

# REDES NEURAIS ARTIFICIAIS NA SOLUÇÃO DE PROBLEMAS INVERSOS EM ENGENHARIA – ESTUDO DE CASO

ROGÉRIO CAMPOS DE OLIVEIRA



**REDES NEURAS ARTIFICIAIS NA SOLUÇÃO DE  
PROBLEMAS INVERSOS EM ENGENHARIA –  
ESTUDO DE CASO**



ROGÉRIO CAMPOS DE OLIVEIRA

**REDES NEURAIAS ARTIFICIAIS NA SOLUÇÃO DE  
PROBLEMAS INVERSOS EM ENGENHARIA –  
ESTUDO DE CASO**

1ª Edição

Quipá Editora  
2024

Copyright © dos autores e autoras. Todos os direitos reservados.

Esta obra é publicada em acesso aberto. O conteúdo dos capítulos, os dados apresentados, bem como a revisão ortográfica e gramatical são de responsabilidade de seus autores, detentores de todos os Direitos Autorais, que permitem o download e o compartilhamento, com a devida atribuição de crédito, mas sem que seja possível alterar a obra, de nenhuma forma, ou utilizá-la para fins comerciais.

Dados Internacionais de Catalogação na Publicação (CIP)

---

O48r      Redes neurais artificiais na solução de problemas inversos em engenharia – estudo de caso / Rogério Campos de Oliveira. — Iguatu, CE : Quipá Editora, 2024.

108. : il.

ISBN 978-65-5376-338-8 [E-book/PDF]

ISBN 978-65-5376-339-5 [Físico]

1. Engenharia. 2. Redes neurais - artificiais. I. Oliveira, Rogério Campos de. II. Título.

CDD 629

---

Obra publicada pela Quipá Editora em junho de 2024

Quipá Editora  
www.quipaeditora.com.br  
@quipaeditora

À Kátia, Rafaela, Rodrigo e Carolina pelo sacrifício e  
compreensão e aos meus Pais pelo apoio.

## **AGRADECIMENTOS**

Aos meus orientadores Prof. Dr. Antônio José da Silva Neto e Prof. Dr. Luis Biondi Neto (*in memoriam*)

Ao Prof. Dr. Júlio Alves Hermínio (*in memoriam*)

Ao Prof. Dr. Francisco Duarte Moura Neto (*in memoriam*)

Aos meus Pais pelo apoio irrestrito

Aos meus professores e colegas do IPRJ

Ao Instituto Federal de Ciência e Tecnologia do Estado do Rio Grande do Norte - IFRN

Ao Instituto Politécnico do Rio de Janeiro – IPRJ/UERJ

À Secretaria Estadual de Tributação do Estado do Rio Grande do Norte – SET/RN

## APRESENTAÇÃO

Este livro fundamenta-se na minha dissertação de mestrado cujo título é *Máquina de Comitê de Redes Neurais Artificiais na Solução do Problema Inverso em Transferência Radiativa* (2010), que tem por objetivo servir de fonte de consulta para alunos e profissionais da área de Matemática Aplicada e Computacional, utilizando máquina de comitê de redes neurais artificiais como ferramenta básica. Apesar de ser um estudo de caso, as técnicas aqui utilizadas, sejam na solução do problema direto ou do problema inverso e a utilização de redes neurais artificiais para obtenção de resultados, podem ser aplicadas a quaisquer áreas do conhecimento que utilizem a teoria de problemas diretos e inversos baseado no bem conhecido algoritmo de Levenberg-Marquardt, assim como o reconhecimento de padrões das redes neurais artificiais - *learning machines* - obtido através de um processo de aprendizado.

No capítulo 1, na introdução, no item 1.1 tratamos do tema, sua relevância e objetivos, citamos as pesquisas realizadas por inúmeros pesquisadores nessas áreas, mostramos também a arquitetura da Máquina de Comitê de Redes Neurais Artificiais utilizada no trabalho de dissertação, que foi considerada pela banca como inovadora, tendo obtido excelentes resultados em comparação com outros métodos utilizados anteriormente.

No capítulo 2, tratamos sobre os problemas direto e inverso em transferência radiativa, no item 2.1 tratamos da solução do problema direto em transferência radiativa, onde apresentamos o modelo para a solução do problema direto em transferência radiativa em meios participantes, isto é, meios absorvedores, emissores e espalhadores isotrópicos homogêneos ou heterogêneos, consistindo no cálculo dos valores da intensidade de radiação, em todo domínio espacial e angular, considerando-se conhecidas a geometria, as propriedades radiativas do meio, o termo fonte e as suas condições de contorno. Em 2.1.1 apresentamos a formulação matemática do problema direto em transferência radiativa, considerada para regime permanente, representando a interação entre a radiação e o meio, e dada pela versão linear da equação de Boltzmann, conhecida também como equação de transporte (ÖZISIK, 1973). A solução do problema direto consiste em obter os parâmetros de radiação 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  $f_2$ , que são: as refletividades difusas  $\rho_1$  e  $\rho_2$  nas superfícies internas do meio, a espessura ótica  $\tau_0$  e o albedo de espalhamento  $\omega$ . Em 2.1.2 apresentamos a 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). No item 2.2 abordamos a solução do problema inverso em transferência radiativa. Aqui o nosso 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  $\vec{Z} = \{\tau_0, \omega, \rho_1, \rho_2\}^T$  a partir dos dados experimentais simulados  $Y_l = I_{exp_l} = I_{calc_l}(\vec{Z}_{exato}) + \sigma r_l$ .

No capítulo 3 apresentamos os fundamentos sobre redes neurais artificiais e máquinas de comitê de redes neurais artificiais e sua aplicação à solução do problema inverso em transferência radiativa. Primeiramente, no item 3.1, apresentamos um histórico sobre redes neurais artificiais e máquinas de comitês de redes neurais artificiais. Dando sequência apresentamos no item 3.2 a fundamentação teórica sobre redes neurais artificiais, onde descrevemos o neurônio biológico, a estrutura do neurônio artificial usado como modelo, os vários tipos de função de transferência, como a função degrau (*threshold function*), a função linear, as funções log-sigmoide e tan-sigmoide, o neurônio artificial de entrada vetorial, a rede neural artificial, os paradigmas de treinamento de redes neurais artificiais, o *Perceptron* e o *Perceptron de Múltiplas Camadas*. No item 3.3, apresentamos a fundamentação teórica sobre máquinas de comitês de redes neurais artificiais, com sua classificação e estruturas: estáticas e dinâmicas. No item 3.4 apresentamos a formulação matemática e a solução do problema inverso em transferência radiativa usando redes neurais artificiais. No item 3.5, apresentamos a modelagem e formulação matemática para a solução do problema inverso em transferência radiativa usando máquinas de comitês de redes neurais artificiais.

No capítulo 4 apresentamos os resultados e discussão, onde abordamos a metodologia aplicada, item 4.1, fazemos um comparativo dos resultados obtidos com as RNA especialistas e com as máquinas de comitês.

Finalmente, no capítulo 5 apresentamos as nossas conclusões sobre os resultados obtidos.

## 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
FORTRAN 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
IPRJ	Instituto Politécnico do Rio de Janeiro
LDA	Linear discriminant analysis
LIDAR	Light detection and ranging
LM	Levenberg-Marquardt
MATLAB	Matrix Laboratory ( <i>software</i> 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)
<i>purelin</i>	Função de ativação tipo linear do Neural Network Toolbox do MATLAB2008®
RNA	Rede neural artificial
SA	Simulated annealing ( <i>recozimento simulado</i> )
SLP	Single layer perceptron

<i>S<sub>M</sub></i>	Método das ordenadas discretas
SVM	Support vector machine
<i>tansig</i>	Função de ativação tipo tangente hiperbólica do Neural Network Toolbox do MATLAB2008®
<i>trainlm</i>	Função de treinamento dom 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 para cada neurônio  $j$  da camada de saída  $J$  da Figura 21.
- $E_n^q$  Erro na estimativa do parâmetro de ordem  $n$ ,  $n = 1, 2, \dots, N_u$  no padrão  $q$ , obtido pelo confronto entre o valor exato  $Z_{n,exato}$  do parâmetro  $n$  e o seu valor estimado  $t_n^q \equiv Z_{n,i}^q$  (para MLP das Figuras 29 e 30).
- $E_j^q$  Erro na estimativa do parâmetro radiativo de ordem  $j$ ,  $j = 1, 2, \dots, N_s$  no padrão  $q$ , obtido pelo confronto entre o valor exato  $Z_{j,exato}$  do parâmetro e o seu valor estimado  $\dot{Z}_j^q$  (Figura 33).
- $f(\cdot)$  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  $x = L$ .
- $f_n$  Coeficientes 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).
- $g(\cdot)$  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, $i = 1, \dots, N$ .
$h(\cdot)$	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.
$I_{expl}$	Intensidade de radiação experimental simulada (o mesmo que $Y_l$ ).
$I_{calc}(\vec{Z})$	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 $\mu$ .
$I^+(0, \mu)$	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 $\tau = 0$ .
$I_m^{Nx}$	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 $p$ .
$(I_m^i)^{p+1}$	Estimativa para a intensidade de radiação na iteração $p + 1$ .
$I_{m\ ef}^1$	Intensidade efetiva de radiação que emerge do meio em $\tau = 0$ .
$I_{m\ ef}^{Nx}$	Intensidade efetiva de radiação que emerge do meio em $\tau = \tau_0$ .

$k_a$	Coeficiente de absorção.
$L$	Espessura física do meio medida ao longo do eixo $x$ .
$N_H$	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 (página 89).
$N_{H1}$ 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 de modelo dado pela 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 $m$ , $m = 1, 2, \dots, N_{H1}$ , $q = 0, \dots, Q-1$ , 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 $\mu$ .
$P_n(\mu')$	Polinômios de Legendre em $\mu'$ .
$q$	Ordem do padrão de treinamento.
$Q$	Limite superior do número de padrões de treinamento $q$ .
$q^i$	Discretização espacial do espalhamento interno isotrópico usando quadratura de Gauss-Legendre.
$r_i$	Número pseudo-randômico no intervalo $[-1,1]$ .
$R$	Saída global da máquina de comitê tipo dinâmica (Figura 28), obtida com a equação (31).
$RNA_i$	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).
$w_{m,l}^{(1)q}$	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, no padrão de treinamento $q$ , $q = 0, \dots, Q-1$ , do MLP das Figuras 29 e 30.
$w_{m,0}^{(1)q}$	Bias do neurônio $m$ da primeira camada oculta (1) do MLP das Figuras 29 e 30, onde $m = 1, 2, \dots, N_{H1}$ e $q = 0, \dots, Q-1$ .
$w_{m,l}^{(1)q+1}$	Valor atual do peso sináptico que conecta o nó $l$ de entrada com o neurônio $m$ da primeira camada oculta (1) na iteração $q + 1$ (Figuras 29 e 30).
$w_{k,m}^{(2)q}$	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 $k = 1, 2, \dots, N_{H2}$ , $m = 1, 2, \dots, N_{H1}$ e $q = 0, \dots, Q-1$ .

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

$x$	Coordenada espacial medida ao longo do eixo $-x$ .
$X$	Vetor das entradas nas Figuras 21, 26 e 28.
$X_j$	Vetores de entradas da máquina de comitê da Figura 27, $j = 1, \dots, N$ .
$Y$	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 $Y_i^q$ das RNA $_i$ , $i = 1, \dots, M$ , $l = 1, 2, \dots, N$ no padrão $q$ , $q = 0, \dots, Q-1$ .
$\vec{Z}$	Vetor dos parâmetros radiativos, incógnitas do problema inverso de transferência radiativa, dado pela equação (24).
$\vec{Z}_{exato}$	Vetor alvo, cujos elementos são os valores desejados para os parâmetros radiativos.
$\vec{Z}^q$	Vetor de estimativas dos parâmetros radiativos, para o padrão $q$ , onde $q = 0, \dots, Q-1$ , (ver Figuras 31, 32 e 33).
$\vec{z}_i^q$	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).
$Z_{n,i}^q$	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).
$\vec{z}_i^{q=0}$	Estimativa inicial para o vetor $\vec{z}_i^q$ .
$\vec{Z}^q$	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 $q$ , $q = 0, \dots, Q-1$ .
$\bar{Z}_n^q$	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).
$z_u^q$	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).
$\vec{z}^q$	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.
$\alpha_n$	Pesos da quadratura de Gauss-Legendre correspondentes aos pontos de colocação, $\mu_n$ , $n = 1, 2, \dots, M$ .
$\beta$	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 $\gamma$ 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 partindo 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 partindo 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 partindo 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 partindo 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.
$\varepsilon$	Tolerância requerida na solução do problema direto.
$\eta$	Taxa de aprendizado.
$\theta$	Ângulo entre o eixo- $x$ positivo e a direção do feixe de radiação.
$\theta'$	Ângulo entre o eixo- $x$ positivo e a direção do feixe de radiação refletida.
$\mu$	Cosseno do ângulo polar $\theta$ .
$\mu'$	Cosseno do ângulo polar $\theta'$ .
$\mu_m$	Variável angular discretizada, $m = 1, \dots, M$ .
$\rho_1$	Refletividade difusa na superfície interna em $x = 0$ .
$\rho_2$	Refletividade difusa na superfície interna em $x = L$ .
$\sigma$	Desvio padrão dos erros de medição.
$\sigma_s$	Coefficiente de espalhamento.
$\tau$	Variável óptica.
$\tau_0$	Espessura óptica do meio.
$\bar{\tau}_0$	Valor máximo escolhido para a $\tau_0$ (limite superior artificial).
$\tau_i$	Variável óptica discretizada, $i = 1, \dots, N_x$ .
$\chi$	Função de ativação da camada de saída $s$ do MLP da Figura 33.
$\chi'(\hat{z}_j^q)$	Derivada da função de ativação $\chi$ 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).
$\omega$	Albedo de espalhamento simples.

## SUMÁRIO

<b>APRESENTAÇÃO</b>	
<b>CAPÍTULO 1</b>	<b>19</b>
INTRODUÇÃO	
<b>CAPÍTULO 2</b>	<b>27</b>
PROBLEMAS DIRETO E INVERSO EM TRANSFERÊNCIA RADIATIVA	
<b>CAPÍTULO 3</b>	<b>40</b>
FUNDAMENTOS SOBRE REDES NEURAIS E MÁQUINAS DE COMITÊS E SUA APLICAÇÃO À SOLUÇÃO DO PROBLEMA INVERSO EM TRANSFERÊNCIARADIATIVA	
<b>CAPÍTULO 4</b>	<b>74</b>
RESULTADOS E DISCUSSÃO	
<b>CAPÍTULO 5</b>	<b>86</b>
CONCLUSÃO	
<b>REFERÊNCIAS BIBLIOGRÁFICAS</b>	<b>88</b>
<b>APÊNDICE A</b>	<b>97</b>
Resultados dos testes efetuados que não foram mostrados no capítulo 4	
<b>SOBRE O AUTOR</b>	<b>108</b>

## CAPÍTULO 1

### INTRODUÇÃO

#### O TEMA: SUA RELEVÂNCIA E OBJETIVOS

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. 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 e 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 ópticas 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 óptica 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 Campos-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 (modo combinado) 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 (absorção + espalhamento) 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 McCornick (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, Maurente, Vielmo 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 para tal problema é aplicar o princípio dividir e conquistar, ou seja, dividir uma tarefa computacional complexa em tarefas computacionais mais simples distribuindo tarefas entre um determinado número de especialistas, aqui 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.

Neste trabalho, utilizam-se RNA e máquinas de comitê de redes neurais artificiais para solucionar o problema inverso de transferência radiativa obtendo estimativas para os parâmetros de

transferência radiativa de um meio unidimensional, absorvedor, emissor e espalhador isotrópico, comparando os resultados obtidos.

As pesquisas com máquinas de comitês 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 comitês 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 maquina 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 (200-), 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 previsão 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 comitês 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 comitês 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), utilizou 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. Campos-Oliveira et al. (2008) propôs uma máquina de comitê de redes neurais artificiais, 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.

A motivação principal para a utilização de máquinas de comitês de redes neurais artificiais na solução do problema inverso em transferência radiativa, deve-se ao fato de além de incluir todos aqueles benefícios no uso de redes neurais artificiais, como por exemplo, não-linearidade, mapeamento de entrada-saída, adaptabilidade, robustez, plasticidade, etc; apresentam, de forma geral, um resultado global na fase de generalização melhor do que aqueles obtidos por cada RNA isoladamente, considerando que ocorre um acréscimo computacional extra pequeno.

O objetivo específico deste trabalho está na formulação e na solução do problema inverso de transferência radiativa para obter estimativas para as propriedades radiativas do meio, i.e. a espessura

óptica  $\tau_0$ , o albedo de espalhamento simples  $\omega$ , e as refletividades difusas  $\rho_1$  e  $\rho_2$  num meio unidimensional, participante e homogêneo a partir do uso de dados experimentais sintéticos,  $\bar{Y}$ , usando para tal, dois modelos distintos de máquinas de comitês de redes neurais artificiais fundamentados no tipo estática que utiliza a média de conjunto (*ensemble*), conforme configuração mostrada na Figura 1.

O primeiro modelo de máquina de comitê de rede neural artificial usado é o caso clássico do tipo estático com a média de conjunto. O segundo modelo usado é uma variação do modelo anterior onde o combinador de saída da máquina combina a média aritmética do conjunto com uma rede neural de saída do tipo MLP, cuja excitação de entrada é justamente a média das saídas das RNA especialistas, i.e, a média do conjunto.

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 comitês 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 máquinas de comitês foram gerados com a solução da versão linear da equação de Boltzmann, que é usada para modelar o problema direto de transferência radiativa (ÖZISIK, 1973)

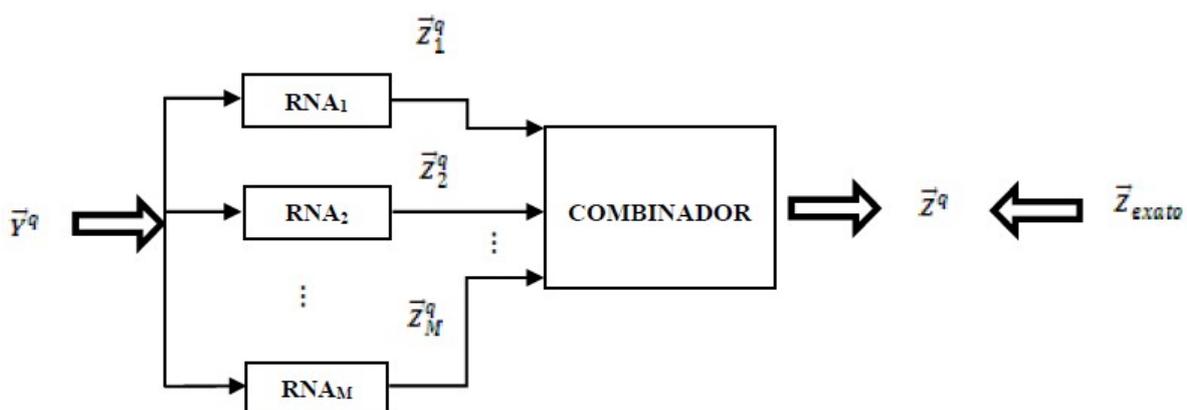


Figura 1: Estrutura da máquina de comitê de redes neurais artificiais.

## CAPÍTULO 2

### PROBLEMAS DIRETO E INVERSO EM TRANSFERÊNCIA RADIATIVA

#### 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, as propriedades radiativas do meio, o termo fonte e as suas condições de contorno.

#### Formulação matemática do problema direto em transferência radiativa

Considere 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  $\tau_0$  é a espessura óptica do meio, caracterizando o comprimento de penetração da radiação no meio.

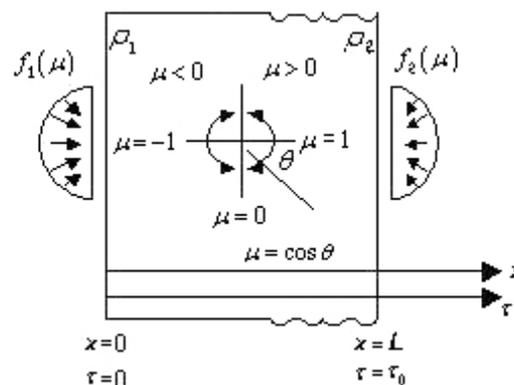


Figura 2: Meio participante unidimensional homogêneo com superfícies internas refletoras difusas e submetido à incidência de radiação externa de intensidades  $f_1$  e  $f_2$  (SILVA-NETO; BECCENERI, 2009, p.18).

A formulação matemática do problema direto de transferência radiativa no meio considerado 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),

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

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

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

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

onde  $I$  representa a intensidade da radiação,  $x$  é a coordenada espacial,  $\mu$  o cosseno do ângulo polar  $\theta$  (ângulo entre o eixo- $x$  positivo e a direção do feixe de radiação),  $\sigma_s$  o coeficiente de espalhamento,  $S(x, \mu)$  uma fonte interna de radiação, e  $\rho_1$  e  $\rho_2$  as refletividades difusas nas superfícies internas do meio, respectivamente em  $x = 0$  e  $x = L$ ,  $p(\mu, \mu')$  é a função de fase, e  $\beta$  é o coeficiente de extinção total.

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

onde  $k_a$  é o coeficiente de absorção.

A função de fase  $p(\mu, \mu')$  usualmente escrita como uma expansão em polinômios de Legendre (ÖZISIK, 1973), é escrita como

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

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 consideramos 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,  $\tau$ , 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  $\beta$  definido na equação (2) e empregando as definições da variável ótica  $\tau$  e da espessura ótica  $\tau_0$ , 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, obtemos 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  $\omega$  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.

### 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 a sua vizinhança não são levados em consideração.

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

Os dados experimentais e as condições de contorno são representados esquematicamente na Figura 3.

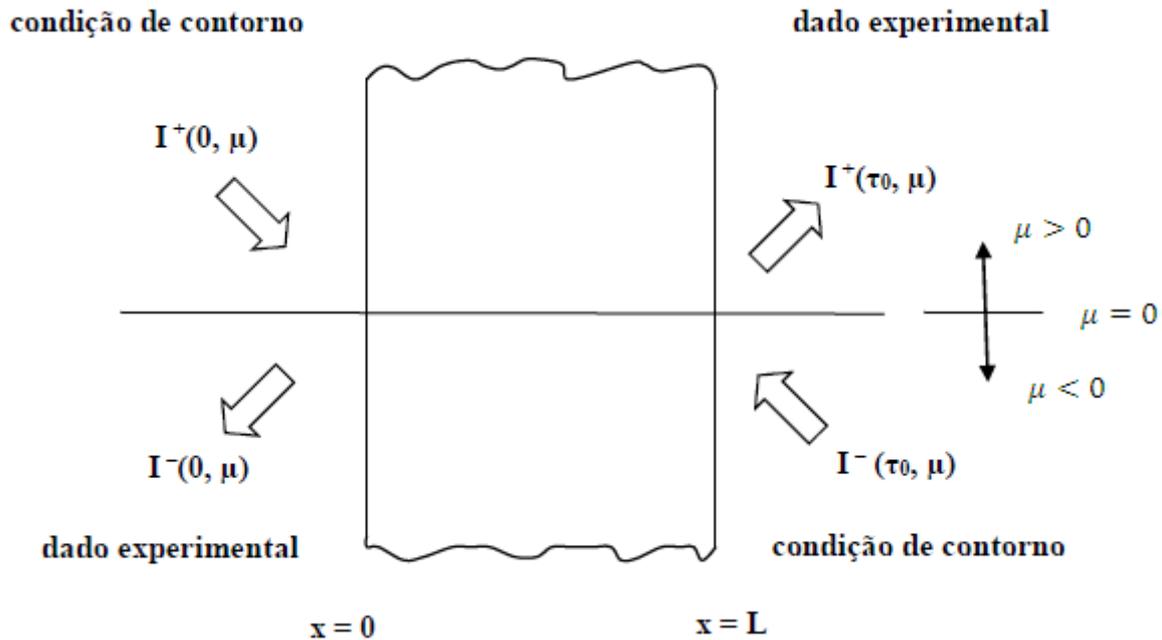


Figura 3: Representação esquemática dos dados experimentais e das condições de contorno do meio (PINHEIRO, 2004, p.21).

Em relação à discretização do domínio de uma equação íntegro-diferencial para obtenção de uma malha computacional Pinheiro (2004, p. 22, 230) afirma :

A solução de uma equação íntegro-diferencial em uma região implica a obtenção dos valores para a variável dependente em cada ponto da região. Computacionalmente, apenas podemos lidar com uma região contínua se determinarmos 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  $(\tau, \mu)$ . Contudo, nada impede que sejam escolhidos alguns pontos dentro da região e apenas neles calcular a solução do problema.

Com base na afirmação acima, 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.

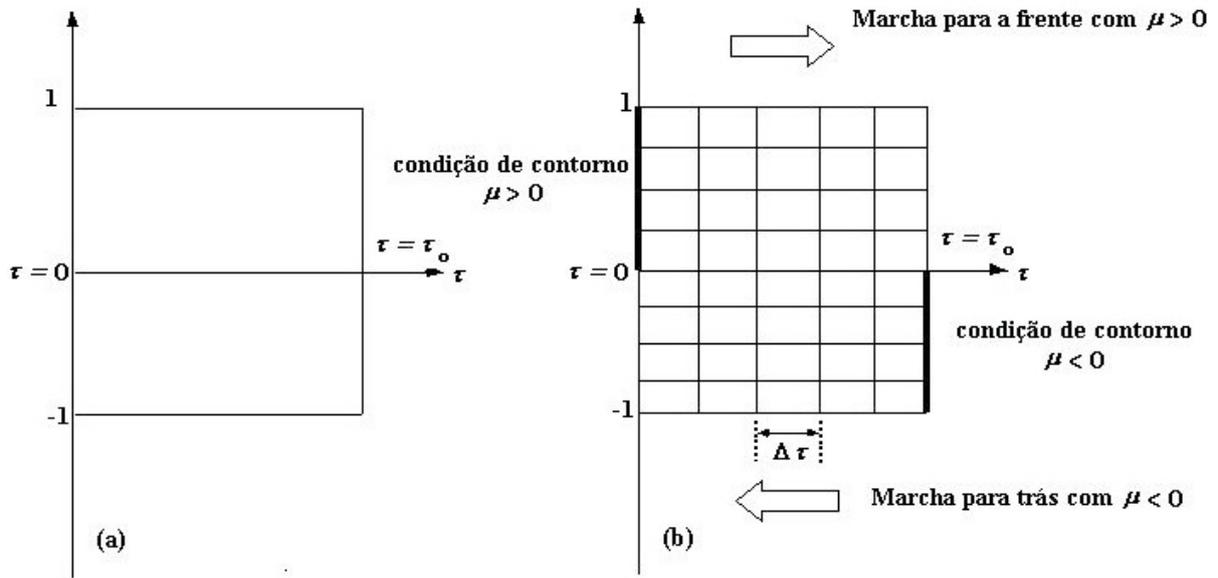


Figura 4: (a) domínio contínuo e (b) malha computacional (Pinheiro, 2004 p.24).

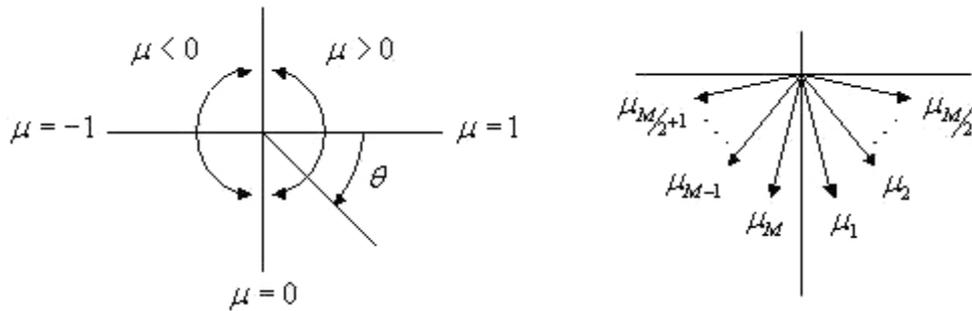


Figura 5: Discretização do domínio angular,  $-1 \leq \mu \leq 1$   
 (SILVA-NETO; BECCENERI, 2009, Cap. 3, p. 24)

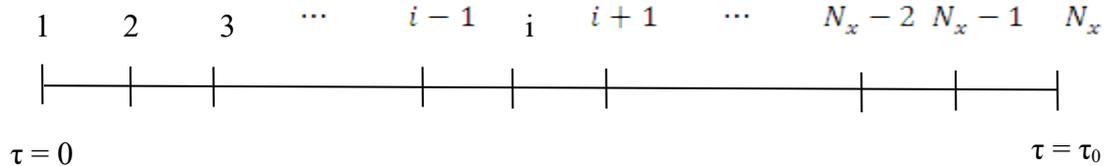


Figura 6: Discretização do domínio espacial para o meio homogêneo unidimensional,  $0 \leq \tau \leq \tau_0$ .

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

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

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)

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

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

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,  $\mu_i$ , e pesos correspondentes,  $\alpha_i$ , para a quadratura Gauss-Legendre com  $M = 20$  (ABRAMOWITZ; STEGUN, 1965).

$\mu$	$\alpha$
$\mu_1 = -\mu_{20} = 0,0765265211$	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

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

e

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

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

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

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

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

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  $i = 1, 2, \dots, N_x - 1$ .

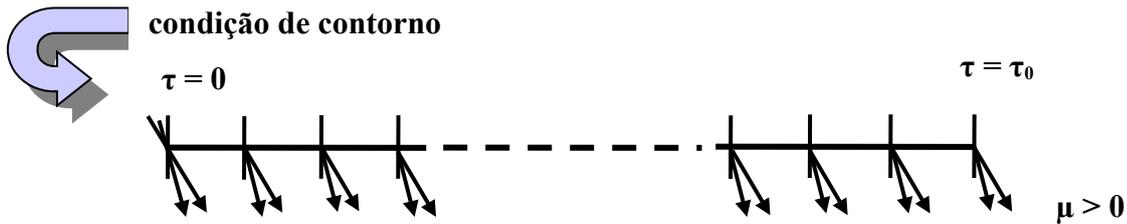


Figura 7: Esquema representativo da marcha para frente.

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  $i = N_x, N_x - 1, \dots, 2$ .

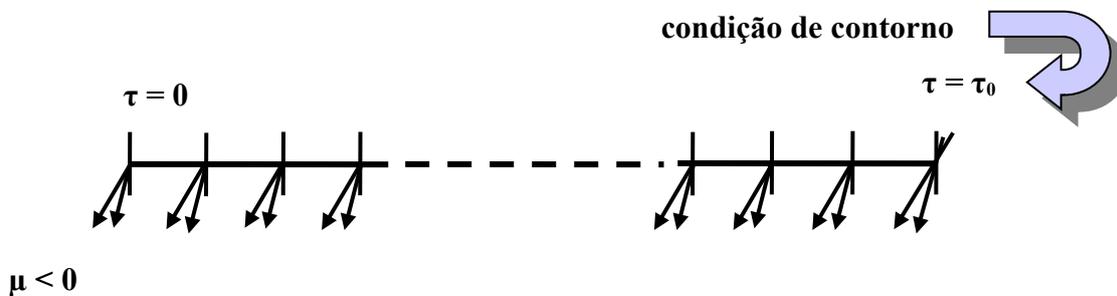


Figura 8: Esquema representativo da marcha para trás.

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  $\varepsilon$  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

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

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

O albedo de espalhamento simples  $\omega$  e as refletividades difusas  $\rho_j$ , com  $j = 1$  ou  $2$ , são grandezas cujos valores estão no intervalo  $[0,1]$ . A espessura ótica  $\tau_0$  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  $\bar{\tau}_0 = 1,0$ . Podemos 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  $\sigma_s$  (SILVA-NETO ; BECCENARI, 2009, Cap. 2, p.23).

### 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 contornos 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.

Considere 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,  $\bar{Y}$ , 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. as 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).

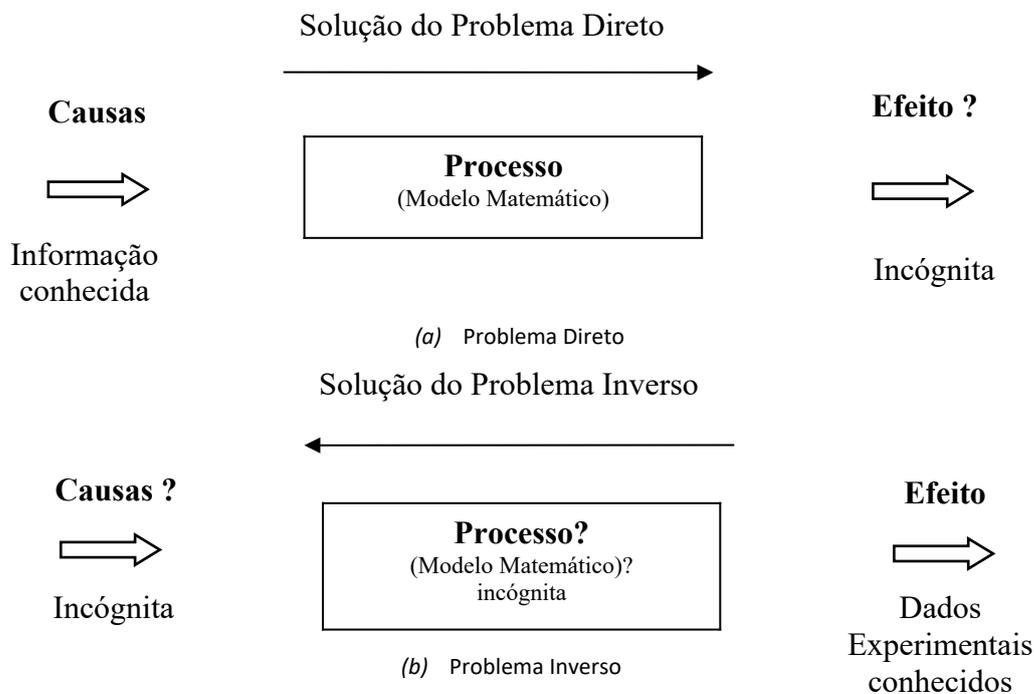


Figura 9: Representação esquemática do problema direto (a) e do problema inverso (b)

[Fonte adaptada de: SILVA-NETO, 2007].

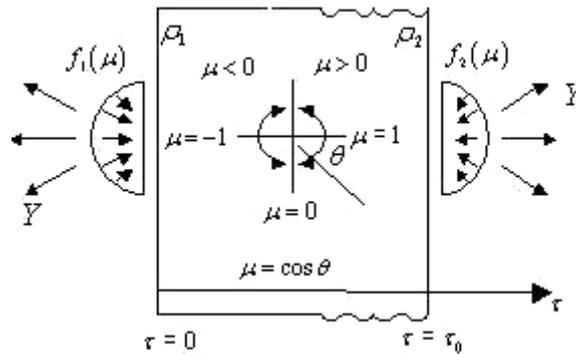


Figura 10: Representação esquemática de um meio participante homogêneo unidimensional e da radiação que emerge do mesmo com intensidade  $\vec{Y}$  e que é medida experimentalmente (SILVA-NETO; CAMPOS-VELHO, 2009, p. 35).

Neste trabalho, como já mencionado anteriormente na seção 1, o nosso 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

$\vec{Y} = \{Y_l\}^T$ ,  $l = 1, 2, \dots, N$ . O superescrito  $T$  é indicativo da transposta.

Como dados experimentais reais não estão disponíveis, é construído um conjunto de dados experimentais sintéticos 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 queremos determinar com a solução do problema inverso,  $\sigma$  simula o desvio padrão dos erros dos dados experimentais, e  $r_l$  é um número pseudo-aleatório gerado no intervalo  $[-1, 1]$ .

## CAPÍTULO 3

### FUNDAMENTOS SOBRE REDES NEURAIS E MÁQUINAS DE COMITÊS 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 comitês, será dissertado um breve histórico e fundamentação teórica sobre redes neurais e máquinas de comitês.

#### **Histórico sobre redes neurais e máquinas de comitês.**

As redes neurais artificiais e as máquinas de comitês de redes neurais artificiais estão ambas inseridas na linha conexionista 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 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, et. al., 2009).

Podemos citar, no contexto, trabalhos de outros pesquisadores de relevada importância na linha conexionista, 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 RNAs e Máquinas de Comitê de Redes Neurais Artificiais, com forte dependência computacional (hardware e software) (WASSERMAN, 1989).

### **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 RNAs 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).

### **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.

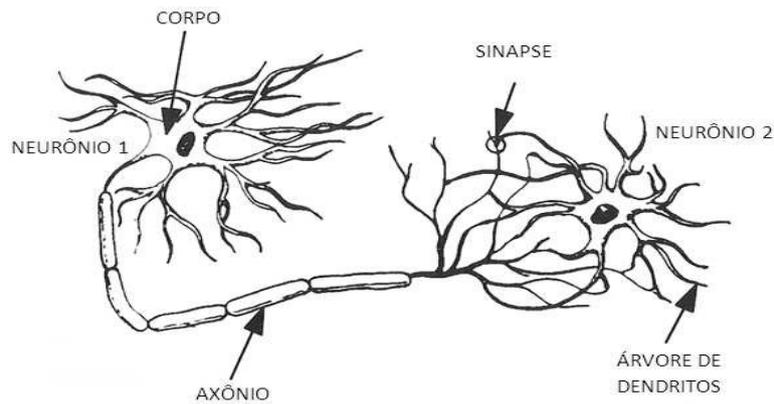


Figura 11: Rede Neural Biológica (Biondi, et. al., 2009, p. 84).

Um neurônio biológico qualquer apresenta a seguinte estrutura: um corpo celular, também chamado de soma, que é responsável pela captação e combinação das informações 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 à 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 et. al., 2009, p. 85).

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

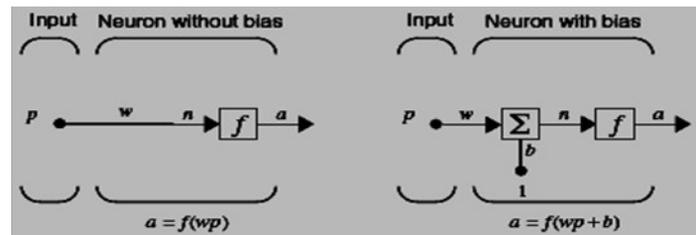


Figura 12: Neurônio artificial com uma única entrada

[Neural Network Toolbox, MATLAB 2008<sup>®</sup>].

(entrada líquida da função de transferência), é então, o único sinal de entrada ou argumento da função de transferência (ou de ativação)  $f$ , que responde com a saída  $a = f(n) = f(wp)$ . À direita da Figura 12, temos 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  $a = f(w \cdot p + b)$ . 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 seqüência são descritas algumas das principais funções de transferência ou função de ativação comumente empregadas no processamento neural artificial.

### Tipos de funções de transferência

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

### 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 classificar decisões.

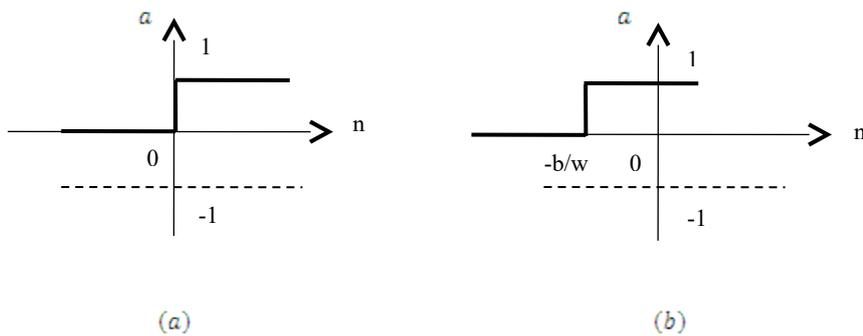


Figura 13: Função de transferência tipo degrau, (a) sem bias e (b) com bias.

### Função linear

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

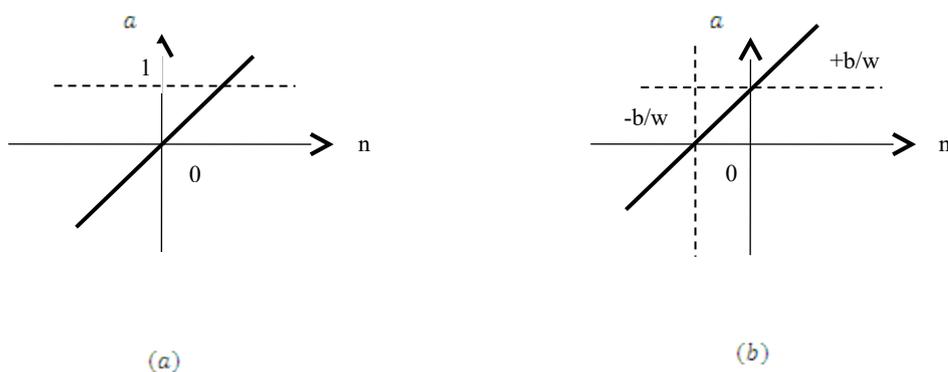


Figura 14: Função de transferência tipo linear, (a) sem bias e (b) com bias.

### Função log-sigmóide

A função log-sigmóide, representada na Figura 15, é definida como uma função estritamente crescente que exibe um comportamento equilibrado entre uma função linear e uma função não-linear

(HAYKIN, 2001, p. 40). Ela é comumente usada em redes neurais com retro-propagação do erro, principalmente devido a 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.

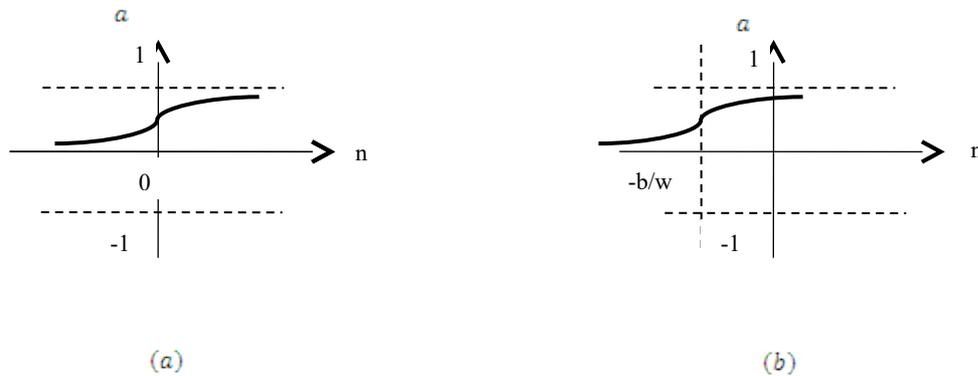


Figura 15: Função de transferência tipo log-sigmóide. (a) sem bias e (b) com bias.

### Função tan-sigmóide

A função de transferência 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  $[-1, 1]$ .

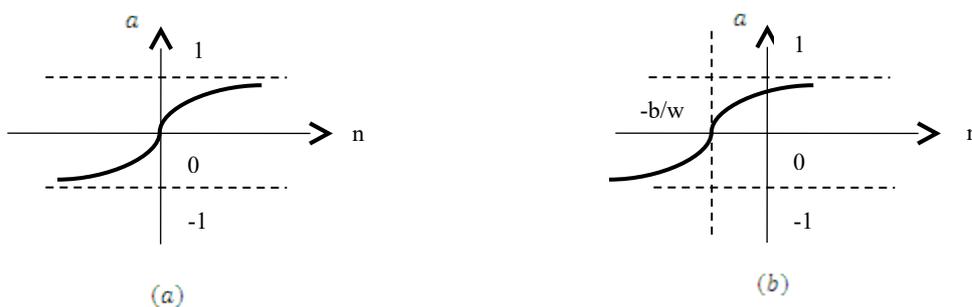


Figura 16: Função de transferência tan-sigmóide: (a) sem bias e (b) com bias.

### O neurônio artificial de entrada vetorial

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

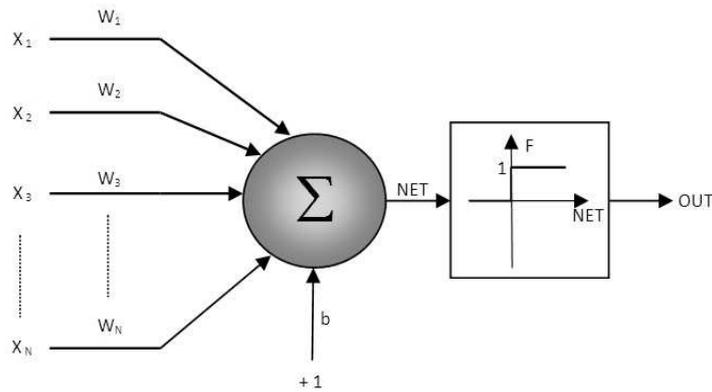


Figura 17: Modelo do Neurônio Artificial de McCulloch e Pitts (Biondi, et. al., 2009, p.87).

McCulloch e Pitts (1943), propuseram um modelo matemático para um neurônio biológico onde o corpo celular é representado por um somatório ( $\Sigma$ ), o axônio por uma função de transferência (ou de ativação) ( $F$ ), geralmente com limiar que pode ser uma das funções descritas na seção terciária 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 inibitórias ou excitatórias (HAYKIN, 2001).

A função NET representa a soma ponderada das entradas  $x_i$ ,  $i = 1, \dots, N$ , pelos respectivos pesos sinápticos ( $NET = \sum x \cdot W + b$ ), 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  $OUT = 0$ .

### 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 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.

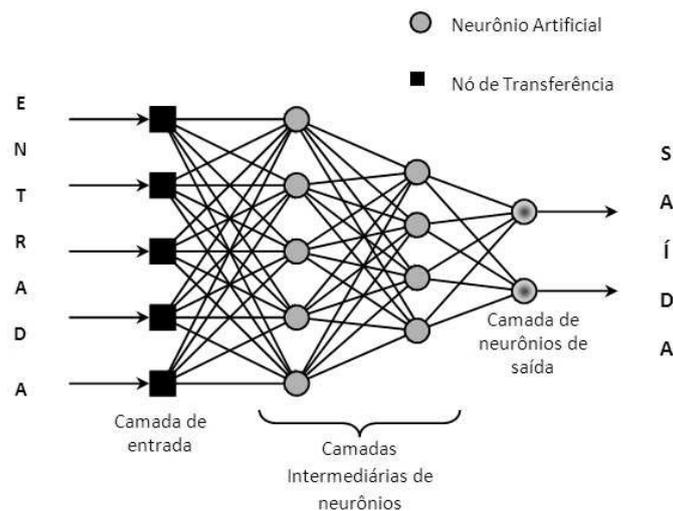


Figura 18: Estrutura de uma rede neural artificial (Biondi, et. al., 2009, p. 86).

O processamento com RNA divide-se em duas fases distintas: fase de treinamento 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 et. al., 2009, p 86).

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 muito bem desenvolvidas na literatura especializada (HAYKIN, 2001). Na seção terciária 3.2.6 discutiremos os dois principais paradigmas de treinamento de RNA.

### **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 na seção terciária 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 supervisionado. A declaração propõe 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.

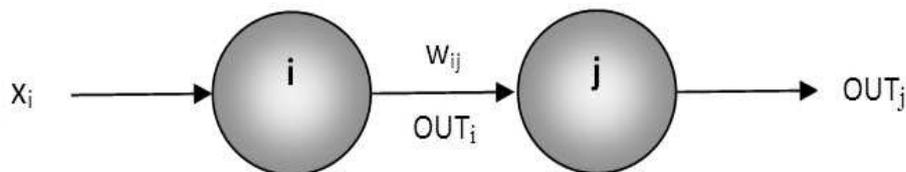


Figura 19: Conexão entre dois neurônios artificiais (Biondi, et. al., 2009, p.89).

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, podemos 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  $\eta$  representa uma constante positiva

denominada taxa de aprendizado e  $(n)$  representa a iteração atual. Com base na regra de Hebb (1949), podemos escrever a equação (26), que permite a atualização dos pesos sinápticos a cada iteração. (BIONDI, et. al., 2009, p.89), (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 exposto, 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 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  $OUT = F(\sum x_t \cdot W + b)$ , sendo, portanto, um processo bastante rápido.

## Perceptron

O Perceptron é a RNA de estrutura mais simples que existe usada com sucesso na classificação de padrões. Proposto por Rosenblatt (1962), 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 (BIONDI, et. al., 2009, p.90), (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 transferência 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  $\Delta W_{ij} = \eta \cdot (T_j - OUT_j) \cdot X_i$ .

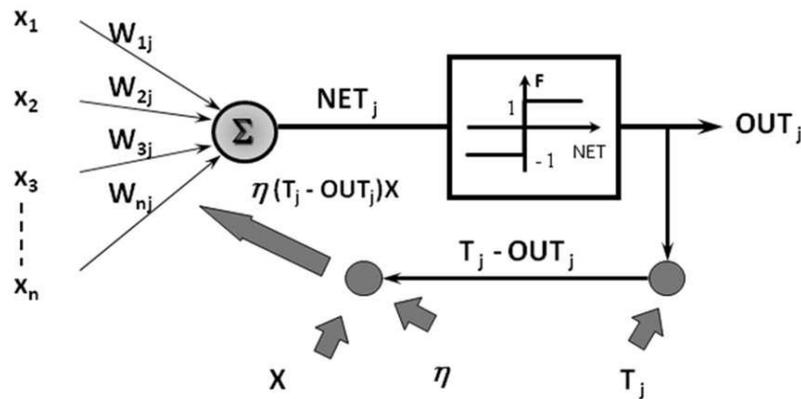


Figura 20: Modelo do Perceptron (Biondi et. al., 2009, p.90).

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  $\eta$  mede o passo a ser adotado pela rede em busca da convergência, i.e., em busca do mínimo global. Se a taxa  $\eta$  for pequena a convergência se torna muito lenta, podendo até paralisar o treinamento devido ao confinamento num mínimo local, ao contrário se a taxa  $\eta$  for muito elevada, o treinamento pode oscilar (BIONDI, et. al., 2009, p.91).

A proposta da criação do Perceptron por Rosenblatt (1962) é a classificação de padrões. Como mencionando anteriormente na seção 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 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.

### **Perceptron de múltiplas camadas**

O Perceptron de Múltiplas Camadas (MLP) é uma generalização da rede Perceptron proposto por Rosenblatt (1962), 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 retro-propagação do erro, ou seja, a regra delta generalizada, regra esta que minimiza o erro usando o método do gradiente descendente.

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.

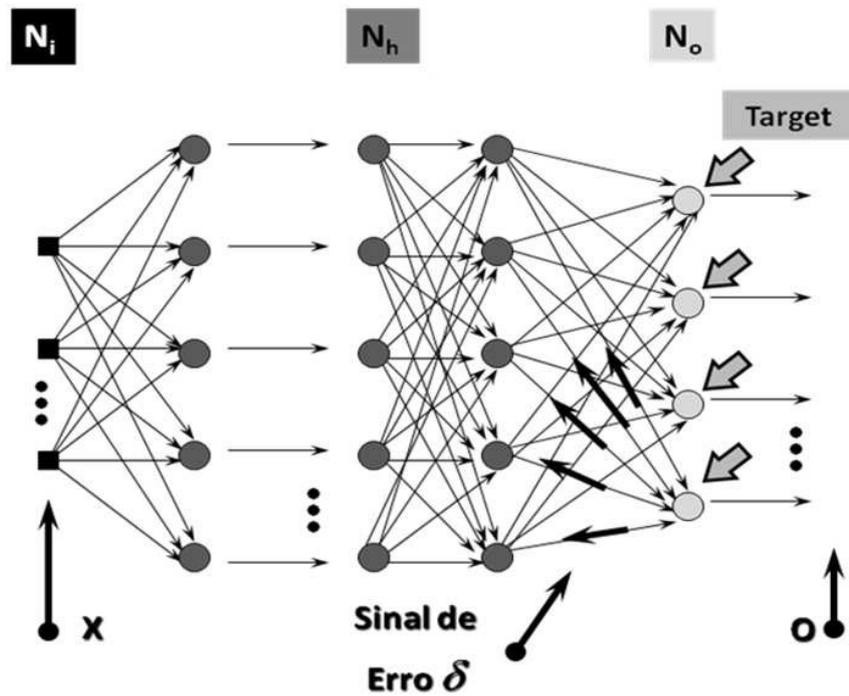


Figura 21: Arquitetura do Perceptron de Múltiplas Camadas (Biondi, et. al., 2009, p. 92).

Os vetores  $\mathbf{X}$ ,  $\mathbf{O}$  e  $\mathbf{T}$  (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 retro-propagação do erro (ver Figura 22), “backpropagation”, BP, é realizado em duas etapas: a primeira é chamada “forward” e a segunda “backward” (JUNIOR; COSTA, 2007). 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. O erro médio quadrático para cada neurônio  $j$  da camada de saída  $J$  e para cada padrão apresentado, é calculado e definido por

$$E = \frac{1}{2} [T_j - O_j]^2 \quad (27)$$

onde  $O_j$  e  $T_j$  são respectivamente a saída calculada e a saída desejada para o neurônio  $j$ , onde  $j = 1, 2, \dots, N_o$ .

A segunda etapa ocorre em sentido contrário. Há uma retro-propagação do sinal de erro  $\delta$  calculado, para cada padrão apresentado, atualizando os pesos e bias das conexões a partir da camada de saída J, propagando-se pelas camadas ocultas até a camada de entrada, pelo uso da regra delta generalizada representada pela equação (28)

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

onde  $\eta$  é a taxa de aprendizado e  $i$  indica indexação referente a camada oculta 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 W de certo vetor de entrada 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).

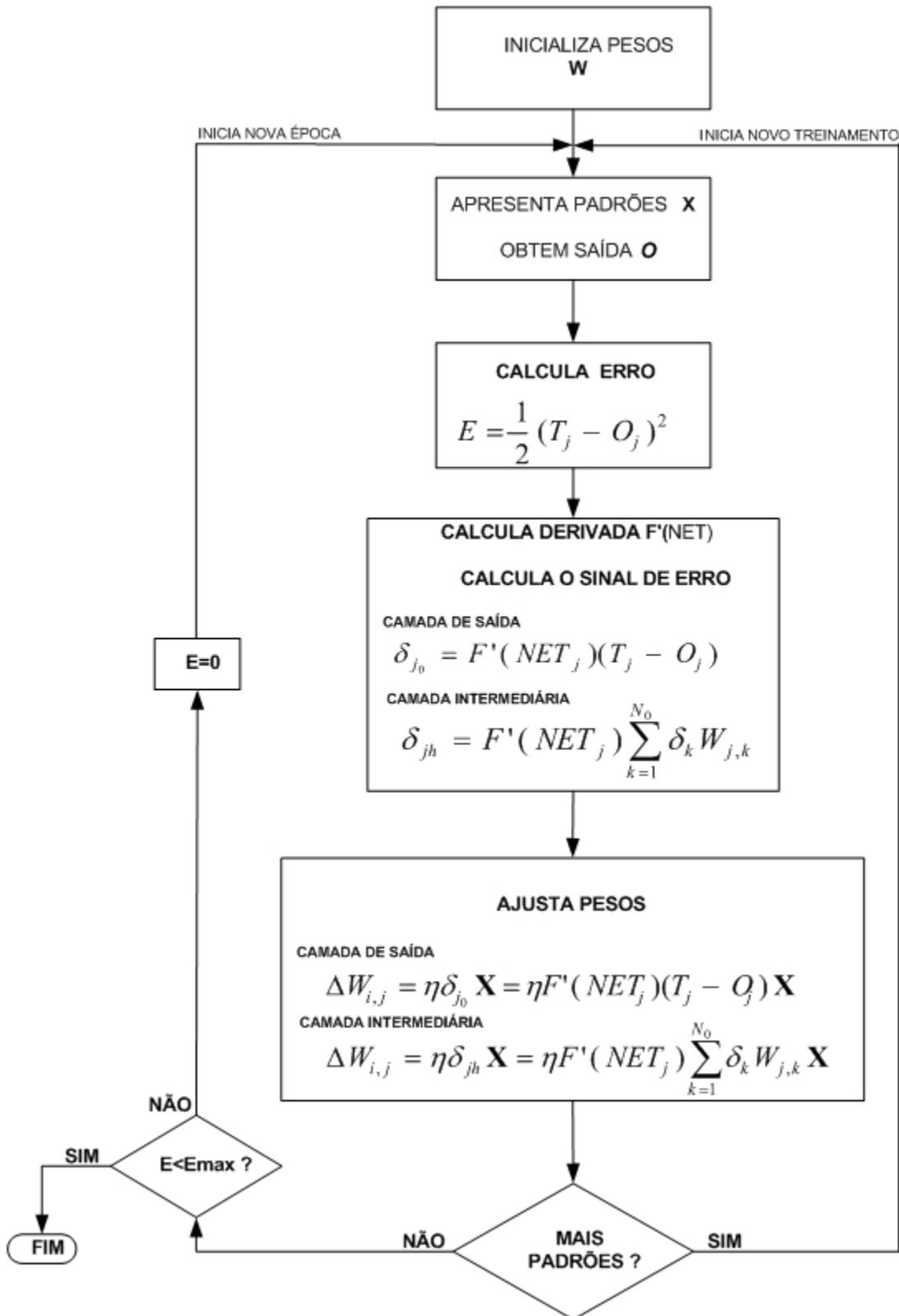


Figura 22: Fluxograma do algoritmo de retro-propagação (BP) (Biondi, et al., 2009, p. 93).

A taxa de aprendizado  $\eta$  é quem 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  $\eta$  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 transferência 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  $\Delta W_{ij}$  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 usada um valor muito alto para a taxa de aprendizado, ocorrerão modificações muito intensas nos pesos, devido a uma alta variação de  $\Delta W_{ij}$ , 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  $\Delta W_{ij}$  de uma iteração para outra.

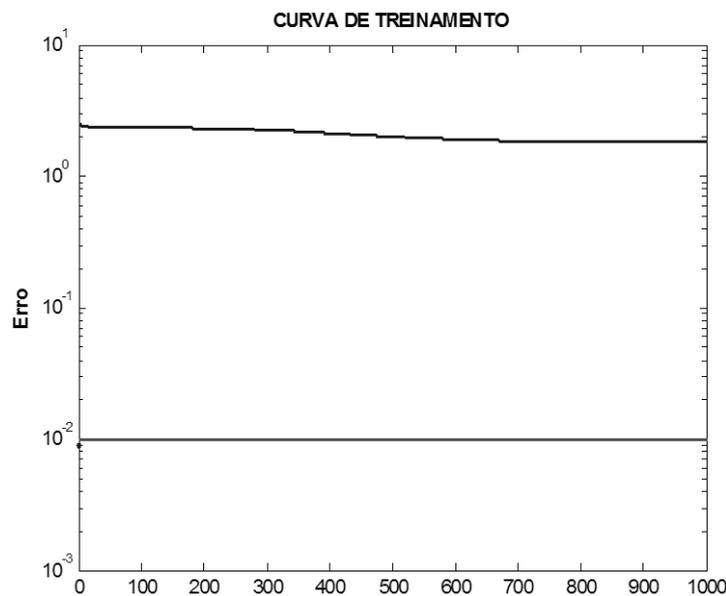


Figura 23: Treinamento com baixa taxa de aprendizado (Biondi, et. al., 2009, p.94).

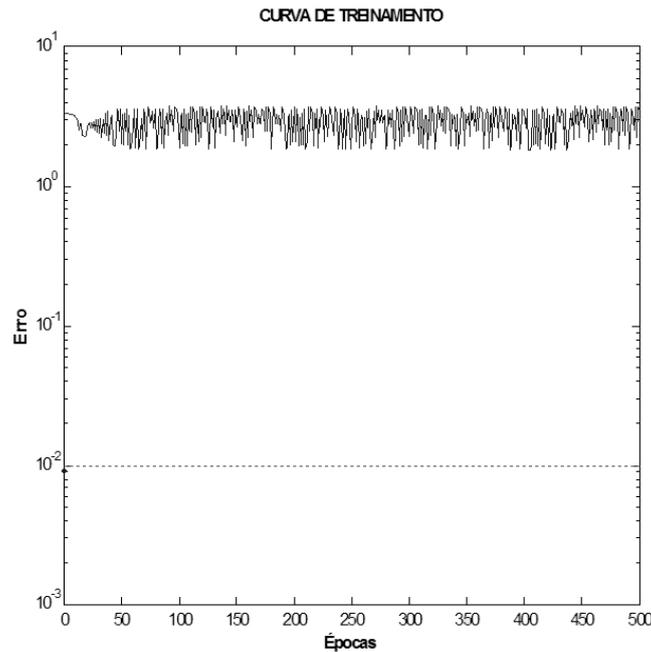


Figura 24: Treinamento com alta taxa de aprendizado (Biondi, et. al., 2009, p.95).

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  $\eta$  sem comprometer a estabilidade da rede. A técnica consiste em amortecer a amplitude de  $\Delta W_{ij}(q)$  a cada iteração após a troca de sinais da derivada  $\partial E(q)/\partial W_{ij}(q)$ , ou seja, após a passagem pelo valor  $W_{ij}(q)$  que resulta em um  $E(q)$  mínimo. A mecânica do amortecimento ocorre quando na iteração em que a parcela  $-\eta \cdot (\partial E(q)/\partial W_{ij}(q))$  muda de sinal, o termo  $\alpha \cdot \eta \cdot \partial E(q-1)/\partial W_{ij}(q-1)$ , ainda preserva o seu antigo sinal e, desta forma,  $\Delta W_{ij}(q)$  é amortecido, gerando um efeito estabilizador e, assim, evitando a oscilação no processo de busca no espaço dos pesos (JUNIOR; COSTA, 2007, p. 58), (BIONDI, et. al., 2009, p.93-95).

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

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

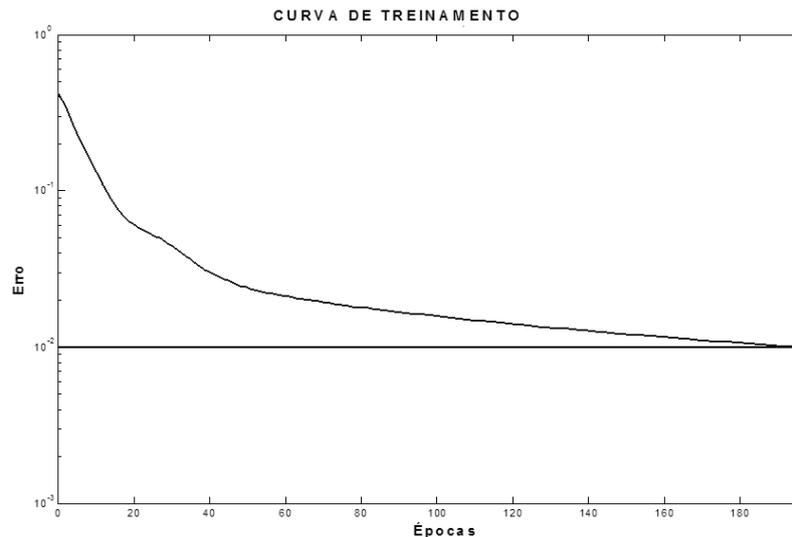


Figura 25: Treinamento com alta taxa de aprendizado e com coeficiente de momento (Biondi, et. al., 2009, p.95).

### Fundamentação teórica sobre máquina 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 princípio “dividir para conquistar”, dividindo problemas computacionalmente complexos em diversos problemas mais simples 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 àquele 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 exatidão das RNAs especialistas e aumentar a sua diversidade, porém sem a garantia de que elas sejam totalmente independentes (ZHOU; WU; TANG, 2002).

### Classificação das máquinas de comitê de redes neurais artificiais

As máquinas de comitês de redes neurais, são aproximadores universais, que podem ser classificadas em duas grandes categorias (HAYKIN, 2001), a saber :

Estruturas Estáticas: Nessa classe de máquinas de comitês 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 de conjunto e o método do reforço. No método da média 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)

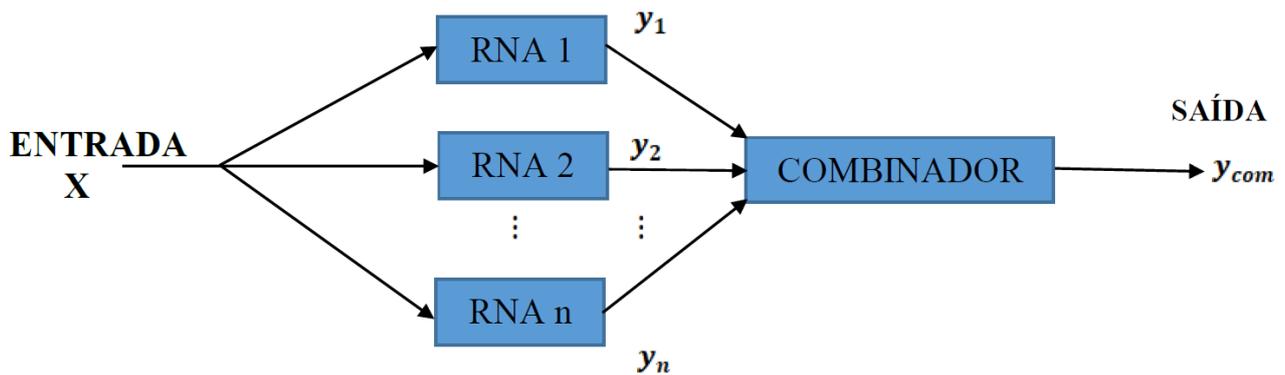


Figura 26: Diagrama em blocos de uma estrutura estática de um comitê usando o método da média.

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).

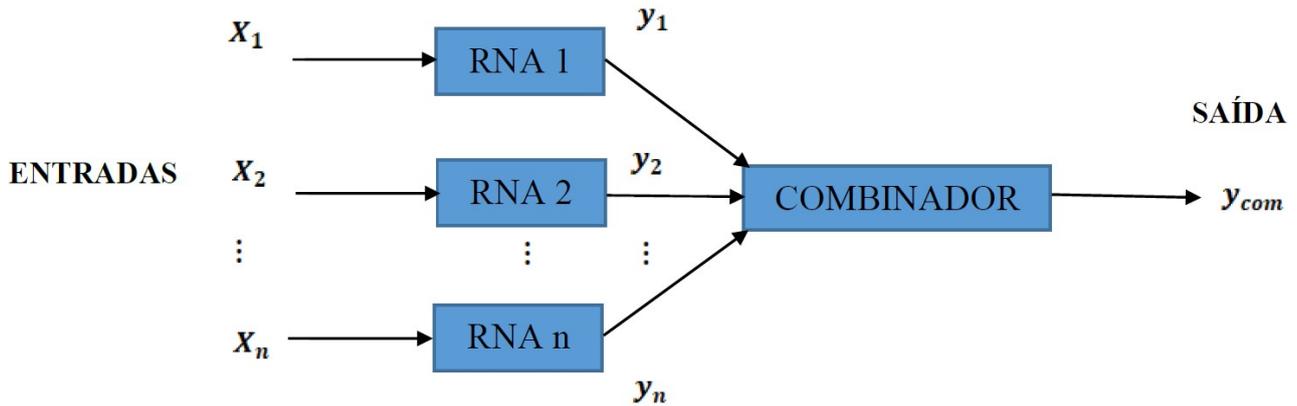


Figura 27: Diagrama em blocos de uma estrutura estática de uma máquina de comitê usando o método do reforço.

Estruturas Dinâmicas: Nesta categoria de máquinas de comitê, cujo método mais conhecido é a mistura de especialistas, o sinal de entrada e as saídas das RNA especialistas são aplicadas a uma estrutura combinadora não linear, que geralmente é uma 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  $NET_i$  representa o produto do vetor peso sináptico pelo vetor de entrada do neurônio especialista. O termo  $NET_j$  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

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

A saída global  $R$  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).

$$R = \sum_{j=1}^i g_i Y_i \quad (31)$$

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.

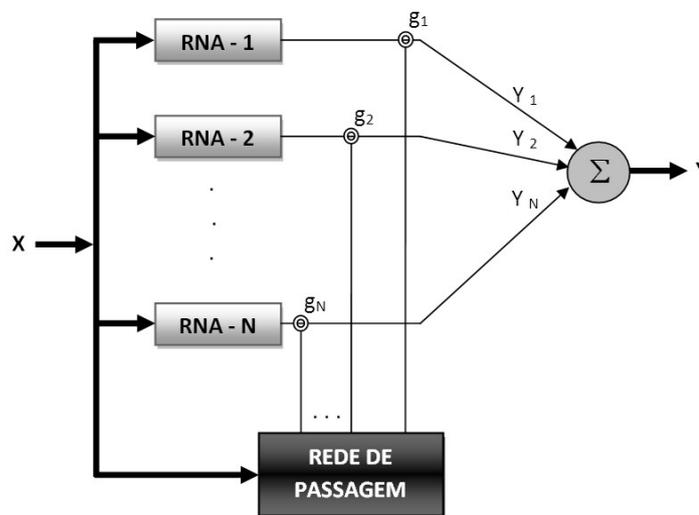


Figura 28: Diagrama de blocos de uma estrutura dinâmica de uma máquina de comitê usando mistura de especialistas (Biondi et al., 2009, p. 98).

### Formulação matemática e solução do problema inverso em transferência radiativa usando redes neurais artificiais

Primeiramente será desenvolvida uma formulação matemática considerando que as redes neurais especialistas que compõem as máquinas de comitês tenham duas camadas ocultas. Esta formulação engloba àquela para redes neurais especialistas com apenas uma camada oculta, cuja formulação será mostrada em seguida.

A Figura 29 representa um MLP, que modela cada uma das redes neurais especialistas, que compõem as máquinas de comitês 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.

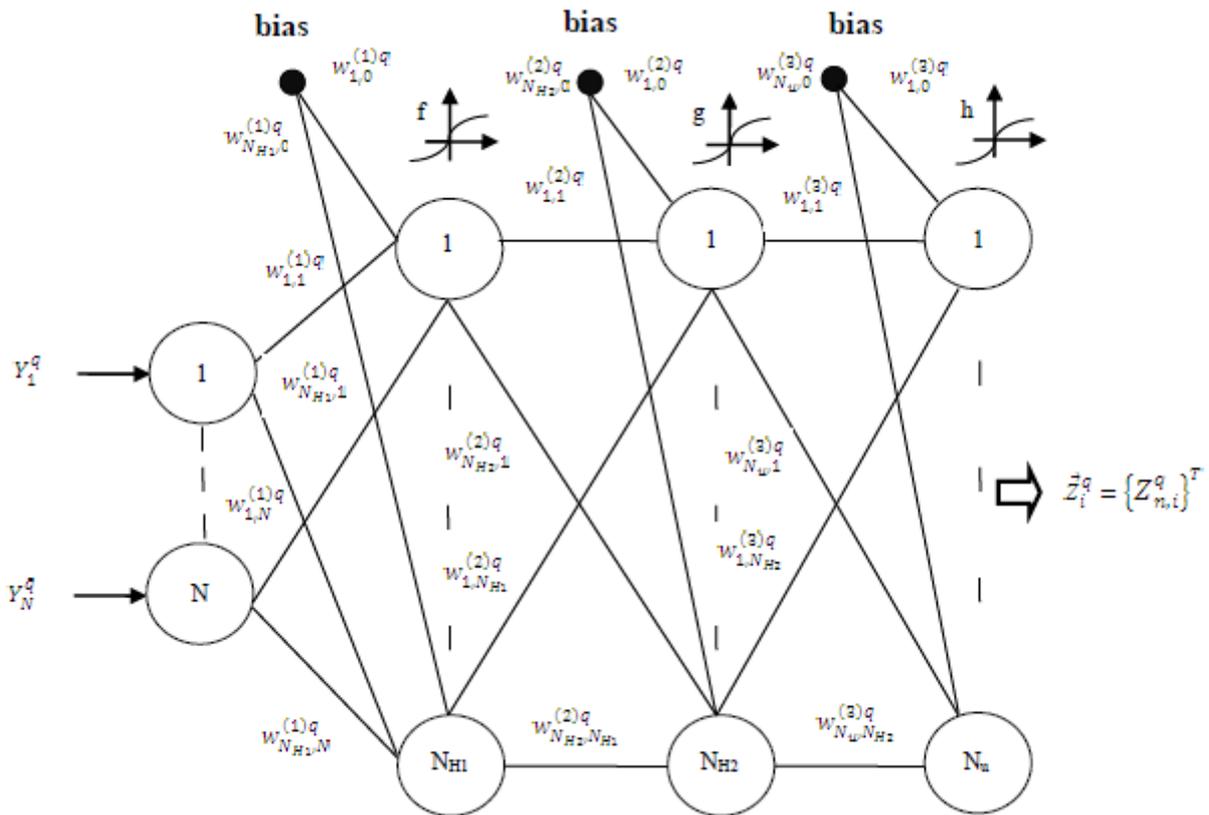


Figura 29: MLP com duas camadas ocultas para o problema inverso.

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 conforme a equação (25) da Seção 2, e fornecidas pela camada de entrada

$$a_m^q = \sum_{l=1}^N w_{m,l}^{(1)q} Y_l^q + w_{m,0}^{(1)q}, \quad m = 1, 2, \dots, N_{H1} \quad (32)$$

onde  $w_{m,l}^{(1)q}$ ,  $m = 1, 2, \dots, N_{H1}$ ,  $l = 1, 2, \dots, N$ , 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

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

onde  $f(\cdot)$  é uma função de ativação. Várias escolhas para a função  $f(\cdot)$  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

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

onde  $w_{k,m}^{(2)q}$ ,  $k = 1, 2, \dots, N_{H2}$ ,  $m = 1, 2, \dots, N_{H1}$ , 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

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

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

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

onde  $w_{n,k}^{(3)q}$ ,  $n = 1, 2, \dots, N_u$ ,  $k = 1, 2, \dots, N_{H2}$ , 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), na seção 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

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

Considerando disponíveis os dados experimentais  $Y_l^q$ ,  $l = 1, 2, \dots, N$ ,  $q = 0, \dots, Q-1$ , observamos na equação (38) que  $t_n^q$ ,  $n = 1, 2, \dots, N_u$ , componentes dos vetores  $\vec{t}_i^q$ , são as estimativas 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<sub>*i*</sub>,  $i = 1, 2, \dots, M$ , aqui inicialmente  $M = 10$  e finalmente  $M = 6$ . Mas antes que possamos usar a equação (38) devemos determinar os pesos sinápticos  $w^{(1)q}, w^{(2)q}$  e  $w^{(3)q}$ .

A determinação dos pesos  $w^{(1)q}, w^{(2)q}$  e  $w^{(3)q}$  é feita pela apresentação de um conjunto de padrões: entradas conhecidas  $Y_l^q$  e saídas desejadas  $\vec{Z}_{exato}$ . Calculam-se então, os pesos que produzam o melhor ajuste entre os valores  $\vec{t}_i^q \equiv \vec{Z}_i^q$  estimados e o vetor de saídas desejado,  $\vec{Z}_{exato}$ . Os padrões usados no treinamento supervisionado das RNA são gerados calculando-se os valores  $I_{calcul}(\vec{Z}_{exato})$  a partir do vetor exato de parâmetros radiativos  $\vec{Z}_{exato}$  com as ordenadas discretas e a solução por diferenças finitas mencionados na seção 2.1.2, e aplicando a equação (25) (CHANDRASEKHAR, 1950) e (SMITH, 1985).

Para a determinação de  $w^{(1)q}, w^{(2)q}$  e  $w^{(3)q}$  é usado o algoritmo de retro-propagação. Parte-se de uma estimativa inicial para os pesos  $w^{(1)q}, w^{(2)q}$  e  $w^{(3)q}$ , com  $q = 0$ , e o vetor das entradas  $\vec{Y}^q = \{Y_l^q\}^T$  é passado à frente através da rede produzindo as estimativas  $\vec{t}_i^{q=0} \equiv \vec{Z}_i^{q=0}$  para as saídas

(parâmetros radiativos), que serão comparadas com as saídas desejadas  $\vec{Z}_{\text{exato}}$  produzindo os erros médios quadráticos,

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

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

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

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

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

onde

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

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

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

são gradientes locais de retro-propagação 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  $\eta$  representa a taxa de aprendizado, que pode assumir diferentes valores entre 0 e 1. Aqui  $\eta = 0,4$ , 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.

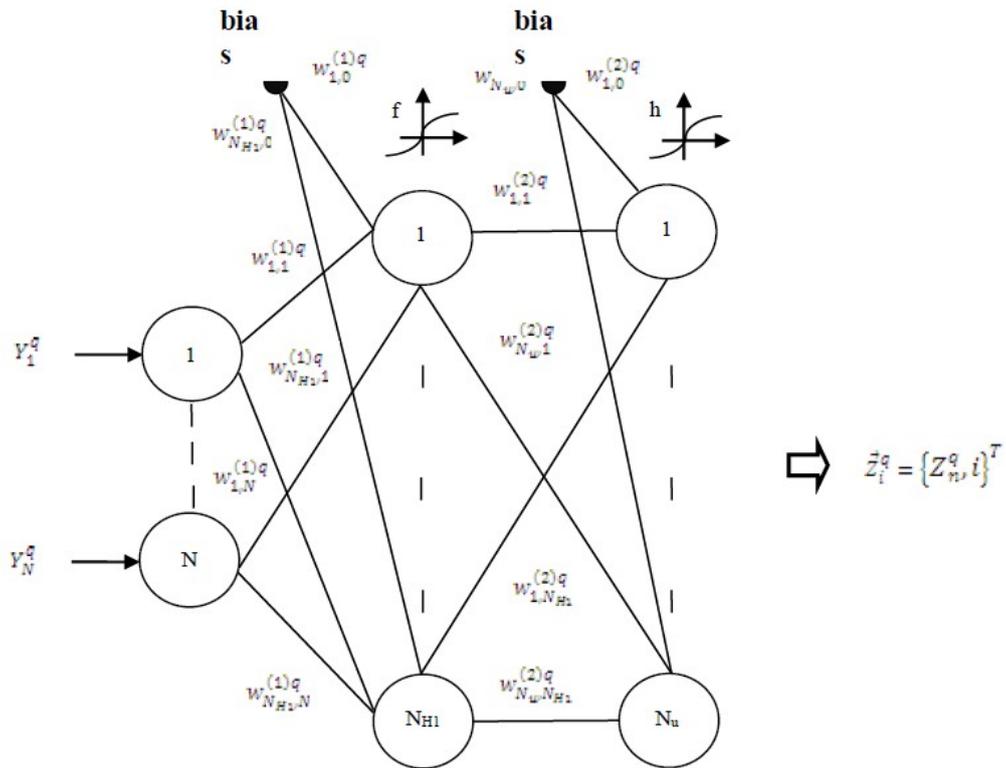


Figura 30 : MLP com uma camada oculta para o problema inverso.

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,  $\text{RNA}_i$ ,  $i = 1, 2, \dots, M$ , terão a seguinte formulação

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

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

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

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

onde

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

com

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

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

onde as equações (43c) e (43e) são os gradientes locais, de retro-propagação 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.

### **Modelagem e formulação matemática para a solução do problema inverso em transferência radiativa usando máquina de comitê de redes neurais artificiais**

A modelagem e a formulação do problema inverso em transferência radiativa utilizando máquina de comitê são apresentadas nesta seção secundária, tendo como modelo básico uma máquina de comitê do tipo estática que utiliza a média de conjunto de acordo com a configuração mostrada na Figura 1, contida na seção primária 1 desta dissertação.

### **Modelagem do problema inverso em transferência radiativa com máquina de comitê de redes neurais artificiais**

Quanto à modelagem têm-se duas configurações que são aqui consideradas: a primeira é o modelo clássico da máquina do tipo estática com a média de conjunto, representado na Figura 31, onde o combinador combina as saídas das RNA especialistas através de sua média aritmética (combinação linear), que será então a saída final da máquina de comitê que será confrontada com o vetor das saídas desejadas (alvo); 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 a saída final da máquina de comitê que será confrontada com o alvo.

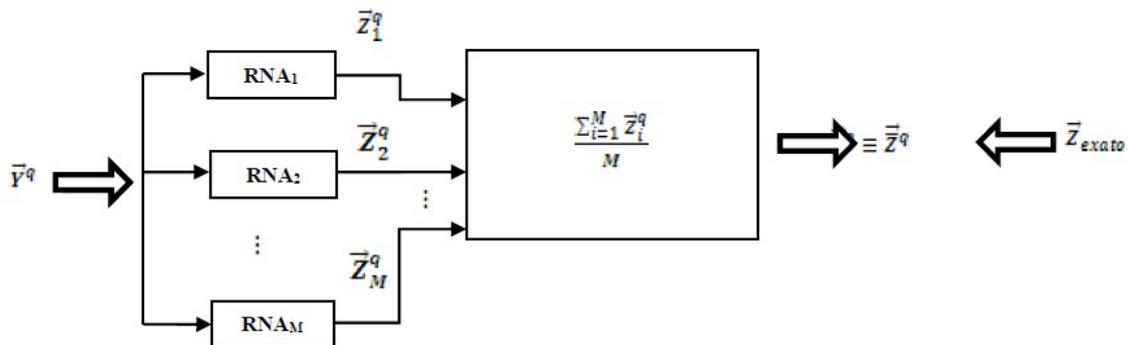


Figura 31: Modelo da máquina de comitê tipo estática onde o combinador efetua apenas a média aritmética das saídas das RNA especialistas.

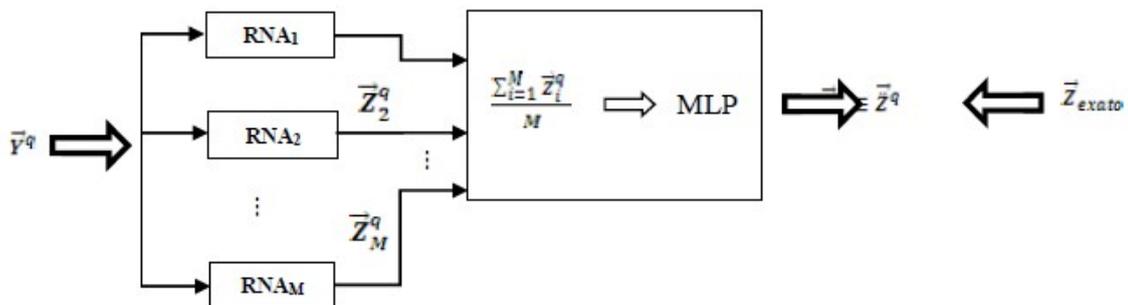


Figura 32: Modelo da máquina de comitê tipo estática com MLP integrado ao combinador.

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, têm topologias idênticas com mesmo número de camadas ocultas, mesmo número de neurônios nas camadas ocultas, mesmas funções de treinamento e de transferência (ou de ativação), e mesma taxa de aprendizado  $\eta = 4$ . 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.

### Formulação do problema inverso em transferência radiativa com máquina de comitê de redes neurais artificiais

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, temos que os vetores  $\vec{Z}_i^q$ , que são as saídas das redes neurais especialistas  $RNA_i$ ,  $i = 1, 2, \dots, M$ , cujas componentes  $Z_{n,i}^q$ ,  $n = 1, 2, \dots, N_u$  são as estimativas para as componentes do vetor  $\vec{Z} \equiv \vec{Z}^q$  da equação (24), são combinados linearmente por meio de uma média aritmética a cada padrão  $q$  apresentado de acordo com a equação (44) e, então o vetor  $\vec{Z}^q$ , obtido através da equação (44), será a saída da máquina de comitê representada de acordo com o modelo da Figura 31.

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

O vetor  $\vec{Z}^q$  é um vetor coluna cujos elementos  $Z_n^q$ ,  $n = 1, 2, \dots, N_u$ , aqui  $N_u = 4$ , é cada um deles a média aritmética dos correspondentes componentes  $Z_{n,i}^q$  dos vetores  $\vec{Z}_i^q$ ,  $i = 1, 2, \dots, M$ , das estimativas dos parâmetros radiativos obtidos pelas redes neurais especialistas a cada padrão  $q$  apresentado.

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 do MLP de saída do combinador da máquina de comitê de redes neurais 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  $Z_n^q$ ,  $n = 1, 2, \dots, N_u$ , fornecidos pela média aritmética das saídas das redes neurais especialistas  $RNA_i$ ,  $i = 1, 2, \dots, M$ , obtida através da equação (44)

$$Z_u^q = \sum_{n=1}^{N_u} w_{u,n}^{q(h^*)} Z_n^q + w_{u,0}^{q(h^*)}, \quad u = 1, 2, \dots, N_h, \quad n = 1, 2, \dots, N_u \quad (45)$$

onde  $w_{u,n}^{(h^*)q}$ ,  $u = 1, 2, \dots, N_h$ ,  $n = 1, 2, \dots, N_u$ ,  $q = 0, 1, 2, \dots, Q-1$ , (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 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)

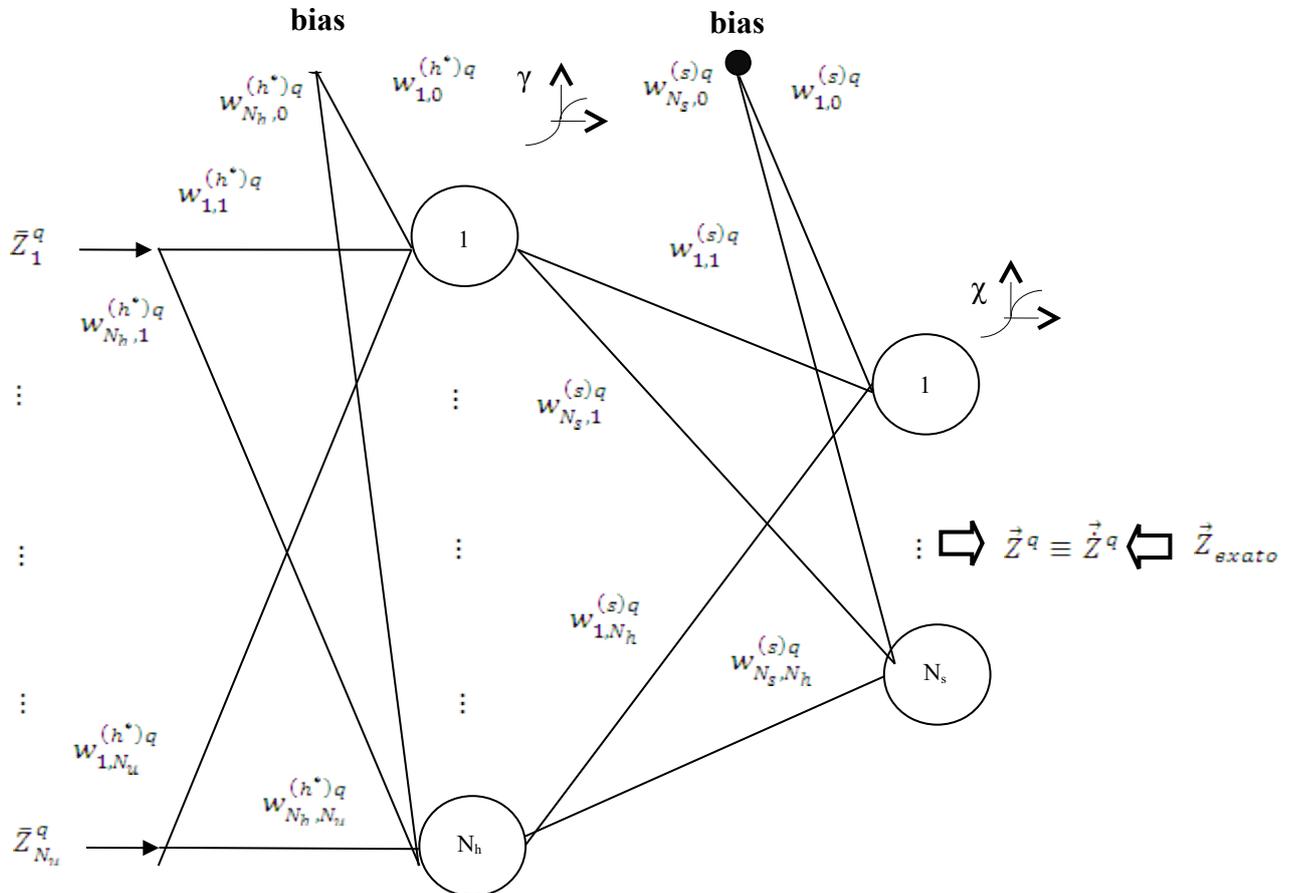


Figura 33: Modelo do MLP do combinador da máquina de comitê.

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 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  $h^*$

$$\hat{Z}_j^q = \sum_{u=1}^{N_h} w_{j,u}^{(s)q} \hat{Z}_u^q + w_{j,0}^{q(s)}, \quad j = 1, 2, \dots, N_s, u = 1, 2, \dots, N_h \quad (47)$$

onde  $w_{j,u}^{(s)q}$ ,  $j = 1, 2, \dots, N_s$ ,  $u = 1, 2, \dots, N_h$ , 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

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

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

As respostas  $\vec{Z}_j^q$ ,  $j = 1, 2, \dots, N_s$ , 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  $\vec{Z}^q \equiv \vec{Z}^q$ , que é comparado com vetor alvo  $\vec{Z}_{exato}$  a cada iteração  $q$ ,  $q = 0, 1, 2, \dots, Q-1$ .

Para a determinação de  $w_{u,n}^{(h^*)q}$  e  $w_{j,u}^{(s)q}$  o procedimento é o mesmo que o utilizado para a obtenção dos pesos  $w^{(1)q}$ ,  $w^{(2)q}$  e  $w^{(3)q}$ , 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 erros médios quadráticos são

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

onde  $Z_{j,exato}$ ,  $j = 1, 2, \dots, N_s$  são os componentes do vetor alvo  $\vec{Z}_{exato}$ , constante para qualquer padrão  $q$ .

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

$$w_{j,u}^{(s)q+1} = w_{j,u}^{(s)q} + \eta \delta_j^{(s)q} \hat{z}_u^q \quad (50a)$$

$$w_{u,n}^{(h^*)q+1} = w_{u,n}^{(h^*)q} + \eta \delta_u^{(h^*)q} \bar{z}_n^q \quad (50b)$$

onde

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

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

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

A taxa de aprendizado  $\eta$ , pode assumir diferentes valores entre 0 e 1. Aqui  $\eta = 0,4$  e  $\chi'$  e  $\gamma'$  representam as derivadas das funções de ativação.

## CAPÍTULO 4

### RESULTADOS E DISCUSSÃO

#### Metodologia aplicada

Os resultados aqui mostrados foram obtidos a partir dos testes de generalização realizados de forma separada e independente, usando o mesmo programa computacional, para os dois modelos de máquina de comitê representados nas Figuras 31 e 32, no capítulo 3, ambos utilizando as mesmas saídas das RNA especialistas como entradas para o combinador de cada uma.

Os resultados obtidos nos testes de generalização realizados são mostrados através de tabelas e gráficos gerados a partir de quatro 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, todas do tipo MLP, cada qual com uma camada oculta;
2. Os resultados obtidos com uma máquina de comitê do tipo estática tendo integrado ao seu combinador uma rede neural do tipo MLP com uma camada oculta cujas excitações de entrada são a média aritmética das saídas das RNA especialistas, todas do tipo MLP, cada qual com uma camada oculta;
3. Os resultados obtidos com uma máquina de comitê do tipo estática cujo combinador realiza a média aritmética das saídas das RNA especialistas, todas do tipo MLP, cada qual com duas camadas ocultas; e
4. Os resultados obtidos com uma máquina de comitê do tipo estática tendo integrado ao seu combinador uma rede neural do tipo MLP com uma camada oculta cujas excitações de entrada são a média aritmética das saídas das RNA especialistas, todas do tipo MLP, cada qual com duas camadas ocultas.

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)

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

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  $Y_i = I_{exp_i} = I_{calc_i}(\vec{Z}_{exato}) + \sigma r_i$ , tem-se:  $r_i \in [-1, 1]$  e o desvio padrão  $\sigma = 0,002$ ;
- Taxa de aprendizado:  $\eta = 0,4$ ;
- Número de neurônios nas camadas ocultas das RNA especialistas:  $N_{H1} = N_{H2} = 8, 10, 12, 14, 16, \text{ ou } 18$ ;
- 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 a Figura 32:  $N_h = 12$ , para todos os testes realizados;
- 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  $\leq 3000$ ;
- Limitantes superiores estabelecidos para o erro médio quadrático (E):  $tol < 10^{-3}$  ou  $tol < 10^{-4}$  (testou-se também com  $10^{-5}$ ).

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  $Y_i$  usados nos testes de generalização, e na geração randômica dos padrões de treinamento  $\{\vec{Y}_q, \vec{Z}_q\}$ , foram utilizados programas e subprogramas do compilador 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 (ver Figura 33) da máquina de comitê, descrita no modelo da Figura 32, foram treinados usando o algoritmo da retro-propagação do erro, o Backpropagation, baseado no método do gradiente decrescente, que minimiza o erro médio quadrático E a cada padrão q apresentado, de acordo com a otimização de

Levenberg-Marquardt, usando a função de treinamento `trainlm`, integrante do Neural Network Toolbox do software MATLAB®. Foram usadas funções de transferência (ou de ativação ou restritiva) tipo tangente hiperbólica (*tansig*), para as camadas ocultas e a linear (*purelin*) para a camada de saída, ambas também integrantes do Neural Network Toolbox do software MATLAB®.

Foram treinadas e depois testadas quanto à generalização, de forma separada e independente, máquinas de comitês de configurações em conformidade com os dois modelos mostrados nas Figuras 31 ou 32, cada uma composta por dez RNA especialistas, tendo cada uma das RNA especialistas uma ou duas camadas ocultas, de acordo com o modelo de MLP mostrado na Figura 30 ou com o modelo do MLP mostrado na Figura 29, sendo usada em cada máquina de comitê apenas um único modelo para todas as RNA especialistas, não havendo combinação de ambos os modelos para a mesma máquina. Todas as RNA especialistas de uma mesma máquina de comitê possuem mesma topologia, mesmas funções de treinamento, mesmas funções de ativação, mesmo número de camadas ocultas, mesmo número de neurônios por camada, etc, ou seja, são absolutamente idênticas.

Para cada modelo de máquina de comitê e para cada valor do albedo de espalhamento simples ( $\omega = 0,1; 0,5$  ou  $0,9$ ), 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, o tempo de treinamento, o número de épocas completadas durante o 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  $ER_{\rho_2}$ , para cada rede neural especialista, com grifo para as seis com os melhores resultados;

3. Tabela contendo as estimativas finais 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.

As figuras mostradas são dos gráficos das curvas de performance obtidas durante os testes de validação, geradas durante o processo de treinamento com o Neural Network Toolbox do software MATLAB®, para a rede neural especialista (com uma ou duas camadas ocultas) de cada modelo de máquina de comitê, que apresenta o menor tempo computacional e o menor número de épocas para  $N_{H1}$  e/ou  $N_{H2} = 14$ , e  $tol = 10^{-3}$  ou  $10^{-4}$ , e para a rede neural do combinador de saída da máquina de comitê com modelo de acordo com a Figura 32.

Serão mostrados, por meio de tabelas e gráficos, na próxima seção secundária apenas os melhores resultados das estimativas para o vetor solução  $\vec{Z} = \{\tau_0, \omega, \rho_1, \rho_2\}$ , obtidas com as RNA especialistas (com uma e/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  $M = 6$ , considerando o limite superior para a tolerância requerida para o erro médio quadrático  $E$  igual a  $10^{-3}$  ou  $10^{-4}$ . Os demais casos serão mostrados em tabelas e gráficos, no Apêndice A, de acordo com o procedimento abaixo descrito.

As tabelas e gráficos mostrando os resultados são apresentados na seguinte ordem:

1. Primeiro para as máquinas de comitês efetuando apenas a média aritmética no combinador e constituídas por RNA especialistas com uma ou duas camadas ocultas;
2. Em seguida, da mesma forma, para máquinas de comitês que combinam a média aritmética mais uma rede neural tipo MLP (ver Figura 33) no combinador de saída.

Foram efetuados testes de generalização com número de neurônios nas camadas ocultas das RNA especialistas variando de 8 a 18 neurônios (em múltiplos de dois), os melhores resultados foram os obtidos com 14 neurônios. Portanto, devido ao número elevado de resultados, a partir deste ponto os resultados apresentados foram os obtidos com redes neurais com 14 neurônios nas suas camadas ocultas.

### **Comparativo dos resultados obtidos com as RNA especialistas e com as máquinas de comitês**

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  $10^{-4}$ , modelos de máquina de comitê conforme representados nas Figuras 31 e 32, e  $M = 10$  ou  $M = 6$ . São mostrados também, na primeira coluna da Tabela 2, os valores esperados para os 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

$N_{H1}$  e/ou  $N_{H2}=14$ ;  $\omega = 0,1; 0,5$  ou  $0,9$  e  $tol = 10^{-3}$  ou  $10^{-4}$ .

$\overline{ER}$ (%)	$N_H$	$tol$	$\omega$	$ER$ (%)	$ER$ (%)	$ER$ (%)	$ER$ (%)
				$MA$ $M=10$	$MA$ $M=6$	$MLP$ $M=10$	$MLP$ $M=6$
$\tau_0 = 1,17$	01	$10^{-3}$	0,1	0,76	0,18	0,11	0,75
$\omega = 17,1$	01	$10^{-3}$	0,1	8,2	6,6	1,2	2,4
$\rho_1 = 7,21$	01	$10^{-3}$	0,1	4,34	2,74	4,18	3,36
$\rho_2 = 4,46$	01	$10^{-3}$	0,1	7,78	3,92	3,78	3,92
$\tau_0 = 1,70$	01	$10^{-3}$	0,5	1,64	1,89	1,38	0,89
$\omega = 1,90$	01	$10^{-3}$	0,5	1,32	0,94	1,46	1,34
$\rho_1 = 6,05$	01	$10^{-3}$	0,5	5,14	1,96	0,7	1,96
$\rho_2 = 3,13$	01	$10^{-3}$	0,5	2,88	1,14	0,62	1,14
$\tau_0 = 13,42$	01	$10^{-3}$	0,9	13,42	13,15	8,47	8,05
$\omega = 2,14$	01	$10^{-3}$	0,9	0,58	1,03	0,76	1,03
$\rho_1 = 6,68$	01	$10^{-3}$	0,9	5,14	3,0	2,74	4,8
$\rho_2 = 1,90$	01	$10^{-3}$	0,9	0,46	1,22	1,32	0,46
$\tau_0 = 0,93$	02	$10^{-3}$	0,1	0,22	0,23	5,94	5,98
$\omega = 4,23$	02	$10^{-3}$	0,1	12,5	12,7	32,1	32,1
$\rho_1 = 2,63$	02	$10^{-3}$	0,1	2,18	0,68	0,56	2,16
$\rho_2 = 4,67$	02	$10^{-3}$	0,1	4,66	4,12	6,6	6,06
$\tau_0 = 1,71$	02	$10^{-3}$	0,5	1,68	1,45	2,53	2,29
$\omega = 1,84$	02	$10^{-3}$	0,5	0,04	0,9	2,9	3,7
$\rho_1 = 2,3$	02	$10^{-3}$	0,5	1,18	1,38	1,06	1,22
$\rho_2 = 3,79$	02	$10^{-3}$	0,5	3,24	2,18	3,4	2,34
$\tau_0 = 7,05$	02	$10^{-3}$	0,9	7,5	7,09	7,45	7,09
$\omega = 1,65$	02	$10^{-3}$	0,9	0,34	0,38	0,22	0,12
$\rho_1 = 5,26$	02	$10^{-3}$	0,9	4,9	3,24	1,78	0,14
$\rho_2 = 2,16$	02	$10^{-3}$	0,9	1,38	1,26	4,46	4,32
$\tau_0 = 0,98$	02	$10^{-4}$	0,1	0,98	1,13	0,85	1,13
$\omega = 9,37$	02	$10^{-4}$	0,1	9,1	5,0	4,6	5,0
$\rho_1 = 6,53$	02	$10^{-4}$	0,1	6,54	6,68	6,38	6,68
$\rho_2 = 3,45$	02	$10^{-4}$	0,1	3,44	3,64	2,24	3,64
$\tau_0 = 1,37$	02	$10^{-4}$	0,5	1,37	0,99	1,35	2,29
$\omega = 0,96$	02	$10^{-4}$	0,5	0,8	0,92	0,9	3,7
$\rho_1 = 1,12$	02	$10^{-4}$	0,5	0,4	0,42	0,14	1,22
$\rho_2 = 1,65$	02	$10^{-4}$	0,5	1,8	1,4	2,12	2,34
$\tau_0 = 3,89$	02	$10^{-4}$	0,9	3,89	2,97	3,89	2,98
$\omega = 0,91$	02	$10^{-4}$	0,9	0,21	0,0	0,11	0,09
$\rho_1 = 1,87$	02	$10^{-4}$	0,9	0,94	0,94	0,04	0,04
$\rho_2 = 1,43$	02	$10^{-4}$	0,9	0,7	0,54	1,18	1,06

Verifica-se na Tabela 2 que os melhores resultados obtidos para as estimativas de  $\vec{z}$  com RNA especialistas, para  $\omega = 0,1$ , do ponto de vista de precisão, considerando o valor esperado  $\overline{ER}$  para o erro relativo referente a cada parâmetro radiativo, foi o caso  $N_H = 2$  e  $tol = 10^{-3}$ , conforme mostrado nas Tabelas 3 e 4.

Tabela 3: Estimativas obtidas para  $N_{H1} = N_{H2} = 14$ ;  $\omega = 0,1$  e  $tol = 10^{-3}$ .

RNA	tempo	época	$\tau_0$	$\omega$	$\rho_1$	$\rho_2$
		s				
1	0:03:23	33	0,9981	0,0796	0,5288	0,5269
2	0:02:52	30	1,0044	0,0756	0,5075	0,5215
3	0:02:12	23	0,9922	0,0772	0,5143	0,5183
4	0:03:02	32	1,0170	0,0812	0,4987	0,5223
5	0:02:23	25	1,0061	0,0854	0,5176	0,5293
6	0:02:51	30	0,9877	0,1012	0,5309	0,5230
7	0:02:34	35	1,0081	0,0983	0,5000	0,5220
8	0:02:54	40	0,9845	0,0902	0,5099	0,5196
9	0:02:23	31	1,0001	0,0849	0,5112	0,5304
10	0:02:39	36	0,9801	0,1011	0,4900	0,5200

Tabela 4: Erros relativos percentuais por rede neural especialista para  $N_{H1} = N_{H2} = 14$ ;  $\omega = 0,1$  e  $tol = 10^{-3}$ .

RNA	$ER\tau_0$	$ER\omega$	$ER\rho_1$	$ER\rho_2$
	(%)	(%)	(%)	(%)
1	0,19	20,4	5,76	5,38
2	0,44	4,88	1,5	4,3
3	0,78	4,56	2,86	3,66
4	1,7	3,76	0,26	4,46
5	0,61	2,92	3,52	5,86
6	1,23	0,24	6,18	4,6
7	0,81	0,34	0,0	4,4
8	1,55	1,96	1,98	3,92
9	0,01	3,02	2,24	6,08
10	1,99	0,22	2,0	4,0
$\overline{ER}$	0,93	4,23	2,63	4,67

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 7 e 8, e do ponto de vista de desempenho computacional (menor tempo de processamento e menor n° de épocas), foi a RNA 3 com apenas 2 minutos e 12 segundos e 23 épocas. Pode-se concluir que para o

caso  $\omega = 0,1$ , as RNA obtiveram muito boas estimativas para os parâmetros radiativos, exceto para RNA 1 que obteve um erro relativo alto na estimação do albedo de espalhamento simples  $\omega$ ,  $ER\omega = 20,4\%$ .

A Figura 34 mostra a curva de performance da RNA 3, acima mencionada.

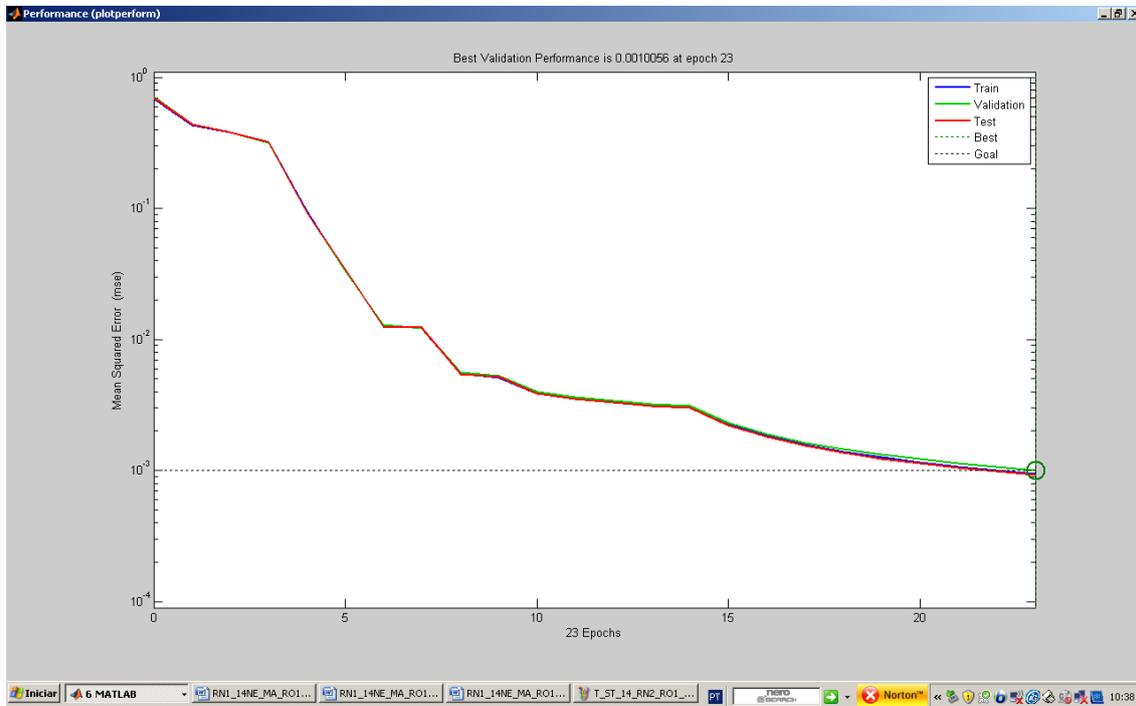


Figura 34: Curva de performance da RNA nº 03,  $N_{H1} = N_{H2} = 14$  ( $\omega = 0,1$ ).

No entanto, observa-se na Tabela 02 que as estimativas para  $\vec{Z}$ , obtidas com as máquinas de comitês para o caso  $N_H = 2$ ,  $\omega = 0,1$  e  $tol = 10^{-3}$  com ambos os modelos de máquina de comitê, não foram bons pois o  $ER_\omega > 5\%$ . Da Tabela 2, conclui-se que o melhor resultado, i.e., o mais preciso, obtido pelas máquinas de comitês para  $\omega = 0,1$ , foi para o caso  $N_H = 1$ ,  $tol = 10^{-3}$  e modelo de máquina conforme o da Figura 32, i.e., média + MLP, e  $M = 6$ . Verifica-se, para este caso, que para todos os parâmetros radiativos,  $ER < 4\%$ , o que se considera um excelente resultado.

Ainda, na Figura 34, observa-se que o ponto ótimo de treinamento da RNA 3 ocorreu com 23 épocas, a partir deste ponto, a rede não treina mais, e passa a “decorar” os resultados, ocorrendo o que se chama de overtraining.

Na Figura 35 é mostrada a curva de performance do treinamento do MLP da máquina de comitê, que obteve as melhores estimativas para o caso  $\omega = 0,1$  (baixo espalhamento).

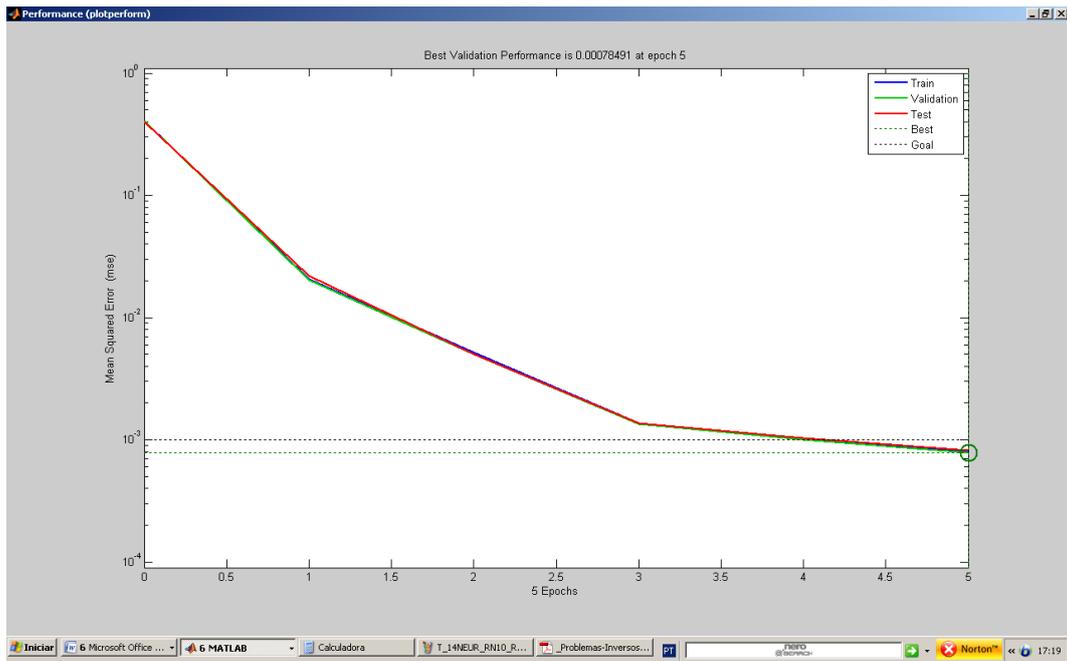


Figura 35: Curva de performance do MLP do combinador,  $N_h = 12$  ( $\omega = 0,1$ ).

Comprova-se na Figura 35, que o desempenho do MLP de saída da máquina de comitê foi bem superior ao desempenho obtido pela melhor RNA especialista para  $\omega = 0,1$ , ou seja, 5 épocas processadas pelo MLP da máquina de comitê, contra 23 épocas da RNA especialista n° 03. Isto comprova a eficiência superior em termos de desempenho do MLP de saída da máquina de comitê em relação à melhor RNA especialista, acrescentando um esforço computacional extra muito pequeno para obtenção, em geral, de melhores estimativas justificando assim a sua utilização.

Da Tabela 2, verifica-se que para  $\omega = 0,5$  (médio espalhamento) do ponto de vista de precisão, considerando o valor esperado  $\overline{ER}$  para o erro relativo, referente a cada parâmetro radiativo, as melhores estimativas obtidas para  $\vec{Z}$  com as RNA especialistas foram aquelas para o caso  $N_H = 2$  e  $tol = 10^{-4}$ , conforme mostrado nas Tabelas 5 e 6.

Tabela 5: Estimativas obtidas para  $N_{H1} = N_{H2} = 14$ ;  $\omega = 0,5$  e  $tol = 10^{-4}$ .

RNA	tempo s	época	$\tau_0$	$\omega$	$\rho_1$	$\rho_2$
<b>1</b>	<b>0:24:41</b>	<b>269</b>	<b>0,9841</b>	<b>0,4911</b>	<b>0,4926</b>	<b>0,5036</b>
<b>2</b>	<b>0:22:05</b>	<b>238</b>	<b>0,9977</b>	<b>0,4877</b>	<b>0,5029</b>	<b>0,5072</b>
<b>3</b>	<b>0:21:53</b>	<b>242</b>	<b>0,9918</b>	<b>0,4978</b>	<b>0,4957</b>	<b>0,5134</b>
<b>4</b>	<b>0:13:17</b>	<b>142</b>	<b>0,9677</b>	<b>0,5011</b>	<b>0,4832</b>	<b>0,5047</b>
<b>5</b>	<b>0:09:30</b>	<b>101</b>	<b>0,9917</b>	<b>0,5002</b>	<b>0,4944</b>	<b>0,5088</b>
<b>6</b>	<b>0:15:10</b>	<b>164</b>	<b>1,0000</b>	<b>0,5002</b>	<b>0,5014</b>	<b>0,5105</b>
<b>7</b>	<b>0:16:52</b>	<b>185</b>	<b>0,9742</b>	<b>0,4886</b>	<b>0,5052</b>	<b>0,5091</b>
<b>8</b>	<b>0:13:57</b>	<b>193</b>	<b>0,9886</b>	<b>0,4997</b>	<b>0,5081</b>	<b>0,5208</b>
<b>9</b>	<b>0:14:08</b>	<b>192</b>	<b>0,9819</b>	<b>0,4909</b>	<b>0,4982</b>	<b>0,5027</b>
<b>10</b>	<b>0:11:34</b>	<b>159</b>	<b>0,9850</b>	<b>0,5024</b>	<b>0,4982</b>	<b>0,5091</b>

Tabela 6: Erros relativos percentuais por rede neural especialista para  $N_{H1} = N_{H2} = 14$ ;  $\omega = 0,5$  e  $tol = 10^{-4}$ .

RNA	$ER_{\tau_0}$ (%)	$ER_{\omega}$ (%)	$ER_{\rho_1}$ (%)	$ER_{\rho_2}$ (%)
<b>1</b>	<b>1,59</b>	<b>1,78</b>	<b>1,48</b>	<b>0,72</b>
<b>2</b>	<b>0,23</b>	<b>2,46</b>	<b>0,58</b>	<b>1,44</b>
3	0,82	0,44	0,86	2,68
4	3,23	0,22	3,36	0,94
<b>5</b>	<b>0,83</b>	<b>0,04</b>	<b>1,12</b>	<b>1,76</b>
<b>6</b>	<b>0</b>	<b>0,04</b>	<b>0,28</b>	<b>2,1</b>
7	2,58	2,28	1,04	1,82
8	1,14	0,06	1,62	4,16
<b>9</b>	<b>1,81</b>	<b>1,82</b>	<b>0,36</b>	<b>0,54</b>
<b>10</b>	<b>1,5</b>	<b>0,48</b>	<b>0,36</b>	<b>0,36</b>
<b><math>\bar{ER}</math></b>	<b>1,37</b>	<b>0,96</b>	<b>1,12</b>	<b>1,65</b>

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

Observa-se na Tabela 6, que o resultado menos preciso foi o obtido pela RNA 8,  $ER_{\rho_2} = 4,16\%$ , o que já é um resultado muito bom, tendo em vista que é menor do que 5%. As demais RNA, na Tabela 6, obtiveram excelentes estimativas para o vetor dos parâmetros radiativos  $\vec{z}$ , podendo-se considerar como a melhor de todas as especialistas a RNA 10, pois para todas as estimativas obtidas para os parâmetros radiativos  $ER \leq 1,5\%$ . Do ponto de vista de desempenho computacional (menor tempo de processamento e menor nº de épocas), a melhor foi a RNA 5 com 9 minutos e 30 segundos, e 101 épocas. Pode-se concluir que para o caso  $\omega = 0,5$ , todas as RNA obtiveram excelentes estimativas para os parâmetros radiativos.

Na Tabela 2, verifica-se que todos os resultados das estimativas obtidas para  $\vec{z}$ , usando ambos os modelos de máquinas de comitês de redes neurais foram excelentes, tanto para  $M = 10$ , como para  $M = 6$ . Porém com uma precisão um pouco melhor para o caso  $N_H = 2$ ,  $\omega = 0,5$  e  $tol = 10^{-4}$ , com o modelo da Figura 31, i.e., usando somente a média, e  $M = 6$ , obtendo para todos os parâmetros radiativos  $ER \leq 1,4\%$ . O pior resultado foi para o caso  $N_H = 1$ ,  $\omega = 0,5$  e  $tol = 10^{-3}$ , também com o modelo da Figura 31, e  $M = 10$ , pois  $ER_{\rho_1} = 5,14\%$ .

Verifica-se ainda que os resultados obtidos com as máquinas de comitês foram equivalentes àqueles obtidos com as RNA especialistas individualmente, neste caso.

Na Figura 36 está representada a curva de performance da RNA especialista com melhor desempenho computacional para  $\omega = 0,5$ .

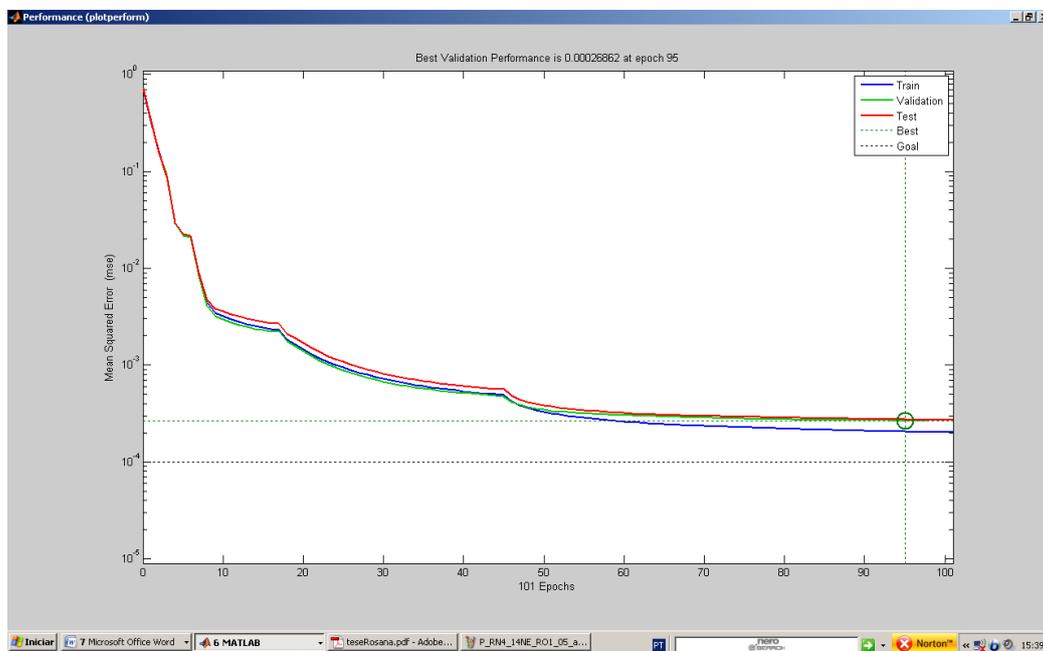


Figura 36: Curva de performance da RNA nº 05,  $N_{H1} = N_{H2} = 14$  ( $\omega = 0,5$ ).

A curva de performance do MLP do combinador da máquina de comitê, de acordo com o modelo da Figura 32, para o caso  $\omega = 0,5$ , com o melhor desempenho, é a mesma que está representada na Figura 35.

Por último da Tabela 2, verifica-se que para  $\omega = 0,9$  (alto espalhamento) do ponto de vista de precisão, considerando o valor esperado  $\overline{ER}$  para o erro relativo, referente a cada parâmetro radiativo, as melhores estimativas obtidas com as RNA especialistas foram às para o caso  $N_H = 2$ ;  $\omega = 0,9$  e  $tol = 10^{-4}$ , conforme mostrado nas Tabelas 7 e 8.

Tabela 7: Estimativas obtidas para  $N_{H1} = N_{H2} = 14$ ;  $\omega = 0,9$  e  $tol = 10^{-4}$ .

RNA	tempo	época	$\tau_0$	$\omega$	$\rho_1$	$\rho_2$
	s					
<b>1</b>	<b>0:24:41</b>	<b>269</b>	<b>0,9712</b>	<b>0,8959</b>	<b>0,4865</b>	<b>0,5033</b>
<b>2</b>	<b>0:22:05</b>	<b>238</b>	<b>0,9486</b>	<b>0,8933</b>	<b>0,4848</b>	<b>0,4883</b>
<b>3</b>	<b>0:21:53</b>	<b>242</b>	<b>0,9627</b>	<b>0,9105</b>	<b>0,4939</b>	<b>0,4955</b>
<b>4</b>	<b>0:13:17</b>	<b>142</b>	<b>0,9250</b>	<b>0,8928</b>	<b>0,5082</b>	<b>0,4954</b>
<b>5</b>	<b>0:09:30</b>	<b>101</b>	<b>0,9607</b>	<b>0,9059</b>	<b>0,4906</b>	<b>0,4897</b>
<b>6</b>	<b>0:15:10</b>	<b>164</b>	<b>0,9677</b>	<b>0,9095</b>	<b>0,5098</b>	<b>0,4821</b>
<b>7</b>	<b>0:16:52</b>	<b>185</b>	<b>0,9632</b>	<b>0,8841</b>	<b>0,5022</b>	<b>0,5034</b>
<b>8</b>	<b>0:13:57</b>	<b>193</b>	<b>0,9592</b>	<b>0,8897</b>	<b>0,4854</b>	<b>0,5022</b>
<b>9</b>	<b>0:14:08</b>	<b>192</b>	<b>0,9560</b>	<b>0,9057</b>	<b>0,5030</b>	<b>0,4958</b>
<b>10</b>	<b>0:11:34</b>	<b>159</b>	<b>0,9964</b>	<b>0,8938</b>	<b>0,4886</b>	<b>0,5096</b>

Tabela 8: Erros relativos percentuais por rede neural especialista para  $N_{H1} = N_{H2} = 14$ ;  $\omega = 0,9$  e  $tol = 10^{-4}$ .

RNA	$ER\tau_0$ (%)	$ER\omega$ (%)	$ER\rho_1$ (%)	$ER\rho_2$ (%)
<b>1</b>	2,88	0,46	2,7	0,66
<b>2</b>	5,14	0,74	3,04	2,34
<b>3</b>	<b>3,73</b>	<b>1,17</b>	<b>1,22</b>	<b>0,9</b>
<b>4</b>	7,5	0,8	1,64	0,92
<b>5</b>	<b>3,93</b>	<b>0,66</b>	<b>1,88</b>	<b>2,06</b>
<b>6</b>	<b>3,23</b>	<b>1,06</b>	<b>1,96</b>	<b>3,58</b>
<b>7</b>	<b>3,68</b>	<b>1,77</b>	<b>0,44</b>	<b>0,68</b>
<b>8</b>	4,08	1,14	2,92	0,44
<b>9</b>	4,4	0,63	0,6	0,84
<b>10</b>	<b>0,36</b>	<b>0,69</b>	<b>2,28</b>	<b>1,92</b>
$\bar{E}$	3,89	0,91	1,87	1,43

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

Observa-se na Tabela 8, que os resultados mais precisos foram os obtidos pelas RNA 1 e 10, e do ponto de vista de desempenho computacional (menor tempo de processamento e menor n° de épocas), foi a RNA 5 com apenas 9 minutos e 30 segundos e 101 épocas. Pode-se concluir que para o caso  $\omega = 0,9$ , as RNA especialistas obtiveram estimativas muito boas para os parâmetros radiativos, exceto a RNA 4 que obteve  $ER\tau_0 = 7,5\%$  ( $> 5\%$ ).

A curva de performance que representa o desempenho da RNA 5 é dada também pela Figura 36.

Os melhores resultados obtidos para as estimativas dos parâmetros radiativos para  $\omega = 0,9$ , usando máquina de comitê, são aqueles que foram conseguidos com as mesmas condições para as

RNA especialistas, i.e. para o caso  $N_H = 2$  e  $tol = 10^{-4}$  (ver Tabela 2). Ambos os modelos de máquinas de comitê, tanto para  $M = 10$ , como para  $M = 6$ , obtiveram excelentes resultados, sendo ligeiramente melhores do que os obtidos pelas RNA, no geral. No entanto, a máquina de comitê com modelo de acordo com o da Figura 31, i.e. com o combinador efetuando apenas a média aritmética das saídas das RNA especialistas, foi a que apresentou os melhores resultados, neste caso.

Foram também realizados testes com  $N_H = 1$  e  $tol = 10^{-4}$ , numa tentativa de obter resultados mais exatos e precisos, utilizando apenas uma camada oculta. Porém, não ocorreu a convergência esperada, pois a ordem de grandeza do erro médio quadrático minimizado não atingiu menos que  $10^{-3}$ , o que só foi conseguido ao se usar  $N_H = 2$ . Testou-se ainda a combinação: duas camadas escondidas,  $N_H = 2$ , e  $tol = 10^{-5}$ . Porém, como aconteceu para o caso  $N_H = 1$  e  $tol = 10^{-4}$ , não ocorreu a convergência esperada, mantendo-se a ordem de grandeza do erro médio quadrático minimizado em torno  $10^{-4}$  para um número muito alto de épocas ( $\leq 3000$ ), acarretando em um custo computacional alto e desnecessário.

Devido ao aspecto altamente estocástico do processamento com RNA e máquinas de comitês, não havendo, portanto, um método determinístico para o seu projeto, que depende em muito da sensibilidade do projetista, a convergência para os resultados esperados só ocorreu após várias tentativas de ajuste, principalmente no quesito números de neurônios versus número de camadas ocultas, variando de 8 à 18 neurônios e de 01 a 03 camada ocultas. Alcançando-se finalmente os melhores resultados com 14 neurônios por camada oculta,  $N_H = 1$  ou 2 e  $tol = 10^{-3}$  ou  $10^{-4}$ .

Conclui-se, com base nos testes aqui mostrados, que tanto as RNA especialistas quanto as máquinas de comitês funcionaram muito bem na estimação dos parâmetros radiativos para a solução do problema inverso em transferência radiativa. Verifica-se ainda que as máquinas de comitês, em geral, apresentaram desempenho igual ou superior aos alcançados pelas RNA especialistas para a estimação dos parâmetros radiativos, e que a seleção das RNA especialistas com os melhores resultados realmente influi positivamente na melhoria dos mesmos.

## CAPÍTULO 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 de 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  $Y_i$ , minimizando o erro médio quadrático  $E$ , a cada padrão  $q$  apresentado, decorrente do confronto entre as saídas obtidas e as saídas desejadas, usando o algoritmo de retro-propagação para reajustar os pesos e bias das camadas ocultas das RNA especialistas, pesos estes que constituem a “memória” das RNA especialistas, e em consequência das máquinas de comitês por elas compostas, caracterizando assim o processo de aprendizado e a sua posterior capacidade de generalização e classificação de padrões.

Os resultados obtidos e descritos na capítulo 4 e no Apêndice A, demonstram que os objetivos, tanto do ponto de vista geral quanto do específico deste trabalho foram alcançados, ou seja, a solução do problema inverso em transferência radiativa usando RNA e máquinas de comitês de RNA, e a comprovação de que as máquinas de comitês, no geral, produzem estimativas na fase de generalização tão boas ou melhores do que as RNA usadas isoladamente, sem no entanto, desperdiçar o esforço computacional gasto anteriormente no treinamento das redes neurais especialistas que compõem as máquinas de comitês com um custo computacional extra pequeno.

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 de conjunto (ver 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 mostrando ter sido uma experiência válida (ver Tabela 2).

Comprova-se através da análise dos valores dos erros relativos devido às estimativas obtidas para cada um dos parâmetros radiativos, pela aplicação de máquinas de comitês, 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 (ver Figuras 31 e 32), efetuando então um novo treinamento, sem porém descartar nenhuma das demais

redes neurais especialistas já treinadas e com os piores resultados, acarretou, de modo geral, na obtenção de resultados mais precisos para as estimativas dos parâmetros radiativos do que aqueles obtidos sem efetuar a seleção das RNA com melhores resultados ( $M = 10$ ).

Finalmente deixamos algumas sugestões para trabalhos futuros aplicando máquinas de comitês 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, Monte Carlo, etc. para estimar os parâmetros radiativos; e
- A hibridização de máquinas de comitês de redes neurais artificiais tipo dinâmicas com métodos de otimização global, como LM, SA, GA, PSO, Monte Carlo, etc. para estimar os parâmetros radiativos.

## REFERÊNCIAS BIBLIOGRÁFICAS

- ABRAMOWITZ, M.; STEGUN, I. A, (Ed.). *Handbook of mathematical functions, with formulas, graphs and mathematical tables*. Nova York: Dover Publications, 1965. 1046p.
- BARICHELLO, L. B.; SIEWERT, C. E.; GARCIA, R. D. M. On inverse boundary condition problems in radiative transfer. *J.l 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.
- 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.
- CAMPOS-OLIVEIRA, R. de et al. Aplicação da técnica de máquina de comitê de redes neurais artificiais para a solução de problemas inversos em transferência radiativa. *XI Encontro de Modelagem Computacional*, Volta Redonda – RJ, não paginado, nov./ 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 self-organizing 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). 75f. Tese (Doutorado) – Universidade do Sul da Flórida, Flórida, 2001.
- CASTRO, C. L. de; BRAGA, A. de P.; ANDRADE, A. V. *Aplicação de um modelo ensemble de redes neurais artificiais para previsão de séries temporais não estacionárias*. XXV Congresso da Sociedade Brasileira de Computação, UNISINOS – São Leopoldo do Sul/RS, 22 a 29 de julho.

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, V. H. F.; SILVA-NETO, A. J. da. A comparison of the one-dimensional radiative transfer problem solutions obtained with the monte carlo method and three variations of the discrete ordinates method. *Proceedings of 19th International Congress of Mechanical Engineering - Cobem 2007*, Brasília-DF.

\_\_\_\_\_; SILVA-NETO, A. J. da; SOEIRO, F. J. C. P. Estimation of optical thickness and single scattering albedo with artificial neural network and a Monte Carlo method. *COBEM 2007, Proceedings of 19th International Congress of Mechanical Engineering*, Brasilia, DF, Brasil. November 5-9, 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, june / 2007.

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

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

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

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, Lisboa: Elsevier B.V., 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*. 'Anais do XVI Congresso Brasileiro de Engenharia Mecânica', Uberlândia, MG, 2001.

FRANCA Francis H. R. et al. Inverse design of thermal systems with dominant radiative transfer. *Advances in heat transfer*, vol. 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*, vol. 14, number 4, p. 423-436(14). June 2006.

FRANÇA, Francis H. R. et al. Application of inverse analysis to illumination design. *Inverse Problems in Science and Engineering*, vol. 17, number 6, Sept. 2009, p. 737-753.

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, (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. *XI Encontro de Modelagem Computacional*. Volta Redonda-RJ. Não paginado, nov. 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*. MIT Press, Cambridge, Mass.: MIT Press, 1988. 637p.

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.

HAYKIN, S. *Redes neurais: princípios e prática*. 2ª Edição. Tradução de Paulo Martins Engel. Porto Alegre, RS: Bookman, 2001, 900 p.

HEBB, D. O. *The organizations of behavior*, New York: J.Wiley, 1949. 335p.

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 – UFSC, 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, [São Paulo], 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: Editora Ciência Moderna Ltda, 2007, 125p.

KAUATI, A. T.; ROBERTY, N. C.; SILVA- NETO, A. J. da. A source detector methodology for the construction and solution of the one-dimensional inverse transport equation. In *Proc. 3rd International Conference on Inverse Problems in Engineering: Theory and Practice*, Port Ludlow, USA, 1999.

\_\_\_\_\_; ROBERTY, N. C.; SILVA- NETO, A. J. da. Uso da técnica fonte-detetor em tomografia com espalhamento anisotrópico. In *Anais do XVII Congresso Brasileiro de Engenharia Biomédica*, Florianópolis, Santa Catarina, 2000.

\_\_\_\_\_; SILVA-NETO, A. J. da; 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. da; 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.

\_\_\_\_\_. Self-organized formation of topologically correct feature maps. *Biological Cybernetics*, vol.43, p. 59-69, 1982.

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

\_\_\_\_\_; 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 *Ocean Optics XIII* (S.G. Ackleson and R. Frouin, Ed.), 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. 135p.

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

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 Mass Transfer*, 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 backward-scattering 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) – Centro de Tecnologia, Programa de Pós – Graduação em Engenharia Elétrica, Universidade Federal do Rio Grande do Norte, Natal – RN, jan./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 cancer 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.

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 *Ocean Optics XII* (J.S. Jaffe, ed.) 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 de Matemática Aplicada e Computacional XXIII, Sociedade Brasileira de Matemática Aplicada e Computacional (this proceedings), 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: Na experimental study*. *Neural Networks*, v. 21, n.9, p. 1363-1379, nov. 2008.

MINSKY, M.; PAPERT, S. *Perceptrons*. Cambridge, Mass.: MIT Press, 1969. 258p.

MONTERO, C. R. F.; ROBERTY, N. C.; SILVA-NETO, A. J. da. 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. da. 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, 1965.137p.

ÖZISIK, M. N. *Radiative transfer and interactions with conduction and convection*. John Wiley & Sons, Nova York, 1973. 575p.

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.

PAULA, M. V.S. de Oliveira et al. User authentication based on human typing pattern with artificial neural networks and support vector machine. In: ENIA - SBC 2005 - Encontro Nacional de Inteligência Artificial, 2005, São Leopoldo - RS. *Anais do ENIA Encontro Nacional de Inteligência Artificial - SBC 2005*, 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 do Rio de Janeiro – UERJ, Nova Friburgo, 2004.

PUMA-VILLANUEVA, W. J.; ZUBEN, F. J. V. *Comitê de máquinas em predição de séries temporais*. EA072 – Inteligência Artificial em Aplicações Industriais, Departamento de Engenharia de Computação e Automação Industrial, UNICAMP, Campinas – SP, 2008. 8p.

ROGERS, T. T.; McCLELLAND, J. L. *Semantic Cognition: A parallel distributed processing approach*. Cambridge, Mass.: MIT Press, 2004. 425p.

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.

\_\_\_\_\_. Principles of neurodynamics. New York: Spartan, 1962. 616p.

RUMELHART, D. E.; HINTON, G. E.; WILLIAM, R. J. *Learning internal representation by error propagation in parallel distributed processing*. D. E. Rumelhart and J. L. McClelland (Ed.), Cambridge, M. A.: 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.  
*Referências bibliográficas*

SEDANO, S. T. C. *Estimação simultânea da temperatura do gás e a emissividade num forno utilizando análise inversa em radiação térmica*. 8º Congresso Ibero Americano de Engenharia Mecânica. Cusco, 23 a 25 de outubro de 2007.

SILVA, L. A. da. *Classificação de nódulos mamográficos utilizando um comitê de redes neurais artificiais*. 2005. 94f. Dissertação (Mestrado em Engenharia) – Escola Politécnica da Universidade de São Paulo, São Paulo, 2005.

\_\_\_\_\_; HERNANADEZ, D. M.; RANGAYYAN, 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. da. *Problemas inversos: aplicações em engenharia e medicina*. Em *Cápsulas de Ciência* (F.D. Moura Neto, Ed.), LIBRE – Liga Brasileira de Editoras, 1ª Edição, p. 57-86, 2007.

\_\_\_\_\_; 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, 122p.

\_\_\_\_\_; CAMPOS -VELHO, H. F. de. *Problemas inversos em transferência radiativa: uma formulação implícita*, p. 33-41, 2009 (no prelo)

SILVA-NETO, A. J. da. Explicit and implicit formulations for inverse radiative transfer problems. *Proc. 5th World Congress on Computational Mechanics, Mini Symposium MS125 -Computational Treatment of Inverse Problems in Mechanics*, Viena, Austria, 2002. (Invited Lecture).

\_\_\_\_\_; MOURA-NETO, F. D. *Problemas inversos: conceitos fundamentais e aplicações*. EdUERJ - Editora da Universidade do Estado do Rio de Janeiro, Rio de Janeiro, 2005.p.?

\_\_\_\_\_; ÖZISIK, M. N. An inverse problem of estimating thermal conductivity, optical thickness, and single scattering albedo of semitransparent medium. In *Proc. 1st International Conference on Inverse Problems in Engineering: Theory and Practice*, Flórida, USA, p. 267–273, 1993.

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

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

\_\_\_\_\_; ROBERTY, N. C. A modelagem da tomografia com espalhamento como um problema inverso. *Anais do IV Fórum Nacional de Ciência e Tecnologia em Saúde*, p. 317–318, 1998a.

\_\_\_\_\_; ROBERTY, N. C. Uma comparação do método de ordenadas discretas com outras técnicas para a solução da equação de transporte unidimensional. *Anais do VII Congresso Brasileiro de Engenharia e Ciências Térmicas II*, Rio de Janeiro, RJ, p. 878–883, 1998b.

\_\_\_\_\_; ROBERTY, N. C. A abordagem fonte-detector na solução de problemas inversos em meios participantes. *Anais do V Congresso de Engenharia Mecânica, Norte-Nordeste, III*, Fortaleza, p. 340-347, 1998c.

\_\_\_\_\_; SOEIRO, F. J. C. P. The solution of an inverse radiative transfer problem with the simulated annealing and levenberg-marquardt methods. In *Anais do Mini-Symposium on Inverse Problems in Medicine, Engineering and Geophysics*, Santos – SP, p. 62–71, 2000.

SILVA-NETO, A. J. da; 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 *Proc. 5th World Congress on Computational Mechanics*, Viena, Austria, 2002.

\_\_\_\_\_; SOEIRO, F. J. C. P. Inverse problem of space dependent albedo estimation with artificial neural networks and hybrid methods. *Proc. 18th International Congress of Mechanical Engineering, COBEM, ABCM*, Ouro Preto, Brazil, 2005.

\_\_\_\_\_; 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 Solid Mechanics Applications - 2nd MIT Conference on Computational Fluid and Solid Mechanics*, Cambridge, EUA, 2003.

SILVA, E.; THOMÉ, A. C. G. *Reconhecimento de caracteres manuscritos utilizando time de redes neurais*. Instituto de Matemática (IM) – Núcleo de Computação Eletrônica (NCE), Universidade Federal do Rio de Janeiro (UFRJ), [Rio de Janeiro], [200-].

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

SOEIRO, F. J. C. P.; CARVALHO, G.; SILVA-NETO, A. J. da. Thermal properties estimation of polymeric materials with the simulated annealing method. In *Anais do 8º Congresso Brasileiro de Engenharia e Ciências Térmicas*, Porto Alegre, RS, 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. da. Solution of inverse radiative transfer problems in two-layer media with artificial neural networks. *14th Inverse Problems in Engineering Seminar*, Ames, EUA, 2006.

\_\_\_\_\_; SOARES, P. O.; SILVA-NETO, A. J. da. Solution of inverse radiative transfer problems with artificial neural networks and hybrids methods. *Proc. 13th Inverse Problems in Engineering Seminar*, Cincinnatti, USA, p. 163-169, 2004.

SOUSA, F. L. de 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. *Springer Series in Statistics*, Springer-Verlag, New York, Inc., Secaucus, NJ, 1982.

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* 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(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. Adaptive switching circuits. *IRE WESCON Convention Record*, p. 96-104, 1960.

WILLSHAW, D. J.; MALSBERG, 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.* Van Nostrand Reinhold, New York, 1989.

WASSERMAN, P. D. *Advanced methods in neural computing.* Van Nostrand Reinhold, [New York], 1993.

## APÊNDICE A – RESULTADOS DOS DEMAIS TESTES EFETUADOS

Neste apêndice são apresentados por meio de tabelas e gráficos, os resultados dos testes efetuados que não foram incluídos no capítulo 4. Estes resultados são mostrados em dois conjuntos de tabelas e gráficos na seguinte ordem:

1. Os resultados obtidos com as RNA especialistas com uma ou duas camadas ocultas. Neste conjunto estão inseridas as Tabelas de 9 a 20 e a Figura 37, sendo que as Tabelas 9, 11, 13, 15, 17 e 19 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 na fase de generalização para cada parâmetro radiativo, para as condições informadas nos títulos das tabelas.

Nas Tabelas 10, 12, 14, 16, 18 e 20 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 37 mostra a curva de performance para a RNA nº 02 para  $N_{HI} = 14$ .

2. Os resultados obtidos com as máquinas de comitês de redes neurais: em primeiro lugar com a máquina com modelo de acordo com a Figura 31, mostrados nas Tabelas 21, 22, 23, 24, 25 e 26; em seguida com a máquina com modelo de acordo com a Figura 32, mostrados nas Tabelas 27, 28, 29, 30, 31, 32, 33 e 34. Em todas as tabelas supracitadas, são mostradas as estimativas obtidas com  $M = 10$  (antes da seleção das melhores RNA) 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 comitês. As Figuras 38 e 39 mostram a curvas de performance do MLP (ver Figura 33) da máquina de comitê cujo modelo é representado na Figura 32, para os casos  $N_h = 12, tol = 10^{-3}$  e  $N_h = 12, tol = 10^{-4}$ .

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 uma das RNA especialistas, 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.

## RESULTADOS OBTIDOS COM AS REDES NEURAIS ESPECIALISTAS ISOLADAMENTE

### Rede neural especialista com uma camada oculta com 14 neurônios

Tabela 9: Estimativas obtidas para  $N_{HI} = 14$ ;  $\omega = 0,1$  e  $tol = 10^{-3}$ .

RNA	tempo	época s	$\tau_0$	$\omega$	$\rho_1$	$\rho_2$
1	0:29:51	494	1.0078	0.0920	0.5503	0.5309
2	0:07:26	117	1.0148	0.0774	0.5673	0.5290
3	0:10:47	171	1.0194	0.0772	0.5376	0.5295
4	0:20:12	333	0.9984	0.0905	0.5347	0.5269
5	0:50:25	843	1.0170	0.1389	0.4672	0.5042
6	0:11:01	180	1.0065	0.0951	0.4979	0.5201
7	0:21:17	345	1.0120	0.1057	0.5363	0.5380
8	0:11:59	189	1,0039	0,0976	0,4758	0,5162
9	0:13:49	223	0,9824	0,0796	0,4871	0,5166
10	0:18:40	300	1,0139	0,0638	0,5627	0,5116

Tabela 10: Erros relativos percentuais por rede neural especialista para  $N_{HI} = 14$ ;  $\omega = 0,1$  e  $tol = 10^{-3}$ .

RNA	$ER\tau_0$ (%)	$ER\omega$ (%)	$ER\rho_1$ (%)	$ER\rho_2$ (%)
1	<b>0,78</b>	<b>8,0</b>	<b>10,06</b>	<b>6,18</b>
2	1,48	22,6	13,46	5,8
3	1,94	22,8	7,52	5,9
4	<b>0,16</b>	<b>9,5</b>	<b>6,94</b>	<b>5,38</b>
5	1,7	38,9	6,56	0,84
6	<b>0,65</b>	<b>4,9</b>	<b>0,42</b>	<b>4,02</b>
7	1,2	5,7	7,26	7,6
8	<b>0,65</b>	<b>2,4</b>	<b>4,84</b>	<b>3,24</b>
9	<b>1,76</b>	<b>20,4</b>	<b>2,58</b>	<b>3,32</b>
10	1,39	36,2	12,54	2,32
$\bar{E}$	1,17	17,14	7,21	4,46

Nota: As destacadas em negrito são as selecionadas.

Tabela 11: Estimativas obtidas para  $N_{HI} = 14$ ;  $\omega = 0,5$  e  $tol = 10^{-3}$ .

RNA	tempo	época	$\tau_0$	$\omega$	$\rho_1$	$\rho_2$
s						
1	0:29:51	494	0.9790	0.4805	0.4547	0.5254
2	0:07:26	117	0.9764	0.4865	0.4860	0.5259
3	0:10:47	171	0.9935	0.4954	0.4529	0.5130
4	0:20:12	333	0.9819	0.5012	0.4582	0.5048
5	0:50:25	843	0.9762	0.4899	0.5100	0.5150
6	0:11:01	180	0.9770	0.5134	0.4642	0.5062
7	0:21:17	345	0.9827	0.4889	0.5127	0.5223
8	0:11:59	189	0,9852	0,4938	0,4899	0,5130
9	0:13:49	223	0,9808	0,4942	0,4363	0,5246
10	0:18:40	300	1,0028	0,4904	0,4780	0,4938

Tabela 12: Erros relativos percentuais por rede neural especialista para  $N_{HI} = 14$ ;  $\omega = 0,5$  e  $tol = 10^{-3}$ .

RNA	$ER\tau_0$	$ER\omega$	$ER\rho_1$	$ER\rho_2$
	(%)	(%)	(%)	(%)
1	2,1	3,9	9,06	5,08
2	<b>2,36</b>	<b>2,7</b>	<b>2,8</b>	<b>5,18</b>
3	0,65	0,92	9,42	2,6
4	1,81	0,24	8,36	0,96
5	<b>2,38</b>	<b>2,02</b>	<b>2,0</b>	<b>3,0</b>
6	<b>2,3</b>	<b>2,68</b>	<b>7,16</b>	<b>1,24</b>
7	<b>1,73</b>	<b>2,22</b>	<b>2,54</b>	<b>4,46</b>
8	<b>1,48</b>	<b>1,24</b>	<b>2,02</b>	<b>2,6</b>
9	1,92	1,16	12,74	4,92
10	<b>0,28</b>	<b>1,92</b>	<b>4,4</b>	<b>1,24</b>
$\bar{E}$	1,7	1,9	6,05	3,13

*Nota: As destacadas em negrito são as selecionadas.*

Tabela 13: Estimativas obtidas para  $N_{HI} = 14$ ;  $\omega = 0,9$  e  $tol = 10^{-3}$ .

RNA	tempo	época	$\tau_0$	$\omega$	$\rho_1$	$\rho_2$
s						
1	0:29:51	494	0.8407	0.8823	0.5149	0.4897
2	0:07:26	117	0.8942	0.9378	0.4607	0.4951
3	0:10:47	171	0.8196	0.9136	0.4659	0.4985
4	0:20:12	333	0.8677	0.9186	0.4670	0.5114
5	0:50:25	843	0.8612	0.8722	0.4858	0.5015
6	0:11:01	180	0.8912	0.8932	0.4959	0.4898
7	0:21:17	345	0.8684	0.8791	0.4255	0.5224
8	0:11:59	189	0.8756	0.8898	0.4832	0,5092
9	0:13:49	223	0,8672	0,8814	0,5236	0,5143
10	0:18:40	300	0,8721	0,8795	0,4207	0,4907

Tabela 14: Erros relativos percentuais por rede neural especialista para  $N_{HI} = 14$ ;  $\omega = 0,9$  e  $tol = 10^{-3}$ .

RNA	$ER\tau_0$ (%)	$ER\omega$ (%)	$ER\rho_1$ (%)	$ER\rho_2$ (%)
1	15,93	1,97	2,98	2,06
<b>2</b>	<b>10,58</b>	<b>4,2</b>	<b>7,86</b>	<b>0,98</b>
3	18,04	1,51	6,82	0,3
<b>4</b>	<b>13,23</b>	<b>2,07</b>	<b>6,6</b>	<b>2,28</b>
5	13,88	3,09	2,84	0,3
<b>6</b>	<b>10,88</b>	<b>0,76</b>	<b>0,82</b>	<b>2,04</b>
<b>7</b>	<b>13,16</b>	<b>2,32</b>	<b>14,9</b>	<b>4,48</b>
<b>8</b>	<b>12,44</b>	<b>1,13</b>	<b>3,36</b>	<b>1,84</b>
9	13,28	2,07	4,72	2,86
<b>10</b>	<b>12,79</b>	<b>2,28</b>	<b>15,86</b>	<b>1,86</b>
$\bar{E}$	13,42	2,14	6,68	1,9

Nota: As destacadas em negrito são as selecionadas

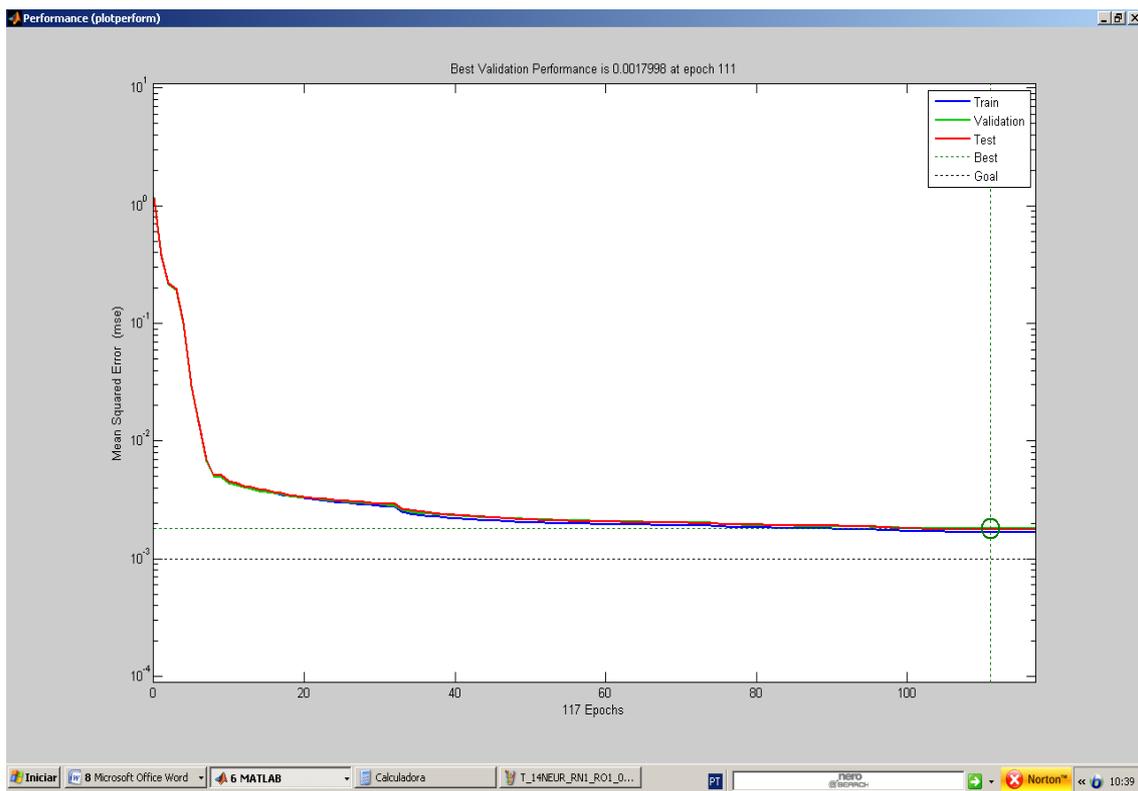


Figura 37: Curva de performance da RNA n° 02,  $N_{HI} = 14$ .

**Rede neural especialista com duas camadas ocultas com 14 neurônios cada**Tabela15: Estimativas obtidas para  $N_{H1} = N_{H2}=14$ ;  $\omega = 0,5$  e  $tol=10^{-3}$ .

RNA	tempo	época	$\tau_0$	$\omega$	$\rho_1$	$\rho_2$
		s				
1	0:03:23	33	0,9601	0,5098	0,4920	0,5313
2	0:02:52	30	0,9882	0,4718	0,4986	0,5120
3	0:02:12	23	0,9841	0,4993	0,4772	0,5193
4	0:03:02	32	0,9933	0,5019	0,4865	0,5215
5	0:02:23	25	1,0016	0,5091	0,5111	0,5105
6	0:02:51	30	0,9746	0,4900	0,4892	0,5221
7	0:02:34	35	0,9670	0,5038	0,5022	0,5141
8	0:02:54	40	0,9977	0,5033	0,4808	0,4862
9	0:02:23	31	0,9785	0,5188	0,4886	0,5111
10	0:02:39	36	0,9867	0,4938	0,5146	0,5337

Tabela 16: Erros relativos percentuais por rede neural especialista para  $N_{H1} = N_{H2} = 14$ ;  $\omega = 0,5$  e  $tol=10^{-3}$ .

RNA	$ER\tau_0$ (%)	$ER\omega$ (%)	$ER\rho_1$ (%)	$ER\rho_2$ (%)
1	3,99	1,96	1,6	6,26
2	1,18	5,64	0,28	2,4
3	1,59	0,14	4,56	3,86
4	<b>0,67</b>	<b>0,38</b>	<b>2,7</b>	<b>4,3</b>
5	<b>0,16</b>	<b>1,82</b>	<b>2,22</b>	<b>2,1</b>
6	2,54	2,0	2,16	4,42
7	3,3	0,76	0,44	2,82
8	0,23	0,66	3,84	2,76
9	2,15	3,76	2,28	2,22
10	1,33	1,24	2,92	6,74
$\bar{E}$	1,71	1,84	2,3	3,79

Nota: As destacadas em negrito são as selecionadas.

Tabela 17: Estimativas obtidas para  $N_{H1} = N_{H2} = 14$ ;  $\omega = 0,9$  e  $tol = 10^{-3}$ .

RNA	tempo	época	$\tau_0$	$\omega$	$\rho_1$	$\rho_2$
		s				
1	0:03:23	33	0,8958	0,9025	0,4926	0,4843
2	0:02:52	30	0,9457	0,8918	0,4523	0,4924
3	0:02:12	23	0,9349	0,9253	0,4510	0,4954
4	0:03:02	32	0,9226	0,8787	0,4571	0,5069
5	0:02:23	25	0,9285	0,9104	0,5061	0,4842
6	0:02:51	30	0,9484	0,9212	0,4720	0,5123
7	0:02:34	35	0,9554	0,8943	0,4768	0,4851
8	0:02:54	40	0,9097	0,8856	0,4881	0,4957
9	0:02:23	31	0,9103	0,9300	0,5027	0,4778
10	0:02:39	36	0,8984	0,8908	0,4558	0,4963

Tabela 18: Erros relativos percentuais por rede neural especialista para  $N_{H1} = N_{H2} = 14$ ;  $\omega = 0,9$  e  $tol = 10^{-3}$ .

RNA	$ER\tau_0$	$ER\omega$	$ER\rho_1$	$ER\rho_2$
	(%)	(%)	(%)	(%)
1	10,42	0,28	1,48	3,14
2	<b>5,43</b>	<b>0,91</b>	<b>9,54</b>	<b>1,52</b>
3	6,51	2,81	9,8	0,92
4	7,74	2,37	8,58	1,38
5	7,15	1,16	1,22	3,16
6	5,16	2,36	5,6	2,46
7	4,46	0,63	4,64	2,98
8	9,03	1,6	2,38	0,86
9	8,97	3,33	0,54	4,44
10	10,16	1,02	8,84	0,74
$\bar{E}$	7,05	1,65	5,26	2,16

Nota: As destacadas em negrito são as selecionadas.

Tabela 19: Estimativas obtidas para  $N_{H1} = N_{H2} = 14$ ;  $\omega = 0,1$  e  $tol = 10^{-4}$ .

RNA	tempo	época	$\tau_0$	$\omega$	$\rho_1$	$\rho_2$
		s				
1	0:24:41	269	0,9921	0,0863	0,4473	0,5159
2	0:22:05	238	0,9961	0,0876	0,4592	0,5191
3	0:21:53	242	0,9993	0,0949	0,4563	0,5163
4	0:13:17	142	0,9857	0,1015	0,4892	0,5265
5	0:09:30	101	0,9971	0,0835	0,4747	0,5072
6	0:15:10	164	0,9855	0,0929	0,4553	0,5233
7	0:16:52	185	0,9855	0,0890	0,4728	0,5137
8	0:13:57	193	0,9884	0,0957	0,4566	0,5135
9	0:14:08	192	0,9880	0,0958	0,4693	0,5156
10	0:11:34	159	0,9844	0,0821	0,4926	0,5212

Tabela 20: Erros relativos percentuais por rede neural especialista para  $N_{H1} = N_{H2} = 14$ ;  $\omega = 0,1$  e  $tol = 10^{-4}$ .

RNA	$ER\tau_0$ (%)	$ER\omega$ (%)	$ER\rho_1$ (%)	$ER\rho_2$ (%)
1	<b>0,79</b>	13,7	10,54	3,18
2	0,39	12,4	8,16	3,82
3	<b>0,07</b>	5,1	8,74	3,26
4	1,43	1,5	2,16	5,3
5	<b>0,29</b>	16,5	5,06	1,44
6	1,45	7,1	8,94	4,66
7	1,45	11	5,44	2,74
8	1,16	4,3	8,68	2,7
9	1,2	4,2	6,14	3,12
10	1,56	17,9	1,48	4,24
$\bar{E}$	0,98	9,37	6,53	3,45

Nota: As destacadas em negrito são as selecionadas.

## RESULTADOS OBTIDOS COM AS MÁQUINAS DE COMITÊS

### Máquina de comitê usando a média aritmética no combinador

Tabela 21: Resultados da máquina de comitê para  $M = 10$  e  $M = 6$ , com  $N_{H1} = 14$ ;  $\omega = 0,1$  e  $tol = 10^{-3}$ .

PARÂMETRO RADIATIVO	RESULTADO $M = 10$	$ER$ (%) $M = 10$	RESULTADO $M = 6$	$ER$ (%) $M = 6$
$\tau_0$	1,0076	0,76	1,0018	0,18
$\omega$	0,0918	8,2	0,0934	6,6
$\rho_1$	0,5217	4,34	0,5137	2,74
$\rho_2$	0,5223	3,78	0,5248	3,92

Tabela 22: Resultados da máquina de comitê para  $M = 10$  e  $M = 6$ , com  $N_{H1} = 14$ ;  $\omega = 0,5$  e  $tol = 10^{-3}$ .

PARÂMETRO RADIATIVO	RESULTADO $M = 10$	$ER$ (%) $M = 10$	RESULTADO $M = 6$	$ER$ (%) $M = 6$
$\tau_0$	0,9836	1,64	0,9811	1,89
$\omega$	0,4934	1,32	0,4953	0,94
$\rho_1$	0,4743	5,14	0,4693	1,96
$\rho_2$	0,5144	2,88	0,5161	1,14

Tabela 23: Resultados da máquina de comitê, para  $M = 10$  e  $M = 6$ , com  $N_{H1} = 14$ ;  $\omega = 0,9$  e  $tol = 10^{-3}$ .

PARÂMETRO RADIATIVO	RESULTADO $M = 10$	ER (%) $M = 10$	RESULTADO $M = 6$	ER (%) $M = 6$
$\tau_0$	0,8658	13,42	0,8685	13,15
$\omega$	0,8948	0,58	0,8907	1,03
$\rho_1$	0,4743	5,14	0,4850	3,0
$\rho_2$	0,5023	0,46	0,5061	1,22

Tabela 24: Resultados da máquina de comitê, para  $M = 10$  e  $M = 6$ , com  $N_{H1} = N_{H2} = 14$ ;  $\omega = 0,5$  e  $tol = 10^{-3}$ .

PARÂMETRO RADIATIVO	RESULTADO $M = 10$	ER (%) $M = 10$	RESULTADO $M = 6$	ER (%) $M = 6$
$\tau_0$	0,9832	1,68	0,9855	1,45
$\omega$	0,5002	0,04	0,5045	0,9
$\rho_1$	0,4941	1,18	0,4931	1,38
$\rho_2$	0,5162	3,24	0,5109	2,18

Tabela 25: Resultados da máquina de comitê, para  $M = 10$  e  $M = 6$ , com  $N_{H1} = N_{H2} = 14$ ;  $\omega = 0,9$  e  $tol = 10^{-3}$ .

PARÂMETRO RADIATIVO	RESULTADO $M = 10$	ER (%) $M = 10$	RESULTADO $M = 6$	ER (%) $M = 6$
$\tau_0$	0,9250	7,5	0,9291	7,09
$\omega$	0,9031	0,34	0,9034	0,38
$\rho_1$	0,4755	4,9	0,4838	3,24
$\rho_2$	0,4931	1,38	0,4937	1,26

Tabela 26: Resultados da máquina de comitê, para  $M = 10$  e  $M = 6$ , com  $N_{H1} = N_{H2} = 14$ ;  $\omega = 0,1$  e  $tol = 10^{-4}$ .

PARÂMETRO RADIATIVO	RESULTADO $M = 10$	ER (%) $M = 10$	RESULTADO $M = 6$	ER (%) $M = 6$
$\tau_0$	0,9902	0,98	0,9887	1,13
$\omega$	0,0909	9,1	0,0950	5,0
$\rho_1$	0,4673	6,54	0,4666	6,68
$\rho_2$	0,5172	3,44	0,5182	3,64

**Máquina de comitê usando a média aritmética mais MLP no combinador**Tabela 27: Resultados da máquina de comitê, para  $M = 10$  e  $M = 6$ , com  $N_{HI} = 14$ ;  $\omega = 0,1$  e  $tol = 10^{-3}$ .

PARÂMETRO RADIATIVO	RESULTADO $M = 10$	ER (%) $M = 10$	RESULTADO $M = 6$	ER (%) $M = 6$
$\tau_0$	0,9989	0,11	0,9925	0,75
$\omega$	0,0988	1,2	0,0976	2,4
$\rho_1$	0,5209	4,18	0,5168	3,36
$\rho_2$	0,5189	3,78	0,5196	3,92

Tabela 28: Resultados da máquina de comitê, para  $M = 10$  e  $M = 6$ , com  $N_{HI} = 14$ ;  $\omega = 0,5$  e  $tol = 10^{-3}$ .

PARÂMETRO RADIATIVO	RESULTADO $M = 10$	ER (%) $M = 10$	RESULTADO $M = 6$	ER (%) $M = 6$
$\tau_0$	1,0138	1,38	1,0089	0,89
$\omega$	0,5073	1,46	0,5067	1,34
$\rho_1$	0,4965	0,7	0,4902	1,96
$\rho_2$	0,5031	0,62	0,5057	1,14

Tabela 29: Resultados da máquina de comitê, para  $M = 10$  e  $M = 6$ , com  $N_{HI} = 14$ ;  $\omega = 0,9$  e  $tol = 10^{-3}$ .

PARÂMETRO RADIATIVO	RESULTADO $M = 10$	ER (%) $M = 10$	RESULTADO $M = 6$	ER (%) $M = 6$
$\tau_0$	0,9153	8,47	0,9195	8,05
$\omega$	0,8932	0,76	0,8907	1,03
$\rho_1$	0,5137	2,74	0,5240	4,8
$\rho_2$	0,4934	1,32	0,4977	0,46

Tabela 30: Resultados da máquina de comitê, para  $M = 10$  e  $M = 6$ , com  $N_{HI} = N_{H2} = 14$ ;  $\omega = 0,1$  e  $tol = 10^{-3}$ .

PARÂMETRO RADIATIVO	RESULTADO $M = 10$	ER (%) $M = 10$	RESULTADO $M = 6$	ER (%) $M = 6$
$\tau_0$	0,9406	5,94	0,9402	5,98
$\omega$	0,1321	32,1	0,1321	32,1
$\rho_1$	0,4972	0,56	0,4892	2,16
$\rho_2$	0,5330	6,6	0,5303	6,06

Tabela 31: Resultados da máquina de comitê, para  $M = 10$  e  $M = 6$ , com  $N_{H1} = N_{H2} = 14$ ;  $\omega = 0,9$  e  $tol = 10^{-3}$ .

PARÂMETRO RADIATIVO	RESULTADO $M = 10$	ER (%) $M = 10$	RESULTADO $M = 6$	ER (%) $M = 6$
$\tau_0$	0,9255	7,45	0,9291	7,09
$\omega$	0,8980	0,22	0,8989	0,12
$\rho_1$	0,4911	1,78	0,4993	0,14
$\rho_2$	0,4777	4,46	0,4784	4,32

Tabela 32: Resultados da máquina de comitê, para  $M = 10$  e  $M = 6$ , com  $N_{H1} = N_{H2} = 14$ ;  $\omega = 0,1$  e  $tol = 10^{-4}$ .

PARÂMETRO RADIATIVO	RESULTADO $M = 10$	ER (%) $M = 10$	RESULTADO $M = 6$	ER (%) $M = 6$
$\tau_0$	0,9915	0,85	0,9887	1,13
$\omega$	0,0954	4,6	0,0950	5,0
$\rho_1$	0,4681	6,38	0,4666	6,68
$\rho_2$	0,5112	2,24	0,5182	3,64

Tabela 33: Resultados da máquina de comitê, para  $M = 10$  e  $M = 6$ , com  $N_{H1} = N_{H2} = 14$ ;  $\omega = 0,5$  e  $tol = 10^{-4}$ .

PARÂMETRO RADIATIVO	RESULTADO $M = 10$	ER (%) $M = 10$	RESULTADO $M = 6$	ER (%) $M = 6$
$\tau_0$	0,9865	1,35	0,9771	2,29
$\omega$	0,4955	0,9	0,5185	3,7
$\rho_1$	0,5007	0,14	0,4939	1,22
$\rho_2$	0,5106	2,12	0,5117	2,34

Tabela 34: Resultados da máquina de comitê, para  $M = 10$  e  $M = 6$ , com  $N_{H1} = N_{H2} = 14$ ;  $\omega = 0,9$  e  $tol = 10^{-4}$ .

PARÂMETRO RADIATIVO	RESULTADO $M = 10$	ER (%) $M = 10$	RESULTADO $M = 6$	ER (%) $M = 6$
$\tau_0$	0,9611	3,89	0,9702	2,98
$\omega$	0,8990	0,11	0,9008	0,09
$\rho_1$	0,4998	0,04	0,4998	0,04
$\rho_2$	0,4941	1,18	0,4947	1,06

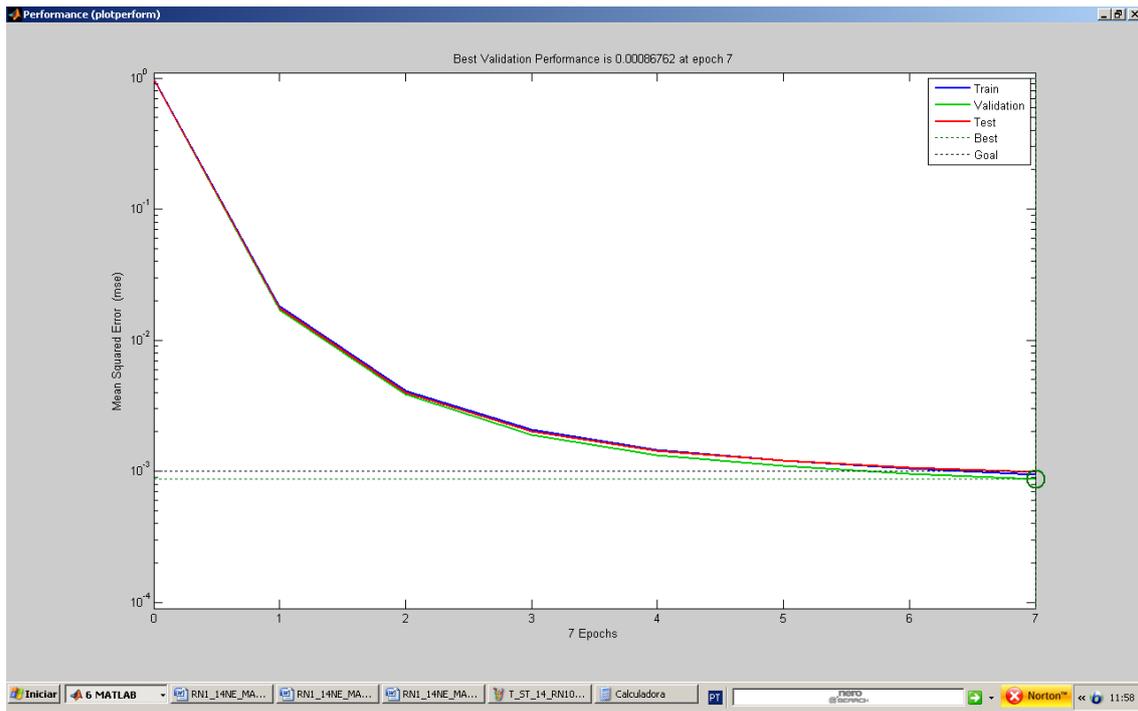


Figura: 38: Curva de performance do MLP do combinador  $N_h = 12$ ,  $tol = 10^{-3}$ .

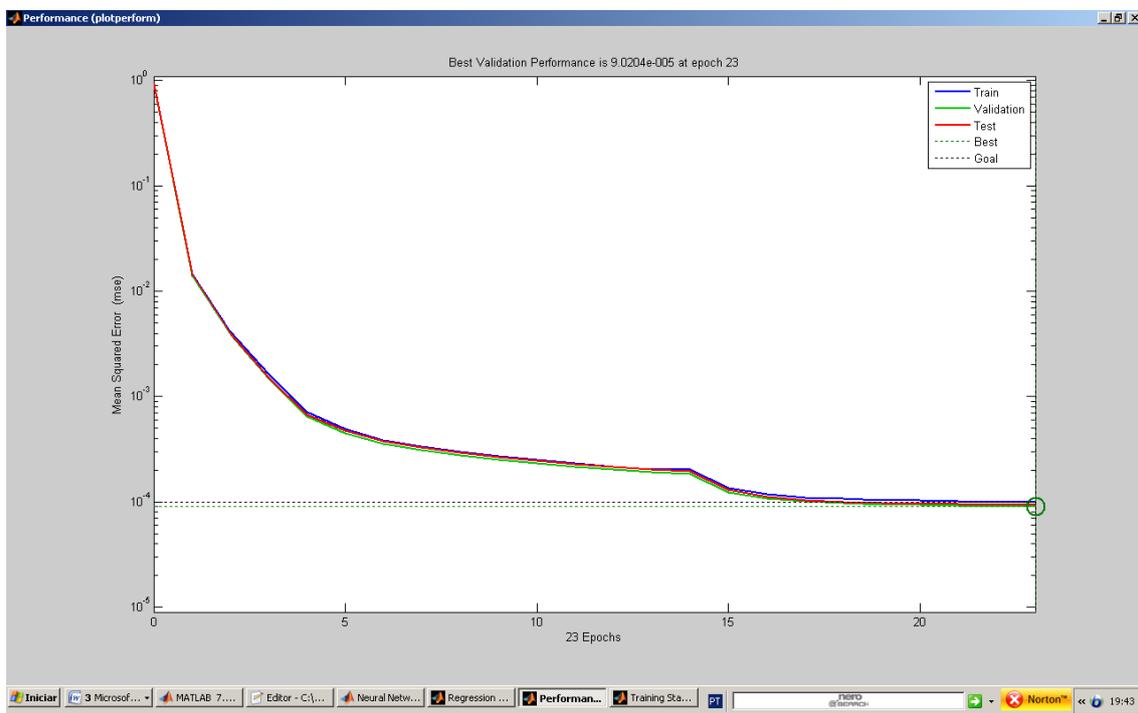


Figura: 39: Curva de performance do MLP do combinador  $N_h = 12$ ,  $tol = 10^{-4}$ .

## **SOBRE O AUTOR**

### **ROGÉRIO CAMPOS DE OLIVEIRA**

Natural de Natal – RN, formado em Engenharia Mecânica pela Universidade Gama Filho – Rio de Janeiro – RJ (julho de 1980), com especialização em Mecânica Computacional (*strictu sensu*) com ênfase em vibrações mecânicas e método dos elementos finitos (MEF) pela Universidade Federal do Rio Grande do Norte (2000) e mestre em Modelagem Computacional na área de Matemática Aplicada e Computacional com ênfase em IA e problemas inversos em engenharia, pelo Instituto Politécnico do Rio de Janeiro – IPRJ/UERJ, 2010, Nova Friburgo - RJ. É professor do quadro permanente, do Instituto Federal de Ciência e Tecnologia do Rio Grande do Norte, IFRN, desde fevereiro de 1995, onde leciona as disciplinas de Resistência dos Materiais I e II e Elementos de Máquinas nos cursos de Engenharia de Energia e Técnico Subsequente de Mecânica, tendo lecionado as seguintes disciplinas ao longo de sua carreira docente: geometria descritiva, desenho mecânico, desenho técnico, desenho geométrico, cálculo diferencial e integral I e II, cálculo numérico, matemática I, II e III e álgebra linear. Como engenheiro mecânico atuou em projetos e cálculo de suportes e análise de tensões e flexibilidade de tubulações nucleares e industriais nos projetos da usina nuclear de Angra II e da COPENE pela Natron Consultoria e Projetos SA, trabalhou na fase II do projeto da PECTEN pela Jakoo Poÿri Consultoria e Projetos SA e no projeto da expansão da RLAM - PETROBRAS em Aratú - BA pela Consulflex Engenharia e Consultoria Ltda. Fora da engenharia é auditor fiscal aposentado do tesouro estadual do Rio Grande do Norte, onde trabalhou de 1996 a 2022 nas áreas de automação comercial e auditoria fiscal.

ISBN 978-655376339-5



9

786553

763395