*, v. 39, p. 4902-4910, 2000.
- McCULLOCH, W. S.; PITTS, W. A logical calculus of ideas immanent in nervous Activity. *Bulletin of Mathematical Biophysics*, v. 5, p.115-133, 1943.
- MINKU, F. L.; LUDERMIR, T. B. Clustering and co-evolution to construct neural network ensembles: an experimental study. *Neural Networks*, v. 21, n. 9, p. 1363-1379, nov. 2008.
- MINSKY, M.; PAPERT, S. *Perceptrons*. Cambridge: MIT Press, 1969. 258 p.
- MONTERO, C. R. F.; ROBERTY, N. C.; SILVA NETO, A. J. Absorption coefficient estimation in heterogeneous media using a domain partition consistent with divergent beams. *Inverse Problems in Engineering*, v. 9, p. 587-617, 2001.
- _____; ROBERTY, N. C.; SILVA NETO, A. J. Reconstruction of a combination of absorption and scattering coefficients with a discrete ordinates method consistent with the source-detector system. *Inverse Prob. Eng.*, v. 12, p. 103-121, 2004.
- NILSSON, N.J. *Learning machines: foundations of trainable pattern-classifying systems*. New York: McGraw-Hill, 1965.137 p.
- ÖZISIK, M. N. *Radiative transfer and interactions with conduction and convection*. New York: John Wiley & Sons, 1973. 575 p.
- PAULA, M. V. S. Oliveira et al. User authentication based on human typing pattern with artificial neural networks and support vector machine. In: ENCONTRO NACIONAL DE INTELIGÊNCIA ARTIFICIAL, 2005, São Leopoldo. *Anais...* . São Leopoldo, 2005.
- PINHEIRO, R. P. F. *Problemas inversos em transferência radiativa para estimativa de propriedades e fontes*. 2004. 119 f. Tese (Doutorado em Modelagem Computacional) – Instituto Politécnico, Universidade do Estado do Rio de Janeiro, Nova Friburgo, 2004.

Referências bibliográficas

- POSTYLYAKOV, O. V. Radiative transfer model MCC++ with evaluation of weighting functions in spherical atmosphere for use in retrieval algorithms. *Adv. Space Res.*, v. 34, p. 721-726, 2004.
- PUMA VILLANUEVA, W. J.; ZUBEN, F. J. V. *Comitê de máquinas em predição de séries temporais*. EA072 – Inteligência Artificial em Aplicações Industriais. Campinas: UNICAMP. Departamento de Engenharia de Computação e Automação Industrial, 2008. 8 p.
- ROGERS, T. T.; McCLELLAND, J. L. *Semantic cognition: A parallel distributed processing approach*. Cambridge: MIT Press, 2004. 425 p.
- ROSENBLATT, F. The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review*, v. 65, n. 6, p.386-408, 1958.
- RUMELHART, D. E.; HINTON, G. E.; WILLIAM, R. J. Learning internal representation by error propagation in parallel distributed processing. In: RUMELHART, D. E.; McCLELLAND, J. L. (Ed.). *Parallel distributed processing: explorations in the microstructure of cognition*, vol. 1: foundations Cambridge: MIT Press, 1986.
- SANCHEZ, R.; McCORMICK, N. J.; YI, H. C. Iterative inverse radiative transfer method to estimate optical thickness and surface albedo. *Transport Theor. Stat.*, v. 19, p. 357-385. 1990.
- SEDANO, S. T. C. Estimção simultânea da temperatura do gás e a emissividade num forno utilizando análise inversa em radiação térmica. In: CONGRESSO IBERO AMERICANO DE ENGENHARIA MECÂNICA, 8., 2007, Cusco. [Anais...]. Cusco, 2007.
- SILVA, L. A. *Classificação de nódulos mamográficos utilizando um comitê de redes neurais artificiais*. 2005. 94 f. Dissertação (Mestrado em Engenharia) – Escola Politécnica da Universidade de São Paulo, São Paulo, 2005.
- _____; HERNANDEZ, D. M.; RANGAYAN, R. M. Classification of breast masses using a committee machine of artificial neural networks (in press). *Journal of Electronic Imaging*, Springfield, v.17, n.1, 013017:1-10, jan.–mar. 2008.
- SILVA NETO, A. J. Problemas inversos: aplicações em engenharia e medicina. In: MOURA NETO, F. D. (Ed.). *Cápsulas de Ciência*. Rio de Janeiro: Quartet, 2007. p. 57-86.
- _____; BECCENERI, J. C. (Eds.) Técnicas de inteligência computacional inspiradas na Natureza: aplicação em problemas inversos em transferência radiativa. *Notas em Matemática Aplicada e Computacional*, Sociedade Brasileira de Matemática Aplicada e Computacional, v. 41, 2009, 122 p.
- SILVA NETO, A. J. ; CAMPOS VELHO, H. F. *Problemas inversos em transferência radiativa: uma formulação implícita*. In: SILVA NETO, A. J.; BECCENERI, J. C. (Eds.). *Técnicas de inteligência computacional inspiradas na natureza: aplicação em problemas inversos em transferência radiativa*. São Carlos, SP: SBMAC, 2009. Cap. 3, p. 33-4

Referências bibliográficas

SILVA NETO, A. J. Explicit and implicit formulations for inverse radiative transfer problems. In: WORLD CONGRESS ON COMPUTATIONAL MECHANICS, 5., 2002, Vienna = MINI SYMPOSIUM MS125 - COMPUTATIONAL TREATMENT OF INVERSE PROBLEMS IN MECHANICS. *Proceedings...* . Viena, Austria, 2002. (Invited Lecture).

_____; MOURA NETO, F. D. *Problemas inversos: conceitos fundamentais e aplicações*. Rio de Janeiro: EdUERJ, 2005.

_____; ÖZISIK, M. N. An inverse problem of estimating thermal conductivity, optical thickness, and single scattering albedo of semitransparent medium. In: INTERNATIONAL CONFERENCE ON INVERSE PROBLEMS IN ENGINEERING: THEORY AND PRACTICE, Florida. *Proceedings...* . Florida, 1993. p. 267–273.

_____; ÖZISIK, M. N. An inverse problem of simultaneous estimation of radiation phase function, albedo and optical thickness. *J. Quant. Spectrosc. Radiat. Transfe*, v. 53, n. 4, p. 397–409, 1995a.

_____. _____. *J. Quant. Spectrosc. Radiat. Transfer* v. 53, n. 4, p.397–409, 1995b.

_____; ROBERTY, N. C. A modelagem da tomografia com espalhamento como um problema inverso. In: FÓRUM NACIONAL DE CIÊNCIA E TECNOLOGIA EM SAÚDE, 4., 1998. *Anais...* 1998a. p. 317–318.

_____; ROBERTY, N. C. Uma comparação do método de ordenadas discretas com outras técnicas para a solução da equação de transporte unidimensional. In: CONGRESSO BRASILEIRO DE ENGENHARIA E CIÊNCIAS TÉRMICAS II, 7., 1998, Rio de Janeiro. *Anais...* . Rio de Janeiro, 1998b. p. 878–883.

_____; ROBERTY, N. C. A abordagem fonte-detetor na solução de problemas inversos em meios participantes. In: CONGRESSO DE ENGENHARIA MECÂNICA, NORTE-NORDESTE III, 5., 1998, Fortaleza. *Anais...* Fortaleza, 1998c. p. 340–347.

_____; SOEIRO, F. J. C. P. The solution of an inverse radiative transfer problem with the simulated annealing and levenberg-marquardt methods. In: MINI-SYMPOSIUM ON INVERSE PROBLEMS IN MEDICINE, ENGINEERING AND GEOPHYSICS, 2000, Santos. *Anais...* . Santos, 2000. p. 62–71.

SILVA NETO, A. J.; SOEIRO, F. J. C. P. Estimation of the phase function of anisotropic scattering with a combination of gradient based and stochastic global optimization methods. In: WORLD CONGRESS ON COMPUTATIONAL MECHANICS, 5., 2002, Vienna. *Proceedings...* Vienna, 2002.

SILVA NETO, A. J.; SOEIRO, F. J. C. P. Inverse problem of space dependent albedo estimation with artificial neural networks and hybrid methods. In: INTERNATIONAL CONGRESS OF MECHANICAL ENGINEERING, 18., 2005, Ouro Preto. *Proceedings...* Ouro Preto: ABCM, 2005.

Referências bibliográficas

SILVA NETO, A. J.; SOEIRO, F. J. C. P. Solution of implicitly formulated inverse heat transfer problems with hybrid methods. In: MINI-SYMPOSIUM INVERSE PROBLEMS FROM THERMAL/FLUIDS AND SOLIDMECHANICS APPLICATIONS = MIT CONFERENCE ON COMPUTATIONAL FLUID AND SOLID MECHANICS, 2., 2003, Cambridge. *Proceedings...* Cambridge, 2003.

SILVA, E.; THOMÉ, A. C. G. *Reconhecimento de caracteres manuscritos utilizando time de redes neurais*. Universidade Federal do Rio de Janeiro. Instituto de Matemática – Núcleo de Computação Eletrônica. [Rio de Janeiro], [2003]. Disponível em: http://www.labic.nce.ufrj.br/downloads/23sbccenia_2003.pdf.

SMITH, G. D. *Numerical solution of partial differential equations-finite difference methods*. Oxford: Clarendon Press, 1985. 350 p.

SOEIRO, F. J. C. P.; CARVALHO, G.; SILVA NETO, A. J. Thermal properties estimation of polymeric materials with the simulated annealing method. In: CONGRESSO BRASILEIRO DE ENGENHARIA E CIÊNCIAS TÉRMICAS, 8., 2000, Porto Alegre. *Anais...* Porto Alegre, 2000.

SOEIRO et al. Using neural networks to obtain initial estimates for the solution of inverse heat transfer problems. *Inverse Problems, Design and Optimization Symposium*, Rio de Janeiro, Brazil, 2004.

_____; SILVA NETO, A. J. Solution of inverse radiative transfer problems in two-layer media with artificial neural networks. In: INVERSE PROBLEMS IN ENGINEERING SEMINAR, 14., 2006, Ames. [Proceedings...]. Ames, 2006.

_____; SOARES, P. O.; SILVA NETO, A. J. Solution of inverse radiative transfer problems with artificial neural networks and hybrids methods. In: INVERSE PROBLEMS IN ENGINEERING SEMINAR, 13., 2004, Cincinnati. *Proceedings...* Cincinnati, 2004. p. 163-169.

SOUSA, F. L. et al. Application of the generalized extremal optimization algorithm to an inverse radiative transfer problem. *Inverse Problems in Science and Engineering*, v. 15, n. 7, 2007.

SPURR, R. et al. Simultaneous retrieval of aerosols and ocean properties: a classic inverse modelling approach. I - analytic Jacobians from the linearized CAODISORT models. *J. Quant. Spectrosc. Radiat. Transfer*, v. 104, p. 428-449, 2007.

SUNDMAN, L. K.; SANCHEZ, R.; McCORMICK, N. J. Ocean optical source estimation with widely spaced irradiance measurements. *Applied Optics*, v. 37, n.18, p. 3793–3803, 1998.

TAO, Z.; McCORMICK, N. J.; SANCHEZ, R. Ocean source and optical property estimation from explicit and implicit algorithm. *Applied Optics*, v. 33, p. 3265–3275, 1994.

VAPNIK, V. *Estimation of dependences based on empirical data*. New York: Secaucus : Springer, 1982. Springer Series in Statistics.

Referências bibliográficas

VERHOEF, W.; BACH, H. Simulation of hyperspectral and directional radiance images using coupled biophysical and atmospheric radiative transfer models. *Remote Sens. Environ.*, v. 87, p. 23-41, 2003.

ZHANG, T.; GORDON, H. R. Retrieval of elements of the columnar aerosol scattering phase matrix from polarized sky radiance over the ocean: simulations. *Applied Optics*, v. 36, p. 7948–7959, 1997.

ZHOU, Z-H; WU, J.; Tang, W. Ensembling neural networks: many could be better than all. *Artif. Intell.*, v.137, n. 1–2, p. 239–263, 2002.

ZHOU, H.-C. et al. An inverse radiative transfer problem of simultaneously estimating profiles of temperature and a radiative parameters from boundary intensity and temperature measurements. *J. Quant. Spectrosc. Radiat. Transfer*, v. 74, p. 605-620, 2002.

WERBOS, P. J. *Beyond regression: new tools for prediction and analysis in the behavioral sciences*. 1974. PhD. Thesis. Harvard University, 1974.

WICK, G. C. Über ebene diffusions problem. *Z. Phys.*, v. 120, p. 702-705, 1943.

WIDROW, B.; HOFF Jr., M. E. Adaptative switching circuits. *IRE WESCON Convention Record*, p. 96-104, 1960.

WILLSHAW, D. J.; MALSBURG, C. von der. How patterned neural connections can be set up by self organization. *Proceedings of the Royal Society London B*, [London],v.194, p.431-445, 1976.

WASSERMAN, P. D. *Neural network computing*. New York: Van Nostrand Reinhold, 1989.

WASSERMAN, P. D. *Advanced methods in neural computing*. [New York]: Van Nostrand Reinhold, 1999

APÊNDICE A – RESULTADOS DOS DEMAIS TESTES EFETUADOS

Neste apêndice são apresentados por meio de tabelas, os resultados dos testes de generalização efetuados que não foram incluídos no capítulo 4 e as curvas de performance geradas no treinamento das RNA especialistas com os melhores desempenhos para cada caso, e do MLP do combinador da máquina de comitê com modelo mostrado na Figura 32 para cada caso. Essas tabelas e curvas de performance são mostrados em dois conjuntos na seguinte ordem:

1. Tabelas com os resultados obtidos com as RNA especialistas com uma ou duas camadas ocultas para $\omega = 0,1, 0,5$ ou $0,9$ e $tol = 10^{-3}$ ou 10^{-4} . Neste conjunto estão inseridas as Tabelas de 9 a 26 e as Figuras 42 e 43, sendo que as Tabelas 9, 11, 13, 15, 17, 19, 21, 23 e 25 mostram cada uma o número de ordem que cada uma das RNA especialistas ocupa na máquina de comitê, o tempo de processamento, o número de épocas completadas e as estimativas obtidas por cada uma das RNA especialistas na fase de generalização para cada parâmetro radiativo, para as condições informadas nos títulos das tabelas. Nas Tabelas 10, 12, 14, 16, 18, 20, 22, 24 e 26 são mostrados os erros relativos percentuais ER ocorridos na estimação de cada parâmetro radiativo e, na última linha de cada tabela, a média para o erro relativo referente a cada um dos parâmetros radiativos. A Figura 42 mostra a curva de performance para a RNA nº 03 com $N_{HL} = 1$, $N_H = 12$, $tol = 10^{-4}$ e $\omega = 0,1, 0,5$ ou $0,9$ e a Figura 43 mostra a curva de performance para a RNA nº 10 com $N_{HL} = 2$, $N_H = 26$, $\omega = 0,1, 0,5$ ou $0,9$ e $tol = 10^{-3}$.
2. Tabelas com os resultados obtidos com as máquinas de comitê de redes neurais: em primeiro lugar com a máquina com modelo de acordo com a Figura 31, mostrados nas Tabelas 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37 e 38; em seguida com a máquina com modelo de acordo com a Figura 32, mostrados nas Tabelas 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49 e 50 e por último os resultados obtidos com as máquinas de comitê misto, de acordo com os modelos das Figuras 31 ou 32, apenas para $M = 6$, mostrados juntos nas Tabelas 51, 52 e 53 para $tol = 10^{-3}$ e/ou 10^{-4} . Em todas as tabelas supracitadas, são mostradas as estimativas obtidas com $M = 10$, antes da seleção das melhores RNA especialistas (exceto para as máquinas CM), e $M = 6$, após a seleção das melhores RNA, além dos respectivos erros

relativos percentuais ER ocorridos na estimação de cada parâmetro radiativo pelas máquinas de comitê.

Como foram utilizadas as mesmas RNA especialistas, após a fase de treinamento, para fornecer as entradas durante os testes de generalização para o combinador de cada um dos dois modelos de máquina de comitê aqui utilizados, os valores das tabelas que mostram as estimativas obtidas para cada RNA especialista, e os valores das tabelas que mostram os respectivos erros relativos por rede neural especialista para cada parâmetro radiativo, são os mesmos para ambos os modelos de máquina de comitê.

1. Resultados obtidos com as redes neurais especialistas isoladamente

1.1. Rede neural especialista com uma camada oculta

Tabela 9: Estimativas obtidas para $N_{HL} = 1$; $\omega = 0,5$ e $tol = 10^{-3}$.

RNA	N_H	tempo	épocas	τ_0	ω	ρ_1	ρ_2
1	8	00:22:00	464	0,9754	0,4662	0,4907	0,5115
2	10	02:57:30	3000	0,9722	0,4886	0,4690	0,5000
3	12	00:03:36	90	0,9694	0,4898	0,4742	0,5119
4	14	00:34:11	307	0,9983	0,4864	0,4666	0,5038
5	16	00:30:26	186	0,9872	0,5094	0,4730	0,5071
6	18	00:15:18	70	0,9895	0,4681	0,4870	0,5219
7	20	00:26:12	103	0,9741	0,5157	0,4674	0,5035
8	22	00:19:41	72	0,9750	0,5045	0,5314	0,5108
9	24	00:04:56	51	0,9940	0,4921	0,4923	0,5126
10	26	00:17:47	53	0,9570	0,4954	0,4958	0,5204

Tabela 10: Erros relativos percentuais por rede neural especialista para $N_{HL} = 1$; $\omega = 0,5$ e $tol = 10^{-3}$.

RNA	$ER\tau_0$ (%)	$ER\omega$ (%)	$ER\rho_1$ (%)	$ER\rho_2$ (%)
1	2,46	6,76	1,86	2,3
2	2,78	2,28	6,2	0
3	3,06	2,04	5,16	2,38
4	0,17	2,72	6,68	0,76
5	1,28	1,88	5,4	1,42
6	1,05	6,38	2,6	4,38
7	2,59	3,14	6,52	0,7
8	2,5	0,9	6,28	2,16
9	0,6	1,58	1,54	2,52
10	4,3	0,92	0,84	4,08
$\overline{ER}$	2,08	2,86	4,31	2,07

Nota: As destacadas em negrito são as selecionadas.

Apêndice A: resultados obtidos com as redes neurais especialistas isoladamente

A curva de performance da RNA nº 09 está mostrada na Figura 35, no capítulo 4.

Tabela 11: Estimativas obtidas para $N_{HL} = 1$; $\omega = 0,1$ e $tol = 10^{-4}$.

RNA	N_H	tempo	épocas	τ_0	ω	ρ_1	ρ_2
1	8	00:31:11	464	0,9508	0,1124	0,5540	0,5304
2	10	03:55:27	3000	0,9964	0,1124	0,5075	0,5294
3	12	00:03:39	90	1,0050	0,1181	0,5061	0,5203
4	14	00:29:09	307	1,0235	0,1103	0,4732	0,5296
5	16	00:11:19	186	1,0060	0,0853	0,5200	0,5126
6	18	00:41:54	570	0,9906	0,0879	0,5182	0,5192
7	20	00:36:23	460	1,0121	0,0669	0,5046	0,5195
8	22	00:13:08	120	1,0119	0,0868	0,5094	0,5221
9	24	00:21:05	201	0,9707	0,1158	0,5087	0,5146
10	26	00:41:35	295	0,9941	0,0848	0,4930	0,5002

Tabela 12: Erros relativos percentuais por rede neural especialista para $N_{HL} = 1$; $\omega = 0,1$ e $tol = 10^{-4}$.

RNA	$ER\tau_0$ (%)	$ER\omega$ (%)	$ER\rho_1$ (%)	$ER\rho_2$ (%)
1	4,92	12,4	10,8	6,08
2	0,36	12,4	1,5	5,88
3	0,5	18,1	1,22	4,06
4	2,35	10,3	5,36	5,92
5	0,6	14,7	4	2,52
6	0,94	12,1	3,64	3,84
7	1,21	33,1	0,92	3,9
8	1,19	13,2	1,88	4,42
9	2,93	15,8	1,74	2,92
10	0,59	15,2	1,4	0,04
ER	1,56	15,73	3,25	3,96

Nota: As destacadas em negrito são as selecionadas.

Tabela 13: Estimativas obtidas para $N_{HL} = 1$; $\omega = 0,5$ e $tol = 10^{-4}$.

RNA	N_H	tempo	épocas	τ_0	ω	ρ_1	ρ_2
1	8	00:31:11	464	0,9754	0,4662	0,4907	0,5115
2	10	03:55:27	3000	0,9722	0,4886	0,4690	0,5000
3	12	00:03:39	90	0,9694	0,4898	0,4742	0,5119
4	14	00:29:09	307	0,9983	0,4864	0,4666	0,5038
5	16	00:11:19	186	0,9872	0,5094	0,4730	0,5071
6	18	00:41:54	570	0,9657	0,5036	0,5199	0,5254
7	20	00:36:23	460	0,9833	0,5128	0,4724	0,5203
8	22	00:13:08	120	0,9956	0,5270	0,4996	0,5042
9	24	00:21:05	201	0,9617	0,4835	0,4977	0,5128
10	26	00:41:35	295	0,9718	0,5257	0,5199	0,5098

Tabela 14: Erros relativos percentuais por rede neural especialista para $N_{HL} = 1$; $\omega = 0,5$ e $tol = 10^{-4}$.

RNA	$ER\tau_0$ (%)	$ER\omega$ (%)	$ER\rho_1$ (%)	$ER\rho_2$ (%)
1	2,46	6,76	1,86	2,3
2	2,78	2,28	6,2	0
3	3,06	2,04	5,16	2,38
4	0,17	2,72	6,68	0,76
5	1,28	1,88	5,4	1,42
6	3,43	0,72	3,98	5,08
7	1,67	2,56	5,52	4,06
8	0,44	5,4	0,08	0,84
9	3,83	3,3	0,46	2,56
10	2,82	5,14	3,98	1,96
$\overline{ER}$	2,19	3,28	3,93	2,14

Nota: As destacadas em negrito são as selecionadas.

Tabela 15: Estimativas obtidas para $N_{HL} = 1$; $\omega = 0,9$ e $tol = 10^{-4}$.

RNA	N_H	tempo	épocas	τ_0	ω	ρ_1	ρ_2
1	8	00:31:11	464	0,8187	0,9199	0,4196	0,5153
2	10	03:55:27	3000	0,8556	0,9134	0,5576	0,5317
3	12	00:03:39	90	0,8405	0,8878	0,4360	0,5457
4	14	00:29:09	307	0,9112	0,9116	0,5137	0,5262
5	16	00:11:19	186	0,9103	0,8928	0,4807	0,5251
6	18	00:41:54	570	0,8899	0,8607	0,5000	0,5143
7	20	00:36:23	460	0,8850	0,8942	0,4995	0,5115
8	22	00:13:08	120	0,8990	0,9244	0,4892	0,4979
9	24	00:21:05	201	0,9215	0,8849	0,4891	0,4986
10	26	00:41:35	295	0,9085	0,9150	0,5055	0,4987

Tabela 16: Erros relativos percentuais por rede neural especialista para $N_{HL} = 1$; $\omega = 0,9$ e $tol = 10^{-4}$.

RNA	$ER\tau_0$ (%)	$ER\omega$ (%)	$ER\rho_1$ (%)	$ER\rho_2$ (%)
1	18,13	2,21	16,08	3,06
2	14,44	1,49	11,52	6,34
3	15,95	1,36	12,8	9,14
4	8,88	1,29	2,74	5,24
5	8,97	0,8	3,86	5,02
6	11,01	4,37	0	2,86
7	11,5	0,64	0,1	2,3
8	10,1	2,71	2,16	0,42
9	7,85	1,68	2,18	0,28
10	9,15	1,67	1,1	0,26
$\overline{ER}$	11,60	1,82	5,25	3,49

Nota: As destacadas em negrito são as selecionadas.

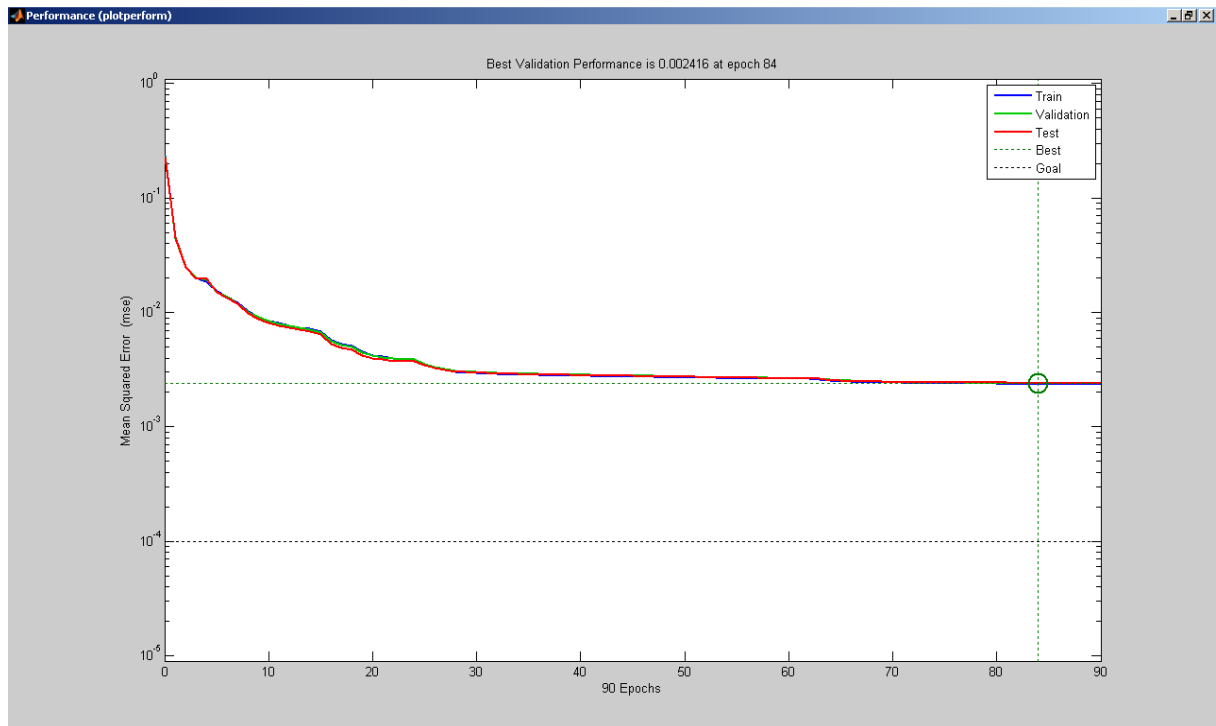


Figura 41: Curva de performance da RNA n° 3 com $N_{HL} = 1$, $N_H = 12$; $\omega = 0,1$, $0,5$ ou $0,9$ e $tol = 10^{-4}$.

1.2. Rede neural especialista com duas camadas ocultas

Tabela17: Estimativas obtidas para $N_{HL} = 2$; $\omega = 0,1$ e $tol = 10^{-3}$.

RNA	N_H	tempo	épocas	τ_0	ω	ρ_1	ρ_2
1	8	00:05:23	154	1,0160	0,1006	0,4935	0,5451
2	10	00:03:21	71	1,0214	0,0623	0,4888	0,5138
3	12	00:02:54	47	0,9903	0,0942	0,5388	0,5183
4	14	00:02:08	27	0,9890	0,0987	0,4859	0,5279
5	16	00:01:38	17	1,0032	0,0730	0,5175	0,5143
6	18	00:04:13	33	0,9939	0,0911	0,5222	0,5163
7	20	00:03:00	19	0,9941	0,0816	0,5272	0,4974
8	22	00:05:29	26	1,0112	0,0659	0,5077	0,5235
9	24	00:05:47	23	1,0112	0,1051	0,5359	0,5141
10	26	00:04:01	13	1,0182	0,0794	0,4966	0,5277

Tabela 18: Erros relativos percentuais por rede neural especialista para $N_{HL} = 2$; $\omega = 0,1$ e $tol = 10^{-3}$.

RNA	$ER\tau_0$ (%)	$ER\omega$ (%)	$ER\rho_1$ (%)	$ER\rho_2$ (%)
1	1,6	0,6	1,3	9,02
2	2,14	37,7	2,24	2,76
3	0,97	5,8	7,76	3,66
4	1,1	1,3	2,82	5,58
5	0,32	27,0	3,5	2,86
6	0,61	8,9	4,44	3,26
7	0,59	18,4	5,44	0,52
8	1,12	34,1	1,54	4,7
9	1,12	5,1	7,18	2,82
10	1,82	20,6	0,68	5,54
$\overline{ER}$	1,14	15,95	3,69	4,07

Nota: As destacadas em negrito são as selecionadas.

Tabela 19: Estimativas obtidas para $N_{HL} = 2$; $\omega = 0,5$ e $tol = 10^{-3}$.

RNA	N_H	tempo	épocas	τ_0	ω	ρ_1	ρ_2
1	8	00:05:23	154	0,9627	0,4796	0,4729	0,5066
2	10	00:03:21	71	1,0131	0,4819	0,4605	0,5110
3	12	00:02:54	47	1,0010	0,5153	0,4911	0,5090
4	14	00:02:08	27	0,9725	0,4876	0,4968	0,4978
5	16	00:01:38	17	0,9834	0,5061	0,4598	0,5105
6	18	00:04:13	33	0,9734	0,4988	0,4840	0,5191
7	20	00:03:00	19	0,9727	0,5007	0,4653	0,5091
8	22	00:05:29	26	0,9810	0,4993	0,5095	0,5252
9	24	00:05:47	23	0,9721	0,5216	0,5147	0,4947
10	26	00:04:01	13	1,0111	0,4919	0,4975	0,5000

Tabela 20: Erros relativos percentuais por rede neural especialista para $N_{HL} = 2$; $\omega = 0,5$ e $tol = 10^{-3}$.

RNA	$ER\tau_0$ (%)	$ER\omega$ (%)	$ER\rho_1$ (%)	$ER\rho_2$ (%)
1	3,73	4,08	5,42	1,32
2	1,31	3,62	7,9	2,2
3	0,1	3,06	1,78	1,8
4	2,75	2,48	0,64	0,44
5	1,66	1,22	8,04	2,1
6	2,66	0,24	3,2	3,82
7	2,73	0,14	6,94	1,82
8	1,9	0,14	1,9	5,04
9	2,79	4,32	2,94	1,06
10	1,11	1,62	0,5	0,0
$\overline{ER}$	7,05	1,65	5,26	2,16

Nota: As destacadas em negrito são as selecionadas.

Tabela 21: Estimativas obtidas para $N_{HL} = 2$; $\omega = 0,9$ e $tol = 10^{-3}$.

RNA	N_H	tempo	épocas	τ_0	ω	ρ_1	ρ_2
1	8	0:05:23	154	0,9230	0,8764	0,4763	0,5130
2	10	0:03:21	71	0,9245	0,8842	0,4577	0,4831
3	12	00:02:54	47	0,9361	0,8920	0,4916	0,4824
4	14	00:02:08	27	0,9229	0,8816	0,4808	0,5100
5	16	00:01:38	17	0,9299	0,8606	0,5118	0,4943
6	18	00:04:13	33	0,9484	0,9013	0,4884	0,5041
7	20	00:03:00	19	0,9336	0,9481	0,5383	0,4901
8	22	00:05:29	26	0,9260	0,9014	0,4829	0,5055
9	24	00:05:47	23	0,9208	0,9100	0,5275	0,4994
10	26	00:04:01	13	0,9199	0,8935	0,4728	0,4928

Tabela 22: Erros relativos percentuais por rede neural especialista para $N_{HL} = 2$; $\omega = 0,9$ e $tol = 10^{-3}$.

RNA	$ER\tau_0$ (%)	$ER\omega$ (%)	$ER\rho_1$ (%)	$ER\rho_2$ (%)
1	7,7	2,62	4,74	2,6
2	7,55	1,76	8,46	3,38
3	6,39	0,89	1,68	3,52
4	7,71	2,04	3,84	2
5	7,01	4,38	2,36	1,14
6	5,16	0,14	2,32	0,82
7	6,64	5,34	7,66	1,98
8	7,4	0,16	3,42	1,1
9	7,92	1,11	5,5	0,12
10	8,01	0,72	5,44	1,44
$\overline{ER}$	7,15	1,92	4,54	1,81

Nota: As destacadas em negrito são as selecionadas.

Tabela 23: Estimativas obtidas para $N_{HL} = 2$; $\omega = 0,1$ e $tol = 10^{-4}$.

RNA	N_H	tempo	épocas	τ_0	ω	ρ_1	ρ_2
1	8	0:05:48	154	1,0160	0,1006	0,4935	0,5451
2	10	0:08:08	172	1,0090	0,0927	0,4664	0,5328
3	12	01:05:55	757	0,9791	0,0957	0,4943	0,5186
4	14	00:27:52	299	0,9828	0,0949	0,4708	0,5163
5	16	00:17:11	157	0,9836	0,0832	0,4835	0,5207
6	18	00:17:58	112	0,9804	0,0847	0,4692	0,5094
7	20	00:20:34	98	0,9900	0,0860	0,4905	0,5064
8	22	00:17:32	68	0,9784	0,0899	0,4541	0,5143
9	24	00:19:56	61	0,9828	0,0875	0,4789	0,5171
10	26	00:19:48	51	0,9845	0,0965	0,4776	0,5127

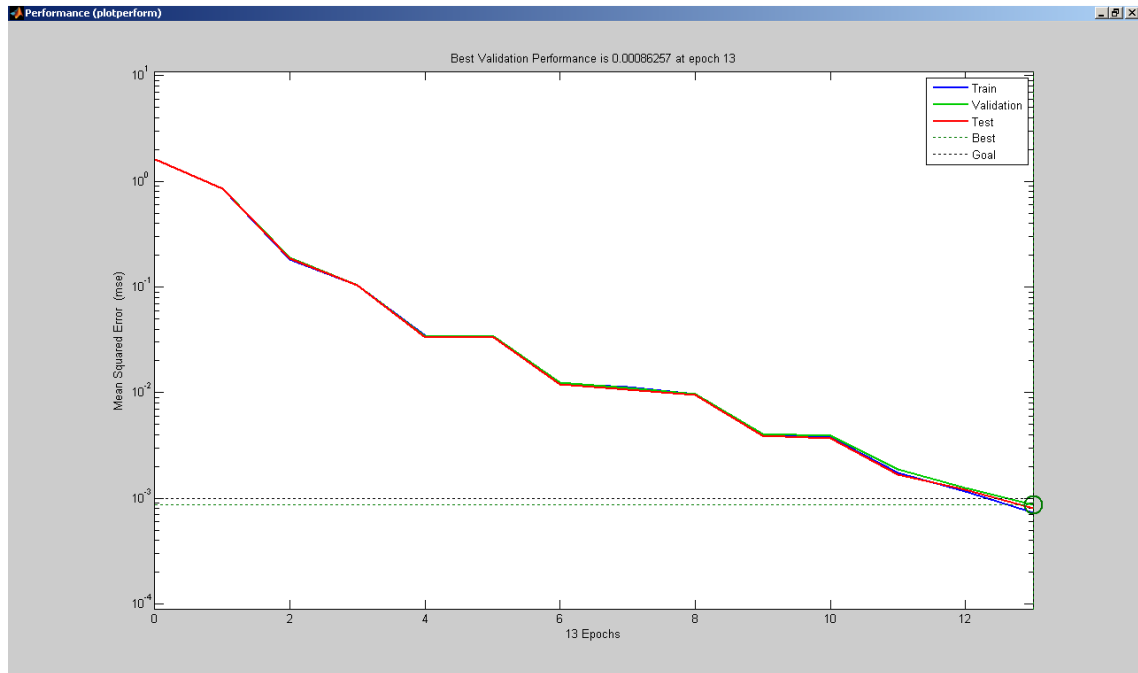


Figura 42: Curva de performance da RNA nº 10 com $N_{HL} = 2$; $N_H = 26$; $\omega = 0,1$, $0,5$ ou $0,9$ e $tol = 10^{-3}$.

Tabela 24: Erros relativos percentuais por rede neural especialista para $N_{HL} = 2$; $\omega = 0,1$ e $tol = 10^{-4}$.

RNA	$ER\tau_0$ (%)	$ER\omega$ (%)	$ER\rho_1$ (%)	$ER\rho_2$ (%)
1	1,6	0,6	1,3	9,02
2	0,9	7,3	6,72	6,56
3	2,09	4,3	1,14	3,72
4	1,72	5,1	5,84	3,26
5	1,64	16,8	3,3	4,14
6	1,96	15,3	6,16	1,88
7	1,0	14,0	1,9	1,28
8	2,16	10,1	9,18	2,86
9	1,72	12,5	4,22	3,42
10	1,55	3,5	4,48	2,54
$\overline{ER}$	1,63	8,95	4,42	3,87

Nota: As destacadas em negrito são as selecionadas.

Tabela 25: Estimativas obtidas para $N_{HL} = 2$; $\omega = 0,9$ e $tol = 10^{-4}$.

RNA	N_H	tempo	épocas	τ_0	ω	ρ_1	ρ_2
1	8	0:05:48	154	0,9230	0,8764	0,4763	0,5130
2	10	0:08:08	172	0,9325	0,8737	0,4744	0,4923
3	12	01:05:55	757	0,9691	0,8910	0,4903	0,5049
4	14	00:27:52	299	0,9999	0,9005	0,4851	0,4997
5	16	00:17:11	147	0,9979	0,8847	0,5025	0,5052
6	18	00:17:58	112	0,9690	0,8996	0,4999	0,5096
7	20	00:20:34	98	0,9581	0,9007	0,5034	0,4947
8	22	00:17:32	68	0,9724	0,8952	0,5181	0,4945
9	24	00:19:56	61	0,9667	0,9026	0,4990	0,4889
10	26	00:19:48	51	0,9576	0,8907	0,5064	0,5091

Tabela 26: Erros relativos percentuais por rede neural especialista para $N_{HL} = 2$; $\omega = 0,9$ e $tol = 10^{-4}$.

RNA	$ER\tau_0$ (%)	$ER\omega$ (%)	$ER\rho_1$ (%)	$ER\rho_2$ (%)
1	7,67	2,62	4,74	2,6
2	6,75	2,92	5,12	1,54
3	3,09	1,0	1,94	0,98
4	0,01	0,056	2,98	0,06
5	0,21	1,7	0,5	1,04
6	3,1	0,044	0,02	1,92
7	4,19	0,078	0,68	1,06
8	2,76	0,53	3,62	1,1
9	3,33	0,29	0,2	2,22
10	4,24	1,03	1,28	1,82
ER	3,53	1,03	2,1	1,43

Nota: As destacadas em negrito são as selecionadas.

Para o caso $N_{HL} = 2$ e $tol = 10^{-4}$ a curva de performance está mostrada na Figura 34 no capítulo 4.

2. Resultados obtidos com as máquinas de comitê

2.1 Máquina de comitê usando apenas a média aritmética no combinador

Tabela 27: Resultados da máquina de comitê para $M = 10$; $M = 6$; com $N_{HL} = 1$; $\omega = 0,1$ e $tol = 10^{-3}$.

PARÂMETRO RADIATIVO	RESULTADO $M = 10$	ER (%) $M = 10$	RESULTADO $M = 6$	ER(%) $M = 6$
τ_0	0,9999	0,01	1,0083	0,83
ω	0,1012	1,2	0,0974	2,6
ρ_1	0,5084	1,68	0,5003	0,06
ρ_2	0,5239	4,78	0,5215	4,3

Tabela 28: Resultados da máquina de comitê para $M = 10$ e $M = 6$, com $N_{HL} = 1$; $\omega = 0,5$ e $tol = 10^{-3}$.

PARÂMETRO RADIATIVO	RESULTADO $M = 10$	ER (%) $M = 10$	RESULTADO $M = 6$	ER(%) $M = 6$
τ_0	0,9792	2,08	0,9758	2,42
ω	0,4916	1,68	0,4967	0,66
ρ_1	0,4847	3,06	0,4893	2,14
ρ_2	0,5104	2,08	0,5105	2,1

Tabela 29: Resultados da máquina de comitê, para $M = 10$ e $M = 6$, com $N_{HL} = 1$; $\omega = 0,9$ e $tol = 10^{-3}$.

PARÂMETRO RADIATIVO	RESULTADO $M = 10$	ER (%) $M = 10$	RESULTADO $M = 6$	ER(%) $M = 6$
τ_0	0,8842	11,58	0,9069	9,31
ω	0,9025	0,28	0,9040	0,44
ρ_1	0,4872	2,56	0,4900	2,0
ρ_2	0,5122	2,44	0,4998	0,04

Tabela 30: Resultados da máquina de comitê, para $M = 10$ e $M = 6$, com $N_{HL} = 1$; $\omega = 0,1$ e $tol = 10^{-4}$.

PARÂMETRO RADIATIVO	RESULTADO $M = 10$	ER (%) $M = 10$	RESULTADO $M = 6$	ER(%) $M = 6$
τ_0	0,9961	0,39	0,9965	0,35
ω	0,0981	1,9	0,0992	0,8
ρ_1	0,5095	1,9	0,5137	2,74
ρ_2	0,5198	3,96	0,5239	4,78

Apêndice A: resultados obtidos com as máquinas de comitê

Tabela 31: Resultados da máquina de comitê, para $M = 10$ e $M = 6$, com $N_{HL} = 1$; $\omega = 0,5$ e $tol = 10^{-4}$.

PARÂMETRO RADIATIVO	RESULTADO $M = 10$	ER (%) $M = 10$	RESULTADO $M = 6$	ER(%) $M = 6$
τ_0	0,9781	2,19	0,9878	1,22
ω	0,4993	0,14	0,4923	1,54
ρ_1	0,4883	2,34	0,4986	0,28
ρ_2	0,5107	2,14	0,5072	1,44

Tabela 32: Resultados da máquina de comitê, para $M = 10$ e $M = 6$, com $N_{HL} = 1$; $\omega = 0,9$ e $tol = 10^{-4}$.

PARÂMETRO RADIATIVO	RESULTADO $M = 10$	ER (%) $M = 10$	RESULTADO $M = 6$	ER(%) $M = 6$
τ_0	0,8840	11,6	0,9754	2,46
ω	0,9005	0,06	0,8957	0,48
ρ_1	0,4891	2,18	0,5024	0,48
ρ_2	0,5165	3,3	0,4987	0,26

Tabela 33: Resultados da máquina de comitê, para $M = 10$ e $M = 6$, com $N_{HL} = 2$; $\omega = 0,1$ e $tol = 10^{-3}$.

PARÂMETRO RADIATIVO	RESULTADO $M = 10$	ER (%) $M = 10$	RESULTADO $M = 6$	ER(%) $M = 6$
τ_0	1,0049	0,49	1,0023	0,23
ω	0,0852	14,8	0,0952	4,8
ρ_1	0,5114	2,28	0,5172	3,44
ρ_2	0,5198	3,96	0,5198	3,96

Tabela 34: Resultados da máquina de comitê, para $M = 10$ e $M = 6$, com $N_{HL} = 2$; $\omega = 0,5$ e $tol = 10^{-3}$.

PARÂMETRO RADIATIVO	RESULTADO $M = 10$	ER (%) $M = 10$	RESULTADO $M = 6$	ER(%) $M = 6$
τ_0	0,9843	1,57	0,9852	1,48
ω	0,4983	0,34	0,5024	0,48
ρ_1	0,4852	2,96	0,5013	0,26
ρ_2	0,5083	1,66	0,5076	1,52

Apêndice A: resultados obtidos com as máquinas de comitê

Tabela 35: Resultados da máquina de comitê, para $M = 10$ e $M = 6$, com $N_{HL} = 2$; $\omega = 0,9$ e $tol = 10^{-3}$.

PARÂMETRO RADIATIVO	RESULTADO $M = 10$	ER (%) $M = 10$	RESULTADO $M = 6$	ER(%) $M = 6$
τ_0	0,9290	7,1	0,9311	6,89
ω	0,8941	0,66	0,8855	1,61
ρ_1	0,4884	2,32	0,4886	2,28
ρ_2	0,4981	0,38	0,5015	0,3

Tabela 36: Resultados da máquina de comitê, para $M = 10$ e $M = 6$, com $N_{HL} = 2$; $\omega = 0,1$ e $tol = 10^{-4}$.

PARÂMETRO RADIATIVO	RESULTADO $M = 10$	ER (%) $M = 10$	RESULTADO $M = 6$	ER(%) $M = 6$
τ_0	0,9887	1,13	0,9916	0,84
ω	0,0912	8,8	0,0951	4,9
ρ_1	0,4779	4,42	0,4761	4,78
ρ_2	0,5193	3,86	0,5233	4,66

Tabela 37: Resultados da máquina de comitê, para $M = 10$ e $M = 6$, com $N_{HL} = 2$; $\omega = 0,5$ e $tol = 10^{-4}$.

PARÂMETRO RADIATIVO	RESULTADO $M = 10$	ER (%) $M = 10$	RESULTADO $M = 6$	ER(%) $M = 6$
τ_0	0,9867	1,33	0,9883	1,17
ω	0,4931	1,38	0,4929	1,42
ρ_1	0,4947	1,06	0,4984	0,32
ρ_2	0,5092	1,84	0,5052	1,04

Tabela 38: Resultados da máquina de comitê, para $M = 10$ e $M = 6$, com $N_{HL} = 2$; $\omega = 0,9$ e $tol = 10^{-4}$.

PARÂMETRO RADIATIVO	RESULTADO $M = 10$	ER (%) $M = 10$	RESULTADO $M = 6$	ER(%) $M = 6$
τ_0	0,9647	3,53	0,9792	2,08
ω	0,8915	0,94	0,8956	0,49
ρ_1	0,4955	0,90	0,4992	0,16
ρ_2	0,5012	0,24	0,5005	0,1

2.2 Máquina de comitê usando a média aritmética mais MLP no combinador

Tabela 39: Resultados da máquina de comitê, para $M = 10$ e $M = 6$, com $N_{HL} = 1$; $\omega = 0,1$ e $tol = 10^{-3}$.

PARÂMETRO RADIATIVO	RESULTADO $M = 10$	ER (%) $M = 10$	RESULTADO $M = 6$	ER(%) $M = 6$
τ_0	0,9986	0,14	0,9997	0,03
ω	0,0991	0,9	0,1004	0,4
ρ_1	0,4997	0,06	0,5003	0,06
ρ_2	0,5093	1,86	0,5058	1,16

Tabela 40: Resultados da máquina de comitê, para $M = 10$ e $M = 6$, com $N_{HL} = 1$; $\omega = 0,5$ e $tol = 10^{-3}$.

PARÂMETRO RADIATIVO	RESULTADO $M = 10$	ER (%) $M = 10$	RESULTADO $M = 6$	ER(%) $M = 6$
τ_0	0,9903	0,97	0,9914	0,86
ω	0,4992	0,16	0,4997	0,06
ρ_1	0,5066	1,32	0,5026	0,52
ρ_2	0,5022	0,44	0,4998	0,04

Tabela 41: Resultados da máquina de comitê, para $M = 10$ e $M = 6$, com $N_{HL} = 1$; $\omega = 0,9$ e $tol = 10^{-3}$.

PARÂMETRO RADIATIVO	RESULTADO $M = 10$	ER (%) $M = 10$	RESULTADO $M = 6$	ER(%) $M = 6$
τ_0	0,9396	6,04	0,9505	4,95
ω	0,8976	0,27	0,8952	0,53
ρ_1	0,5019	0,38	0,5082	1,64
ρ_2	0,5061	1,22	0,5056	1,12

Tabela 42: Resultados da máquina de comitê, para $M = 10$ e $M = 6$, com $N_{HL} = 1$; $\omega = 0,1$ e $tol = 10^{-4}$.

PARÂMETRO RADIATIVO	RESULTADO $M = 10$	ER (%) $M = 10$	RESULTADO $M = 6$	ER(%) $M = 6$
τ_0	0,9990	0,1	0,9999	0,01
ω	0,0989	1,1	0,0992	0,8
ρ_1	0,5138	2,76	0,4981	0,38
ρ_2	0,5084	1,68	0,5083	1,66

Apêndice A: resultados obtidos com as máquinas de comitê

Tabela 43: Resultados da máquina de comitê, para $M = 10$ e $M = 6$, com $N_{HL} = 1$; $\omega = 0,5$ e $tol = 10^{-4}$.

PARÂMETRO RADIATIVO	RESULTADO $M = 10$	ER (%) $M = 10$	RESULTADO $M = 6$	ER(%) $M = 6$
τ_0	0,9958	0,42	0,9972	0,28
ω	0,5030	0,6	0,5008	0,16
ρ_1	0,4955	0,9	0,5011	0,22
ρ_2	0,5031	0,62	0,4998	0,04

Tabela 44: Resultados da máquina de comitê, para $M = 10$ e $M = 6$, com $N_{HL} = 1$; $\omega = 0,9$ e $tol = 10^{-4}$.

PARÂMETRO RADIATIVO	RESULTADO $M = 10$	ER (%) $M = 10$	RESULTADO $M = 6$	ER(%) $M = 6$
τ_0	0,9313	6,87	0,9770	2,3
ω	0,8857	1,59	0,8907	1,03
ρ_1	0,4933	1,34	0,5024	0,88
ρ_2	0,5085	1,7	0,4987	0,5

Tabela 45: Resultados da máquina de comitê, para $M = 10$ e $M = 6$, com $N_{HL} = 2$; $\omega = 0,1$ e $tol = 10^{-3}$.

PARÂMETRO RADIATIVO	RESULTADO $M = 10$	ER (%) $M = 10$	RESULTADO $M = 6$	ER(%) $M = 6$
τ_0	0,9957	0,43	0,9991	0,09
ω	0,0964	3,6	0,1022	2,2
ρ_1	0,5126	2,52	0,5078	1,56
ρ_2	0,5173	3,46	0,5087	1,74

Tabela 46: Resultados da máquina de comitê, para $M = 10$ e $M = 6$, com $N_{HL} = 2$; $\omega = 0,5$ e $tol = 10^{-3}$.

PARÂMETRO RADIATIVO	RESULTADO $M = 10$	ER (%) $M = 10$	RESULTADO $M = 6$	ER(%) $M = 6$
τ_0	1,0018	0,18	1,0006	0,06
ω	0,4970	0,6	0,5005	0,1
ρ_1	0,4966	0,68	0,4989	0,22
ρ_2	0,5044	0,88	0,5004	0,08

Apêndice A: resultados obtidos com as máquinas de comitê

Tabela 47: Resultados da máquina de comitê, para $M = 10$ e $M = 6$, com $N_{HL} = 2$; $\omega = 0,9$ e $tol = 10^{-3}$.

PARÂMETRO RADIATIVO	RESULTADO $M = 10$	ER (%) $M = 10$	RESULTADO $M = 6$	ER(%) $M = 6$
τ_0	0,9663	3,37	0,9610	3,9
ω	0,9021	0,23	0,8923	0,86
ρ_1	0,4973	0,54	0,4997	0,06
ρ_2	0,4965	0,7	0,5031	0,62

Tabela 48: Resultados da máquina de comitê, para $M = 10$ e $M = 6$, com $N_{HL} = 2$; $\omega = 0,1$ e $tol = 10^{-4}$.

PARÂMETRO RADIATIVO	RESULTADO $M = 10$	ER (%) $M = 10$	RESULTADO $M = 6$	ER(%) $M = 6$
τ_0	0,9883	1,17	0,9925	0,75
ω	0,0925	7,5	0,0952	4,8
ρ_1	0,4794	4,12	0,4993	0,14
ρ_2	0,5158	3,16	0,5018	0,36

Tabela 49: Resultados da máquina de comitê, para $M = 10$ e $M = 6$, com $N_{HL} = 2$; $\omega = 0,5$ e $tol = 10^{-4}$.

PARÂMETRO RADIATIVO	RESULTADO $M = 10$	ER (%) $M = 10$	RESULTADO $M = 6$	ER(%) $M = 6$
τ_0	0,9883	1,17	0,9925	0,75
ω	0,4974	0,52	0,4992	0,16
ρ_1	0,4972	0,56	0,4994	0,12
ρ_2	0,4970	0,6	0,4985	0,3

Tabela 50: Resultados da máquina de comitê, para $M = 10$ e $M = 6$, com $N_{HL} = 2$; $\omega = 0,9$ e $tol = 10^{-4}$.

PARÂMETRO RADIATIVO	RESULTADO $M = 10$	ER (%) $M = 10$	RESULTADO $M = 6$	ER(%) $M = 6$
τ_0	0,9669	3,31	0,9857	1,43
ω	0,8912	0,98	0,8997	0,03
ρ_1	0,5007	0,14	0,5003	0,06
ρ_2	0,4984	0,32	0,4999	0,02

2.3 Máquina de comitê misto usando a média aritmética ou MLP no combinador

Tabela 51: Resultados da máquina de comitê misto, para $M = 6$, com $N_{HL} = 1$ e/ou 2; $\omega = 0,1$ e $tol = 10^{-3}$ e/ou 10^{-4} .

PARÂMETRO RADIATIVO	MÉDIA $M = 6$ $tol = 10^{-3}$ e/ou 10^{-4}	ER (%) $M = 6$ $tol = 10^{-3}$ e/ou 10^{-4}	MLP $M = 6$ $tol = 10^{-3}$	ER(%) $M = 6$ $tol = 10^{-3}$
τ_0	0,9950	0,5	0,9975	0,25
ω	0,0980	2,0	0,1009	0,9
ρ_1	0,4910	1,8	0,5013	0,26
ρ_2	0,5167	3,34	0,5112	2,24

Tabela 52: Resultados da máquina de comitê misto, para $M = 6$, com $N_{HL} = 1$ e/ou 2; $\omega = 0,5$ e $tol = 10^{-3}$ e/ou 10^{-4} .

PARÂMETRO RADIATIVO	MÉDIA $M = 6$ $tol = 10^{-3}$ e/ou 10^{-4}	ER (%) $M = 6$ $tol = 10^{-3}$ e/ou 10^{-4}	MLP $M = 6$ $tol = 10^{-3}$	ER(%) $M = 6$ $tol = 10^{-3}$
τ_0	0,9959	0,41	0,9982	0,18
ω	0,4968	0,64	0,4955	0,9
ρ_1	0,4954	0,92	0,4958	0,84
ρ_2	0,5052	1,04	0,4970	0,6

Tabela 53: Resultados da máquina de comitê misto, para $M = 6$, com $N_{HL} = 1$ e/ou 2; $\omega = 0,9$ e $tol = 10^{-3}$ e/ou 10^{-4} .

PARÂMETRO RADIATIVO	MÉDIA $M = 6$ $tol = 10^{-3}$ e/ou 10^{-4}	ER (%) $M = 6$ $tol = 10^{-3}$ e/ou 10^{-4}	MLP $M = 6$ $tol = 10^{-3}$	ER(%) $M = 6$ $tol = 10^{-3}$
τ_0	0,9792	2,08	0,9886	1,14
ω	0,8956	0,49	0,8884	1,29
ρ_1	0,4992	0,16	0,5008	0,16
ρ_2	0,5005	0,1	0,5006	0,12

Livros Grátis

(<http://www.livrosgratis.com.br>)

Milhares de Livros para Download:

[Baixar livros de Administração](#)

[Baixar livros de Agronomia](#)

[Baixar livros de Arquitetura](#)

[Baixar livros de Artes](#)

[Baixar livros de Astronomia](#)

[Baixar livros de Biologia Geral](#)

[Baixar livros de Ciência da Computação](#)

[Baixar livros de Ciência da Informação](#)

[Baixar livros de Ciência Política](#)

[Baixar livros de Ciências da Saúde](#)

[Baixar livros de Comunicação](#)

[Baixar livros do Conselho Nacional de Educação - CNE](#)

[Baixar livros de Defesa civil](#)

[Baixar livros de Direito](#)

[Baixar livros de Direitos humanos](#)

[Baixar livros de Economia](#)

[Baixar livros de Economia Doméstica](#)

[Baixar livros de Educação](#)

[Baixar livros de Educação - Trânsito](#)

[Baixar livros de Educação Física](#)

[Baixar livros de Engenharia Aeroespacial](#)

[Baixar livros de Farmácia](#)

[Baixar livros de Filosofia](#)

[Baixar livros de Física](#)

[Baixar livros de Geociências](#)

[Baixar livros de Geografia](#)

[Baixar livros de História](#)

[Baixar livros de Línguas](#)

[Baixar livros de Literatura](#)
[Baixar livros de Literatura de Cordel](#)
[Baixar livros de Literatura Infantil](#)
[Baixar livros de Matemática](#)
[Baixar livros de Medicina](#)
[Baixar livros de Medicina Veterinária](#)
[Baixar livros de Meio Ambiente](#)
[Baixar livros de Meteorologia](#)
[Baixar Monografias e TCC](#)
[Baixar livros Multidisciplinar](#)
[Baixar livros de Música](#)
[Baixar livros de Psicologia](#)
[Baixar livros de Química](#)
[Baixar livros de Saúde Coletiva](#)
[Baixar livros de Serviço Social](#)
[Baixar livros de Sociologia](#)
[Baixar livros de Teologia](#)
[Baixar livros de Trabalho](#)
[Baixar livros de Turismo](#)