

INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA DO ESPÍRITO  
SANTO  
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA METALÚRGICA E DE  
MATERIAIS

**HELDER VAGO**

**ESTIMATIVA DA SIGNIFICÂNCIA DE  
VARIÁVEIS FÍSICAS INDEPENDENTES NO PROCESSO DE LAMINAÇÃO A  
QUENTE DE AÇO C-Mn, POR MEIO DE REDES NEURAIS ARTIFICIAIS.**

Vitória  
2014

**HELDER VAGO**

**ESTIMATIVA DA SIGNIFICÂNCIA DE  
VARIÁVEIS FÍSICAS INDEPENDENTES NO PROCESSO DE LAMINAÇÃO A  
QUENTE DE AÇO C-Mn, POR MEIO DE REDES NEURAIS ARTIFICIAIS.**

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Metalúrgica e de Materiais do Instituto Federal de Educação, Ciência e Tecnologia do Espírito Santo, como requisito parcial para obtenção do Título de Mestre em Engenharia Metalúrgica e de Materiais.

Orientador: Prof. Dr. Marcelo Lucas  
Pereira Machado

Vitória  
2014

**HELDER VAGO**

**ESTIMATIVA DA SIGNIFICÂNCIA DE  
VARIÁVEIS FÍSICAS INDEPENDENTES NO PROCESSO DE LAMINAÇÃO A  
QUENTE DE AÇO C-Mn, POR MEIO DE REDES NEURAIS ARTIFICIAIS.**

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Metalúrgica e de Materiais do Instituto Federal de Educação, Ciência e Tecnologia do Espírito Santo, como requisito parcial para obtenção do Título de Mestre em Engenharia Metalúrgica e de Materiais.

Aprovado em 14 de fevereiro de 2014.

**COMISSÃO EXAMINADORA**



**Prof. Dr. Marcelo Lucas Pereira Machado.**

Instituto Federal de Educação, Ciência e Tecnologia do Espírito Santo.

Orientador.



**Prof. Dr. José Denti Filho.**

Universidade Federal de Espírito Santo.



**Prof. Dr. Leonardo Cabral Gontijo.**

Instituto Federal de Educação, Ciência e Tecnologia do Espírito Santo.

## DECLARAÇÃO DO AUTOR

Declaro, para fins de pesquisa acadêmica, didática e técnico-científica, que a presente Dissertação pode ser parcialmente utilizada desde que se faça referência à fonte e ao autor.

Vitória, 14 de fevereiro de 2014.

  
HELDER VAGO

*"In memoriam" de meus pais,  
Sérvio Vago e Iolanda Rossi Vago,  
pela dedicação por me dar uma vida digna.*

## **AGRADECIMENTOS**

ao Prof. Dr. Marcelo Lucas Pereira Machado, a credibilidade dada à proposta do projeto, a dedicação, a disponibilidade e a maturidade na atividade de orientação;

ao Deus,  
ao Jesus Cristo,  
ao Divino Espírito Santo,  
à Santíssima Virgem Maria,  
à comunhão de todos os Santos,  
à comunhão de todos os Anjos e  
aos Irmãos de Luz,  
as bênçãos recebidas;

o amor incondicional de meus cães  
Landa, Kumah e Shari.

## RESUMO

Uma das técnicas de se obter o modelo de um processo é a modelagem caixa-branca que exige que se conheça bem o sistema a ser modelado, como as leis Físicas, Químicas ou de outras áreas científicas pertinentes ao mesmo, o que torna este tipo de modelagem complexa ou impraticável. A identificação de sistemas é uma outra forma de modelagem. Uma das características desta técnica é que pouco ou nenhum conhecimento prévio do sistema é necessário, o que a define como modelagem caixa-preta ou modelagem empírica. Os trabalhos efetuados em identificação de sistemas utilizando Redes Neurais Artificiais tem mostrado boa eficiência, alta velocidade de resposta e confiabilidade em seus resultados em diversas áreas, inclusive quando utilizadas para mapeamento de modelos para a predição de propriedades mecânicas de aço laminado a quente. O treinamento de uma Rede Neural Artificial, a partir dos padrões de entrada-saída, é capaz de memorizar e generalizar o processo em estudo proporcionando a capacidade de promover a análise da influência de Variáveis Independentes nas Variáveis Dependentes (sensibilidade) de um processo. Inspirado nessa proposta, este estudo primeiramente proporciona uma revisão bibliográfica sobre a utilização das Redes Neurais Artificiais no mapeamento de funções, afim de obter base científica na orientação da seleção de redes. Essa revisão de literatura claramente revela que as Redes Neurais Artificiais são utilizadas por vários pesquisadores na área de mapeamento de funções que, inclusive, orientam algumas arquiteturas para essa aplicação. A seguir é apresentada uma técnica de sensibilidade e um modelo caixa-branca para a laminação a quente de aço C-Mn. Também é estabelecida a metodologia para a verificação do comportamento das Redes Neurais Artificiais no mapeamento de processos; a validação da técnica de Sensibilidade como instrumento para determinar a significância das variáveis independentes de um processo e, por fim, a aplicação dessa técnica ao processo de Laminação de Tiras a Quente de aço C-Mn, confrontando-a com um Modelo caixa-branca. Os resultados deste estudo claramente indicam que as Redes Neurais Artificiais representam boas ferramentas para mapeamento de funções pois conseguem armazenar uma identificação satisfatória de um modelo, preservando sua generalização, podendo ser simulado a qualquer momento. A técnica de Sensibilidade da variável

independente é representativa da significância desta variável no processo. A variável de Sensibilidade de maior valor indica a que tem maior representação na variável dependente em estudo e que justificaria, por exemplo, investimentos em seu controle. Qualquer variável independente do processo pode ser inserida na análise de Sensibilidade, mesmo que essa não influencie na saída em estudo. Os resultados também confirmam que a técnica de Sensibilidade, via Rede Neural Artificial, aplicada a processos metalúrgicos mostra ser uma ótima ferramenta para identificar os parâmetros mais influentes e os parâmetros de menor efeito sobre a força de laminação, que pode impactar consideravelmente nas diretrizes de investimentos e projetos; bem como na orientação operacional, proporcionando redução de custos e melhoria da eficiência do processo.

Palavras-chave: Rede neural artificial. Sensibilidade de variáveis. Mapeamento de funções. Identificação de processo. Laminação de tiras a quente.



## ABSTRACT

One of the techniques to obtain a model of a process is the white-box modeling which requires well known system to be modeled, such as physical, chemical laws or other relevant scientific areas related to the subject; which makes this type of modeling complex or impractical. System identification is another form of modeling. One of the characteristics of this technique is that little or no prior knowledge of the system is necessary, what defines it as a black-box modeling or empirical modeling. The research carried out in identification systems using Artificial Neural Networks has shown good efficiency, high response speed and reliability in its results in several areas, including when used for mapping models for prediction of mechanical properties of hot rolled steel. The training of an Artificial Neural Network, from the input-output standards, is able to memorize and generalize the process under study providing the ability to promote the analysis of the influence of independent variables on dependent variables (sensitivity) of a process. Inspired by this proposal, this study first provides a literature review on the use of Artificial Neural Networks in mapping functions, in order to obtain scientific basis in guiding the selection of networks. This literature review clearly shows that Artificial Neural Networks are used by several researchers in the area of mapping functions which govern some architectures for this application. Following, a technique of sensitivity and a white box model for hot rolling of C-Mn steel are presented. The methodology for verification of the performance of Artificial Neural Networks in the mapping process is also established; validation of the technique of Sensitivity as an instrument to determine the significance of independent variables of a process and, finally, the application of this technique to the process of hot strip mill of C-Mn steel, confronting it with a white-box Model. The results of this study clearly indicate that Artificial Neural Networks are good tools for mapping functions because they can store a satisfactory identification of a model while preserving its generalization, which can be simulated at any time. The technique of sensitivity of the independent variable is representative of the significance of this variable in the process. The variable sensitivity of greater value indicates the highest representation in the dependent variable under study and that would justify, for example, investments in its control. Any independent variable of the process can be inserted in the sensitivity analysis, even if this does not influence on

output under study. The results also confirm that the technique of sensitivity, via Artificial Neural Network, applied to metallurgical processes proves to be a great tool to identify the most influential parameters and parameters of minor effect on the strength of lamination, which can impact significantly on the guidelines of investments and projects; as well as on operational guidance, providing cost reduction and improvement of process efficiency.

Keywords: Artificial neural network. Sensitivity of variables. Mapping functions. Process identification. Hot strip mill.

## LISTA DE FIGURAS

Figura 1 - Representação de um neurônio artificial identificando seus elementos básicos...	30
Figura 2 - Representação de uma RNA: entradas "xi"; 1ª camada intermediária com "n1" neurônios; 2ª camada intermediária com "n2" neurônios; camada de saída com número de neurônios igual à quantidade de saídas.....	32
Figura 3 - Arquiteturas das redes: (a) feedforward de camada simples (b) feedforward multicamadas (c) recorrente (d) reticulada.....	33
Figura 4 - Diagrama de blocos das fases de: treinamento/aprendizagem (a); e operação (b). .....	39
Figura 5 - Situações envolvendo underfitting e overfitting.....	42
Figura 6 - Caracterização de Underfitting pelo software: a RNA tem poucos neurônios.....	43
Figura 7 - Caracterização de Overfitting pelo software: a RNA tem muitos neurônios.....	43
Figura 8 - Superfície de erro para neurônios com funções de ativação não lineares apresentando Mínimos Locais e Mínimo Global.....	44
Figura 9 - Diagrama em blocos da identificação de sistemas.....	49
Figura 10- Equipamentos constituintes do Processo de Laminação de Tiras a Quente: Laminador de Desbaste (a), Trem acabador de 6 passes(b).....	53
Figura 11 - Diagrama do processo identificando as variáveis independentes temperatura (Te), redução (r) e velocidade (v); e a variável dependente força de laminação (F).....	71
Figura 12 - Arquitetura da RNA proposta.....	79
Figura 13 - Performance da Rede.....	80
Figura 14 - Parâmetros de treinamento default do software.....	80
Figura 15 - Correlação das saídas da rede "y" e das saídas originais "d" para treinamento, teste, validação e todas as saídas.....	82
Figura 16 - Comparação gráfica entre as saídas desejadas "d" e as saídas "y" fornecidas pela RNA.....	83
Figura 17 - Comparação gráfica entre as novas saídas "y" correspondentes às alterações, pelo acréscimo de 5% na entrada "x1", e as saídas originais "d".	84
Figura 18 - Comparação gráfica entre as novas saídas "y" correspondentes às alterações, pelo acréscimo de 5% na entrada "x2", e as saídas originais "d".	84

Figura 19 - Comparação gráfica entre as novas saídas "y" correspondentes às alterações, pelo acréscimo de 5% na entrada "x3", e as saídas originais "d". .....	85
Figura 20 - Sentido da relação entre a força de laminação e a temperatura no passe 1 .....	95
Figura 21 - Sentido da relação entre a força de laminação e a redução no passe 1 .....	95
Figura 22 - Sentido da relação entre a força de laminação e a velocidade no passe 1 .....	96
Figura 23 - Parâmetros de treinamento para todas RNAs utilizadas nesta fase de mapeamento. ....	97
Figura 24 - Tipos de convergência das RNAs a partir da relação entre número de padrões x número de neurônios. ....	98
Figura 25 - Comparação gráfica entre os valores de força de laminação: saídas desejadas "d" e as saídas fornecidas pela RNA "y", mapeando o Passe 1.....	101
Figura 26 - Comparação gráfica entre os valores de força de laminação: saídas desejadas "d" e as saídas fornecidas pela RNA "y", mapeando o Passe 2.....	101
Figura 27 - Comparação gráfica entre os valores de força de laminação: saídas desejadas "d" e as saídas fornecidas pela RNA "y", mapeando o Passe 3.....	102
Figura 28 -Comparação gráfica entre os valores de força de laminação: saídas desejadas "d" e as saídas fornecidas pela RNA "y", mapeando o Passe 4.....	102
Figura 29 - Comparação gráfica entre os valores de força de laminação: saídas desejadas "d" e as saídas fornecidas pela RNA "y", mapeando o Passe 5.....	102
Figura 30 - Comparação gráfica entre os valores de força de laminação: saídas desejadas "d" e as saídas fornecidas pela RNA "y", mapeando o Passe 6.....	103
Figura 31 - SENSN das Variáveis Independentes, por passe, via Modelo.....	105
Figura 32 - SENSN das Variáveis Independentes, por passe, via RNA. ....	106

## LISTA DE TABELAS

Tabela 1 - Performances dos treinamentos das RNAs, mapeando a função F1, com variações das funções de ativação e do número de neurônios na camada intermediária: 3 e 7. ....	75
Tabela 2 - Performances dos treinamentos das RNAs, mapeando a função F2, com variações das funções de ativação e do número de neurônios na camada intermediária: 3 e 7. ....	75
Tabela 3 - Performances dos treinamentos das RNAs, mapeando a função F3, com variações das funções de ativação e do número de neurônios na camada intermediária: 3 e 7. ....	76
Tabela 4 - Performances dos treinamentos das RNAs, mapeando a função F4, com variações das funções de ativação e do número de neurônios na camada intermediária: 3 e 7. ....	76
Tabela 5 - Identificação das topologias das RNAs treinadas, que apresentaram performances com erro na ordem de 10 <sup>-2</sup> ou menor, no mapeamento das funções em análise. ....	78
Tabela 6 - Sensibilidade Média da variável independente "x1". ....	86
Tabela 7 - Sensibilidade Média da variável independente "x2". ....	86
Tabela 8 - Sensibilidade Média da variável independente "x3". ....	86
Tabela 9 - Valores de Sensibilidade Média e Sensibilidade Média Normalizada das variáveis independentes obtidas por meio de RNAs. ....	87
Tabela 10 - Valores de Sensibilidade Média e Sensibilidade Média Normalizada obtidas por meio da função F3 aplicada às variáveis alteradas, uma a cada vez. ....	88
Tabela 11 - Valores de Sensibilidade Média Normalizada obtidas a partir da função F3 e por meio da RNA. ....	89
Tabela 12 - Erro Médio Absoluto das Saídas da RNA Treinada. ....	90
Tabela 13 - Erro Médio Absoluto das Saídas da RNA simulada com x1=0. ....	90
Tabela 14 - Erro Médio Absoluto das Saídas da RNA simulada com x2=0. ....	91
Tabela 15 - Erro Médio Absoluto das Saídas da RNA simulada com x3=0. ....	91
Tabela 16 - Valores dos erros, em %, para o treinamento e para cada situação de não existência da variável independente no processo. ....	91

Tabela 17 - Valores de Sensibilidades Médias Normalizadas, via RNA, e Erros Médios Absolutos apresentados por cada Variável Independente, considerando-as ZERADAS, uma a cada vez. ....	92
Tabela 18 - Faixas de cada variável, do banco de dados, utilizado para os ensaios, por passe. ....	93
Tabela 19 - Valores de SENS e SENSN, por passe, de cada variável independente encontrados por meio da aplicação da Modelagem. ....	94
Tabela 20 - Número de neurônios das camadas intermediárias das RNAs aprovadas, por passe. ....	100
Tabela 21 - Valores de Sensibilidades Médias e Sensibilidades Médias Normalizadas para todos os passes do processo de LTQ de Aço C-Mn, via RNA. ....	104
Tabela 22 - Valores de Sensibilidade Média Normalizada para as variáveis independentes obtidas via Modelo e por meio das RNAs. ....	105

## LISTA DE ABREVIACOES

EMA - Erro Medio Absoluto

EQM - Erro Quadratico Medio

IA - Inteligencia Artificial

LIN - Linear

LOG - Logstica

LTD - Laminao de Tiras a Quente

PMC - Perceptron MultiCamadas

RNA - Rede Neural Artificial

SENS - Sensibilidade Media

SENSN - Sensibilidade Media Normalizada

TANH - Tangente Hiperblica

TEM - Tenso de Escoamento Media

TEMc - Tenso de Escoamento Media Corrigida

TEMm - Tenso de Escoamento Media Modificada

## LISTA DE SÍMBOLOS

[C]: porcentagem de Carbono no aço C-Mn

d: distância entre passes

$d_{ant}$ : diâmetro do grão austenítico anterior à laminação

$d_i$ : valor original (desejado) da saída de um processo

$d_{if}$ : diferença individual de cada padrão

$d_{imáx}$ : valor máximo do range da saída original  $d_i$

$d_{imín}$ : valor mínimo do range da saída original  $d_i$

$d_{rec}$ : tamanho do grão austenítico completamente recristalizado após deformação

e: erro entre saídas de processo, desejada e fornecida

E: módulo de Young para o cilindro de trabalho

F: força de laminação

$F_{med}$ : força de laminação medida no processo

$g(u)$ : função de ativação

h: espessuras da peça laminada, tanto na sua entrada no processo, quanto em cada passe de laminação

$h_1$ : espessura de entrada do material na laminação

$h_2$ : espessura de saída do material na laminação

$h_n$ : espessura do material laminado, no ângulo neutro

$\Delta h$ : diferença entre as espessuras de entrada e saída no passe

[Mn]: porcentagem de Manganês no aço C-Mn

n: número de entradas da Rede Neural Artificial

N : quantidade de neurônios da camada escondida

$N_{max}$ : quantidade máxima de neurônios da camada escondida

P: número de padrões apresentado ao treinamento/aprendizagem da RNA

Q: fator geométrico da equação de Sims

r: grau de redução do material laminado

R: raio original do cilindro de laminação

$R'$ : raio do cilindro com deformação durante a laminação

s: número de saídas da RNA,

t: tempo de aplicação da deformação

$t_{din}$ : tempo de recristalização dinâmica



$t_p$ : tempo entre os passes  
 $t_{0,5}$ : tempo para alcançar 50% de recristalização  
 $T$ : temperatura do material no passe de laminação  
 $T_e$ : temperatura de entrada do aço no passe de laminação  
 $u$ : potencial de ativação  
 $v$ : velocidade periférica do cilindro no passe  
 $V$ : velocidade do cilindro  
 $V_e$ : velocidade de saída da tira  
 $V_r$ : velocidade tangencial do cilindro  
 $w_i$  - pesos sinápticos  
 $w_i$  antigo: valor antigo do peso sináptico  
 $w_i$  novo: valor novo do peso sináptico  
 $W$ : largura da peça laminada  
 $x_i$ : valor da entrada em um neurônio  
 $x_{ia}$ : valor atual da entrada  $x_i$  a ser normalizada  
 $x_{imáx}$ : valor máximo do range da entrada  $x_i$   
 $x_{imin}$ : valor mínimo do range da entrada  $x_i$   
 $x_{in}$ : valor normalizado da entrada  $x_{ia}$   
 $X_{ant}$ : fração anterior de recristalização do material  
 $X_{din}$ : fração de recristalização dinâmica do material  
 $y$ : sinal de saída de uma Rede Neural Artificial  
 $y_{ia}$ : valor atual da saída  
 $y_{in}$ : valor normalizado da saída  $y_{ia}$   
 $Z$ : parâmetro Zenner-Hollomon  
 $\epsilon$ : deformação efetiva  
 $\dot{\epsilon}$ : taxa de deformação  
 $\epsilon_a$ : deformação acumulada  
 $\epsilon_{aant}$ : deformação acumulada anterior  
 $\epsilon_c$ : deformação crítica  
 $\epsilon_r$ : deformação redundante  
 $\alpha$ : ângulo de contato do material com o cilindro  
 $\sigma$ : desvio padrão  
 $\sigma_s$ : tensão de escoamento no regime estacionário

$\Sigma$ : combinador linear

$\theta$ : limiar de ativação ou bias

$\lambda$ : fator de deslizamento

$\varphi_n$ : ângulo neutro

$\vartheta$ : razão de Poisson do material do cilindro de trabalho

$\eta$ : taxa de aprendizagem da Rede Neural Artificial.

## SUMÁRIO

<b>1 INTRODUÇÃO</b> .....	<b>21</b>
<b>2 JUSTIFICATIVA</b> .....	<b>25</b>
<b>3 OBJETIVOS</b> .....	<b>28</b>
3.1 OBJETIVO GERAL .....	28
3.2 OBJETIVOS ESPECÍFICOS: .....	28
<b>4 REFERENCIAL TEÓRICO</b> .....	<b>29</b>
4.1 REDE NEURAL ARTIFICIAL (RNA).....	29
<b>4.1.1 O Neurônio</b> .....	<b>29</b>
<b>4.1.2 Arquitetura das RNAs</b> . ....	<b>31</b>
<b>4.1.3 Topologia das RNAs</b> . ....	<b>33</b>
<i>4.1.3.1 Função de Ativação</i> .....	<i>34</i>
<b>4.1.4 Treinamento</b> .....	<b>36</b>
<b>4.1.5 Aprendizagem</b> .....	<b>37</b>
<b>4.1.6 Tipos de RNAs</b> .....	<b>39</b>
<i>4.1.6.1 Perceptron de Multicamadas (PMC)</i> .....	<i>40</i>
<b>4.1.7 Aproximação de Funções</b> .....	<b>48</b>
4.2 SENSIBILIDADE DE VARIÁVEL INDEPENDENTE .....	49
<b>4.2.1 Sensibilidade Média</b> . ....	<b>50</b>
<b>4.2.2 Sensibilidade Média Normalizada</b> . ....	<b>51</b>
4.3. LAMINAÇÃO A QUENTE DE AÇO C-Mn.....	51
<b>4.3.1 Breve descrição do processo de Laminação de Tiras a Quente (LTQ)</b> . ....	<b>52</b>
<b>4.3.2 Modelagem do Processo</b> . ....	<b>53</b>
<i>4.3.2.1 Equação de Sims</i> .....	<i>54</i>
<i>4.3.2.2 Tensão de Escoamento Média Modificada</i> .....	<i>56</i>
<i>4.3.2.3 Dados necessários do processo para modelagem caixa-branca</i> . ....	<i>62</i>

<b>5 METODOLOGIA</b> .....	<b>64</b>
5.1 RNA MAPEANDO FUNÇÕES SIMULADORAS .....	64
5.2 VALIDAÇÃO DA TÉCNICA DE SENSIBILIZAÇÃO DE VARIÁVEIS INDEPENDENTES.....	67
5.2.1 Sensibilidades por meio de simulação na RNA.....	67
5.2.2 Sensibilidades diretamente da Função selecionada.....	68
5.2.3 Validação da Técnica de Sensibilização por meio de RNAs. ....	68
5.2.4 Variáveis Independentes de Sensibilidades próximas a Zero (0). ....	68
5.3. SIGNIFICÂNCIA DAS VARIÁVEIS FÍSICAS INDEPENDENTES DO PROCESSO DE LAMINAÇÃO A QUENTE DE AÇO C-Mn. ....	69
5.3.1 Definição e seleção das variáveis para análise. ....	69
5.3.2 Aplicação da técnica de Sensibilidade ao processo por Modelagem, .....	71
5.3.3 Aplicação da técnica de Sensibilidade ao processo via RNA. ....	72
<b>6 APRESENTAÇÃO E DISCUSSÃO DOS RESULTADOS</b> .....	<b>74</b>
6.1 RNAs MAPEANDO FUNÇÕES SIMULADORAS. ....	74
6.2 VALIDAÇÃO DA TÉCNICA DE SENSIBILIZAÇÃO DE VARIÁVEIS INDEPENDENTES.....	78
6.2.1 Sensibilidades por meio de simulação na RNA.....	79
6.2.2 Sensibilidades diretamente da Função selecionada.....	87
6.2.3 Validação da Técnica de Sensibilização por meio de RNAs. ....	88
6.2.4 Variáveis Independentes de Sensibilidades próximas a Zero (0). ....	89
6.3 APLICAÇÃO DA TÉCNICA DE SENSIBILIDADE NA LTQ .....	92
6.3.1. Seleção dos dados de entrada.....	93
6.3.2 Técnica de Sensibilização via Modelagem.....	93
6.3.3 Técnica de Sensibilização via RNA.....	95
6.3.3.1 Conhecimentos Preliminares.....	95
6.3.3.2 Parâmetros de treinamento. ....	96
6.3.3.3 Número de Neurônios na camada intermediária e número de Padrões.....	98

6.3.3.4 RNAs mapeando o processo de LTQ do aço C-Mn. ....	99
6.3.3.5. Técnica de Sensibilidade aplicada ao processo de LTQ via RNA. ....	103
<b>6.3.4 Comparação entre as duas formas de aplicação da técnica de Sensibilidade: Modelagem x RNA.....</b>	<b>104</b>
<b>7 CONCLUSÕES .....</b>	<b>107</b>
<b>8 SUGESTÕES PARA TRABALHOS FUTUROS. ....</b>	<b>108</b>
<b>REFERÊNCIAS.....</b>	<b>109</b>

## 1 INTRODUÇÃO

A conjuntura econômica mundial globalizada proporcionou intercâmbio de informações e globalização de produtos e consumidores, fazendo com que um mesmo produto seja desejado e consumido em diversas partes do mundo. Este fato provocou a concorrência entre as empresas que exigiu suas adequações quanto à necessidade de atingir maior número de mercados; adaptação de suas estruturas de custos e de seus sistemas produtivos para a melhoria contínua da produtividade; instalação e desenvolvimento de inovação tecnológica entre outros (TUBINO, 2000).

Segundo Campos et al. (2006), "Uma das áreas tecnológicas fundamentais para se aumentar a rentabilidade das unidades é a de controle, automação e otimização de processos industriais". Dentre as diversas atividades desta área destaca-se a operacionalização de certas variáveis pertencentes ao sistema industrial, geralmente desejando-se mantê-las dentro de limites operacionais desejáveis. Estas variáveis podem ser as denominadas dependentes, que sofrem os efeitos do tratamento do processo provocados pelas variáveis independentes, resultado de uma inter-relação causal no processo industrial. Estes efeitos podem ser observados nas medições das variáveis dependentes (GRESSLER, 2004).

A inter-relação causal entre as variáveis pode ser expressa por meio de um modelo matemático de aproximação do processo real, representado por uma equação ou conjunto de equações, que caracteriza o conhecimento dos aspectos essenciais do processo. Existem diversas técnicas de se obter o modelo de um processo. Uma delas é a modelagem caixa-branca, utilizando métodos clássicos de modelagem de processos, que exige que se conheça bem o sistema a ser modelado, como as leis Físicas, Químicas ou de outras áreas científicas pertinentes ao mesmo (GARCIA, 2005).

A identificação de sistemas é uma outra forma de modelagem. Uma das características desta técnica é que pouco ou nenhum conhecimento prévio do sistema é necessário, o que a identifica como modelagem caixa-preta ou modelagem empírica (AGUIRRE, 2007).

Norgaard et al. (2000) também define como modelagem caixa-preta o tipo de identificação que baseia-se exclusivamente em dados, supondo que tenha-se nenhum ou apenas pouco conhecimento sobre a fenomenologia do sistema.

Podem haver vários motivos para se estabelecer as descrições matemáticas de um sistema dinâmico: concepção de sistema de predição, simulação, controle e detecção de falhas, entre outros. Se o conjunto associado com a construção de um modelo usando as leis da Física, Química, Economia, etc., é considerado muito abrangente, a técnica de identificação de sistema é, naturalmente, de maior interesse.

Segundo Aguirre (2007), "[ ] frequentemente não se conhece todas as equações envolvidas no funcionamento de um determinado sistema" ou seria impraticável levantar tais equações, tornando, assim, muito interessante e relevante o estudo de técnicas de identificação.

Os trabalhos efetuados em identificação de sistemas utilizando Redes Neurais Artificiais tem mostrado boa eficiência, alta velocidade de resposta e confiabilidade em seus resultados em diversas áreas tecnológicas, financeiras, alimentícias, entre outras.

Lee e Lee (2002) iniciaram um período de aplicação de Redes Neurais Artificiais buscando a melhoria da predição da força de laminação no processo de laminação a quente devido a variações observadas nas espessuras do material laminado.

A seguir Son et al. (2004) propõem um modelo, por meio de Rede Neural Artificial, que apresentou resultados melhores do que a modelagem analítica usada como comparativo, confirmando em sua pesquisa a aplicação das redes na predição da força de laminação.

Son et al. (2005) estudaram, ainda, a utilização da combinação da Rede Neural Artificial operando off-line e on-line na predição da força de laminação. A Rede Neural Artificial, de arquitetura *feedforward*, tinha uma camada intermediária com

função de ativação tangente sigmoide e função de treinamento de Levenberg-Marquardt, que apresentou excelentes resultados.

Diversas técnicas tem sido utilizadas para mapeamento de modelos para a predição de propriedades mecânicas de aço laminado a quente, destacando-se a Análise de Regressão Linear Múltipla e as Redes Neurais Artificiais.

A Análise de Regressão é uma técnica utilizada para modelar a relação entre duas ou mais variáveis através da minimização da soma dos quadrados da diferença entre os valores preditos e os valores reais (OLIVEIRA & MODENESI, 2008).

No trabalho de Jones et al. (2005), uma comparação foi feita na aplicação das técnicas de regressão múltipla linear, regressão múltipla não linear e redes neurais não lineares para várias famílias de aço laminados a quente. A abordagem da rede neural não linear, em análise caixa-preta, devidamente estruturada forneceu modelos de predição mais precisos do que o uso de regressão múltipla linear ou não linear.

Resultado idêntico foi obtido por Moussaoui et al. (2006) que constatou que o modelo utilizando Rede Neural Artificial foi capaz de prever a força de laminação com mais precisão do que o modelo analítico de Alexander-Ford para laminação a quente.

Azimzadegan et al. (2013) desenvolveram uma Rede Neural Artificial, de arquitetura *feedforward* de múltiplas camadas, para mapear e prever as propriedades de aço de baixa liga, usados em gasodutos, em função de sua composição química e dos tratamentos térmicos sofridos. Os valores da predição da rede estavam de acordo com os valores experimentais e concluíram que a Rede Neural Artificial provou ser uma ferramenta poderosa na otimização de muitos processos industriais.

Por exemplo, em estudo da predição de otimização de comprimentos de parafusos usados em escavações de rochas a utilização de Rede Neural Artificial por Hasanzadehshoiiili et al. (2012) apresentou resultados superiores aos métodos estatísticos convencionais. A arquitetura da rede era *feedforward*.



O treinamento de uma Rede Neural Artificial, a partir dos padrões de entrada-saída, é capaz de memorizar e generalizar o processo em estudo. A característica de generalização proporciona à Rede Neural Artificial a capacidade de promover a análise da significância de Variáveis Independentes do processo, ou seja, a influência destas variáveis nas Variáveis Dependentes.

Pode-se utilizar técnicas de Sensibilização de Variáveis Independentes que fornecem informações de variações na variável dependente frente a variações em variável independente. Hasanzadehshooilli et al. (2012) usou o Método de Amplitude de Cosseno (CAM) para análise de Sensibilidade em seu estudo. Outra técnica de Sensibilização pode ser efetivada por meio de simulações em uma Rede Neural Artificial treinada para mapeamento de funções.

Inspirado nessa proposta, este estudo primeiramente proporciona uma revisão bibliográfica sobre a utilização das Redes Neurais Artificiais no mapeamento de funções, afim de obter base científica na orientação da seleção de redes. A revisão de literatura claramente revela que as Redes Neurais Artificiais são utilizadas por vários pesquisadores na área de mapeamento de funções que, inclusive, orientam algumas arquiteturas para essa proposta.

Este estudo está organizado da seguinte forma: a Seção 4 apresenta uma revisão teórica sobre Redes Neurais Artificiais, uma apresentação da técnica de sensibilidade a ser adotada neste estudo e um modelo caixa-branca para a laminação a quente de aço C-Mn. Na seção 5 são apresentados os métodos aplicados para: a verificação do comportamento das Redes Neurais Artificiais no mapeamento de processo; a validação da técnica de sensibilidade como instrumento para determinar a significância das variáveis independentes de um processo e, por fim, a aplicação desta técnica ao processo de Laminação de Tiras a Quente de aço C-Mn. Os resultados verificados na aplicação destes métodos e suas respectivas discussões são apresentados na seção 6. Na seção 7 encontram-se as conclusões verificadas no presente estudo. São, também, apresentadas sugestões para trabalhos futuros na seção 8.

## 2 JUSTIFICATIVA

Segundo Norgaard et al. (2000), um problema significativo na modelagem de processos é o fato de simplesmente ser impossível abarcar todos os tipos de conhecimentos a priori que possam existir. Outro problema é que mesmo que o conhecimento sobre o sistema esteja disponível, é difícil utilizá-lo no que se refere a uma descrição de tempo contínuo do sistema, isto é, em termos de equações diferenciais.

Independentemente do fato de que todos os sistemas a princípio são não lineares, a maior parte da literatura sobre a identificação de sistemas trata da identificação de sistemas lineares. Algumas das razões mais importantes para isso são:

- a) muitos sistemas podem ser bem descritos por um modelo linear e, quando não é razoável utilizar um único para toda a gama de funcionamento, às vezes é possível identificar diferentes modelos lineares para diferentes regimes da faixa de operação;
- b) a partir de uma perspectiva computacional é menos complicado de executar;
- c) a análise é menos complicada do ponto de vista estatístico;
- d) é muito mais simples projetar controladores para um sistema descrito por um modelo linear.

Então, as ferramentas matemáticas disponíveis para análise de sistemas não lineares normalmente utilizam técnicas de linearização, transformando as características não lineares em pequenas tarefas lineares. "Apesar de muitas vezes eficiente, essa abordagem não retrata a realidade dos sistemas físicos, podendo resultar em soluções subótimas" (BRAGA et al., 2007).

O argumento de Aguirre (2007) reforça que a análise de modelos não lineares exige a utilização de algoritmos muito complexos e que estes apresentam certos regimes dinâmicos que modelos lineares não conseguem representar.

Norgaard et al. (2000) explica que transferir o conhecimento para uma descrição discretizada geralmente é muito difícil e, portanto, este conhecimento é frequentemente perdido no processo de discretização.

Na comparação de técnicas de identificação é difícil dizer que uma irá superar outra antes que ambas tenham sido avaliadas. Ao se buscar uma técnica que execute razoavelmente bem e de forma prática uma identificação, certo tipos de Redes Neurais Artificiais parecem ser uma excelente escolha. Em especial, a rede Perceptron Multicamadas ganhou uma imensa popularidade pois possui uma capacidade impressionante de aproximação de funções (NORGAARD et al., 2000).

Pela definição de Haykin (2001), "Uma Rede Neural Artificial é uma máquina que é projetada para modelar".

As Redes Neurais Artificiais são processadores simples constituídos de neurônios que tem a tendência natural para armazenar conhecimento experimental e torná-lo disponível para o uso. Estes neurônios apresentam a possibilidade de serem não lineares, dando esta característica à Rede Neural Artificial a que pertencem; propriedade muito importante para processamento de mecanismos físicos inerentemente não lineares.

Uma Rede Neural Artificial representa um mapeamento entre os seus espaços de entrada e de saída: pode ser tanto uma relação linear, quanto uma relação altamente não linear. O que determina o mapeamento são a sua estrutura, organização dos neurônios, suas conexões e a função associada a cada neurônio (CAMPOS et al., 2004).

Braga et al. (2007) também cita a importância das Redes Neurais Artificiais como ferramenta para controle de sistemas devido à sua capacidade de modelar com fidelidade ações não lineares; e ainda acrescenta que as Redes Neurais Artificiais são capazes de atuar como mapeadoras universais de funções multivariáveis.

Norgaard et al. (2000) acrescenta que, além das áreas de identificação de sistemas e controle, as Redes Neurais Artificiais tem sido aplicadas em diversas outras, tais como: medicina, processamento de imagens, financeira, reconhecimento de voz entre outras.

Já Silva et al. (2010), cita que, conforme testemunhados em diversos periódicos científicos, as Redes Neurais Artificiais tem sido utilizadas em aplicações de medicina, biologia, finanças e economia, ecologia, farmácia, acústica, indústria de alimentos, setores automotivos e aeroespaciais.

As citações anteriores proporcionam maior segurança ao se propor as Redes Neurais Artificiais como mapeadoras de funções, principalmente não lineares, e modeladoras de processos por método de identificação empírica.

### 3 OBJETIVOS

#### 3.1 OBJETIVO GERAL

Estimar a significância de variáveis físicas independentes (temperatura no passe, grau de redução da espessura e velocidade de laminação) sobre a força de laminação, no processo de laminação de tiras a quente do aço C-Mn; por meio de identificação de sistemas, mapeados por Redes Neurais Artificiais.

#### 3.2 OBJETIVOS ESPECÍFICOS:

- a) identificar, caracterizar e selecionar as Redes Neurais Artificiais quanto às suas aplicações como mapeadoras de funções; e verificar suas performances na identificação de sistemas;
- b) validar a Técnica de Sensibilização de Variáveis Independentes, utilizando Rede Neural Artificial, por meio de uma expressão aritmética, simulando um processo; e de dados sintéticos;
- c) estimar e validar as significâncias das variáveis físicas independentes com relação à força de laminação no processo de laminação de tiras a quente do aço C-Mn, utilizando a técnica de sensibilização simulada em Redes Neurais Artificiais.

## 4 REFERENCIAL TEÓRICO

### 4.1 REDE NEURAL ARTIFICIAL (RNA)

"Um dos motivos primários era criar um programa de computador que fosse capaz de aprender por experiência" (NORGAARD et al., 2000).

RNAs são modelos computacionais que, baseados em informações, possuem a capacidade de aquisição e manutenção do conhecimento. Uma vez ajustadas, as RNAs passam a ter a habilidade de produzir saídas adequadas para entradas que não estavam presentes no seu treinamento, desde que tais entradas sejam oriundas do mesmo tipo de população alvo do problema. Esta nada mais é do que a característica de generalização da rede, ou seja, da capacidade de fornecer respostas coerentes a padrões ou casos novos (BRAGA et al., 1998).

Piuleac et al. (2010) comenta que o estabelecimento de uma RNA consiste das seguintes etapas:

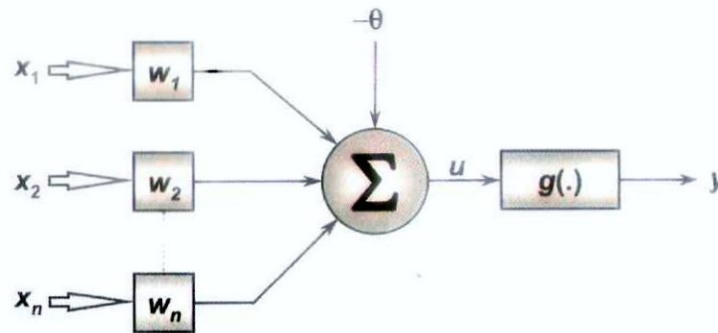
- a) coleta de dados, que compõem o conjuntos de padrões de ensaio para treinamento (formação) da rede e teste;
- b) desenvolvimento da topologia da rede;
- c) treinamento e, finalmente;
- d) estabelecimento do desempenho do modelo da RNA, comparando-se a predição (saídas) da rede para padrões não utilizados no treinamento.

As RNAs são definidas como um conjunto de unidades de processamento, os neurônios artificiais, que são interligadas por interconexões, as sinapses artificiais, representadas por vetores ou matrizes de pesos sinápticos (SILVA et al., 2010).

#### 4.1.1 O Neurônio.

Silva et al. (2010) caracteriza o neurônio artificial como constituído de sete elementos básicos que estão apresentados na Figura 1 e definidos a seguir:

**Figura 1** - Representação de um neurônio artificial identificando seus elementos básicos.



Fonte: SILVA et al. (2010).

- sinais de entrada  $\{x_1, x_2, \dots, x_n\}$ , recebidos de uma aplicação específica e representam os valores assumidos por uma variável. Antes do treinamento, os dados de entrada são normalizados para evitar que uma variável de entrada tenha um efeito maior no treinamento, em comparação com as outras (PIULEAC et al, 2010). YAP et al. (2012) caracteriza esta normalização conforme expressão:

$$X_{in} = \frac{X_{ia} - X_{imín}}{X_{imáx} - X_{imín}} \quad (1)$$

onde:

$x_{ia}$  é o valor atual da entrada  $x_i$  a ser normalizada,

$x_{imín}$  é o valor mínimo do range da entrada  $x_i$ ,

$x_{imáx}$  é o valor máximo do range da entrada  $x_i$  e

$x_{in}$  é o valor normalizado da entrada  $x_{ia}$ ; que estará na faixa de 0 a 1.

- pesos sinápticos  $\{w_1, w_2, \dots, w_n\}$ , ponderam as variáveis de entrada permitindo quantificar suas relevâncias;

- combinador linear ( $\Sigma$ ), soma os valores que foram ponderados pelos pesos;

- limiar de ativação ou bias ( $\theta$ ), aumenta ou diminui a entrada na função de ativação, deslocando o campo local do neurônio (HAYKIN, 2001), aumentando o número de graus de liberdade da RNA o que tende a ampliar sua capacidade de ajustes de dados;

- potencial de ativação ( $u$ ), soma do resultado do combinador linear e limiar de ativação;

- função de ativação  $\{g\}$ , limita a saída do neurônio dentro de um intervalo de valores;
- sinal de saída  $\{y\}$ , valor final produzido pelo neurônio. Este valor é "desnormalizado" a partir dos valores mínimo e máximo das saídas originais, conforme expressão:

$$Y_{ia} = Y_{in} * (d_{im\acute{a}x} - d_{im\grave{i}n}) + d_{im\grave{i}n} \quad (2)$$

onde:

$y_{in}$  é o valor normalizado da saída  $y_{ia}$ , que estará na faixa de 0 a 1;

$d_{im\grave{i}n}$  é o valor mínimo do range da saída original  $d_i$ ,

$d_{im\acute{a}x}$  é o valor máximo do range da saída original  $d_i$  e

$y_{ia}$  é o valor atual da saída "desnormalizada".

SILVA et al. (2010), utilizando as identificações das variáveis, estabelece as expressões relacionadas ao funcionamento do neurônio, onde "n" é o número de entradas:

$$u = \sum_{i=1}^n w_i \cdot x_i - \theta \quad (3)$$

$$y = g(u) \quad (4)$$

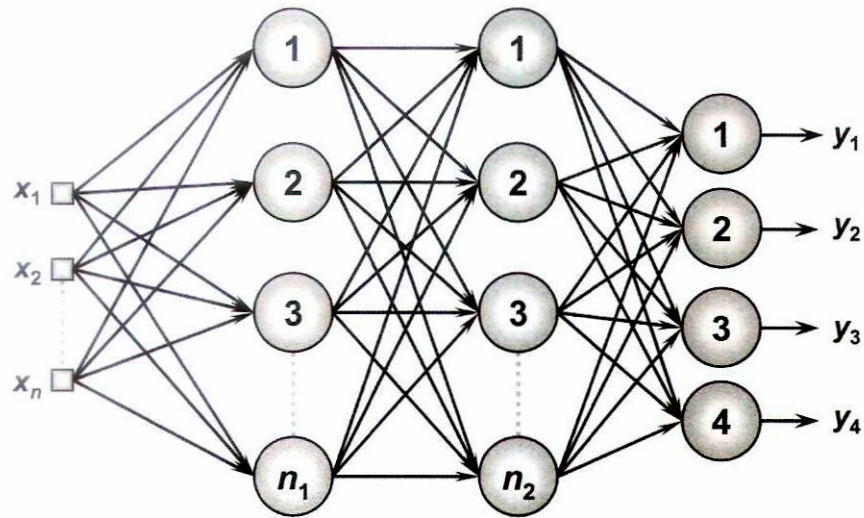
#### 4.1.2 Arquitetura das RNAs.

A arquitetura de uma RNA considera a disposição dos neurônios, as formas de interligação destes e a constituição de suas camadas.

Uma RNA pode ser representada conforme Figura 2, onde podem ser vistas suas partes componentes: camada de entrada, camadas intermediárias e camada de saída com seus respectivos neurônios.



**Figura 2** - Representação de uma RNA: entradas " $x_i$ "; 1ª camada intermediária com " $n_1$ " neurônios; 2ª camada intermediária com " $n_2$ " neurônios; camada de saída com número de neurônios igual à quantidade de saídas.



Fonte: SILVA et al. (2010).

As principais camadas de uma RNA são:

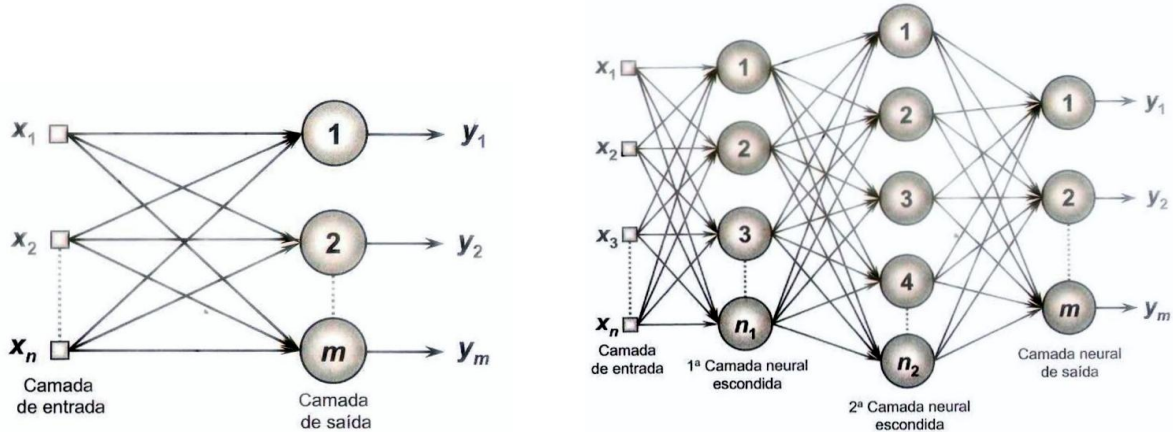
- a) a de entrada, constituída pelos próprios padrões de entrada, para recebimento das informações;
- b) as intermediárias, que extraem as características associadas ao processo a ser inferido; e
- c) a de saída, para produção e apresentação dos resultados finais da rede.

Silva et al. (2010) divide as principais arquiteturas de redes em:

- a) redes *feedforward* de camada simples, com uma camada de entrada e uma de saída, empregadas em problemas envolvendo classificação de padrões e filtragem linear; representada na Figura 3(a);
- b) *feedforward* de múltiplas camadas, com uma ou mais camadas intermediárias (ou ocultas, ou escondidas, ou *hidden*) empregadas em aproximação de funções, classificação de padrões, identificação de sistemas, otimização, robótica, controle de processos; representada na Figura 3(b);
- c) recorrentes, que possuem pelo menos um laço de realimentação, onde a saída de um neurônio alimenta outro neurônio da mesma camada ou de uma anterior, empregadas em processamento dinâmico de informações; representada na Figura 3(c);

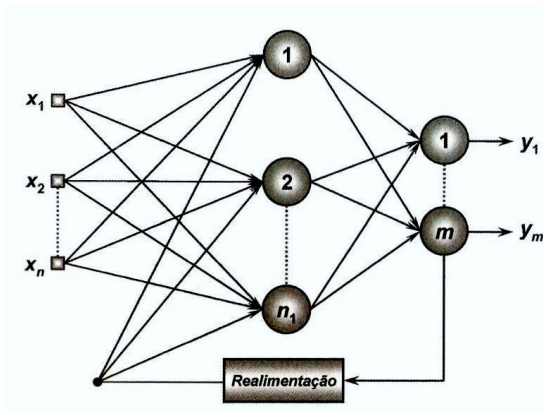
- d) reticuladas, onde os sinais de entradas são inseridos em todos os neurônios; empregadas para reconhecimento de padrões, otimização de sistemas; representada na Figura 3(d).

**Figura 3** - Arquiteturas das redes: (a) feedforward de camada simples (b) feedforward multicamadas (c) recorrente (d) reticulada.

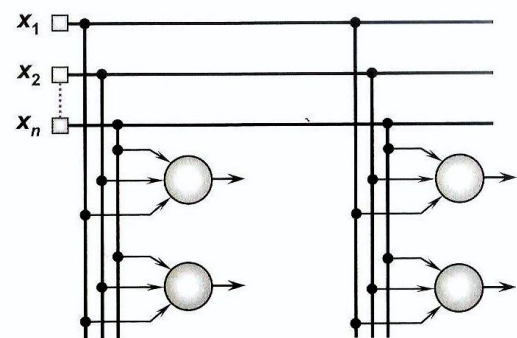


(a)

(b)



(c)



(d)

Fonte: SILVA et al. (2010).

Campos et al. (2004) comenta que uma rede em malha aberta (estática) representa geralmente uma relação algébrica não linear entre suas entradas e saídas.

#### 4.1.3 Topologia das RNAs.

Segundo Aguirre (2007), a determinação da topologia de uma RNA envolve escolher:

- a) as variáveis de entrada;
- b) o número de neurônios;
- c) o número de camadas;
- d) o tipo da função de ativação; e
- e) as conexões necessárias.

Um das técnicas estatísticas mais utilizadas para seleção das melhores topologias é a validação cruzada (*cross-validation*) cujo propósito é avaliar a performance da rede: a RNA candidata é treinada com um conjunto de dados e sua validação é testada por um outro conjunto de dados diferente daquele usado no ajuste de seus parâmetros internos. O conjunto total de padrões disponíveis é aleatoriamente dividido em duas partes, isto é, subconjunto de treinamento e subconjunto de teste (ou validação) (SILVA et al., 2010).

No software MATLAB®, versão 7.12.0.635 (R2011a), verifica-se a divisão dos padrões, por meio da ferramenta "nntool" em três grupos: treinamento, com 70% dos padrões; validação, com 15% dos padrões; e teste, com 15% dos padrões. O treinamento é usado para ajustar os pesos e os limiares de ativação. Os padrões de validação são utilizados para verificação da convergência das saídas e ajuste dos pesos e dos limiares de ativação. Também são referência para parar o treinamento mais cedo se o desempenho da RNA piorar. Os padrões de teste são usados como uma verificação adicional de que a RNA apresenta uma boa generalização, mas não tem efeito algum sobre o treinamento.

#### 4.1.3.1 Função de Ativação

Existem diversos tipos de funções de ativação que servem para limitar a saída do neurônio e são selecionadas para utilização conforme a complexidade de sistema. Podem ser divididas em dois grupos principais: funções parcialmente diferenciáveis e funções totalmente diferenciáveis.

Funções de ativação parcialmente diferenciáveis: função degrau, função degrau bipolar ou função sinal e função rampa simétrica.

Funções de ativação totalmente diferenciáveis: função logística, função tangente hiperbólica, função gaussiana e função linear.

Sendo "u" o sinal de entrada da função de ativação, cada função é definida da seguinte forma (SILVA et al., 2010):

Função degrau:

$$g(u) = \begin{cases} 1, & \text{se } u \geq 0 \\ 0, & \text{se } u < 0 \end{cases}$$

Função degrau bipolar:

$$g(u) = \begin{cases} 1, & \text{se } u \geq 0 \\ -1, & \text{se } u < 0 \end{cases}$$

Rampa Simétrica:

$$g(u) = \begin{cases} a & \text{se } u > a \\ u & \text{se } -a \leq u \leq a \\ -a & \text{se } u < -a \end{cases}$$

Logística:

$$g(u) = \frac{1}{(1 + e^{-\beta u})}$$

Tangente hiperbólica:

$$g(u) = \frac{(1 - e^{-\beta u})}{(1 + e^{-\beta u})}$$

Gaussiana:

$$g(u) = e^{-(u-c)^2/(2\sigma^2)}$$

Linear:

$$g(u) = u$$

A função de ativação linear é aplicada na utilização de RNAs como aproximadoras universais de funções, visando-se mapear o comportamento entre as variáveis de entrada e saída de processos. Se a função linear for utilizada como ativação da camada de saída, tem-se que o respectivo neurônio de saída realizará tão somente

uma combinação linear das funções de ativação implementadas pelos neurônios da camada anterior (SILVA et al., 2010).

#### 4.1.4 Treinamento.

O processo de treinamento de uma RNA consiste na aplicação de passos ordenados para sintonização dos pesos sinápticos e limiares de seus neurônios, objetivando a generalização de soluções a serem produzidas pelas suas saídas, cujas respostas são representativas do sistema físico em que estas estão mapeando (SILVA et al., 2010).

Cada apresentação completa das amostras do subconjunto de treinamento visando o ajuste do vetor de peso "w" é chamada de "época". As épocas são repetidas até que o erro desejável se torne pequeno ou um critério de parada seja satisfeito (PIULEAC et al., 2010). O número máximo de épocas para um determinado treinamento pode ser estabelecido como critério de parada quando a precisão, ou seja, o erro solicitado é inalcançável (*early stopping*).

Os treinamentos podem ser de diferentes formas: supervisionados, cada padrão de amostras do treinamento é composto pelos sinais de entrada e suas correspondentes saídas; não supervisionados, não existe as respectivas saídas desejadas; *off-line*, os ajustes efetuados nos vetores de pesos das redes e em seus limiares só são efetivados após a apresentação de todo o conjunto de treinamento; *on-line*, os padrões são apresentados um por vez. O treinamento supervisionado é o método de treinamento mais comum (DEHGHANI & NEKAHI, 2010). O treinamento supervisionado de uma RNA consiste na modificação iterativa dos pesos sinápticos da rede para minimizar o Erro Quadrático Médio (EQM) entre a resposta desejada, a partir dos dados de entrada, e a saída obtida no neurônio de saída.

O conjunto de padrões para treinamento deve ser grande o suficiente e conter vários exemplos de todos os casos operacionais possíveis e o conjunto de teste deve conter pelo menos um exemplo de cada caso real possível, de forma a validar a rede (CAMPOS et al., 2006).

#### 4.1.5 Aprendizagem.

"Um dos destaques mais relevantes das RNAs está na capacidade de aprender a partir da apresentação de amostras que exprimem o comportamento do sistema" (SILVA et al., 2010).

O conjunto dos passos ordenados, citado na seção anterior 4.1.4, visando o treinamento da rede é denominado de Algoritmo de Aprendizagem. Ao longo de sua aplicação, a rede será capaz de extrair características discriminantes do sistema a ser mapeado por intermédio de amostras que foram retiradas do contexto deste sistema.

Existem diversas formas de aprendizagem que são explicadas por Haykin (2001), tais como: correção de erro, baseada em memória, hebbiana, competitiva, boltzmann... As funções de aprendizagem são expressas por meio de equações, que são transformadas em algoritmos.

Braga et al. (2007) explica que o aprendizado supervisionado se aplica à pesquisa de mapeamento entre padrões de entrada e de saída. Os exemplos mais conhecidos de algoritmos de aprendizagem supervisionada são a Regra Delta e sua generalização para redes de múltiplas camadas, o algoritmo *Backpropagation*.

Braga et al. (2007) ainda comenta que o exemplo mais típico de aprendizado supervisionado é o de correção de erros: cada padrão da entrada "i" tem a sua saída de rede "y<sub>i</sub>" comparada com a desejada "d<sub>i</sub>" que fornece informações sobre a direção de ajuste dos pesos. A minimização dessa diferença é incremental buscando minimizar o erro "e" entre essas respostas:

$$e = d_i - y_i \tag{5}$$

A atualização dos pesos por correção dos erros, utilizada pelos algoritmos de aprendizagem, pode ser representada de forma genérica por (BRAGA et al., 2007):

$$W_i \text{ novo} = W_i \text{ antigo} + \eta \cdot e \cdot X_i \quad (6)$$

onde:

$W_i \text{ novo}$ : valor novo do peso sináptico;

$W_i \text{ antigo}$ : valor antigo do peso sináptico;

$\eta$ : taxa de aprendizagem da rede,

em geral:  $0 < \eta < 1$ ,

e para redes multicamadas:  $0,25 \leq \eta \leq 0,75$ , inclusive (SILVA et al., 2010);

$e$ : erro;

$x_i$ : valor de entrada do neurônio correspondente.

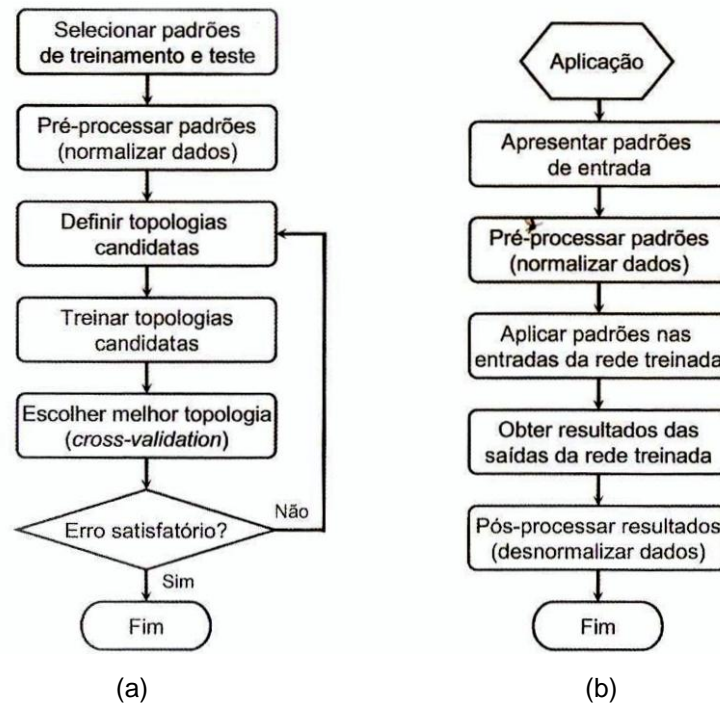
O livro do SILVA et al. (2010), citado na referência, apresenta o algoritmo detalhado na página 65, e demonstrado nas páginas 77 e 78.

Vale observar que o erro é diretamente mensurável, ou seja, as saídas são fornecidas ao treinamento, o que caracteriza aprendizagem supervisionada.

A Regra Delta define uma parcela dessa expressão: "O ajuste feito em um peso sináptico de um neurônio é proporcional ao produto do sinal de erro pelo sinal de entrada da sinapse em questão" (HAYKIN, 2001).

As fases de treinamento/aprendizagem e operação (simulação) são apresentadas na Figura 4, por meio de diagramas de blocos que orientam os diversos passos para seleção e utilização de uma RNA treinada.

**Figura 4** - Diagrama de blocos das fases de: treinamento/aprendizagem (a); e operação (b).



Fonte: SILVA et al. (2010).

#### 4.1.6 Tipos de RNAs.

A combinação dos diversos tipos de arquitetura, topologia, treinamento e aprendizagem caracterizam diferentes tipos de RNAs, tais como: Perceptron, Adaline, Perceptron MultiCamadas, Funções de Base Radial, Recorrentes de Hopfield, Auto-organizáveis de Kohonen, *Counter Propagation*, entre outras.

Silva et al. (2010) demonstra que uma topologia de RNA Perceptron MultiCamadas, constituída de apenas uma camada neural intermediária com a função de ativação do tipo logística ou tangente hiperbólica, é capaz de mapear qualquer função contínua (pg. 133-134).

Devido a esta citação de Silva et al. (2010) e às diversas citações de autores, tais como: Norgaard et al. (2000), Son et al. (2005), Hasanzadehshooilli et al. (2012) e Azimzadegan et al. (2013), especial atenção será dada à RNA Perceptron MultiCamadas.



#### 4.1.6.1 Perceptron de Multicamadas (PMC)

As RNAs PMC, um tipo de arquitetura *feedforward*, são caracterizadas pela presença de pelo menos uma camada intermediária de neurônios, situada entre a camada de entrada e a de saída. São consideradas uma das arquiteturas mais versáteis quanto à aplicabilidade (SILVA et al., 2010).

As Figuras 2 e 3(b), anteriores, são exemplos de RNAs com camadas intermediárias.

Devido às características não lineares inerentes ao mapeamento entre camadas de RNAs *feedforward*, estas se constituem em ferramentas bastante apropriadas para a modelagem e o controle de sistemas (BRAGA et al., 2007).

Segundo Campos et al. (2004) os parâmetros que caracterizam uma rede multicamadas são:

- a) número de camadas,
- b) número de neurônios por camada,
- c) escolha das conexões,
- d) tipo de funções de ativação dos neurônios e
- e) peso de cada uma das conexões.

Conforme Livingstone (2009), a escolha da arquitetura da rede PMC envolve decidir o número de camadas intermediárias a usar, o número de neurônios nestas camadas e número de neurônios na camada de saída. A princípio, na prática, tem-se verificado que uma camada intermediária é o suficiente e o número de neurônios na camada de saída é igual ao número de variáveis dependentes em análise.

De acordo com o Teorema de Cybenko (1989), qualquer função contínua pode ser uniformemente aproximada por uma RNA, de apenas uma camada interna, e com uma função de ativação sigmoideal contínua não linear arbitrária.

Silva et al. (2010) esclarece que a especificação de topologia de rede PMC mais apropriada para mapear um problema específico é usualmente efetuada de forma empírica, pois devem ser considerados:

- a) o algoritmo de aprendizagem utilizado;
- b) a forma como as matrizes de peso foram iniciadas;
- c) a complexidade do problema a ser mapeado; e
- d) a qualidade do conjunto de treinamento disponível.

Outras regras utilizadas em especificação topológica, considerando um PMC constituído de apenas uma única camada escondida, estão relacionadas ao número de neurônios desta camada.

Um método muito citado em estudos é o de Kolmogorov (HECHT-NIELSEN, 1991) que determina o número de neurônios que deve ter a camada intermediária:

$$\mathbf{N} = 2.n + 1 \quad (7)$$

onde:

N : quantidade de neurônios da camada intermediária, e  
n: número de entradas da RNA.

Blum (1992) trata do número máximo de neurônios que deve conter a camada intermediária da RNA do tipo PMC, da seguinte forma:

$$\mathbf{N}_{\max} = \mathbf{P} / [10 . (n+s)] \quad (8)$$

onde:

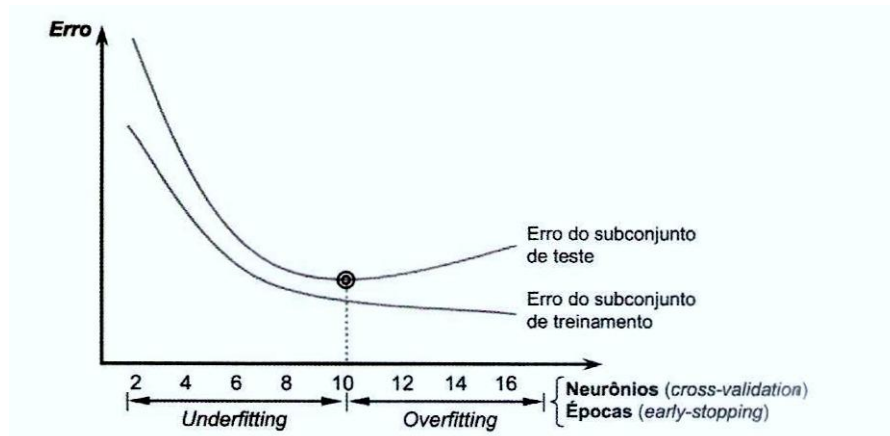
$N_{\max}$ : quantidade máxima de neurônios da camada intermediária,  
P: número de padrões apresentado ao treinamento/aprendizagem da RNA,  
n: número de entradas da RNA, e  
s: número de saídas da RNA.

O aumento indiscriminado de neurônios assim como o incremento de camadas intermediárias não asseguram a generalização apropriada da rede PMC. Tais ações podem provocar uma situação onde a rede acaba memorizando as saídas, apresentando um EQM bem baixo durante a fase de aprendizado, mas muito alto na fase de generalização. Esta condição de memorização excessiva é conhecida como *overfitting*. De forma contrária, frente à precisão requerida, uma rede PMC com

número muito reduzido de neurônios pode ser insuficiente para a extração e armazenamento das características do processo configurando assim uma situação de debilidade mental denominada *underfitting*. Neste caso, tanto o EQM na fase de aprendizado quanto o na fase de generalização serão bem significativos.

A Figura 5 mostra as situações de *underfitting* e *overfitting* em função do número de neurônios da camada intermediária em uma rede PMC. No exemplo, uma RNA com menos de 10 neurônios na camada intermediária apresenta valores de erro que iniciam altos (*underfitting*) e, à medida que é aumentado o número de neurônios, passa a diminuir possibilitando melhor aprendizado da RNA e capacidade de generalização de resultados. O aumento deste número de neurônios acima de 10 proporciona uma redução no EQM da fase de treinamento, pois a RNA passa a memorizar as saídas desejadas, mas não torna-se satisfatória na generalização quando apresenta EQM crescente à medida que aumenta o número de neurônios citado (*overfitting*). Uma generalização insatisfatória gera resultados de simulações na RNA não coerentes com os resultados reais do processo mapeado.

**Figura 5** - Situações envolvendo *underfitting* e *overfitting*.

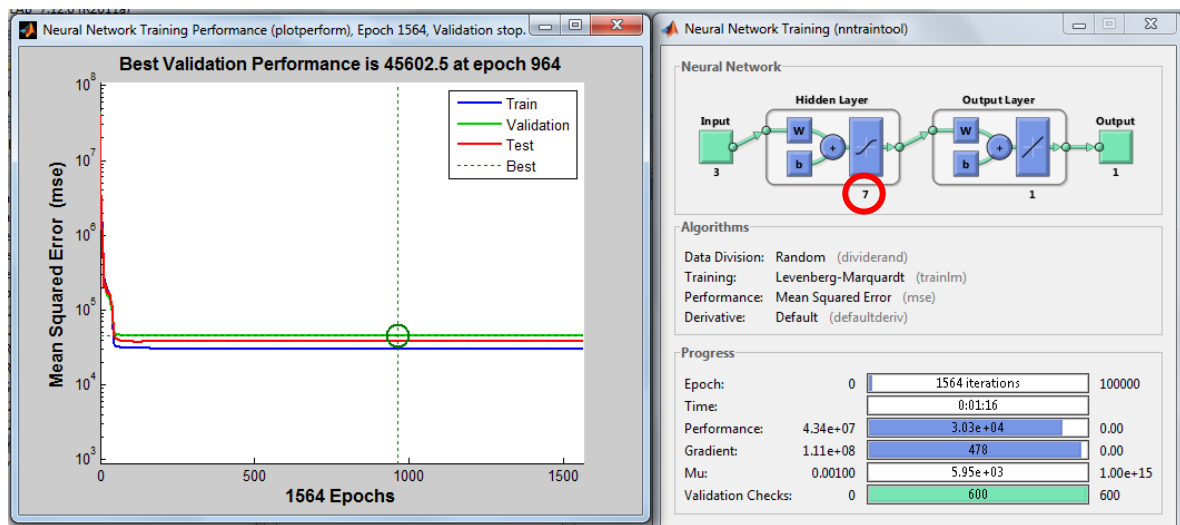


Fonte: SILVA et al. (2010).

O software MATLAB® apresenta, em sua janela de performance da RNA, os comportamentos dos 3 grupos da *cross-validation*: treinamento (azul), validação (verde) e teste (vermelho). A Figura 6 apresenta esta janela, para uma RNA de 7 neurônios identificado na janela ao lado desta, onde pode ser observado que os EQMs, tanto do treinamento quanto da validação, apresentam valores altos,

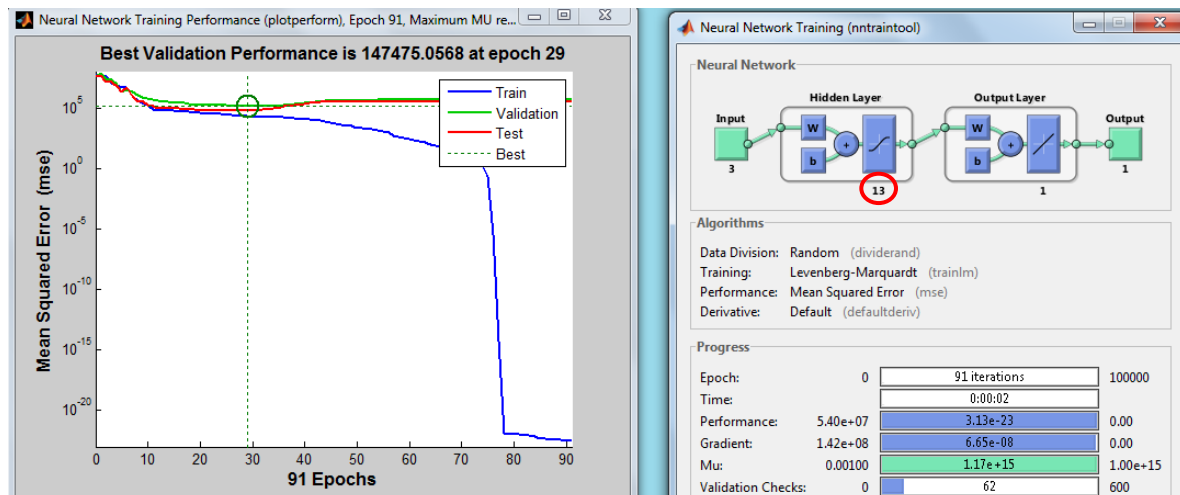
aproximadamente  $4,5 \cdot 10^4$ , caracterizando uma situação de *underfitting*, ou seja, baixo número de neurônios. Situação diferenciada pode ser vista na Figura 7 onde o EQM de treinamento é baixo ( $<10^{-20}$ ), indicando boa memorização, e o EQM de validação é alto ( $>1,4 \cdot 10^5$ ) que indica baixa capacidade de generalização. O fato é que a RNA tem muitos neurônios, neste caso 13, para a análise, caracterizando o *overfitting*: alto número de neurônios.

**Figura 6 -** Caracterização de Underfitting pelo software: a RNA tem poucos neurônios.



Fonte: MATLAB®

**Figura 7 -** Caracterização de Overfitting pelo software: a RNA tem muitos neurônios.



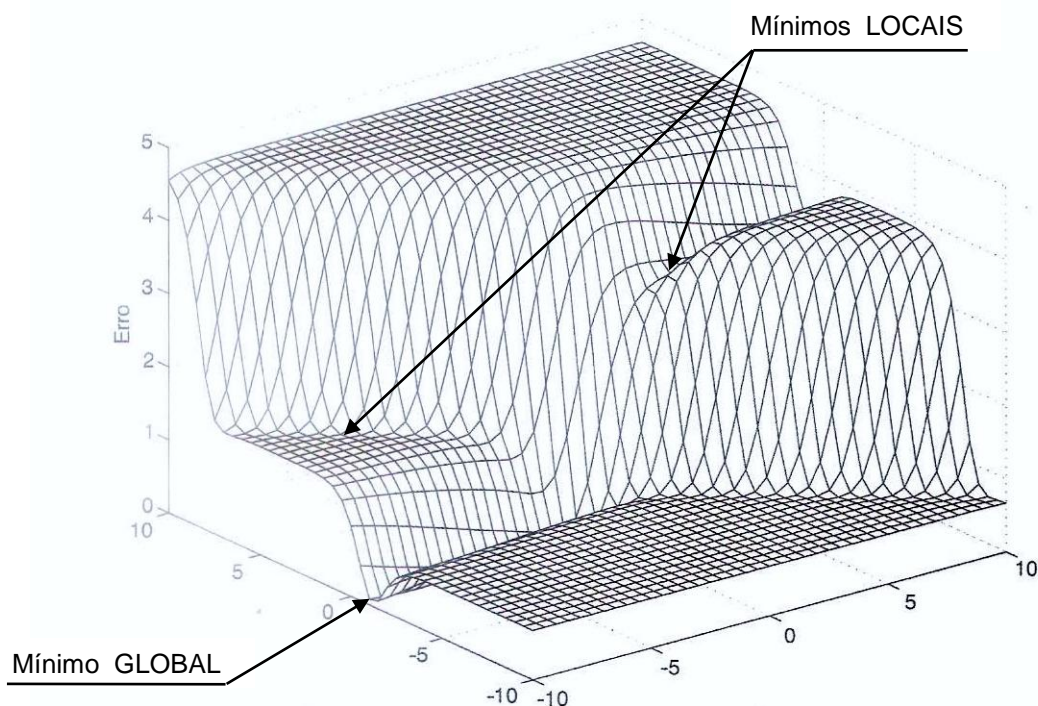
Fonte: MATLAB®

Braga et al. (2007) orienta a utilização de funções não lineares nas camadas intermediárias para resolver problemas de maior ordem pois a utilização de funções lineares em múltiplas camadas resultaria em uma rede de camada única devido ao

fato de transformações lineares sucessivas caracterizarem uma única transformação linear. Ainda ressalta que funções lineares na saída são utilizadas em problemas de aproximação de funções.

Para o caso de RNAs que utilizam funções de ativação não lineares a superfície de erros poderá ter uma forma que apresenta vários mínimos locais além do mínimo global, como mostrado na Figura 8, onde podem ser verificados dois pontos de erros em mínimos locais, que não caracterizam a melhor performance da RNA, e um ponto de mínimo global, melhor situação de operação da RNA.

**Figura 8** - Superfície de erro para neurônios com funções de ativação não lineares apresentando Mínimos Locais e Mínimo Global.



Fonte: BRAGA et al. (2007).

O objetivo do aprendizado por correção de erros é, a partir de um ponto qualquer da superfície de erro, mover-se na direção do erro de mínimo global; porém existe uma tendência de convergência para um mínimo local devido à matriz de pesos inicial do treinamento, pois a grande maioria dos algoritmos de aprendizagem é baseada em métodos de gradiente descendente: o vetor de pesos é ajustado em direção contrária ao vetor gradiente. Maiores detalhes podem ser encontrados em BRAGA et

al., 2007, página 41. Para evitar esse tipo de convergência para um mínimo local treina-se cada topologia candidata mais de uma vez, com diferentes matrizes de pesos iniciais, geradas aleatoriamente.

Silva et al. (2010) explica que o ajuste dos pesos e dos limiares de cada um dos neurônios da rede PMC é efetuado utilizando-se o processo de treinamento supervisionado e o algoritmo *backpropagation*. A operação deste algoritmo divide-se em duas fases:

- a) a primeira é denominada de "propagação adiante" (*forward*), quando obtém-se a resposta da rede por meio das entradas fornecidas e dos pesos e limiares aleatórios iniciais. Estas respostas são comparadas com os valores desejados fornecidos e obtidos seus respectivos valores de erros;
- b) a partir desses erros aplica-se a segunda fase denominada "propagação reversa" (*backward*) quando são executados todos os ajustes de pesos e limiares dos neurônios. Estes ajustes começam pela camada de saída, minimizando os valores dos erros. Posteriormente são feitos os ajustes das camadas intermediárias, uma a uma no sentido da saída para a entrada, a partir de estimativas dos erros das saídas produzidos pelos neurônios da camada imediatamente posterior, que já foram previamente ajustados.

Segundo Braga (2007) o algoritmo *backpropagation* padrão requer que todos os padrões de treinamento sejam apresentados centenas ou até milhares de vezes tornando-o muito lento limitando sua utilização prática.

O método de Levenberg-Marquardt, baseado no método dos mínimos quadrados para modelos não lineares, com uma taxa de aprendizado variável " $\mu$ ", que ajusta a taxa de convergência deste algoritmo, é uma das variações do método *backpropagation* com o objetivo de tornar o processo de convergência mais eficiente. Este é um método de gradiente descendente (SILVA et al., 2010, pg. 117-120).

Conforme o manual do software MATLAB®, o algoritmo de Levenberg-Marquardt apresenta um fator " $\mu$ " que é um número escalar acrescido ao método de Newton, para fazer com que a velocidade da redução dos EQMs seja mais rápida, pois o

método de Newton apresenta bom desempenho somente na região de erros mínimos. Este valor adaptativo de " $\mu$ " sofre variações durante a operação do algoritmo, podendo ser incrementado ou decrementado por fatores previamente estabelecidos, *mu\_inc* ou *mu\_dec*, no software MATLAB®, respectivamente.

Os parâmetros de treinamento do algoritmo de Levenberg-Marquardt, identificados no software MATLAB®, são:

- a) o valor inicial de " $\mu$ " (*mu*);
- b) o fator de decréscimo de " $\mu$ " (*mu\_dec*);
- c) o fator de acréscimo de " $\mu$ " (*mu\_inc*);
- d) o valor máximo permitido para o parâmetro " $\mu$ " (*mu\_max*);
- e) o valor de performance do gradiente, derivada da curva do EQM, a ser atingido (*min\_grad*);
- f) o número máximo de épocas a ser repetido o valor do desempenho da RNA, pela validação do treinamento (*max\_fail*);
- g) o valor do EQM (performance da RNA) solicitado a ser atingido no treinamento da RNA (*goal*);
- h) o tempo máximo para efetivar o treinamento (*time*);
- i) o número máximo de épocas a serem apresentadas no treinamento (*epochs*).

Segundo o manual do software do MATLAB®, o treinamento é parado quando:

- a) o número máximo de épocas solicitado for atingido (*early stopping*);
- b) o tempo máximo para efetivar o treinamento for excedido (*time*);
- c) o desempenho da RNA atingir a meta, que é o valor de EQM (*goal*);
- d) o gradiente de desempenho cair abaixo do valor estabelecido de *min\_grad*;
- e) *mu* exceder *mu\_max*;
- f) o número de épocas de desempenho de validação constante aumentar mais do que o valor estabelecido para *max\_fail*.

A função de treinamento de Levenberg-Marquardt muitas vezes é o mais rápido algoritmo de *backpropagation* na caixa de ferramentas do softwares e é altamente recomendado como uma primeira escolha de algoritmo supervisionado (MATLAB®).

Mukherjee et al. (2012), em estudo que compara três métodos de treinamento, afirma que o Levenberg-Marquardt converge mais rápido e proporciona um modelo preditivo não linear mais preciso.

Gorni et al. (2009) em seu estudo de previsão de propriedades mecânicas em bobinas laminadas a quente, usou RNA *feedforward* com uma camada intermediária com função de ativação Tangente Hiperbólica, e camada de saída com função de ativação Linear. Em testes preliminares foi concluído que consegue-se maior precisão preditiva ao se incluir várias variáveis de saída, ao invés de se criar uma RNA específica para cada uma, pois a presença simultânea dessas "[ ] intensifica as correlações existentes entre elas e as variáveis independentes". O resultado geral de modelagem obtido mostrou que usando-se RNAs obteve-se melhor precisão de cálculo do que equações obtidas por regressão linear múltipla, o que comprovou a maior capacidade das RNAs em lidar com relações não lineares entre as variáveis envolvidas.

Santos et al. (2006), em estudo da modelagem de temperatura no processo de fresamento, verificou o bom desempenho da RNA por ter proporcionado margem de erro muito baixa para valores estimados de temperatura em um processo de usinagem altamente não linear como o fresamento. Foi utilizada uma rede PMC de 5 neurônios na camada de entrada, 7 neurônios na camada intermediária e 1 na camada de saída com funções de ativação Tangente Hiperbólica e Logística nas camadas intermediária e de saída, respectivamente. Foram apresentados 110 padrões (conjunto de entrada-saída) para um treinamento supervisionado. O índice de correlação obtido pelas saídas da RNA foi de 0,915, satisfatório segundo o autor.

Bagheripoor & Bisadi (2013) desenvolveram um modelo (identificação) via RNA para predição de força e torque na laminação de alumínio. Utilizaram uma arquitetura tipo *feedforward*, treinamento via *Levenberg-Marquardt* e algoritmo *backpropagation* "[ ] devido às suas aplicações de sucesso em vários problemas práticos e suas simplicidades nas instalações ". A topologia da RNA foi obtida de forma empírica com variação do número de camadas intermediárias e número de neurônios. As variáveis de entrada utilizadas foram temperatura inicial no passe, porcentagem de redução na espessura, velocidade de laminação e coeficiente de atrito na área de



contato. As previsões da RNA apresentaram baixos erros quando comparados aos valores práticos.

#### 4.1.7 Aproximação de Funções.

Considera-se um mapeamento de entrada-saída descrito pela relação  $d = f(x)$ , onde o vetor "x" é a entrada e o vetor "d" é a saída, e supõe-se que a função  $f(.)$  seja desconhecida mas um conjunto de padrões (entradas, saídas), seja fornecido.

Pretende-se então projetar uma RNA que descreva uma função  $F(.)$ , mapeada pelos padrões fornecidos, que seja uma aproximação da função desconhecida  $f(.)$ , de forma que

$$|F(x)-f(x)| < \xi, \quad \text{para todo } x$$

onde " $\xi$ " é um número positivo pequeno que representa o erro aproximativo.

Este problema de aproximação é típico de aprendizagem supervisionada com " $x_i$ " desempenhando o papel de vetor de entrada e " $d_i$ " o de resposta desejada. Sendo a saída " $y_i$ " da RNA produzida em resposta ao vetor de entradas " $x_i$ ", o vetor de sinal de erro " $\xi_i$ " será obtido pela diferença entre " $d_i$ " (associada à entrada " $x_i$ ") e a saída da RNA " $y_i$ ". Esse sinal de erro é usado para ajustar os parâmetros livres da RNA (pesos e limiares) de forma a minimizar a diferença quadrática entre as saídas do sistema desconhecido e da RNA (HAYKIN, 2001).

O desempenho (performance) da RNA é fornecido por meio do EQM de acordo com a média do quadrado dos erros (MATLAB®), conforme expressão (MUKHERJEE, 2012):

$$EQM = \frac{1}{P} \cdot \sum_{j=1}^P (y_j - d_j)^2 \quad (9)$$

onde:

P: número de padrões,

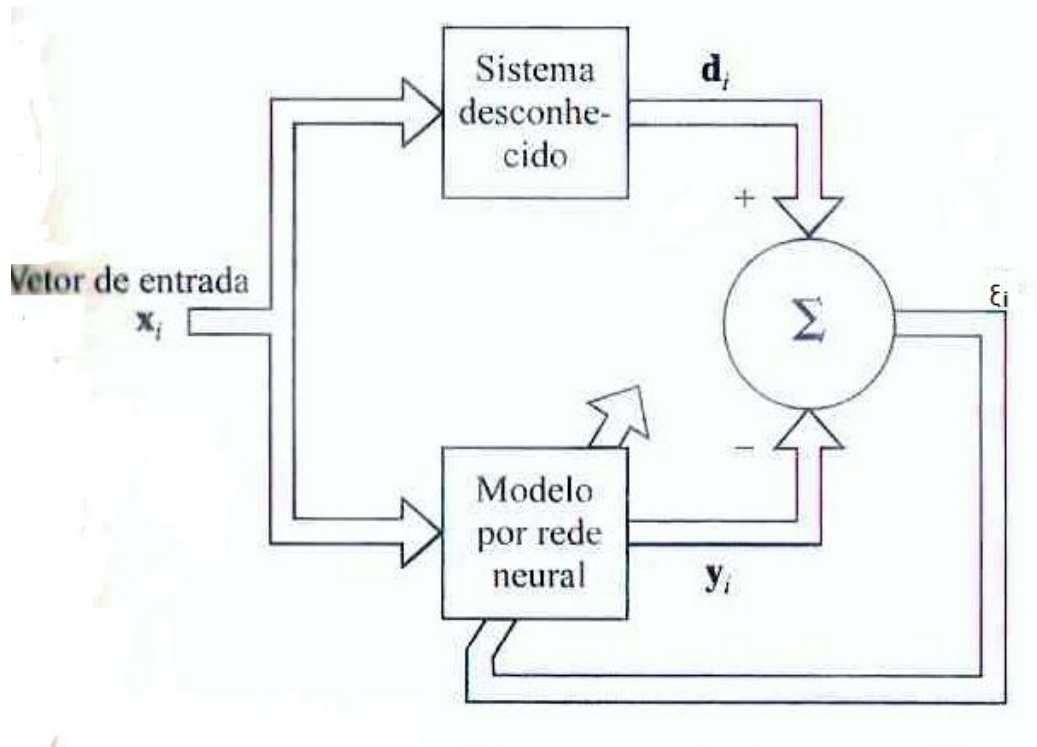
$y_j$ : valor da saída "j" da RNA treinada;

$d_j$ : valor original da saída "j".

Erros pequenos indicam uma boa arquitetura, isto é, uma topologia de RNA com chances de bom treinamento e boa predição de resultados (PIULEAC et al., 2010).

A Figura 9 mostra o diagrama de blocos para essa metodologia: o treinamento da RNA visa mapear o sistema desconhecido por meio de seus padrões de entradas "x" e saídas "d". As respostas "y" da RNA em treinamento serão comparadas com as saídas "d", resultando os erros " $\xi$ " a serem reduzidos.

**Figura 9** - Diagrama em blocos da identificação de sistemas.



Fonte: HAYKIN (2001).

## 4.2 SENSIBILIDADE DE VARIÁVEL INDEPENDENTE

Conhecer as variáveis independentes mais influentes e as de menor efeito sobre determinado(s) parâmetro(s) de um processo, impacta consideravelmente nas diretrizes de investimentos e projetos; bem como na orientação operacional, proporcionando redução de custos e melhoria da eficiência do processo.

Segundo Gorni et al. (2009) a análise da sensibilidade é feita para cada variável independente do sistema. Ela estima a participação da variável independente em alterações nas variáveis dependentes.

A proposta da técnica de análise da sensibilidade das variáveis independentes de um processo será efetivada por meio da obtenção de duas características de sensibilização:

- a) Sensibilidade Média (SENS) e
- b) Sensibilidade Média Normalizada (SENSN).

A verificação da dispersão da variável dependente com cada uma das variáveis independentes pode fornecer comportamentos que caracterizam conhecimentos preliminares do processo sem que haja necessidade de domínio dos conceitos fenomenológicos do mesmo. Os valores de sensibilidade apresentam sinais positivos ou negativos dependendo desta relação de variação:

- a) sinais negativos de sensibilidade ( $SENS < 0$ ) indicam que acréscimos da variável independente acarretam decréscimos na variável dependente; e
- b) sinais positivos ( $SENS > 0$ ) indicam variações diretamente proporcionais entre estas variáveis, dependente e independente.

#### 4.2.1 Sensibilidade Média.

A técnica consiste em fazer com que, uma a cada vez, a variável independente " $x_i$ " escolhida seja acrescida em 5% nos seus valores, que substituem os valores anteriores. Valores acrescidos que ultrapassem o valor máximo da faixa da entrada deverão ser desprezados pois, conforme citado na seção 4.1, as entradas devem ser oriundas do mesmo tipo de população alvo do problema.

Este novo conjunto de dados de entrada é então simulado. É obtida a diferença entre os novos valores de saída " $y$ " obtidos na simulação e os valores de saída desejados " $d$ ". Finalmente determina-se a média dessas diferenças obtendo-se a Sensibilidade Média ( $SENS_i$ ) da variável independente " $x_i$ " em questão, conforme expressão:

$$\mathbf{SENSi} = \frac{1}{P} \sum_{j=1}^P (y_j - d_j) \quad (10)$$

onde:

P: número de padrões;

y: valor da saída simulada com a alteração da respectiva entrada "x<sub>i</sub>";

d: valor original da saída do sistema, sem alteração nos padrões.

Observa-se, pela expressão (10), que a relação de proporcionalidade ou não da influência da variável independente sobre a dependente fica determinado pelo sinal aritmético do resultado da subtração.

Este processo é feito para cada variável independente do sistema, caracterizando uma derivação parcial da saída do processo em função de cada variável independente. Quanto maior é a Sensibilidade Média, maior é o efeito da variável independente considerada. A unidade de medida desta SENS é a da própria grandeza de saída em análise do processo.

#### 4.2.2 Sensibilidade Média Normalizada.

Uma outra forma de apresentar esta influência é pela Sensibilidade Média Normalizada, que pode ser obtida pela expressão:

$$\mathbf{SENSNi} = \left( \mathbf{SENSi} / \sum_{i=1}^e |\mathbf{SENSi}| \right) \cdot 100 \quad (11)$$

onde:

e: número de variáveis independentes.

Esta SENSN é expressa em %, caracterizando a influência da variável independente junto às outras influências das variáveis independentes em análise e também apresenta sinal aritmético.

#### 4.3. LAMINAÇÃO A QUENTE DE AÇO C-Mn.

Dentre as inúmeras possibilidades de aplicação das RNAs, uma das mais interessantes baseia-se no estudo da determinação das propriedades mecânicas de produtos siderúrgicos.

Son et al. (2005) comenta que as indústrias do aço, na busca de produtividade e qualidade, acabam por automatizar suas plantas utilizando técnicas de Inteligência Artificial (IA). Modelos matemáticos dos processos de laminação para predição apresentam grande complexidade e exigem grande demanda de tempo dificultando suas aplicações. As RNAs surgem, então, como uma ferramenta da IA para facilitar as predições necessárias, especialmente a força de laminação.

O comportamento do processo de laminação a quente é caracterizado pela interação de muitos componentes, um grande número de influências e não linearidades. O processo e seus sub processos são de natureza repetitiva: processos similares são executados a cada passe (SCHLANG et al., 2001).

#### 4.3.1 Breve descrição do processo de Laminação de Tiras a Quente (LTQ).

O processo industrial que trata este estudo inicia com placas de aço C-Mn que são reaquecidas num forno a aproximadamente 1200°C. Após desenfundadas, sofrem a remoção de óxidos superficiais pela descarepação primária, por meio de jatos de água de alta pressão, quando são transferidas até o laminador de desbaste, esquema apresentado na Figura - 10 (a). Neste, as placas recebem redução de espessura chegando na medida final de 20 mm a 40 mm. A largura das placas é estabelecida por um laminador de bordas verticais.

Estas placas, agora chamadas de esboço, são bobinadas num *coil box*, onde a temperatura do esboço fica homogênea. Após o desbobinamento no *coil box* nova descarepação, secundária, ocorre para remoção de óxidos.

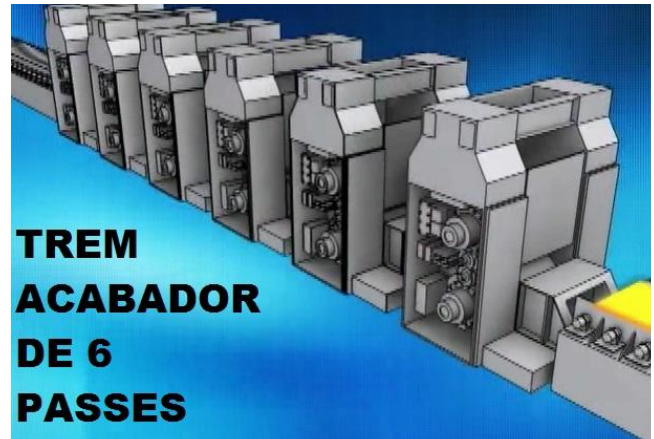
Este esboço é, então, transformado em tira no trem acabador, constituído de 6 cadeias (passes) de laminação, esquema apresentado na Figura - 10 (b), que promovem a redução final para espessuras entre 16 mm e 1,2 mm, em um processo

automatizado. Em seguida, as tiras são resfriadas numa mesa de resfriamento laminar, passando, a seguir, para as bobinadeiras.

**Figura 10-** Equipamentos constituintes do Processo de Laminação de Tiras a Quente: Laminador de Desbaste (a), Trem acabador de 6 passes(b).



(a)



(b)

FONTE: vídeo do LTQ da ArcelorMittal

#### 4.3.2 Modelagem do Processo.

As exigências do mercado estão cada vez mais rigorosas para produtos de alta qualidade em laminação a quente. A exigência mais solicitada é a da precisão na espessura de um produto no processo de laminação. Boa precisão na espessura está altamente relacionada com boa precisão na Força de Laminação. Assim, busca-se uma predição cada vez mais precisa desta força. A maneira convencional de prever a força de laminação é através de modelo matemático (LEE e CHOI, 2004).

O estudo acerca da força de laminação é justificado pela relação desta com a Tensão de Escoamento Média (TEM). Vários modelos levam em consideração as propriedades metalúrgicas do aço obtidas por meio da TEM, considerando os efeitos de recristalização estática e dinâmica.

Machado (2005), em sua Tese, por meio da análise fenomenológica do processo de laminação a quente de diferentes tipos de aço, apresenta modelos (caixa-branca) que geraram resultados compatíveis e satisfatórios na obtenção da Força de

Laminação em bobinas de diferentes tipos de aço. Está descrito, a seguir, o modelo relativo ao aço C-Mn que será utilizado neste estudo.

#### 4.3.2.1 Equação de Sims

Para o caso de aço C-Mn, o modelo citado apresenta a relação entre a força de separação do cilindro ou Força de Laminação (F) e a TEM expressa por meio da equação de Sims, apresentada a seguir:

$$F = \left( TEM_m \cdot W \cdot \sqrt{R' - \Delta h} \right) \cdot \frac{Q}{9800} \quad [\text{ton}] \quad (12)$$

onde:

TEM<sub>m</sub>: Tensão de Escoamento Média modificada [MPa],

W: largura da peça laminada [mm],

R': raio do cilindro com deformação durante a laminação [mm],

Δh: diferença entre as espessuras de entrada e saída no passe [mm],

Q: fator geométrico da equação de Sims.

A expressão da TEM obtida por Misaka e Yoshimoto, expressão (13), considera a porcentagem de Carbono no aço C-Mn e recristalização somente estática, conforme a seguir:

$$TEM = 1,15,9,8 \cdot e^{\left\{0,126 - 1,75 \cdot [C] + 0,594 \cdot [C]^2 + \frac{2851 + 2968 \cdot [C] - 1120 \cdot [C]^2}{T}\right\}} \cdot \epsilon_a^{0,21} \cdot \dot{\epsilon}^{0,13} \quad [\text{MPa}] \quad (13)$$

onde:

[C]: porcentagem de Carbono no aço,

ε<sub>a</sub>: deformação acumulada,

đ : taxa de deformação [s<sup>-1</sup>],

T: temperatura do material no passe de laminação [K]

Correção foi feita a esta expressão para que considerasse a porcentagem de Manganês no aço C-Mn, gerando a expressão (14) de TEM corrigida (TEM<sub>c</sub>):

$$\mathbf{TEMc = TEM \cdot (0,768 + 0,137 \cdot [Mn])} \quad \mathbf{[MPa]} \quad (14)$$

onde:

[Mn]: porcentagem de Manganês no aço C-Mn.

Para que fosse também considerada a fração de recristalização dinâmica do material ( $X_{din}$ ), que ocorre quando a deformação acumulada no material é maior que determinada deformação crítica, esta expressão sofreu uma nova alteração (15), gerando a TEM modificada ( $TEM_m$ ), que é utilizada para o cálculo da Força "F" na expressão (12):

$$\mathbf{TEMm = TEMc \cdot (1 - X_{din}) + 1,14 \cdot \sigma_{ss} \cdot X_{din}} \quad \mathbf{[MPa]} \quad (15)$$

onde:

$\sigma_{ss}$ : Tensão de Escoamento no regime estacionário [MPa].

Outros valores necessários para o cálculo de "F" na expressão (12) são:

a) o raio do cilindro com deformação durante a laminação:

$$\mathbf{R' = R \cdot \left[ 1 + \left( \frac{C \cdot F_{med} \cdot 1000}{W \cdot \Delta h} \right) \right]} \quad \mathbf{[mm]} \quad (16)$$

onde:

R: raio original do cilindro de laminação [mm],

$F_{med}$ : força de laminação medida no processo [ton],

$C = [16 \cdot (1 - \nu^2)] / \pi \cdot E$ , onde:  $\nu$ : razão de Poisson do material do cilindro de trabalho (= 0,3) e

E: módulo de Young para o cilindro de trabalho (= 21092,1 Kgf/mm<sup>2</sup>).

Para o material do cilindro utilizado,  $C = 2,2 \cdot 10^{-4}$  (mm<sup>2</sup>/Kgf).

b) a diferença entre as espessuras de entrada e saída no passe:



$$\Delta h = h_1 - h_2 \quad (17)$$

onde:

h1: espessura de entrada do material na laminação [mm],

h2: espessura de saída do material na laminação [mm].

c) o fator geométrico da equação de Sims:

$$Q = 0,5 \cdot \sqrt{\frac{1-r}{r}} \cdot \left\{ \pi \cdot \tan^{-1} \sqrt{\frac{r}{1-r}} - \sqrt{\frac{R'}{h_2}} \cdot \ln \left[ \left( \frac{h_n}{h_2} \right)^2 \cdot (1-r) \right] \right\} - \frac{\pi}{4} \quad (18)$$

onde:

r: grau de redução do material laminado,

h<sub>n</sub>: espessura do material laminado no ângulo neutro.

d) grau de redução do material laminado:

$$r = \frac{h_1 - h_2}{h_1} \quad (19)$$

e) espessura do material laminado, no ângulo neutro:

$$h_n = 2 \cdot R' \cdot (1 - \cos \varphi_n) + h_2 \quad (20)$$

onde:

$\varphi_n$ : ângulo neutro.

f) ângulo neutro:

$$\varphi_n = \sqrt{\frac{h_2}{R'}} \cdot \tan \left[ \frac{\pi}{8} \cdot \sqrt{\frac{h_2}{R'}} \cdot \ln(1-r) + 0,5 \cdot \tan^{-1} \sqrt{\frac{r}{1-r}} \right] \quad (21)$$

#### 4.3.2.2 Tensão de Escoamento Média Modificada.

Como a TEMm é utilizada pra o cálculo da força "F", torna-se, então, necessário o conhecimento dos fatores envolvidos na sua expressão (15):

$$\mathbf{TEMm} = \mathbf{TEMc} \cdot (\mathbf{1} - \mathbf{Xdin}) + \mathbf{1,14} \cdot \mathbf{\sigma_{ss}} \cdot \mathbf{Xdin}.$$

Relacionados a esta TEMm estão:

- a) a fração de recristalização dinâmica do material:

$$\mathbf{Xdin} = \mathbf{1} - \mathbf{e}^{[-0,693 \cdot (\frac{\mathbf{t_{din}}}{\mathbf{t_{0,5}}})^{1,1}]} \quad (22)$$

onde:

$t_{din}$ : tempo de recristalização dinâmica,

$t_{0,5}$ : tempo para alcançar 50% de recristalização.

- b) a tensão de escoamento no regime estacionário:

$$\mathbf{\sigma_{ss}} = \mathbf{7,2} \cdot (\mathbf{Z})^{0,09} \quad (23)$$

onde:

Z: parâmetro Zenner-Hollomon.

- c) a TEMc que é obtida a partir da expressão (14), que, por sua vez, está em função da TEM, expressão (13).

- d) o tempo de recristalização dinâmica:

$$\mathbf{t_{din}} = \frac{\mathbf{\epsilon_a} - \mathbf{\epsilon_c}}{\dot{\mathbf{\epsilon}}} \quad (24)$$

onde:

$\epsilon_c$ : deformação crítica.

- e) o tempo para alcançar 50% de recristalização:

- estática:

$$t_{0,5} = 2,3 \cdot 10^{-15} \cdot \varepsilon_a^{-2,5} \cdot d_{ant}^2 \cdot e^{\left(\frac{230000}{8,314 \cdot T}\right)} \quad (25)$$

onde:

$d_{ant}$ : diâmetro do grão austenítico anterior à laminação [ $\mu\text{m}$ ].

- dinâmica:

$$t_{0,5} = 0,4 \cdot Z^{-0,8} \cdot e^{\left(\frac{240000}{8,314 \cdot T}\right)} \quad (26)$$

f) o parâmetro Zenner-Hollomon:

$$Z = \dot{\varepsilon} \cdot e^{\left(\frac{375000}{8,314 \cdot T}\right)} \quad (27)$$

g) a deformação acumulada:

$$\varepsilon_a = (\varepsilon_r + \varepsilon) \cdot (1 - X_{ant}) \varepsilon_{a_{ant}} \quad (28)$$

onde:

$\varepsilon_r$ : deformação redundante,

$\varepsilon$ : deformação efetiva,

$X_{ant}$ : fração anterior de recristalização do material,

$\varepsilon_{a_{ant}}$ : deformação acumulada anterior.

h) a deformação crítica:

$$\varepsilon_c = 5,6 \cdot 10^{-4} \cdot d_{ant}^{0,3} \cdot \left[ \dot{\varepsilon} \cdot e^{\left(\frac{300000}{8,314 \cdot T}\right)} \right]^{0,17} \quad (29)$$

i) a taxa de deformação:

$$\dot{\varepsilon} = \frac{\varepsilon}{t} \quad (30)$$

onde:

t: tempo de aplicação da deformação [s].

j) a deformação redundante:

$$\varepsilon_r = 0,5 \cdot \operatorname{tg}\left(\frac{\alpha}{2}\right) \quad (31)$$

onde:

$\alpha$ : ângulo de contato do material com o cilindro.

k) a deformação efetiva:

$$\varepsilon = \frac{2}{\sqrt{3}} \cdot \ln\left(\frac{h_1}{h_2}\right) \quad (32)$$

l) a fração de recristalização do material.

- para recristalização estática:

$$X = 1 - e^{-0,693 \cdot \left(\frac{t_{ip}}{t_{0,5}}\right)} \quad (33)$$

onde:

$t_{ip}$ : tempo entre os passes [s].

- para recristalização dinâmica:

$$X = 1 - e^{-0,693 \cdot \left(\frac{t_{ip}}{t_{0,5}}\right)^{1,1}} \quad (34)$$

m) o tempo de aplicação da deformação:

$$t = \frac{\alpha \cdot 60}{2 \cdot \pi \cdot V} \quad (35)$$

onde:

V: velocidade do cilindro [rpm].

n) o ângulo de contato do material com o cilindro:

$$\alpha = \cos^{-1}\left(1 - \frac{\Delta h}{2 \cdot R'}\right) \quad (36)$$

o) o tempo entre os passes:

$$t_{ip} = \frac{d \cdot 60}{V_e} \quad (37)$$

onde:

d: distância entre passes [mm],

V<sub>e</sub>: velocidade de saída da tira [mm/min].

p) a velocidade de saída da tira:

$$V_e = V_r \cdot (1 + \lambda) \quad (38)$$

onde:

V<sub>r</sub>: velocidade tangencial do cilindro,

λ: fator de deslizamento.

q) a velocidade tangencial do cilindro:

$$V_r = 2 \cdot \pi \cdot R' \cdot V \quad (39)$$

r) o fator de deslizamento:

$$\lambda = \left( \frac{2 \cdot R'}{h^2} \cdot \cos \varphi n - 1 \right) \cdot (1 - \cos \varphi n) \quad (40)$$

s) diâmetro do grão austenítico (d):

- para  $X < 0,95$  (recristalização incompleta):

$$d = d_{\text{rec}} \cdot X^{\frac{4}{3}} + d_{\text{ant}} \cdot (1 - X)^2 \quad (41)$$

onde:

$d_{\text{rec}}$ : tamanho do grão austenítico completamente recristalizado após deformação.

- para  $X=1$  (recristalização completa):

- estática:

- $t_{ip} > 1$ :

$$d = [d_{\text{rec}}^7 + 1,5 \cdot 10^{27} \cdot (t_{ip} - 4,32 \cdot t_{0,5}) \cdot e^{\left(\frac{-400000}{8,314 \cdot T}\right)}]^{1/7} \quad (42)$$

- $t_{ip} < 1$ :

$$d = [d_{\text{rec}}^2 + 4 \cdot 10^7 \cdot (t_{ip} - 4,32 \cdot t_{0,5}) \cdot e^{\left(\frac{-113000}{8,314 \cdot T}\right)}]^{0,5} \quad (43)$$

- dinâmica:

- $t_{ip} > 1$ :

$$d = [d_{\text{rec}}^7 + 8,2 \cdot 10^{25} \cdot (t_{ip} - 2,65 \cdot t_{0,5}) \cdot e^{\left(\frac{-400000}{8,314 \cdot T}\right)}]^{1/7} \quad (44)$$

- $t_{ip} < 1$ :

$$d = [d_{\text{rec}}^2 + 1,2 \cdot 10^7 \cdot (t_{ip} - 2,65 \cdot t_{0,5}) \cdot e^{\left(\frac{-113000}{8,314 \cdot T}\right)}]^{0,5} \quad (45)$$

t) o tamanho do grão austenítico completamente recristalizado após deformação:

- para recristalização estática:

$$d_{\text{rec}} = 1,1 \cdot d_{\text{ant}}^{0,67} \cdot \varepsilon_a^{-0,67} \quad (46)$$

- para recristalização dinâmica:

$$d_{\text{rec}} = 1370 \cdot \varepsilon_a^{-0,13} \cdot e^{\left(\frac{-45000}{8,314 \cdot T}\right)} \quad (47)$$

Desta forma, Machado (2005), prediz, por meio de modelagem, o comportamento da Força de Laminação a partir de parâmetros e características do sistema; ficando bem estabelecida esta força como sendo a saída de um processo onde verifica-se várias variáveis de entrada. Informações mais detalhadas são apresentadas na Tese citada.

#### 4.3.2.3. Dados necessários do processo para modelagem caixa-branca.

Para a operação do modelo caixa-branca selecionado pode-se verificar que torna-se necessário o conhecimento de alguns valores do sistema de LTQ de aço C-Mn em estudo. São eles:

- Composição química do aço:
  - [C]: porcentagem de Carbono no aço,
  - [Mn]: porcentagem de Manganês no aço.
- Características do produto:
  - W: largura da peça laminada [mm],
  - h: espessuras da peça laminada, tanto na sua entrada no processo, quanto em cada passe de laminação [mm],
  - T: temperatura do material laminado em cada passe de laminação [K].
- Características do equipamento:
  - R: raio original dos cilindros de laminação [mm]
  - v: velocidades dos cilindros [rpm],
  - d: distância entre passes [mm].

d) Característica do processo:

$F_{med}$ : força de laminação medida em cada passe do processo [ton].

Este modelo busca, a partir da Equação de Sims, prever a Força de Laminação, o que a caracteriza como a variável de saída (variável dependente) do processo de laminação. Pode-se concluir que as outras variáveis representam as entradas do processo, variáveis independentes.



## 5 METODOLOGIA.

Neste estudo particular a Metodologia será dividida em 3 fases. São elas:

- a) seleção da melhor topologia de RNA para mapeamento de funções;
- b) validação da técnica de sensibilização de variáveis independentes utilizando RNA;
- c) obtenção da significância das variáveis físicas independentes do processo de laminação a quente de aço C-Mn.

### 5.1 RNA MAPEANDO FUNÇÕES SIMULADORAS

A proposta é verificar se uma RNA consegue, realmente, mapear uma função utilizando, como recurso estratégico, expressões aritméticas e dados sintéticos, como simuladores de processo.

A seleção da melhor topologia ao atendimento no mapeamento de funções será obtida por meio de diversos ensaios em RNAs com variações em sua topologia, conforme a seguir:

- a) arquitetura: PMC (Perceptron MultiCamadas);
- b) número de camadas: 3 (entrada, intermediária e saída);
- c) número de entradas: 3 ( $n=3$ );
- d) número de saídas: 1 ( $s=1$ );
- e) padrões: a partir de dados sintéticos;
- f) funções simuladoras: 4 funções aritméticas de configurações diversas;
- g) número de neurônios da camada intermediária: 3 e 7;
- h) funções de ativação em 3 combinações: linear (LIN) na camada intermediária e linear na camada de saída, caracterizada por LIN/LIN; logística (LOG) na camada intermediária e linear na camada de saída, caracterizada por LOG/LIN; tangente hiperbólica (TANH) na camada intermediária e linear na camada de saída, caracterizada por TANH/LIN;
- i) número de treinamentos por RNA: 5;
- j) funções utilizadas do software:
  - função de treinamento: *Levenberg-Marquardt (LM)*;
  - função de aprendizagem: *gradiente descendente (GDM)*;

- função de performance: *erro quadrático médio (EQM)*;

Conforme orientação adquirida na literatura do referencial teórico, será utilizada a arquitetura de rede PMC (*feedforward*) com uma camada escondida.

As características do sistema de modelagem que devem ser atendidas pela RNA são:

- a) a análise será estática de um processo com padrões de entrada-saída, ou seja, haverá treinamento supervisionado para determinação da modelagem e serão fornecidos os padrões com as saídas "d", calculadas para cada expressão aritmética simuladora;
- b) o treinamento será off-line em malha aberta implicando à não necessidade de realimentação na rede, nem da dinâmica do processo;
- c) o trabalho será feito a partir do software MATLAB®, versão 7.12.0.635 (R2011a). Este software foi escolhido devido à sua disponibilidade e larga escala de utilização em diversas áreas de pesquisa e operação.

Serão utilizadas quatro expressões aritméticas, representando modelos lineares (F1) e não lineares de diferentes graus de complexidade (F2), (F3) e (F4), para verificar a capacidade de resposta das RNAs para níveis diferentes de dificuldades no modelo. Os padrões serão no formato  $(x_1, x_2, x_3, d)$ , ou seja, três entradas e uma saída desejada, utilizados nas funções e caracterizados como dados sintéticos (GOMES, 2008).

As funções utilizadas para o treinamento da redes serão:

$$F1: \quad d = 15.x_1 + 2.x_2 + 10.x_3$$

$$F2: \quad d = x_1^2 + 3.x_2 + 10.\sqrt{x_3}$$

$$F3: \quad d = x_1^3 + \sqrt{x_2} + x_3^2$$

$$F4: \quad d = e^{x_1/5} + x_2/4 + 4.x_3$$

onde "d", valor desejado, é a aplicação da respectiva função aritmética ao conjunto de entradas  $(x_1, x_2, x_3)$ .

O número total de padrões será de 100 ( $P=100$ ) para cada função: 70 para treinamento, 15 para validação e 15 para teste, conforme distribuição para validação cruzada do software a ser utilizado (MATLAB®).

As faixas de variações das variáveis de entrada, dados sintéticos, serão estipuladas de forma a equivaler às variações possíveis encontradas em modelos práticos, e que possam proporcionar diferenciação nas análises finais:

$$20 \leq x_1 \leq 30;$$

$$1300 \leq x_2 \leq 1500;$$

$$80 \leq x_3 \leq 100.$$

Os números de neurônios da camada intermediária a serem analisados atendem aos métodos de cálculos citados no referencial teórico:

- a) Segundo Blum (1992),  $N_{\max} = P / [10.(n+s)] = 100 / [10.(3+1)] = 2,5$ ; arredondado para 3 neurônios;
- b) Segundo Kolmogorov (1991),  $N = 2.n + 1 = 2.3 + 1 = 7$  neurônios.

Objetiva-se com este procedimento verificar a influência do número de neurônios da camada intermediária nas respostas das RNAS.

Para verificar a melhor resposta quanto às funções de ativação, conforme orientação da literatura, serão utilizadas funções linear e não lineares na camada intermediária; e função linear na camada de saída, da seguinte forma:

- a) linear (LIN) na camada intermediária e linear na camada de saída, caracterizada por LIN/LIN;
- b) logística (LOG) na camada intermediária e linear na camada de saída, caracterizada por LOG/LIN;
- c) tangente hiperbólica (TANH) na camada intermediária e linear na camada de saída, caracterizada por TANH/LIN.

A função de ativação linear foi fixada na camada de saída baseado em comentário da seção 4.1.3.1.

Para compensar o problema dos mínimos locais e global cada RNA será treinada cinco vezes na tentativa de obter um erro satisfatório (*cross-validation*).

Serão consideradas topologias de performance satisfatória aquelas cujas RNAs apresentarem erros da ordem de  $10^{-2}$  ou menor do que esta ordem (PIULEAC et al, 2010).

Por fim, nesta etapa será selecionada:

- a) uma RNA de topologia satisfatória;
- b) a função mais complexa com relação ao mapeamento;

que serão utilizadas no método de Sensibilização.

## 5.2 VALIDAÇÃO DA TÉCNICA DE SENSIBILIZAÇÃO DE VARIÁVEIS INDEPENDENTES.

Para validar a técnica de Sensibilização de Variáveis Independentes três etapas deverão ser cumpridas:

- a) obtenção dos valores das sensibilidades das três variáveis independentes por meio de simulação na RNA selecionada e treinada, conforme parâmetros da função selecionada;
- b) obtenção dos valores das sensibilidades das três variáveis independentes diretamente da função selecionada, em questão;
- c) comparação dos respectivos valores de Sensibilidades encontrados nos dois formatos de obtenção.

### 5.2.1 Sensibilidades por meio de simulação na RNA.

A partir da RNA selecionada e treinada, conforme função selecionada, resultados da seção 5.1, aplicar-se-á a técnica de Sensibilização de Variáveis Independentes a cada uma de suas três entradas:

- a) cada amostra independente " $x_i$ " a ser testada terá seu valor acrescido de 5% uma a cada vez;
- b) juntamente com o restante do padrão de entrada, será simulada na RNA;
- c) será obtida a saída "y" da RNA correspondente à entrada alterada;

- d) serão calculadas as diferenças entre os valores de "y" e os valores de saídas originais "d", obtidos pelas expressões aritméticas ( $= y - d$ );
- e) a média aritmética dessas diferenças é o valor da SENS para cada variável independente, conforme expressão (10), da seção 4.2.1;
- f) a SENSN será obtida conforme expressão (11) da seção 4.2.2.

### 5.2.2 Sensibilidades diretamente da Função selecionada.

A partir da função selecionada aplicar-se-á a técnica de Sensibilização de Variáveis Independentes a cada uma de suas três entradas:

- a) cada amostra independente " $x_i$ " a ser testada terá seu valor acrescido de 5%, uma a cada vez;
- b) juntamente com o restante dos seus respectivos dados de entrada será calculado o valor de "y" na função em questão;
- c) serão calculadas as diferenças entre os valores de "y" e os valores de saídas originais "d" ( $= y - d$ );
- d) a média aritmética dessas diferenças é o valor da SENS para cada variável independente, conforme expressão (10), da seção 4.2.1;
- e) a SENSN será obtida conforme expressão (11) da seção 4.2.2.

### 5.2.3 Validação da Técnica de Sensibilização por meio de RNAs.

Os valores encontrados nas seções 5.2.1 e 5.2.2 serão comparados para certificação da real representação do ensaio de Sensibilização como ferramenta para obtenção da significância de variáveis independentes dos processos.

### 5.2.4 Variáveis Independentes de Sensibilidades próximas a Zero (0).

Pode ocorrer de alguma Variável Independente apresentar Sensibilidade próxima a ZERO, então, a título de investigação das Sensibilidades encontradas, uma nova simulação será feita usando a RNA já treinada:

- a) cada amostra independente " $x_i$ " a ser testada terá seu valor ZERADO uma a cada vez;
- b) juntamente com o restante do padrão de entrada, será simulada na RNA;

- c) será obtida a saída "y" da RNA correspondente à entrada ZERADA;
- d) considerando "d" como a saída original da função em análise, será calculado o valor do Erro Médio Absoluto (EMA) (MUKHERJEE & ROUTROY, 2012) para cada variável independente, por meio de:

$$EMA = \frac{1}{P} \sum_{i=1}^P \frac{|d_i - y_i|}{d_i} * 100 \quad (48)$$

P: número de padrões.

Este erro representa a variação da saída da rede em função da não existência (valor = 0) da variável independente em análise; e representa uma forma alternativa de análise desta variação da saída.

Supõe-se, desta forma, que a variável ZERADA não participa do processo ou que não tem significância para a saída.

Caso exista alguma identificação dos resultados de Sensibilidade próxima a ZERO e EMA da respectiva variável independente próximo a ZERO, tem-se a indicação de que qualquer variável independente do processo pode ser inserida na análise de significância, mesmo que a mesma não influencie na saída em estudo.

### 5.3. SIGNIFICÂNCIA DAS VARIÁVEIS FÍSICAS INDEPENDENTES DO PROCESSO DE LAMINAÇÃO A QUENTE DE AÇO C-Mn.

Esta fase de obtenção da significância das variáveis do processo será dividida em 3 partes:

- a) definição e seleção das variáveis físicas independentes a serem analisadas; aplicação da técnica de Sensibilidade ao processo:
- b) por Modelagem,
- c) por RNA.

#### 5.3.1 Definição e seleção das variáveis para análise.

- a) estabelecer dentre as variáveis do processo de LTQ, definidas na seção 4.3.2.3, quais serão relevantes para análise de Sensibilidade;
- b) selecionar os dados utilizando a técnica dos "6 sigmas".

As amostras utilizadas nesta fase do estudo são provenientes de bobinas de tiras de aço C-Mn fabricadas pelo processo de LTQ, comentado na seção 4.3.1.

Conforme verificado na seção 4.3.2.3, a variável dependente do processo em estudo é a Força de Laminação; e como variáveis independentes temos diversas categorias: químicas, de produto, de equipamento e de processo.

Como proposta deste estudo, a técnica de sensibilização será aplicada a variáveis físicas independentes, ou seja, a composição química do aço será simplesmente considerada como entrada no modelo.

As peças processadas na laminação serão de largura constantes com tolerâncias admissíveis para 1025 mm.

Algumas características dos equipamentos permanecerão constantes durante a operação, com seus valores originais: diâmetro dos cilindros e distância entre os passes.

Desta forma, restaram como variáveis independentes:

- a) espessuras da peça laminada, tanto na sua entrada no processo, quanto em cada passe de laminação;
- b) temperatura do material laminado, tanto de entrada como em cada passe de laminação;
- c) velocidades dos cilindros em cada passe.

A partir destas, serão consideradas como variáveis independentes de entrada, para análise de suas Sensibilidades, as características do modelo:

- a) a temperatura de entrada do aço "  $T_e$  " no passe, em  $^{\circ}\text{C}$ ;
- b) o grau de redução na espessura do aço "  $r$  " no passe, obtida pela expressão:

$r = (h_1 - h_2)/h_1$  onde "h1" é a espessura de entrada do aço e "h2" é a de saída no passe, adimensional, conforme expressão (19); e

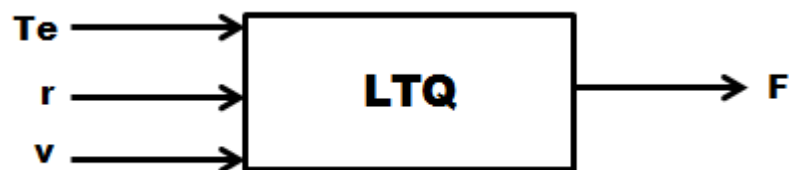
c) a velocidade periférica do cilindro "v" no passe, em m/s.

Estas grandezas são fornecidas pelo sistema de supervisão do processo de LTQ.

A variável de saída será a força de laminação "F" no passe, em kN, que é o valor fornecido pelo processo.

Desta forma, o diagrama das entradas e saída do processo está representado na Figura 11: 3 entradas e 1 saída.

**Figura 11** - Diagrama do processo identificando as variáveis independentes temperatura (Te), redução (r) e velocidade (v); e a variável dependente força de laminação (F).



FONTE: o Autor.

A seleção dos dados será feita pela técnica dos "6 Sigmas", eliminando os padrões que estiverem fora da faixa: valor médio  $\pm 3\sigma$  (desvio padrão).

### 5.3.2 Aplicação da técnica de Sensibilidade ao processo por Modelagem,

Todo o modelo tratado na seção 4.3.2 será instalado numa planilha de Excel, que fornecerá a Força de Laminação "F" a partir dos dados de processo conforme estudo fenomenológico apresentado. Esta planilha estará hábil a possibilitar variações nas variáveis de entrada estabelecidas e fornecer as respectivas alterações na saída.

Uma amostragem de 10 bobinas será utilizada para receber a técnica de Sensibilidade, ou seja, para cada bobina:

a) cada amostra independente "x<sub>i</sub>" ("Te" ou "r" ou "v") a ser testada terá seu valor acrescido de 5%, uma a cada vez;



- b) juntamente com o restante dos seus respectivos dados de entrada será calculado o valor de "y" no modelo;
- c) serão calculadas as diferenças entre os valores de "y" e os valores de saídas originais "d" ( $= y - d$ ), para cada variável independente e para cada bobina.

Para o grupo das 10 bobinas, será obtida a média aritmética dessas diferenças que representa o valor da SENS para a variável independente, conforme expressão (10), da seção 4.2.1. A SENSN será obtida conforme expressão (11) da seção 4.2.2.

### 5.3.3 Aplicação da técnica de Sensibilidade ao processo via RNA.

Estabelecidas as variáveis independentes e a dependente, os padrões para treinamento das RNAs terão a seguinte composição ( $T_e, r, v, F$ ).

Serão estabelecidos os números iniciais de neurônios e de padrões por meio dos métodos de Kolmogorov, expressão (7) e Blum, expressão (8), da seção 4.1.6.1. O número de padrões depende da complexidade do processo a ser modelado (YANG et al., 2004).

A partir da RNA treinada, que identifica o sistema, uma a cada vez, a variável independente " $x_i$ " escolhida será:

- a) acrescida em 5% nos seus valores que substituem os valores anteriores usados no treinamento da rede;
- b) este novo conjunto padrão de entrada ( $T_e, r, v$ ) é então simulado na RNA já treinada;
- c) é obtida a diferença entre os novos valores de saída "y" obtidos na simulação e os valores de saída desejados "d" ( $= y - d$ );
- d) determina-se a média dessas diferenças obtendo-se a  $SENS_i$  da variável independente " $x_i$ " em questão, conforme expressão (10), da seção 4.2.1;
- e) a SENSN será obtida conforme expressão (11) da seção 4.2.2.

Uma análise será feita a partir da correlação da variável independente com cada variável dependente, para cada passe, a título de proporcionar conhecimentos preliminares do processo principalmente no sentido da variação entre estas. Os

valores de sensibilidade apresentam sinais positivos ou negativos dependendo desta relação de variação ser diretamente ou inversamente proporcional, respectivamente.

Por outro lado, serão utilizadas, neste trabalho, RNAs:

- a) com arquitetura do tipo *feedforward* com camada de entrada, uma camada intermediária e camada de saída;
- b) com número de neurônios da camada intermediária obtido por pesquisa empírica da topologia que apresente performance satisfatória;
- c) cujo modelo deve prever o melhor resultado para novos dados (generalização);
- d) com função de ativação Tangente Hiperbólica usada nos neurônios da camada intermediária, enquanto que a da camada de saída será Linear;
- e) com todos os neurônios da rede ligados a um limiar de ativação (bias);
- f) com algoritmo de Levenberg-Marquardt empregado para o treinamento;
- g) com a utilização dos dados de entrada-saída conforme método de validação cruzada (*cross-validation*): 70% para o treinamento/aprendizado, 15% para validação e 15% para o teste, escolhidos de forma aleatória.

## 6 APRESENTAÇÃO E DISCUSSÃO DOS RESULTADOS

### 6.1 RNAs MAPEANDO FUNÇÕES SIMULADORAS.

Foram treinadas as RNAs: de arquitetura PMC (*feedforward*);

- a) com 3 camadas;
- b) 3 entradas;
- c) 1 saída;
- d) 100 padrões do tipo  $(x_1, x_2, x_3, d)$ ;
- e) cinco vezes cada uma;
- f) com uma camada intermediária de 3 ou 7 neurônios;
- g) funções de ativação combinadas de três formas distintas: LIN/LIN, LOG/LIN, TANH/LIN;
- h) para as quatro funções simuladoras de processo F1, F2, F3 e F4.

As funções utilizadas do software MATLAB® foram:

- a) função de treinamento: *Levenberg-Marquardt (LM)*;
- b) função de aprendizagem: *gradiente descendente (GDM)*;
- c) função de performance: *erro quadrático médio (MSE)*.

As performances apresentadas nas Tabelas 1, 2, 3 e 4, respectivamente às funções F1, F2, F3 e F4, mostram os EQM, que caracterizam a performance da RNA, encontrados para cada combinação de funções de ativação, número de neurônios da RNA e para cada treinamento efetuado.

Os valores de EQM mostrados são os verificados no treinamento das RNAs e os respectivos números de épocas, ou *early-stopping* (=1000), quando foram atingidos.

Pela Tabela 1, observa-se que todas as funções de ativação utilizadas atendem a modelos lineares (F1) devido aos baixos erros verificados em todos os 5 treinamentos para qualquer função de ativação na camada intermediária. Os erros foram da ordem de  $10^{-26}$  a  $10^{-2}$ . Ainda evidenciado na Tabela 1, nota-se que para modelos lineares (F1) as funções de ativação não lineares, LOG e TANH, exigem mais treinamento para a redução dos EQM (maior número de épocas) do que as

funções exclusivamente lineares; em muitas situações chegou-se ao *early stopping* (=1000 épocas). Os erros obtidos, da ordem de  $10^{-9}$  a  $10^{-2}$ , foram maiores do que os verificados para funções de ativação somente lineares, da ordem de  $10^{-26}$  a  $10^{-16}$ .

**Tabela 1** - Performances dos treinamentos das RNAs, mapeando a função F1, com variações das funções de ativação e do número de neurônios na camada intermediária: 3 e 7.

Nº NEUR.	Nº TREIN.	LIN/LIN ERROS	ÉPO-CAS	LOG/LIN ERROS	ÉPO-CAS	TANH/LIN ERROS	ÉPO-CAS
3	1	$1,4 \cdot 10^{-26}$	4	$2,5 \cdot 10^{-5}$	194	$1,2 \cdot 10^{-9}$	1000
3	2	$7,3 \cdot 10^{-25}$	5	$5,4 \cdot 10^{-4}$	1000	$3,2 \cdot 10^{-5}$	1000
3	3	$2,0 \cdot 10^{-16}$	5	$3,6 \cdot 10^{-4}$	756	$7,0 \cdot 10^{-3}$	142
3	4	$3,9 \cdot 10^{-23}$	5	$6,9 \cdot 10^{-6}$	1000	$1,3 \cdot 10^{-6}$	1000
3	5	$4,2 \cdot 10^{-17}$	4	$2,1 \cdot 10^{-6}$	1000	$2,4 \cdot 10^{-9}$	1000
7	1	$9,4 \cdot 10^{-19}$	3	$2,4 \cdot 10^{-5}$	1000	$4,3 \cdot 10^{-8}$	1000
7	2	$5,0 \cdot 10^{-22}$	4	$1,6 \cdot 10^{-2}$	279	$2,7 \cdot 10^{-6}$	1000
7	3	$8,5 \cdot 10^{-17}$	4	$2,7 \cdot 10^{-6}$	1000	$6,1 \cdot 10^{-4}$	1000
7	4	$4,6 \cdot 10^{-18}$	3	$5,9 \cdot 10^{-8}$	505	$1,3 \cdot 10^{-4}$	1000
7	5	$2,5 \cdot 10^{-19}$	3	$7,6 \cdot 10^{-5}$	1000	$2,6 \cdot 10^{-2}$	50

Fonte: Autor

**Tabela 2** - Performances dos treinamentos das RNAs, mapeando a função F2, com variações das funções de ativação e do número de neurônios na camada intermediária: 3 e 7

Nº NEUR.	Nº TREIN.	LIN/LIN ERROS	ÉPO-CAS	LOG/LIN ERROS	ÉPO-CAS	TANH/LIN ERROS	ÉPO-CAS
3	1	91	5	0,08	737	0,12	138
3	2	98,2	4	0,1	344	70,7	10
3	3	60,7	3	31,1	10	0,1	191
3	4	88,3	2	0,2	177	0,06	322
3	5	105,9	3	0,09	257	0,28	325
7	1	36,4	3	0,06	265	0,14	331
7	2	51,7	2	0,06	227	0,09	14
7	3	98	3	0,16	300	0,06	180
7	4	59,1	2	0,33	73	0,07	38
7	5	72,6	2	0,11	165	0,16	231

Fonte: Autor

**Tabela 3** - Performances dos treinamentos das RNAs, mapeando a função F3, com variações das funções de ativação e do número de neurônios na camada intermediária: 3 e 7.

Nº NEUR.	Nº TREIN.	LIN/LIN ERROS	ÉPO- CAS	LOG/LIN ERROS	ÉPO- CAS	TANH/LIN ERROS	ÉPO- CAS
3	1	$2,8 \cdot 10^5$	2	5,3	446	0,08	1000
3	2	$2,6 \cdot 10^5$	2	3,6	466	6,7	480
3	3	$5,0 \cdot 10^5$	5	4,4	279	2,8	330
3	4	$3,8 \cdot 10^5$	2	95	91	5	842
3	5	$3,2 \cdot 10^5$	5	58,9	164	0,14	1000
7	1	$2,5 \cdot 10^5$	5	7,8	569	0,2	1000
7	2	$5,2 \cdot 10^5$	2	0,12	665	16,2	141
7	3	$4,9 \cdot 10^5$	2	0,21	932	0,03	1000
7	4	$3,1 \cdot 10^5$	3	0,08	492	3,4	628
7	5	$5,2 \cdot 10^5$	2	0,19	1000	0,71	1000

Fonte: Autor

**Tabela 4** - Performances dos treinamentos das RNAs, mapeando a função F4, com variações das funções de ativação e do número de neurônios na camada intermediária: 3 e 7.

Nº NEUR.	Nº TREIN.	LIN/LIN ERROS	ÉPO- CAS	LOG/LIN ERROS	ÉPO- CAS	TANH/LIN ERROS	ÉPO- CAS
3	1	812	2	0,1	300	0,15	356
3	2	868	3	0,15	169	0,32	52
3	3	891	4	0,06	235	0,07	1000
3	4	751	1	0,42	38	0,1	146
3	5	979	3	0,08	164	0,14	448
7	1	812	3	0,2	148	0,55	96
7	2	439	1	0,2	61	0,37	77
7	3	369	5	0,06	297	0,01	174
7	4	794	4	0,12	215	0,45	20
7	5	879	2	0,21	213	0,34	43

Fonte: Autor

A partir das Tabelas 1, 2, 3 e 4, pode-se verificar que os diversos treinamentos realizados mostram que realmente o aprendizado por correção de erros passa por valores de mínimos locais, pois foram encontrados valores de erros variáveis para uma mesma função, nos diversos treinamentos efetuados. Por exemplo, no caso da função F3, na Tabela 3, na combinação de funções de ativação TANH/LIN, para topologia de 7 neurônios, foram obtidos valores de erros que variaram de  $3 \cdot 10^{-2}$  até 16,2. Observa-se que em diferentes treinamentos obtém-se diferentes performances

da rede. Os valores menores de erro tendem a aproximarem-se do mínimo global, ou assumir o mesmo.

Os modelos não lineares F2, F3 e F4 não tiveram performance satisfatórias quando treinados numa topologia de função de ativação linear (LIN) na camada intermediária, conforme Tabelas 2, 3 e 4, respectivamente, pois foram verificados grandes valores de erros, na ordem de  $10^1$  a  $10^5$ . Isto confirma que a função de ativação linear não possibilita aos neurônios mapearem processos não lineares.

Pela Tabela 3 verifica-se que, para a combinação de funções de ativação LOG/LIN, a RNA atende ao modelo F3 somente com topologia de 7 neurônios - para topologia de 3 neurônios os erros apresentados foram de 3,6 a 58,9. Vários outros treinamentos foram tentados, neste caso, com RNA de 3 neurônios e não houve apresentação de boa performance por parte da rede. Logo, o número de neurônios na camada intermediária influencia no treinamento da rede.

As funções de ativação Logística (LOG) e Tangente Hiperbólica (TANH), em camadas intermediárias de 7 neurônios - conforme método de Kolmogorov - equação (7) da seção 4.1.6.1, atendem aos modelos não lineares de diferentes complexidades, como F2, F3 e F4, devido a algumas performances apresentarem resultados de erros satisfatórios nos treinamentos nas combinações LOG/LIN e TANH/LIN. Logo, os neurônios com a função de ativação não linear são capazes de mapear processos não lineares.

Partindo do estabelecido na seção 5.1 quanto ao valor do erro ser da ordem de grandeza de  $10^{-2}$  ou menor do que esta (os valores que atendem estão apresentados em vermelho nas Tabelas), podemos verificar que:

- a) pela Tabela 1, todas as topologias atendem à função linear F1;
- b) pela Tabela 2, as topologias LIN/LIN não atendem à função não linear F2, sendo atendida pelas outras topologias;
- c) pela Tabela 3,
  - todas as topologias LIN/LIN; e
  - LOG/LIN de 3 neurôniosnão atendem à função não linear F3;

- d) pela Tabela 4, as topologias LIN/LIN não atendem à função não linear F4, sendo atendida pelas outras topologias.

A partir destas verificações, a Tabela 5 apresenta um resumo da identificação das topologias das RNAs que apresentaram performances com erro da ordem de  $10^{-2}$  ou abaixo deste valor. A indicação "X" na Tabela 5 identifica o proposto atendimento da topologia ao mapeamento da respectiva função.

**Tabela 5** - Identificação das topologias das RNAs treinadas, que apresentaram performances com erro na ordem de  $10^{-2}$  ou menor, no mapeamento das funções em análise.

FUNÇÕES	TOPOLOGIAS					
	LIN/ LIN		LOG/ LIN		TANH/ LIN	
	3	7	3	7	3	7
F1	X	X	X	X	X	X
F2			X	X	X	X
F3				X	X	X
F4			X	X	X	X

Fonte: Autor

Pode-se verificar, então, que as topologias que atendem às solicitações a todas as funções são:

- LOG/LIN de 7 neurônios;
- TANH/LIN de 3 neurônios; e
- TANH/LIN de 7 neurônios.

Por fim, é escolhida, então, a topologia de RNA de:

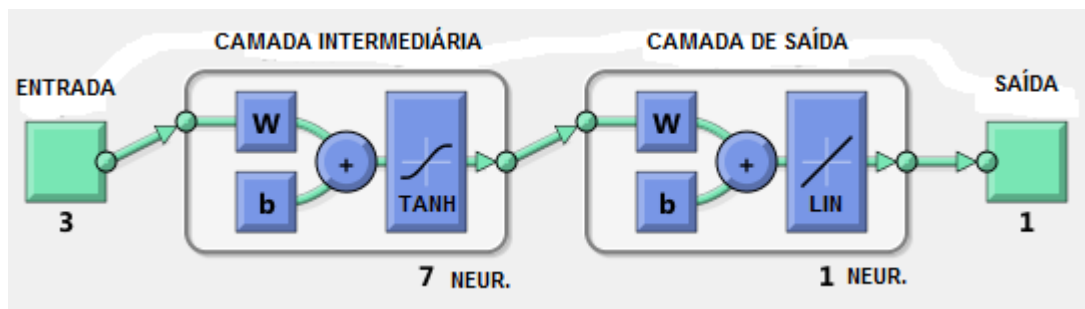
- função de ativação TANH/LIN, por ser satisfatória a todos os mapeamentos quanto ao tamanho do erro, independente do número de neurônios;
- de 7 neurônios, conforme Método de Kolmogorov - equação (7), seção 4.1.6.1;
- e a função F3, que apresentou maior complexidade para mapeamento, para simulações nas análises de sensibilização das variáveis independentes.

## 6.2 VALIDAÇÃO DA TÉCNICA DE SENSIBILIZAÇÃO DE VARIÁVEIS INDEPENDENTES.

### 6.2.1 Sensibilidades por meio de simulação na RNA.

Então, a arquitetura resultante, da seção 6.1, da RNA selecionada, PMC, é mostrada na Figura 12, apresentando uma camada escondida com 7 neurônios e função de ativação Tangente Hiperbólica capaz de mapear a não linearidade da função F3; e função de ativação Linear na camada de saída com um neurônio.

**Figura 12** - Arquitetura da RNA proposta.



Fonte: MATLAB® / Autor

Conforme estabelecido na metodologia, a RNA foi treinada com 100 padrões  $(x_1, x_2, x_3, d)$  correspondentes à função:

$$F3: d = x_1^3 + \sqrt{x_2} + x_3^2,$$

e as funções utilizadas do software, também, foram:

- função de treinamento: *Levenberg-Marquardt (LM)*;
- função de aprendizagem: *gradiente descendente (GDM)*;
- função de performance: *erro quadrático médio (MSE)*.

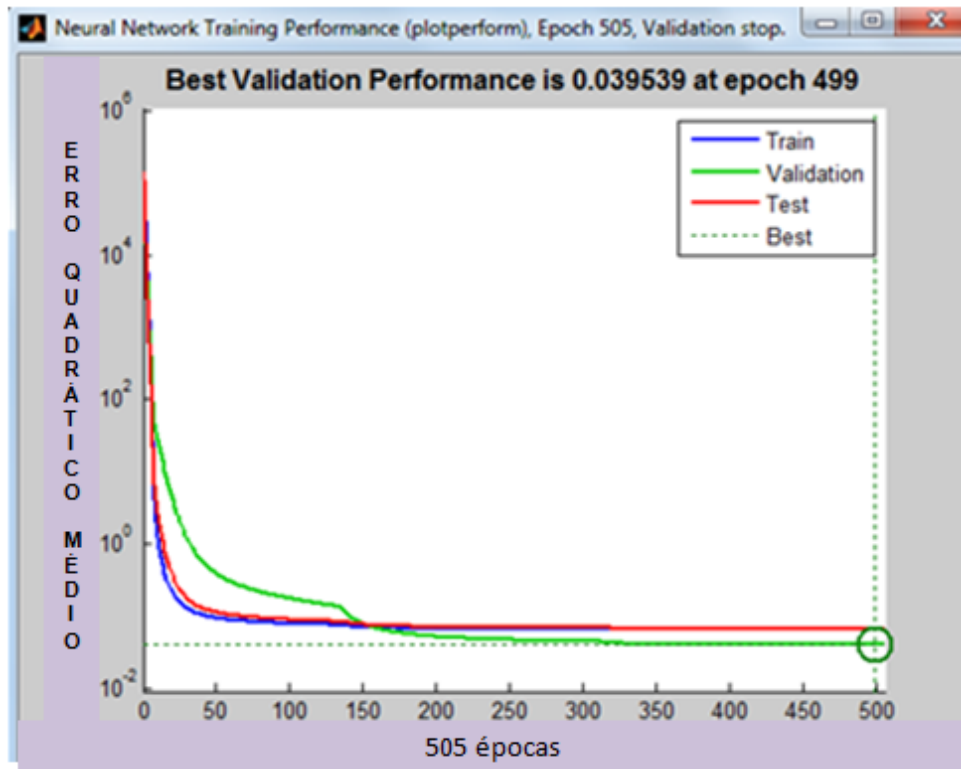
Os resultados, para cada treinamento, são fornecidos pelo software, tais como: performance da rede; erro de correlação no treinamento, teste e validação; e parâmetros de treinamento, que são apresentados nas Figuras 13, 14 e 15, respectivamente.

A Figura 13 apresenta o comportamento de treinamento e validação com consequente ajuste de pesos e teste. Observa-se, nesta, que o menor valor do EQM (*Best Validation Performance*) obtido foi  $3,9539 \cdot 10^{-2}$ ; que a validação do treinamento



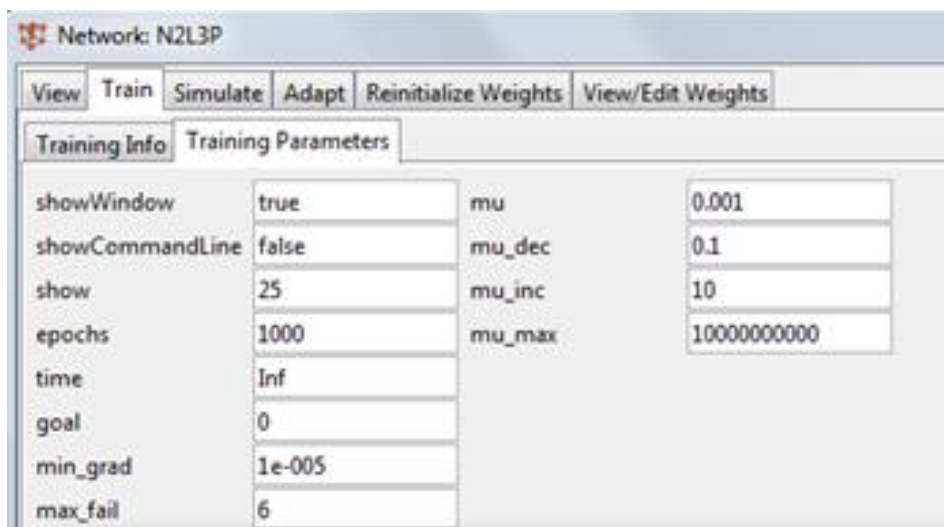
atingiu a melhor performance na época 499 repetindo-se até à 505 ( $max\_fail = 6$ , na figura 14).

Figura 13 - Performance da Rede.



Fonte: MATLAB®.

Figura 14 - Parâmetros de treinamento default do software.



Fonte: MATLAB®.

Os dados da Figura 14, já comentados na seção 4.1.6.1, são:

- a) o valor inicial de " $\mu$ " ( $mu=0,001$ );
- b) o fator de decréscimo de " $\mu$ " ( $mu\_dec=0,1$ );
- c) o fator de acréscimo de " $\mu$ " ( $mu\_inc=10$ );
- d) o valor máximo permitido para o parâmetro " $\mu$ " ( $mu\_max=10^{10}$ );
- e) o valor de performance do gradiente, derivada da curva do EQM, a ser atingido ( $min\_grad=10^{-5}$ );
- f) o número máximo de épocas a ser repetido o valor do desempenho da RNA, pela validação do treinamento ( $max\_fail=6$ );
- g) o valor do EQM (performance da RNA) solicitado a ser atingido no treinamento da RNA ( $goal=0$ );
- h) o tempo máximo para efetivar o treinamento, em segundos ( $time=infinito$ );
- i) o número máximo de épocas a serem apresentadas no treinamento ( $epochs=1000$ );

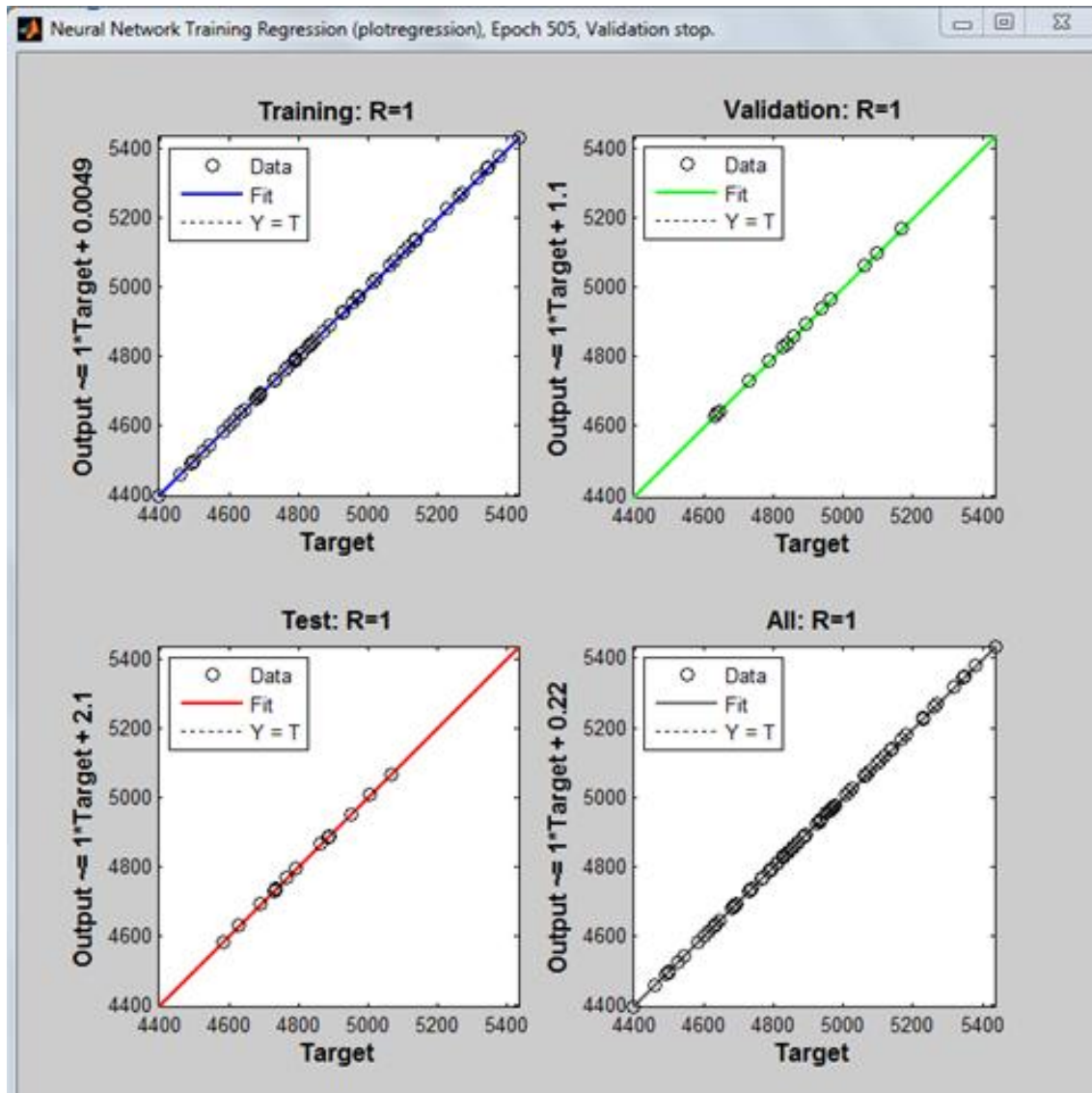
default do software.

Outros parâmetros apresentados nesta janela, Figura 14, são:

- a) *showWindowdown* (*true*): habilita mostrar a janela de treinamento quando está treinando uma RNA;
- b) *showCommandLine* (*show*): mostra a linha de comando que estabelece o número de épocas durante o treinamento;
- c) *show* (*25*): estabelece a periodicidade de indicação do número de épocas na janela de treinamento ( a cada 25 épocas).

A Figura 15 mostra a correlação dos resultados de saídas da rede treinada "y" (*output*) e dos valores desejados "d" (*target*) fornecidos ao treinamento. A apresentação é feita para as saídas dos grupos de padrões de treinamento, validação, teste e todos os padrões disponibilizados. Observa-se que o valor do coeficiente de correlação em todos os casos foi 1 ( $R=1$ ) o que implica numa ótima performance da RNA no mapeamento da função em análise.

**Figura 15** - Correlação das saídas da rede "y" e das saídas originais "d" para treinamento, teste, validação e todas as saídas.

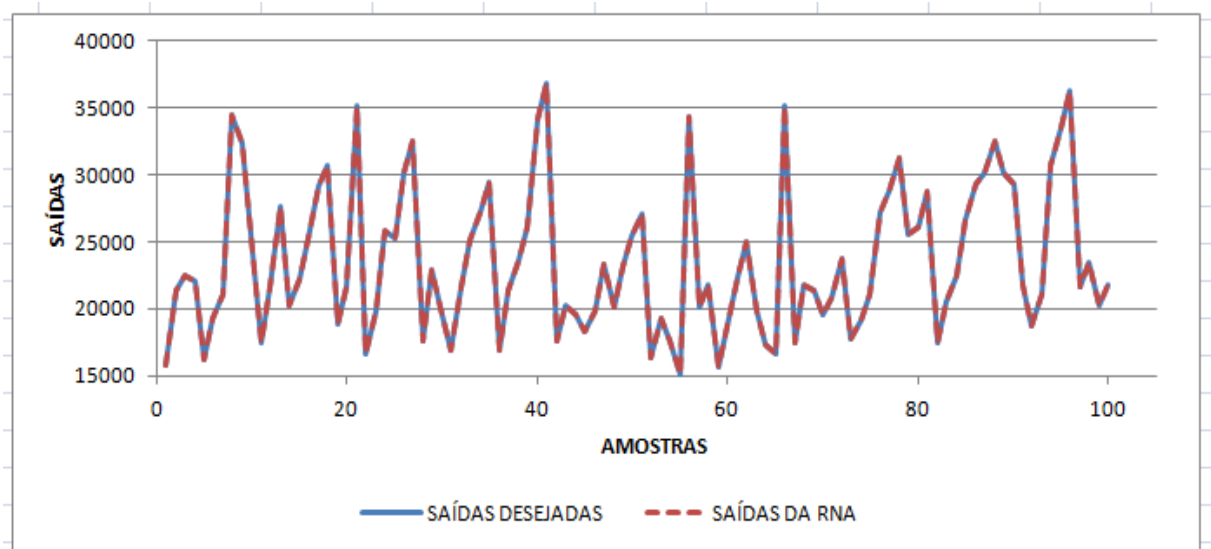


Fonte: MATLAB®.

A comparação das saídas desejadas "d" e as fornecidas "y" pela RNA estão apresentadas na Figura 16, onde observa-se a grande proximidade das respostas da RNA com os valores desejados.

Este fato evidencia que a RNA teve a capacidade de aprender o modelo, ou seja, de identificar a função F3; o que confirma a boa performance deste tipo de topologia no mapeamento de funções. Por fim, as RNAs são capazes de mapear funções.

**Figura 16** - Comparação gráfica entre as saídas desejadas "d" e as saídas "y" fornecidas pela RNA.



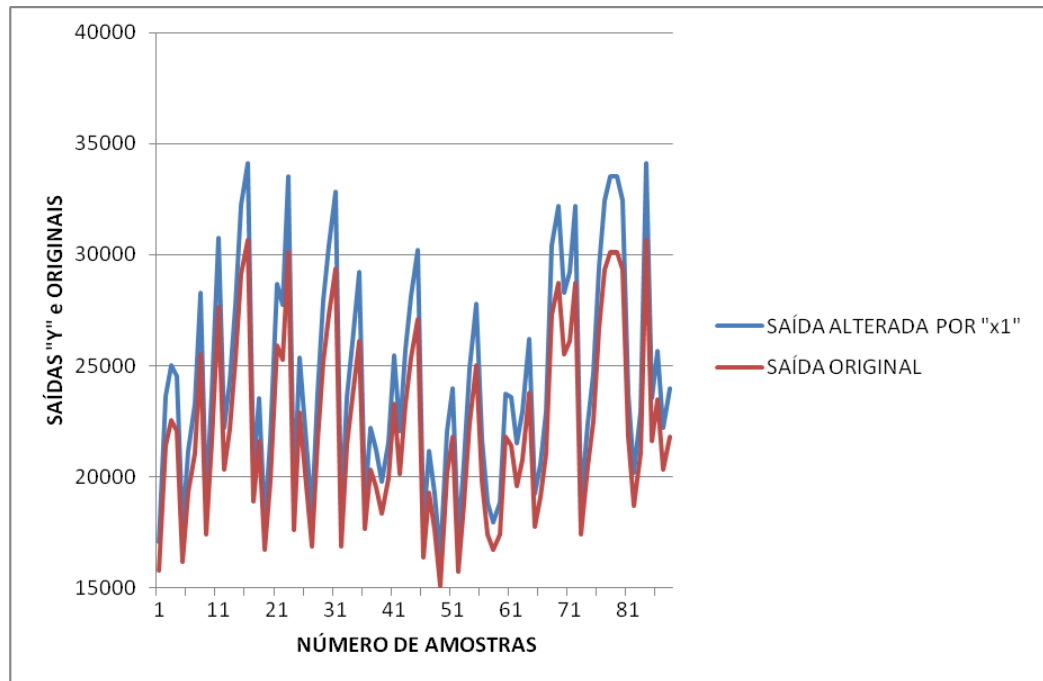
Fonte: Autor

A seguir, conforme procedimento descrito na seção 5.2.1, alterações foram provocadas nas variáveis independentes, uma a uma, e os novos padrões foram simulados na RNA já treinada anteriormente:

- a) cada amostra independente " $x_i$ " a ser testada teve seu valor acrescido de 5% uma a cada vez;
- b) juntamente com o restante do padrão de entrada, foi simulada na RNA;
- c) foi obtida a saída "y" da RNA correspondente à entrada alterada.

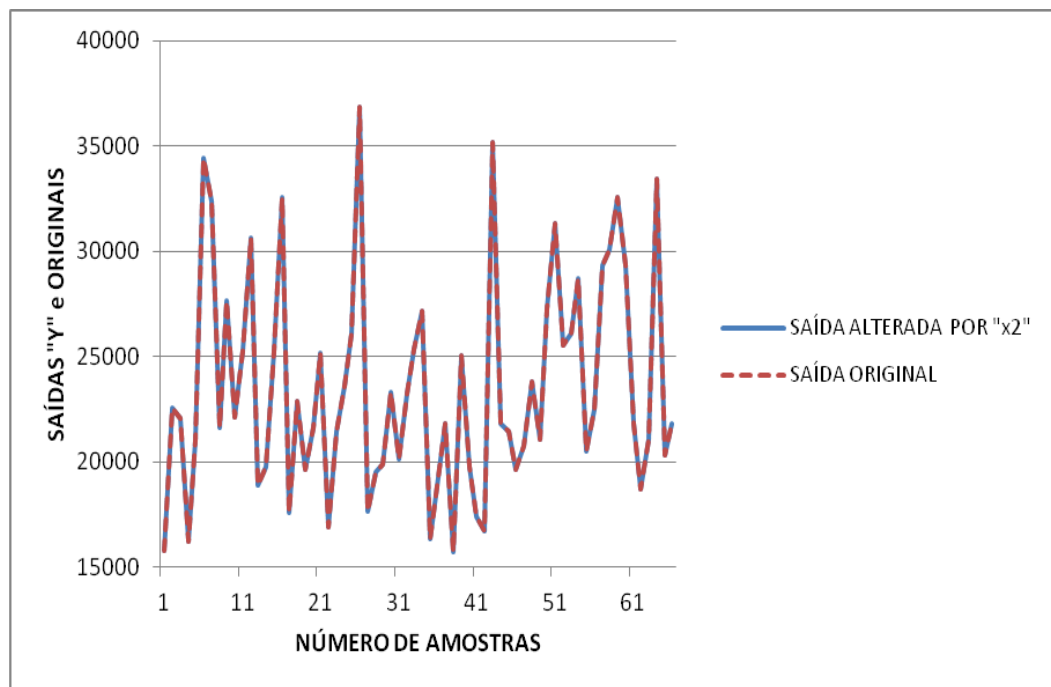
As Figuras 17, 18 e 19, mostram as comparações gráficas dos resultados obtidos das novas saídas simuladas "y" com as alterações das respectivas entradas  $x_1$ ,  $x_2$  e  $x_3$ , aumentadas em 5%, e as saídas originais "d".

**Figura 17** - Comparação gráfica entre as novas saídas "y" correspondentes às alterações, pelo acréscimo de 5% na entrada "x1", e as saídas originais "d".



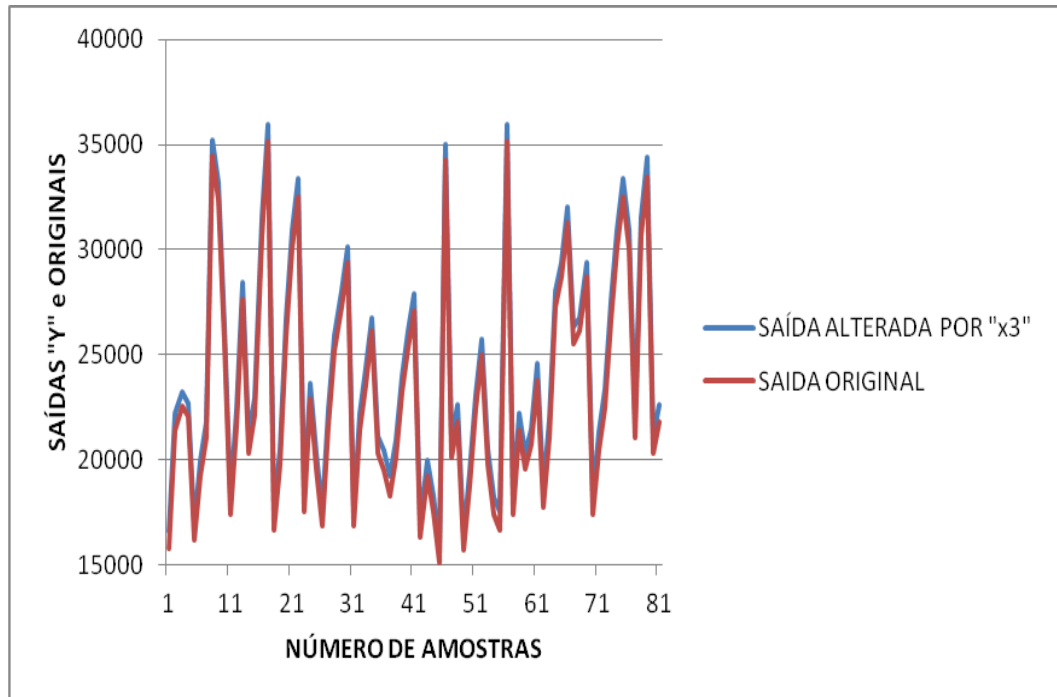
Fonte: Autor

**Figura 18** - Comparação gráfica entre as novas saídas "y" correspondentes às alterações, pelo acréscimo de 5% na entrada "x2", e as saídas originais "d".



Fonte: Autor

**Figura 19** - Comparação gráfica entre as novas saídas "y" correspondentes às alterações, pelo acréscimo de 5% na entrada "x3", e as saídas originais "d".



Fonte: Autor

Verifica-se que a variação na variável dependente ocorreu com menor intensidade pelo acréscimo da variável "x<sub>2</sub>", conforme Figura 18, e com maior intensidade pelo acréscimo da variável "x<sub>1</sub>", conforme Figura 17. A variação da variável "x<sub>3</sub>" provocou variação intermediária na variável dependente, conforme Figura 19.

A partir dos resultados do procedimento anterior, foram obtidas as SENS<sub>i</sub> de cada variável independente em análise, conforme seção 4.2.1, expressão (10), onde "i" é a entrada do padrão. Os valores dessas Sensibilidades estão apresentados nas Tabelas 6, 7 e 8, correspondentes às variações das entradas x<sub>1</sub>, x<sub>2</sub> e x<sub>3</sub>, respectivamente, da seguinte forma:

- a) são apresentados os cinco primeiros padrões de amostras utilizados nos respectivos casos para a função F3, na RNA selecionada;
- b) os padrões cujas entradas, após alteração, excediam o limite superior de sua faixa, conforme item 4.2.1, foram retirados do ensaio, ou seja,  $x_1 > 30$ ,  $x_2 > 1500$  e  $x_3 > 100$ .

Nas Tabelas são apresentados:

- a) os padrões utilizados com alterações em " $x_i$ ",
- b) as saídas originais (d),
- c) as saídas alteradas da rede treinada (y),
- d) as diferenças individuais de cada padrão ( $d_{if} = y - d$ ) e
- e) a média aritmética destas diferenças de todos os padrões apresentados em cada análise, que é a própria  $SENS_i$ .

**Tabela 6** - Sensibilidade Média da variável independente " $x_1$ ".

$1,05.x_1$	$x_2$	$x_3$	d	y	$d_{if}$	$SENS_1$
21	1300	88	15780	17040	1260	
25,2	1500	87	21432	23611	2179	
26,3	1340	83	22551	25014	2463	
26,3	1360	80	22062	24525	2463	
22,1	1420	83	16188	17648	1460	
:	:	:	:	:	:	2218

Fonte: Autor

**Tabela 7** - Sensibilidade Média da variável independente " $x_2$ ".

$x_1$	$1,05.x_2$	$x_3$	D	y	$d_{if}$	$SENS_2$
20	1365	88	15780	15781	0,8329	
25	1407	83	22551	22552	1,0653	
25	1428	80	22062	22063	1,2122	
21	1491	83	16188	16189	0,9974	
24	1386	85	21085	21086	0,8911	
:	:	:	:	:	:	0,9377

Fonte: Autor

**Tabela 8** - Sensibilidade Média da variável independente " $x_3$ ".

$x_1$	$x_2$	$1,05.x_3$	D	y	$d_{if}$	$SENS_3$
20	1300	92,4	15780	16573	793	
24	1500	91,35	21432	22208	776	
25	1340	87,15	22551	23257	706	
25	1360	84	22062	22718	656	
21	1420	87,15	16188	16894	706	
:	:	:	:	:	:	793

Fonte: Autor

Os valores de SENS<sub>N</sub> para cada variável "x<sub>i</sub>" foram calculados a partir das SENS obtidas anteriormente, por meio da expressão (11), seção 4.2.2, e todos estão apresentados na Tabela 9.

**Tabela 9** - Valores de Sensibilidade Média e Sensibilidade Média Normalizada das variáveis independentes obtidas por meio de RNAs .

VARIÁVEL	SENS	SENSN(%)
X1	2218	73,63
X2	0,9377	0,03
X3	793	26,33

Fonte: Autor

Pelas Tabela 9 verifica-se que a simulação do modelo F3 (  $d = x_1^3 + \sqrt{x_2} + x_3^2$  ) mostra que a variável independente "x<sub>1</sub>" tem maior significância ao provocar maiores variações na variável dependente visto que sua Sensibilidade Média (SENS<sub>1</sub>=2218) e sua Sensibilidade Média Normalizada (SENSN<sub>1</sub>=73,63%) foram as de maior valor. Este fato evidencia a importância de flutuações desta variável no processo em cuja modelagem está elevada à potência 3 (x<sub>1</sub><sup>3</sup>). A variável "x<sub>2</sub>", apesar de ter maior valor absoluto em sua faixa de variação (1300 ≤ x<sub>2</sub> ≤ 1500), seção 5.1, acabou por ter pouca significância neste modelo pois sua Sensibilidade Média (SENS<sub>2</sub>=0,9377) e sua Sensibilidade Média Normalizada (SENSN<sub>2</sub>=0,03%) apresentaram os menores valores entre as variáveis independentes. O modelo F3 em análise aplica a função de radiciação na variável "x<sub>2</sub>" ( $\sqrt{x_2}$ ) o que faz com que sua influência em variações na variável dependente seja pequena. Desta forma, fica caracterizada a variável independente "x<sub>1</sub>" como a crítica deste processo; seguida pela variável "x<sub>3</sub>" e, por fim, a variável de menor significância "x<sub>2</sub>".

#### 6.2.2 Sensibilidades diretamente da Função selecionada.

A partir da função F3 selecionada anteriormente foi aplicada a técnica de Sensibilização de Variáveis Independentes a cada uma de suas três entradas:

- com o acréscimo de 5% a cada variável independente por vez;
- obtenção da saída "y" pela função F3 aplicada ao novo conjunto de dados;
- cálculo das diferenças entre os valores de "y" e os valores de saídas originais "d" ( = y - d );



d) cálculo da SENS para cada variável independente obtida conforme expressão (10), da seção 4.2.1;

e) cálculo da SENSN obtida conforme expressão (11), da seção 4.2.2.

Os valores das Sensibilidades verificadas nesta fase estão apresentados na Tabela 10.

**Tabela 10** - Valores de Sensibilidade Média e Sensibilidade Média Normalizada obtidas por meio da função F3 aplicada às variáveis alteradas, uma a cada vez.

VARIÁVEL	SENS	SENSN (%)
X1	2437,67	74,62
X2	0,92	0,03
X3	828	25,35

Fonte: Autor

Pela Tabela 10 pode-se observar que a função simuladora F3 estabelece a sequência de significância das variáveis independentes como sendo a "x<sub>1</sub>" de maior significância, "x<sub>3</sub>" de representação intermediária e "x<sub>2</sub>" como a variável de menor representativa na variável dependente.

### 6.2.3 Validação da Técnica de Sensibilização por meio de RNAs.

Os valores encontrados pela aplicação da técnica de Sensibilização utilizando a função F3, apresentados na seção 6.2.2, representam a real significância das variáveis independentes em variações (interferências) na variável dependente.

Os valores de sensibilidades apresentados na seção 6.2.1 foram obtidos por meio da aplicação desta técnica em uma RNA usando suas características de mapeadora de funções e de generalização.

Na Tabela 11 estão apresentados os valores de SENSN para as duas situações de simulação: por meio da Função F3 e pela RNA.

**Tabela 11** - Valores de Sensibilidade Média Normalizada obtidas a partir da função F3 e por meio da RNA.

VARIÁVEL	SENSN (%)	
	FUNÇÃO F3	RNA
X1	74,62	73,63
X2	0,03	0,03
X3	25,35	26,33

Fonte: Autor.

Pode-se verificar, pela Tabela 11, que os resultados das duas formas de obtenção da Sensibilidade das variáveis independentes apresentam a mesma sequência de significância destas: mais relevante  $x_1$ , seguida por  $x_3$ , e, por fim, a de menor significância,  $x_2$ .

Obtendo-se o EMA, para cada variável, por meio da diferença dos valores, com relação ao valor correspondente à função F3, verifica-se 0,0133 para a variável " $x_1$ ", 0 para a variável " $x_2$ " e 0,039 para a variável " $x_3$ ", compatíveis com a performance da RNA (0,039539).

Esta constatação valida a efetiva operacionalidade da técnica de Sensibilidade de Variáveis Independentes por meio de RNAs mapeando processos. A complexidade do processo é determinante do grau de análise via RNA.

#### 6.2.4 Variáveis Independentes de Sensibilidades próximas a Zero (0).

Os resultados dos valores de Sensibilidade da seção 6.2.3 mostraram a variável independente " $x_2$ " com SENSN próxima a Zero (= 0,03). Foi investigado, então, o efeito da não existência (valor = 0) de cada variável independente sobre a variável dependente, utilizando a RNA treinada. Cada variável, uma a cada vez, participou de todos os padrões com valor igual a ZERO (= 0).

Os resultados foram apresentados por meio do EMA obtido nesta simulação em comparação com a situação original, conforme seção 5.2.4, e apresentado nas Tabelas 13, 14 e 15.

Para uma análise comparativa foi obtido este mesmo erro para as saídas do treinamento da rede, apresentado na Tabela 12, a fim de garantir que este não seja de grandeza equivalente aos outros. Nesta Tabela são apresentados:

- os 5 primeiros padrões utilizados no treinamento da RNA ( $x_1, x_2, x_3$ );
- as saídas originais (d),
- as saídas da rede treinada (y),
- os erros individuais de cada padrão ( $e = [d-y]/d$ ); e
- o EMA de todos os padrões, conforme expressão (12) da seção 5.2.4.

**Tabela 12** - Erro Médio Absoluto das Saídas da RNA Treinada

$x_1$	$x_2$	$x_3$	d	Y	e	EMA
20	1300	88	15780	15779,057	0,00006	
24	1500	87	21432	21431,528	0,000022	
25	1340	83	22551	22550,799	0,000009	
25	1360	80	22062	22062,282	0,000013	
21	1420	83	16188	16187,957	0,000003	
:	:	:	:	:	:	1,1E-05

Fonte: Autor

Nas Tabelas 13, 14 e 15 são apresentados:

- os 5 primeiros padrões utilizados na simulação da RNA ( $x_1, x_2, x_3$ );
- as saídas originais (d),
- as saídas alteradas da rede treinada (y),
- os erros individuais de cada padrão ( $e = [d-y]/d$ ) e
- o EMA de todos os padrões apresentados em cada análise, conforme expressão (12) da seção 5.2.4.

**Tabela 13** - Erro Médio Absoluto das Saídas da RNA simulada com  $x_1=0$ .

$x_1$	$x_2$	$x_3$	d	Y	e	EMA
0	1300	88	15780	6261	0,6033	
0	1500	87	21432	6093	0,7157	
0	1340	83	22551	5414	0,7599	
0	1360	80	22062	4929	0,7766	
0	1420	83	16188	5416	0,6654	
:	:	:	:	:	:	0,7069

Fonte: Autor

**Tabela 14** - Erro Médio Absoluto das Saídas da RNA simulada com  $x_2=0$ 

$x_1$	$x_2$	$x_3$	d	Y	e	EMA
20	0	88	15780	15762	0,0012	
24	0	87	21432	21411	0,001	
25	0	83	22551	22533	0,0008	
25	0	80	22062	22045	0,0008	
21	0	83	16188	16169	0,0012	
:	:	:	:	:	:	0,00089

Fonte: Autor

**Tabela 15** - Erro Médio Absoluto das Saídas da RNA simulada com  $x_3=0$ 

$x_1$	$x_2$	$x_3$	d	Y	e	EMA
20	1300	0	15780	6955	0,5593	
24	1500	0	21432	12784	0,4035	
25	1340	0	22551	14587	0,3532	
25	1360	0	22062	14587	0,3388	
21	1420	0	16188	8216	0,4924	
:	:	:	:	:	:	0,4073

Fonte: Autor

O resumo dos valores dos EMA para o treinamento e para cada situação de não existência da variável independente no processo é apresentado na Tabela 16.

**Tabela 16** - Valores dos erros, em %, para o treinamento e para cada situação de não existência da variável independente no processo.

TREINAMENTO	VARIÁVEL INDEPENDENTE "ZERADA"		
	X1	X2	X3
0,0011	70,7	0,0892	40,7

Fonte: Autor

A Tabela 16 evidencia que o erro de treinamento, de 0,0011%, é muito menor do que os erros obtidos com a alteração das variáveis independentes de forma que não teria influência significativa nos outros erros; e que a variável " $x_2$ " tem pouca representatividade em variações na variável dependente.

Pela tabela 17 pode-se comparar os resultados das SENS<sub>N</sub>, obtidas via RNA na seção 6.2.1, e dos erros obtidos com a simulação de variáveis independentes "zeradas".

**Tabela 17** - Valores de Sensibilidades Médias Normalizadas, via RNA, e Erros Médios Absolutos apresentados por cada Variável Independente, considerando-as ZERADAS, uma a cada vez.

VARIÁVEL INDEPENDENTE	SENSN (%)	EMA (%)
	RNA	VAR. ZERADA
X1	73,63	70,7
X2	0,03	0,0892
X3	26,33	40,7

Fonte: Autor

Pelo fato da variável independente " $x_2$ " ter apresentado a menor SENS<sub>N2</sub> (=0,03%), buscou-se verificar sua real influência na variável dependente pela simulação de sua eliminação no processo (valor=0), apresentada na Tabela 14. O valor do EMA obtido para as saídas nessa nova situação com relação aos valores originais foi bem pequeno (=0,0892%), compatível com a Menor Sensibilidade Média Normalizada verificada (=0,03%). De outra feita "eliminou-se" a variável independente " $x_1$ " verificando-se um erro de 70,7%, conforme Tabela 17, na variável dependente, quando sua sensibilidade apresentada foi de SENS<sub>N1</sub>=73,63%, ou seja, realmente comprova as significâncias grande para a variável " $x_1$ " e desprezível para a variável " $x_2$ ".

Verifica-se, desta forma, que a técnica de Sensibilidade identificou uma variável, " $x_2$ " que participa do processo mas que não tem significância para a saída em análise, ou seja, esta técnica permite que sejam arroladas todas as variáveis que se tenha interesse que sejam analisadas em um processo e o resultado da técnica indicará se a variável independente é significativa ou desprezível para a variável dependente em análise. Por fim, este fato indica que qualquer variável independente do processo pode ser inserida na análise de Sensibilidade, mesmo que a mesma não influencie na saída em estudo.

### 6.3 APLICAÇÃO DA TÉCNICA DE SENSIBILIDADE NA LTQ

### 6.3.1. Seleção dos dados de entrada.

As variáveis independentes selecionadas na seção 5.3.1 foram:

- "Te" temperatura de entrada no passe de laminação (°C);
- "r" grau de redução da espessura da tira no passe (adimensional);
- "v" velocidade periférica do cilindro no passe de laminação (m/s).

Os dados das amostras foram avaliados por meio da técnica dos "6 Sigmas", sendo eliminadas as amostras (bobinas) que tinham algum dado que não pertencia ao respectivo intervalo de valores limitados por "*valor médio - 3.σ*" a "*valor médio + 3σ*". Desta feita, foram eliminados pontos discrepantes destas variáveis que poderiam comprometer a confiabilidade do banco de dados (OLIVEIRA & MODENESI, 2008).

A Tabela 18 apresenta as faixas de cada variável dos padrões (Te, r, v, F) utilizados como dados para a modelagem e para a identificação dos processo.

**Tabela 18** - Faixas de cada variável, do banco de dados, utilizado para os ensaios, por passe.

PASSE	Te (°C)		r		v (m/s)		F (kN)	
	Min	Max	Min	Max	Min	Max	Min	Max
1	938,6	1019,9	0,230	0,374	1,38	1,88	8983	17114
2	924	996,7	0,218	0,371	1,84	2,99	9255	16244
3	907,8	975,8	0,179	0,302	2,28	4,17	8337	14545
4	889,2	953,8	0,127	0,275	2,63	5,55	6890	13777
5	872,7	933,9	0,132	0,222	3,09	7,03	5940	11338
6	857,6	915,2	0,092	0,164	3,51	8,38	4634	8781

FONTE: Autor.

### 6.3.2 Técnica de Sensibilização via Modelagem.

O modelo tratado na seção 4.3.1, teve seu estudo fenomenológico, apresentado na seção 4.3.2, instalado em uma planilha do Excel, onde pode-se obter a Força de Laminação "F" a partir dos dados das bobinas de aço C-Mn resultantes de uma LTQ. A partir de dados de processo de 10 bobinas, provocou-se variações em cada uma das variáveis independentes estabelecidas aplicando a técnica de Sensibilidade, ou seja, para cada bobina:

- a) cada amostra independente " $x_i$ " (" $T_e$ " ou " $r$ " ou " $v$ ") teve seu valor acrescido de 5%, uma a cada vez;
- b) juntamente com o restante dos seus respectivos dados de entrada foi calculado o valor de " $y$ " no modelo;
- c) foram calculadas as diferenças entre os valores de " $y$ " e os valores de saídas originais " $d$ " ( $= y - d$ ), para cada variável independente e para cada bobina.

Para o grupo das 10 bobinas, foi obtida a média aritmética dessas diferenças que representa o valor da SENS da variável independente, conforme expressão (10), da seção 4.2.1. A SENSN foi obtida conforme expressão (11), da seção 4.2.2.

Os valores de SENS e SENSN de cada variável independente encontrados por meio da aplicação da Modelagem às 10 bobinas estão apresentados na Tabela 19, por passe.

**Tabela 19** - Valores de SENS e SENSN, por passe, de cada variável independente encontrados por meio da aplicação da Modelagem.

PASSE	SENS			SENSN		
	$T_e$	$r$	$v$	$T_e$	$r$	$v$
1	-111,8	65,4	7,6	-61	35	4
2	-137,0	79,5	9,1	-61	35	4
3	-100,7	59,2	6,8	-60	36	4
4	-81,5	45,3	5,4	-62	34	4
5	-62,8	31,5	4,1	-64	32	4
6	-49,2	22,5	3,2	-66	30	4

FONTE: Autor

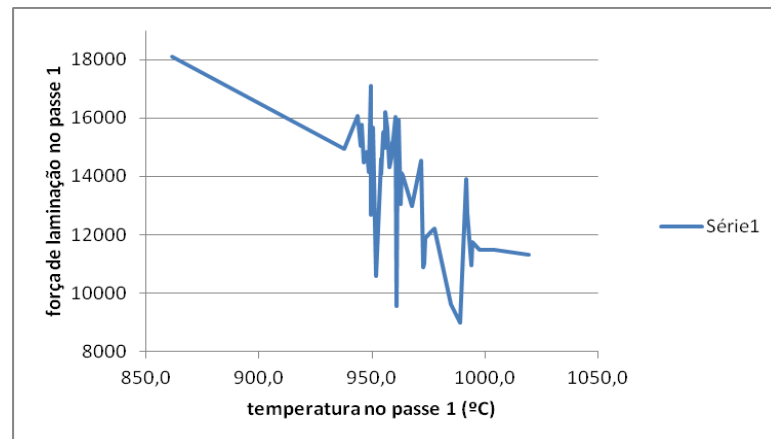
Observa-se que, pela SENSN da Tabela 19, a variável independente mais significativa é a " $T_e$ ", seguida pela " $r$ " e, finalmente, a menos representativa a " $v$ ", em todos os passes da laminação. Pode, também, ser verificada a relação da força de laminação com as variáveis de entrada: o aumento de temperatura no passe reduz a força de laminação, fato indicado pelo sinal negativo (-) das SENS e SENSN; e, de forma contrária, o aumento do grau de redução da espessura do material ou o aumento da velocidade periférica do cilindro provocam aumentos na força de laminação, indicados pelo sinal positivo (+) das SENS e SENSN.

### 6.3.3 Técnica de Sensibilização via RNA.

#### 6.3.3.1 Conhecimentos Preliminares.

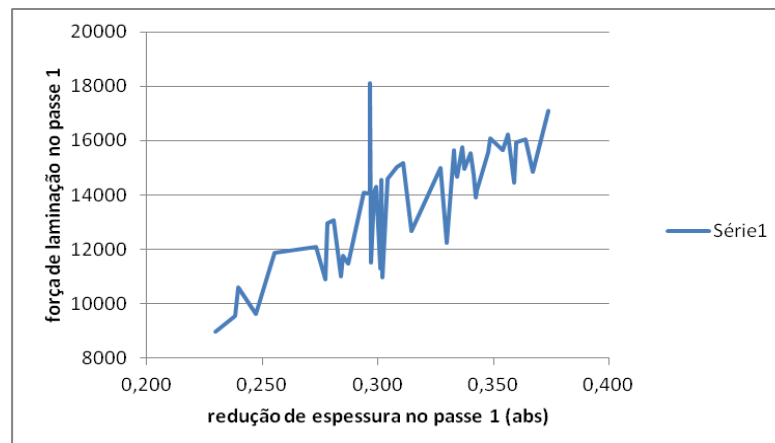
A relação da variável dependente com cada variável independente foi obtida por meio de um gráfico de dispersão, estabelecendo a expectativa dos sinais aritméticos de cada sensibilidade. Como referência é apresentada esta correlação para o 1º passe da laminação, que pode ser visto nas Figuras 20, para a relação da saída com a temperatura; Figura 21, relação da saída com o grau de redução; e Figura 22, relação da saída com a velocidade.

**Figura 20** - Sentido da relação entre a força de laminação e a temperatura no passe 1



Fonte: Autor

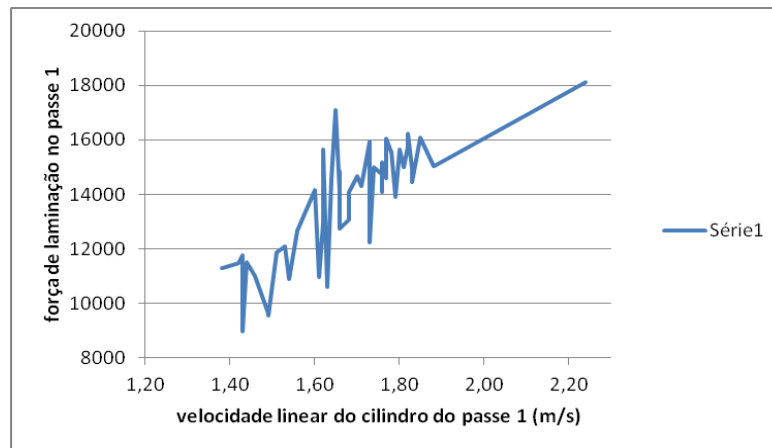
**Figura 21** - Sentido da relação entre a força de laminação e a redução no passe 1



Fonte: Autor



**Figura 22** - Sentido da relação entre a força de laminação e a velocidade no passe 1



Fonte: Autor

Pelas Figuras 20, 21 e 22 pode-se verificar que a variação da força tem sentido contrário à variação da temperatura; e mesmo sentido de variação que o grau de redução e a velocidade do cilindro do passe em análise. Outras dispersões foram verificadas nos outros passes e apresentaram comportamentos idênticos a estes do passe 1, conforme afirmativa de Schlang et al. (2001), comentada na seção 4.3, que processos similares são executados a cada passe. Espera-se, então, obter valores de Sensibilidade, para todos os passes:

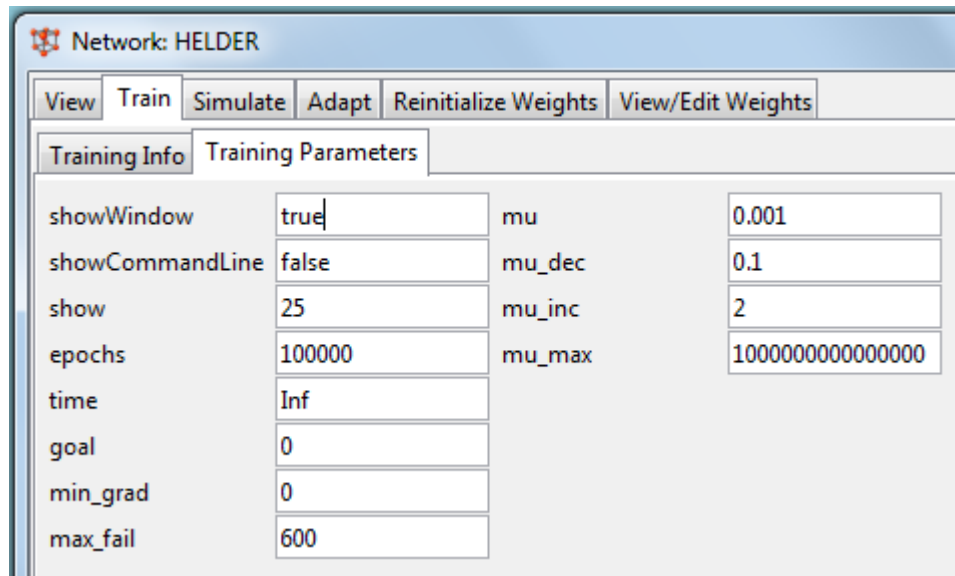
- a) negativos para a variável temperatura ( $SENS_{T_e} < 0$ );
- b) positivos para o grau de redução ( $SENS_r > 0$ ) e
- c) positivos para a velocidade ( $SENS_v > 0$ ).

Esta é uma característica que deve ser atendida pela generalização da RNA, ou seja, RNAs que não atenderem a este conhecimento preliminar foram descartadas por não terem boa capacidade de generalização.

### 6.3.3.2 Parâmetros de treinamento.

A título de informação a Figura 23 apresenta os parâmetros de treinamento utilizados para todas as RNAs nesta fase de mapeamento.

**Figura 23** - Parâmetros de treinamento para todas RNAs utilizadas nesta fase de mapeamento.



Fonte: MATLAB®.

Os dados da Figura 23, já comentados na seção 4.1.6.1, são:

- a) o valor inicial de " $\mu$ " ( $mu=0,001$ );
- b) o fator de decréscimo de " $\mu$ " ( $mu\_dec=0,1$ );
- c) o fator de acréscimo de " $\mu$ " ( $mu\_inc=2$ );
- d) o valor máximo permitido para o parâmetro " $\mu$ " ( $mu\_max=10^{15}$ );
- e) o valor de performance do gradiente, derivada da curva do EQM, a ser atingido ( $min\_grad=0$ );
- f) o número máximo de épocas a ser repetido o valor do desempenho da RNA, pela validação do treinamento ( $max\_fail=600$ );
- g) o valor do EQM (performance da RNA) solicitado a ser atingido no treinamento da RNA ( $goal=0$ );
- h) o tempo máximo para efetivar o treinamento, em segundos ( $time=infinito$ );
- i) o número máximo de épocas a serem apresentadas no treinamento ( $epochs=100000$ );
- j) *showWindow* (*true*): habilita mostrar a janela de treinamento quando está treinando uma RNA;
- k) *showCommandLine* (*show*): mostra a linha de comando que estabelece o número de épocas durante o treinamento;

- l) *show* (25): estabelece a periodicidade de indicação do número de épocas na janela de treinamento ( a cada 25 épocas).

Os valores default do software precisaram ser alterados devido à complexidade do sistema a ser mapeado.

### 6.3.3.3 Número de Neurônios na camada intermediária e número de Padrões.

As RNAs a serem tratadas tem 3 entradas e 1 saída.

Pelo método de Kolmogorov, expressão (7) da seção 4.1.6.1, o número de neurônios recomendado é 7: foram ensaiadas diversas RNAs e nenhuma apresentou validação (convergiu) com 7 neurônios, o que indica ser um número insuficiente de neurônios para a complexidade do processo.

O método de Blum, expressão (8) da seção 4.1.6.1, estabelece que o número máximo de neurônios depende do número de padrões. Utilizando os padrões do 2º passe, vários ensaios foram feitos para diferentes números de padrões e de diferentes números de neurônios de forma a orientar o estabelecimento destes.

**Figura 24** - Tipos de convergência das RNAs a partir da relação entre número de padrões x número de neurônios.

Nº PA- DRÕES	NÚMERO DE NEURÔNIOS												
	7	8	9	10	11	12	13	14	15	16	17	18	19
462	.	.	.	.	.	*	***	***	***	***	***	***	***
504	.	.	.	.	*	*	*	***	***	***	***	***	***
546	.	.	.	*	*	*	*	*	***	***	***	***	***
588	.	.	.	*	*	*	*	*	*	***	***	***	***
630	.	.	.	.	*	*	*	*	*	*	***	***	***
672	.	.	.	.	*	*	*	*	*	*	*	***	***
714	.	.	.	.	*	*	*	*	*	*	*	*	***
	<b>LEGENDA</b>												
	.	não converge											
	*	converge e atende ao método de Blum											
	***	converge e não atende ao método de Blum											

FONTE: Autor

A Figura 24 apresenta as relações de convergência (boa performance) conseguidas a partir do número de padrões x número de neurônios, onde estão evidenciadas as relações que não convergiram, ou convergiram atendendo ao método de Blum, ou convergiram não atendendo ao método de Blum. As quantidades de padrões foram selecionadas em múltiplo de 42, que é o número de padrões por lote disponível; e limitadas para evitar *overfitting*.

Conforme pode ser visto pela Figura 24, para que a RNA tenha uma boa faixa de possibilidade operacional de número de neurônios e estabilidade com o número de padrões, a utilização de 672 padrões é a mais adequada. O número de neurônios fica a ser obtido:

- a) de forma empírica;
- b) a partir de diversas topologias;
- c) que apresentem boa validação no treinamento;
- d) de boa performance; e
- e) capacidade de generalização;

de cada RNA, esperando que esteja entre 11 e 17 neurônios, conforme faixa para 672 padrões na Figura 24.

#### 6.3.3.4 RNAs mapeando o processo de LTQ do aço C-Mn.

Nesta fase do projeto obteve-se 6 RNAs, uma para cada passe de laminação, com arquitetura do tipo *feedforward*, com camada de entrada, camada intermediária e camada de saída. O número de neurônios da camada intermediária de cada RNA foi obtido de forma empírica selecionado entre de 11 a 17, conforme seção 6.3.3.3. A função de ativação usada nos neurônios da camada intermediária foi a Tangente Hiperbólica, enquanto que a da camada de saída foi a Linear. Todos os neurônios da RNA são ligados a um limiar de ativação (bias). O algoritmo de *Levenberg-Marquardt* foi usado no treinamento. Os 672 padrões fornecidos foram distribuídos em 75% para treinamento, 15% para validação e 15% para teste, escolhidos de forma aleatória, conforme método de validação cruzada (*cross-validation*). Esta divisão é feita de forma automática pela ferramenta "nntool" do software MATLAB® utilizado. O treinamento da RNA terminava quando atendia às solicitações feitas nos parâmetros de treinamento, da seção 6.3.3.2, com conseqüente validação.

Para a pesquisa empírica do número de neurônios da camada intermediária, para cada passe foram treinadas 7 RNAs, com variação no número de neurônios de 11 a 17 nesta camada, 5 vezes cada uma. Foram obtidas as SENSs para cada variável independente, em cada passe, para cada topologia de RNA. As RNAs que não atendiam aos conhecimentos preliminares obtidos na seção 6.3.3.1, sinais aritméticos das Sensibilidades, foram descartadas, pois ficava caracterizada a falta de generalização da respectiva RNA. Então, as RNAs aprovadas, uma por passe, apresentaram:

- a) boa performance com EQM entre as ordens de  $10^{-23}$  e  $10^{-25}$ ; e
- b) número de neurônios da camada intermediária que foram capazes de proporcionar boa generalização às RNAs.

Este número de neurônios da camada intermediária de cada RNA aprovada, por passe, está apresentado na Tabela 20.

**Tabela 20** - Número de neurônios das camadas intermediárias das RNAs aprovadas, por passe.

<b>PASSE</b>	<b>Nº NEURÔNIOS</b>
1	14
2	12
3	16
4	16
5	14
6	14

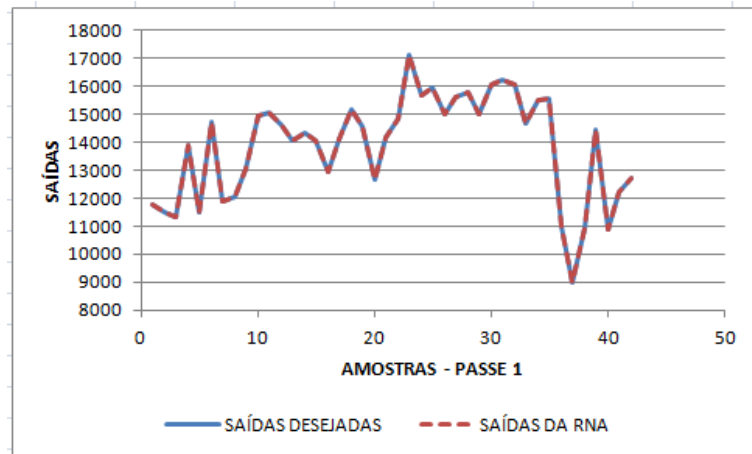
FONTE: Autor

Pode ser verificado que, conforme expectativa segundo Figura 24, o número de neurônios da camada intermediária de cada RNA aprovada atende ao cálculo de Blum.

As saídas "y" fornecidas pela RNAs treinadas e aprovadas, para cada passe, e os respectivos valores desejados (originais) "d" estão registrados de forma comparativa na Figura 25 (passe 1); Figura 26 (passe 2); Figura 27, (passe 3); Figura 28 (passe 4); Figura 29 (passe 5); e Figura 30 (passe 6). Estas Figuras 25, 26, 27, 28, 29 e 30 mostram a boa performance das RNAs treinadas, por fornecerem saídas "y"

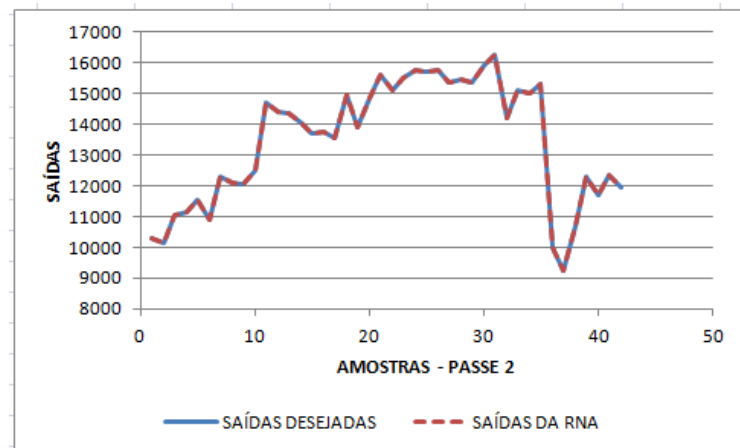
coincidentes com as saídas "d" originais do processo. Este fato evidencia que uma RNA bem dimensionada, quanto à sua arquitetura e topologia, pode mapear satisfatoriamente um processo. Esta é a característica de capacidade de memorização da RNA a partir do treinamento efetivado.

**Figura 25** - Comparação gráfica entre os valores de força de laminação: saídas desejadas "d" e as saídas fornecidas pela RNA "y", mapeando o Passe 1



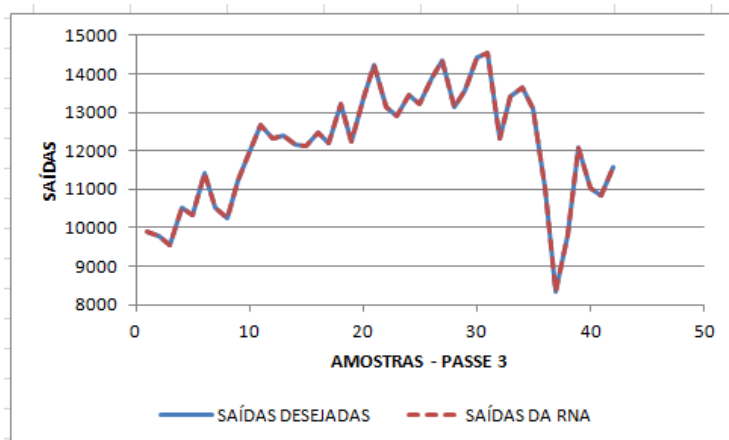
Fonte: Autor

**Figura 26** - Comparação gráfica entre os valores de força de laminação: saídas desejadas "d" e as saídas fornecidas pela RNA "y", mapeando o Passe 2



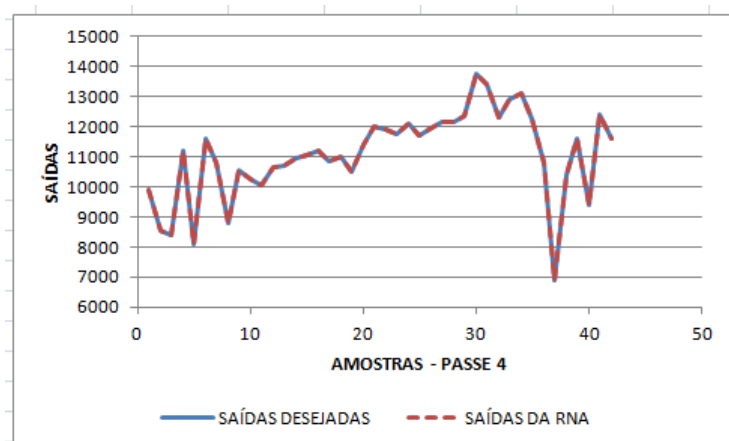
Fonte: Autor

**Figura 27** - Comparação gráfica entre os valores de força de laminação: saídas desejadas "d" e as saídas fornecidas pela RNA "y", mapeando o Passe 3



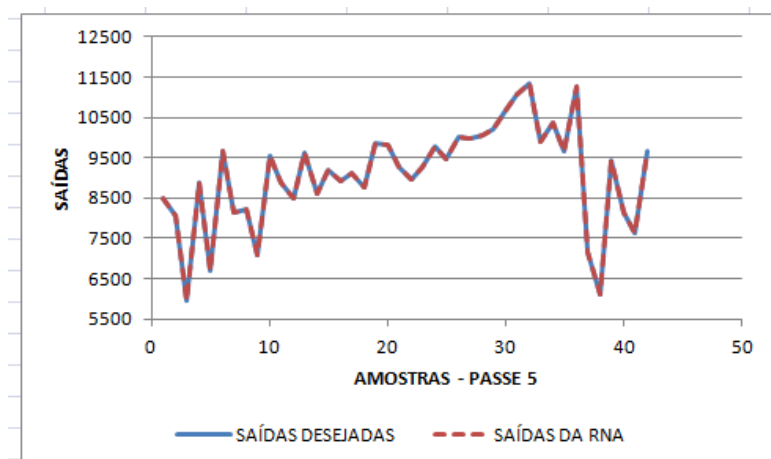
Fonte: Autor

**Figura 28** - Comparação gráfica entre os valores de força de laminação: saídas desejadas "d" e as saídas fornecidas pela RNA "y", mapeando o Passe 4



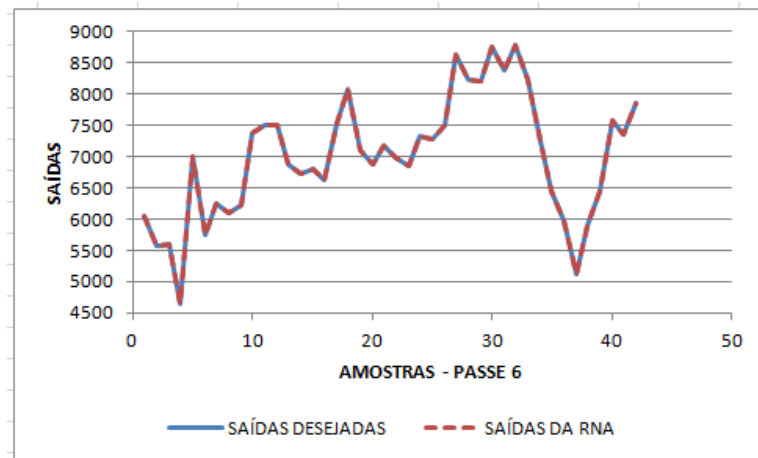
Fonte: Autor

**Figura 29** - Comparação gráfica entre os valores de força de laminação: saídas desejadas "d" e as saídas fornecidas pela RNA "y", mapeando o Passe 5.



Fonte: Autor

**Figura 30** - Comparação gráfica entre os valores de força de laminação: saídas desejadas "d" e as saídas fornecidas pela RNA "y", mapeando o Passe 6



Fonte: Autor

#### 6.3.3.5. Técnica de Sensibilidade aplicada ao processo de LTQ via RNA.

A partir da RNA treinada, que identifica o sistema, uma a cada vez, a variável independente " $x_i$ " escolhida foi:

- acrescida em 5% nos seus valores que substituem os valores anteriores usados no treinamento da rede;
- este novo conjunto padrão de entrada foi então simulado na RNA já treinada;
- foi obtida a diferença entre os novos valores de saída " $y$ " obtidos na simulação e os valores de saída desejados " $d$ " ( $= y - d$ );
- determinou-se a média dessas diferenças obtendo-se a  $SENS_i$  da variável independente " $x_i$ " em questão, conforme expressão (10), da seção 4.2.1;
- a  $SENSN$  foi obtida conforme expressão (11) da seção 4.2.2.

A Tabela 21 apresenta os valores, para cada passe da  $SENS$  e da  $SENSN$  das variáveis independentes do processo, onde: " $T_e$ " é a temperatura de entrada do aço no passe, em  $^{\circ}C$ ; " $r$ " é o grau de redução na espessura do aço no passe, adimensional; e " $v$ " é a velocidade periférica do cilindro no passe, em m/s.

Observa-se que, pela  $SENSN$  da Tabela 21, a variável independente mais significativa é a " $T_e$ ", seguida pela " $r$ " e, finalmente a menos representativa a " $v$ ", em todos os passes da laminação.



**Tabela 21** - Valores de Sensibilidades Médias e Sensibilidades Médias Normalizadas para todos os passes do processo de LTQ de Aço C-Mn, via RNA.

PASSE	SENS			SENSN (%)		
	Te	r	v	Te	r	v
1	-1403	334	59	-78	19	3
2	-2362	392	91	-83	14	3
3	-724	324	118	-62	28	10
4	-1651	308	75	-81	15	4
5	-1354	531	88	-69	27	4
6	-279	141	47	-60	30	10

Fonte: Autor

Nessa Tabela 21 também fica evidente a relação de variação da força de laminação com as variáveis de entrada:

- a) as variações na força de laminação são inversamente proporcionais às variações da temperatura do material laminado no passe, verificado pelo sinal negativo (-) das SENS e SENSN, o que indica que acréscimos nesta temperatura provocam reduções na força de laminação;
- b) as variações na força de laminação são diretamente proporcionais às variações do grau de redução da espessura do material laminado e da velocidade periférica do cilindro de laminação no passe, verificado pelo sinal positivo (+) das SENS e SENSN, o que indica que acréscimos nestas variáveis provocam acréscimos na força de laminação; como era esperado pelos conhecimentos preliminares obtidos pelos estudos de correlação na seção 6.3.3.1.

#### 6.3.4 Comparação entre as duas formas de aplicação da técnica de Sensibilidade: Modelagem x RNA.

A Tabela 22 confronta os valores da SENSN das variáveis independentes, em análise, do processo de LTQ, obtidas via Modelo e por meio de RNAs.

Os resultados também podem ser confrontados por meio das Figuras 31 e 32, que apresentam os gráficos dos valores das SENSNs das Variáveis Independentes por passe, obtidas por meio do Modelo e via RNA, respectivamente.

Pode-se observar, por meio da Tabela 22 ou das Figuras 31 e 32, que todas as duas formas de aplicação da técnica de Sensibilização apresentaram a mesma sequência de significância das variáveis independentes com relação à Força de Laminação:

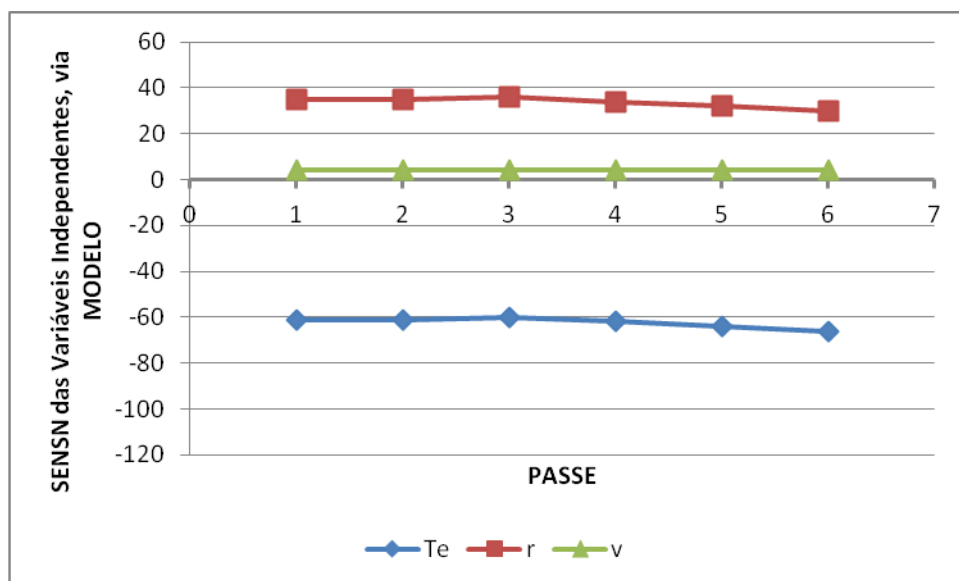
- a) a variável independente mais significativa é a "Te";
  - b) seguida pela "r" e, finalmente
  - c) a menos representativa a "v",
- em todos os passes da laminação.

**Tabela 22** - Valores de Sensibilidade Média Normalizada para as variáveis independentes obtidas via Modelo e por meio das RNAs.

PASSE	SENSN (%)					
	MODELO			RNA		
	Te	r	v	Te	r	v
1	-61	35	4	-78	19	3
2	-61	35	4	-83	14	3
3	-60	36	4	-62	28	10
4	-62	34	4	-81	15	4
5	-64	32	4	-69	27	4
6	-66	30	4	-60	30	10

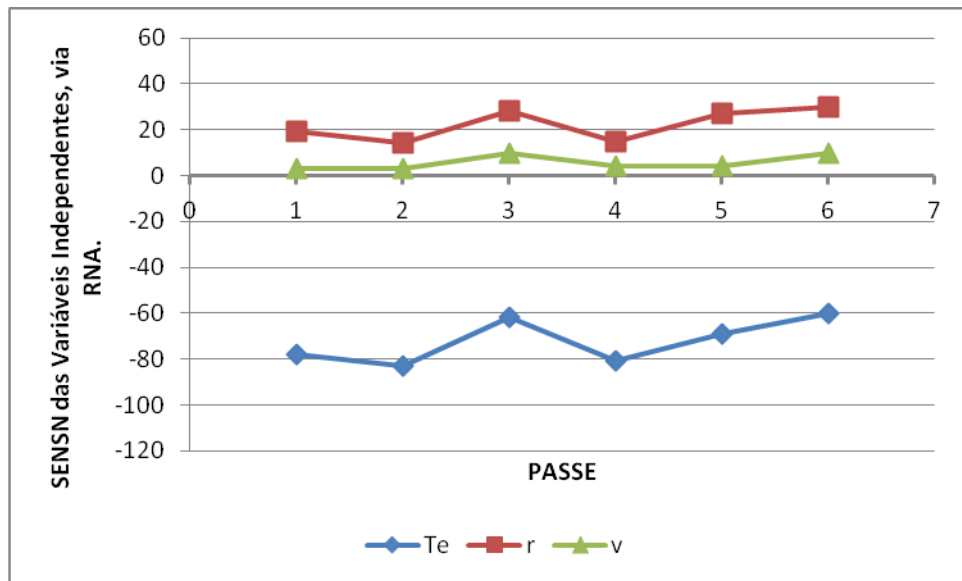
FONTE: Autor.

**Figura 31** - SENSN das Variáveis Independentes, por passe, via Modelo.



FONTE: Autor

**Figura 32** - SENSN das Variáveis Independentes, por passe, via RNA.



FONTE: Autor

A significância obtida por meio do Modelo vem confirmar o resultado idêntico da significância das variáveis independentes obtida via RNA. Este fato determina que a técnica da Sensibilidade aplicada a um processo mapeado por meio de RNAs proporciona resultados confiáveis e de processamento mais rápido sem tanta complexidade no conhecimento/domínio do processo em análise como é exigido pela modelagem caixa branca.

## 7 CONCLUSÕES.

As funções de ativação não lineares do tipo Tangente Hiperbólica e Logística podem ser utilizadas na camada intermediária e a Linear na camada de saída apresentando boa performance para mapeamento de funções não lineares via RNA.

As RNAs, de arquitetura PMC, representam boas ferramentas para mapeamento de funções pois conseguem armazenar, via algoritmo e de forma oculta, uma identificação satisfatória de um modelo, preservando sua generalização, podendo ser simulado a qualquer momento.

A técnica de Sensibilidade da variável independente é representativa da significância desta variável no processo. Esta Sensibilidade mostra tanto a estimativa da variável independente mais representativa, quanto a estimativa da de menor significância. A variável de Sensibilidade de maior valor indica a que tem maior representação na variável dependente em estudo e que justificaria, por exemplo, investimentos em seu controle. Qualquer variável independente do processo pode ser inserida na análise de Sensibilidade, mesmo que a mesma não influencie na saída em estudo.

A técnica de Sensibilidade aplicada a processos metalúrgicos identifica as significâncias das variáveis independentes proporcionando maior facilidade da compreensão das relações entrada-saída dos processos sem grandes conhecimentos dos fenômenos metodológicos dos mesmos.

As RNAs mostraram ser uma ótima ferramenta para identificar os parâmetros mais influentes e os parâmetros de menor efeito sobre a força de laminação, que pode impactar consideravelmente nas diretrizes de investimentos e projetos; bem como na orientação operacional, proporcionando redução de custos e melhoria da eficiência do processo.

## 8 SUGESTÕES PARA TRABALHOS FUTUROS.

Outros assuntos complementares podem ser explorados e desenvolvidos na área do mapeamento do processo de laminação:

- a) como a temperatura é o componente mais relevante dentre as entradas independentes físicas, pode-se fazer um estudo mais aprofundado do comportamento da força de laminação a partir da análise/modelo térmico do processo de LTQ de aço C-Mn, por meio de RNA, especificando mais detalhadamente como ocorrem as relações da força de laminação x comportamento térmico;
- b) neste estudo não foi considerada a composição química do aço como variável independente do processo. Pode-se desenvolver uma análise do comportamento da força de laminação a partir da composição química de aços microligados, buscando a estimativa das significâncias dos elementos químicos nesta força.

## REFERÊNCIAS

- AGUIRRE, L. A. **Introdução à identificação de sistemas: técnicas lineares e não lineares aplicadas a sistemas reais**. 3. ed. Belo Horizonte: Editora UFMG, 2007.
- AZIMZADEGAN, T.; KHOEINI, M. ETAAT, M.; ALIREZA, K. An artificial neural network model for impact properties in X70 pipeline steels. **Neural computing & application**, n. 23, p. 1473-1480, 2013.
- BAGHERIPOOR, M.; BISADI, H. Application of artificial neural network for the prediction of roll force and roll torque in hot strip rolling process. **Applied mathematical modellin**, n. 37, p. 4593-4607, 2013.
- BRAGA, A. de P.; CARVALHO, A. C. P. de L. F. de ; LUDERMIR, T. B. **Fundamentos de redes neurais artificiais**. Rio de Janeiro: DCC/IM, COOPE/Sistemas, NCE/UFRJ, 1998.
- BRAGA, A. de P.; CARVALHO, A. C. P. de L. F. de ; LUDERMIR, T. B. **Redes neurais artificiais: teoria e aplicações**. 2. ed. Rio de Janeiro: LTC, 2007.
- BLUM, A., **Neural networks in C++**. New York: John Wiley & Sons Inc., 1992.
- CAMPOS, M. M. de; SAITO, K. **Sistemas inteligentes em controle e automação de processos**. Rio de Janeiro: Editora Ciência Moderna Ltda., 2004.
- CAMPOS, M. C. M. M. de; TEIXEIRA, H. C. G. **Controles típicos de equipamentos e processos industriais**. São Paulo: Editora Blucher, 2006.
- CYBENKO, G. Approximation by superpositions of a sigmoidal function. **Mathematics of control, signals, and systems**, v. 2, n. 4, p. 303-314, 1989.
- DEGHANI, K.; NEKAHI, A. Artificial neural network to predict the effect of thermomechanical treatments on bake hardenability of low carbon steels. **Materials and design**, n. 31,p. 2224-2229, 2010.
- GARCIA, C. **Modelagem e simulação de processos industriais e de sistemas eletromecânicos**. 2. ed. São Paulo: Editora da Universidade de São Paulo, 2005.
- GOMES, G. S. da S.; LUDERMIR, T. B. Redes neurais artificiais com funções de ativação complemento log-log e probit para aproximar funções na presença de observações extremas. **Learning and nonlinear models, Revista da sociedade brasileira de redes neurais (SBRN)**, v. 6, n. 2, p.142-153, 2008.
- GORNI, A. A.; SILVA, M. R. S. da; SILVEIRA, J. H. D. da. Previsão de propriedades mecânicas de bobinas laminadas a quente através de redes neurais. **Tecnologia em metalurgia, materiais e mineração**, São Paulo, v. 5, n. 4, p. 229-234, 2009.
- GRESSLER, L. A. **Introdução à pesquisa: projetos e relatórios**. 2.ed. São Paulo: Loyola, 2004.

HASANZADEHSHOOIILI, H.; LAKIROUHANI, A.; MEDZVIECKAS, J. Superiority artificial of neural networks over statistical methods in prediction of the optimal length of rock bolts. **Journal of civil engineering and management**, v.18. n. 5, p.655-661, 2012.

HAYKIN, S. **Redes neurais: princípios e práticas**. 2. ed. Porto Alegre: Bookman, 2001.

HECHT-NIELSEN, R. **Neurocomputing**. New York: Addison-Wesley Publishing Company, 1991.

JONES, D. M.; WATTSON, J.; BROWN, K. J. Comparison of hot rolled steel mechanical property prediction models using linear multiple regression, non linear multiple regression and non linear artificial neural network. **Ironmaking & steelmaking**, v. 32, n. 5, p. 435-442, 2005.

LEE, D. M.; CHOI, S.G. Application of on-line adaptable neural network for the rolling force setup of a plate mill. **Engineering applications of artificial intelligence**, n. 17, p. 557-565, 2004.

LEE, D. e LEE, Y. Application of neural network for improving accuracy of roll force model in hot rolling mill. **Control engineering practice**, n. 10, p.473-478, 2002.

LIVINGSTONE, D. **A practical guide to scientific data analysis**. Reino Unido: Wiley, 2009.

MACHADO, M. L. P. **Modelo termomicroestrutural para determinação de variáveis operacionais e metalúrgicas para utilização em simulação e/ou controle na laminação a quente de aços carbono e microligados**. 2005. 380f. Tese (Doutorado em Engenharia Elétrica) - Programa de Pós Graduação em Engenharia Elétrica, Universidade Federal do Espírito Santo, Vitória, 2005.

MOUSSAOUI, A.; SELAIMIA, Y.; ABBASSI, H. A. Hybrid hot strip rolling force prediction using a bayesian trained artificial neural network and analytical models. **American journal of applied sciences**, n. 3, p. 1885-1889, 2006.

MUKHERJEE, I.; ROUTROY, S. Comparing the performance of neural networks developed by using Levenberg-Marquardt and Quasi-Newton with the gradient descent algorithm for modelling a multiple response grinding process. **Expert systems with applications**, n. 39, p. 2397-2407, 2012.

NORGAARD, M.; RAVN, O.; POULSEN, N. K.; HANSEN, L. K. **Neural networks for modelling and control of dynamic systems: a practitioner's handbook**. Grã Bretanha: Springer-Verlag London Limited, 2000.

OLIVEIRA, A. P.; MODENESI, P. J. Modelo de previsão das propriedades mecânicas de perfis estruturais laminados a quente: uma abordagem em redes neurais artificiais. **45º Seminário de laminação, processo e produtos laminados e revestidos**, Pernambuco, ABM, 2008.

PIULEAC, C. G.; RODRIGO, M. A.; CAÑIZARES, P.; CURTEANU, S.; SÁEZ, C. Ten steps modeling of electrolysis processes by using neural networks. **Environmental Modelling & Software**, n. 25, p. 74-81, 2010.

SANTOS, A. L. B.; JUNIOR, S. S. da C.; SILVA, M. B.; SOUTO, U. B. Modelagem via redes neurais para a estimativa da temperatura de corte em ferramentas de metal duro no processo de fresamento frontal. **RECIE**, Uberlândia, v. 15, n. 1/2, p. 83-89, 2006.

SCHLANG, M.; LANG, B.; POPPE, T.; RUNKLER, T.; WEINZIERL, K. Current and future development in neural computation in steel processing. **Control engineering practice**, n. 9, p. 975-986, 2001.

SILVA, I. N. da; SPATTI, D. H.; FLAUZINO, R. A. **Redes neurais artificiais para engenharia e ciências aplicadas**. São Paulo: Artliber, 2010.

SON, J. S.; LEE, D. M.; KIM, I. S.; CHOI, S. K. A study on genetic algorithm to select architecture of a optimal neural network in the hot rolling process. **Journal of materials processing technology**, n. 153, p. 643-648, 2004.

\_\_\_\_\_. A study on on-line learning neural network for prediction for rolling force in hot rolling mill. **Journal of materials processing technology**, n. 164, p. 1612-1617, 2005.

TUBINO, D. F. **Manual de planejamento e controle da produção**. 2. ed. São Paulo: Atlas, 2000.

YANG, Y. Y.; LINKENS, D. A.; TALAMANTES-SILVA, J. Roll load prediction: data collection, analysis and neural network modelling. **Journal of materials processing technology**, n. 152, p. 304-315, 2004.

YAP, W. K.; KARRI, V. Emissions predictive modelling by investigating various neural network models. **Expert systems with applications**, n. 39, p. 2421-2426, 2012.