

UNIVERSIDADE DE SÃO PAULO
ESCOLA DE ENGENHARIA DE SÃO CARLOS
ÁREA DE PÓS-GRADUAÇÃO EM ENGENHARIA MECÂNICA

Serviço de Pós-Graduação EESC/USP

EXEMPLAR REVISADO

Data de entrada no Serviço... 14 / 06 / 05

Ass.: *lesan*

IDENTIFICAÇÃO DO FUNCIONAL DA RESPOSTA
AEROELÁSTICA VIA REDES NEURAIIS ARTIFICIAIS

Ana Paula Carvalho da Silva Ferreira

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

Orientador: Prof. Dr. Flávio Donizeti Marques

São Carlos

2005



DEDALUS - Acervo - EESC



31100052198

FOLHA DE JULGAMENTO

Candidato: Engenheira **ANA PAULA CARVALHO DA SILVA FERREIRA**

Dissertação defendida e julgada em 23-03-2005 perante a Comissão Julgadora:




Prof. Dr. **FLÁVIO DONIZETTI MARQUES (Orientador)**
(Escola de Engenharia de São Carlos/USP)

APROVADO




Prof. Dr. **ANDRÉ CARLOS PONCE DE LEON FERREIRA**
(Instituto de Ciências Matemáticas e de Computação/USP)

Aprovado



Prof. Dr. **ROBERTO GIL ANNES DA SILVA**
(Centro Técnico Aeroespacial/CTA)

Aprovada



Prof. Associado **JONAS DE CARVALHO**
Coordenador do Programa de Pós-Graduação em Engenharia
Mecânica



Profa. Titular **MARIA DO CARMO CALIJURI**
Presidente da Comissão de Pós-Graduação

Agradecimentos

Ao Prof. Flávio Donizeti Marques pela oportunidade de realizar a pesquisa, pela orientação e por estar sempre disposto a ajudar.

Aos colegas do LADINC por me receberem de braços abertos e por estarem sempre prontos a colaborar.

A todos os professores com os quais tive oportunidade de aprender e a todos os funcionários que, de forma direta ou indireta, contribuíram para a realização da pesquisa.

Ao CNPq pelo apoio financeiro à pesquisa.

Ao meu esposo pelo apoio e incentivo.

Finalmente à Deus, quem acredito ser o grande responsável por todas as minhas conquistas.

CONTEÚDO

Resumo	p. v
Abstract	p. vi
Lista de Símbolos	p. vii
1 Introdução	p. 1
1.1 Introdução.	p. 1
1.2 Modelagem da resposta aeroelástica.	p. 4
1.3 Modelos de Ordem Reduzida.	p. 8
1.4 Redes neurais artificiais.	p. 11
1.5 Objetivos do trabalho.	p. 14
1.6 Organização da Dissertação.	p. 15
2 Redes Neurais Artificiais	p. 16
2.1 Introdução.	p. 16
2.2 Breve histórico sobre redes neurais artificiais.	p. 16
2.3 Modelo matemático de um neurônio.	p. 18
2.4 Arquiteturas de redes neurais artificiais.	p. 21

	iii
2.5	Treinamento de redes neurais artificiais. p. 23
2.5.1	Algoritmo de retropropagação do erro. p. 25
2.5.2	Algoritmo de Levenberg-Marquardt. p. 31
3	Identificação do funcional da resposta aeroelástica. p. 39
3.1	Introdução. p. 39
3.2	Aparato experimental. p. 39
3.3	Rede neural para representar o funcional aeroelástico. p. 45
3.3.1	Análise dos parâmetros de treinamento da rede neural. p. 45
3.3.2	Valores iniciais de pesos e bias p. 46
3.3.3	Arquiteturas para a rede. p. 48
3.3.4	Número de atrasos usados no vetor de entrada. p. 49
3.3.5	Tamanho das janelas temporais. p. 50
3.3.6	Parâmetro μ p. 52
3.3.7	Parâmetro μ_{inc} p. 53
3.3.8	Parâmetro μ_{dec} p. 55
3.4	Identificação da resposta aeroelástica linear. p. 57
3.4.1	Treinamento. p. 57
3.4.2	Generalização. p. 61
3.5	Identificação da resposta aeroelástica não linear. p. 65
3.5.1	Treinamento. p. 65
3.5.2	Generalização. p. 69

4 Conclusões e sugestões para trabalhos futuros.

iv
p. 80

Referências

p. 82

RESUMO

FERREIRA, A. P. C. S. (2005). *Identificação do Funcional da Resposta Aeroelástica Via Redes Neurais Artificiais*. São Carlos, 2005. 97p. Dissertação (Mestrado) - Escola de Engenharia de São Carlos, Universidade de São Paulo.

Identificação e predição do comportamento aeroelástico representa um grande desafio para a análise e controle de fenômenos aeroelásticos adversos. A modelagem aeroelástica requer informações tanto sobre a dinâmica estrutural quanto sobre o comportamento aerodinâmico não estacionário. No entanto, a maioria das metodologias disponíveis atualmente são baseadas no desacoplamento entre o modelo estrutural e o modelo aerodinâmico não estacionário. Conseqüentemente, métodos alternativos são bem vindos na área de pesquisa aerolástica. Entre os métodos alternativos está o funcional multicamada, que fornece uma rigorosa representação matemática apropriada para modelagem aeroelástica e pode ser obtido através de redes neurais artificiais. Esse trabalho apresenta uma aplicação desse método, consistindo de um procedimento de identificação baseado em redes neurais artificiais que representam o funcional da resposta aeroelástica. O modelo neural foi treinado usando o algoritmo de Levenberg-Marquardt, o qual tem sido considerado um método de otimização muito eficiente. Ele combina a garantia de convergência do método do gradiente e o alto desempenho do método de Newton, sem a necessidade de calcular as derivadas de segunda ordem. Um modelo de asa ensaiado em túnel de vento foi usado para fornecer a resposta aeroelástica. A asa foi fixada a uma mesa giratória e um motor elétrico lhe fornecia o movimento de incidência. Essa representação aeroelástica funcional foi testada para diversas condições operacionais do túnel de vento. Os resultados mostraram que o uso de redes neurais na identificação da resposta aeroelástica é um método alternativo promissor, o qual permite uma rápida avaliação da resposta aerolástica do modelo.

Palavras chave: Aeroelasticidade, Funcionais multicamadas, Redes neurais, Algoritmo de Levenberg-Marquardt.

ABSTRACT

FERREIRA, A.P.C.S. (2005). *Identification of the Functional Aeroelastic Response by Artificial Neural Networks*. São Carlos, 2005. 97p. Dissertação (Mestrado) - Escola de Engenharia de São Carlos, Universidade de São Paulo.

Identification and prediction of aeroelastic behavior presents a significant challenge for the analysis and control of adverse aeroelastic phenomena. Aeroelastic modeling requires information from both structural dynamics and unsteady aerodynamic behavior. However, the majority of methodologies available today are based on the decoupling of structural model from the unsteady aerodynamic model. Therefore, alternative methods are mostly welcome in the aeroelastic research field. Among the alternative methods there is the multi-layer functional (MLF), that allows a rigorous mathematical framework appropriate for aeroelastic modeling and can be realized by means of artificial neural networks. This work presents an identification procedure based on artificial neural networks to represent the motion-induced aeroelastic response functional. The neural network model has been trained using the Levenberg-Marquardt algorithm that has been considered a very efficient optimization method. It combines the guaranteed convergence of steepest descent and the higher performance of the Newton's method, without the necessity of second derivatives calculation. A wind tunnel aeroelastic wing model has been used to provide motion-induced aeroelastic responses. The wing has been fixed to a turntable, and an electrical motor provides the incidence motion to the wing. This aeroelastic functional representation is then tested for a range of the wind tunnel model operational boundaries. The results showed that the use of neural networks in the aeroelastic response identification is a promising alternative method, which allows fast evaluation of aeroelastic response model.

Keywords: Aeroelasticity, Multi-layer functionals, Neural networks, Levenberg-Marquardt algorithm.

LISTA DE SÍMBOLOS

A matriz hessiana;

a_p atrasos na entrada;

a_y atrasos na saída;

ab distância da origem de coordenadas ao eixo elástico do perfil aerodinâmico;

b_i bias;

b vetor de bias;

CG centro de gravidade;

e vetor de erro;

$e.a.$ eixo elástico;

$F(x)$ índice de desempenho;

FM funcional multicamada;

f função de ativação;

g gradiente;

H função quadrática genérica;

I matriz identidade;

I_α momento de inércia;

J matriz jacobiana;

k_w constante de rigidez para deslocamento vertical;

k_α constante de rigidez para deslocamento torsional;

L força de sustentação do perfil aerodinâmico;

L_i funcional linear;

M momento de arfagem do perfil aerodinâmico;

m massa;

n potencial de ativação;

P pressão aerodinâmica;

p vetor de entradas para os neurônios;

S sensibilidades da rede neural;

t vetor da saídas desejadas para a rede neural;

U velocidade do escoamento;

u vetor de entradas;

v matriz que engloba o erro de todos os neurônios da última camada para todos os pares de treinamento da rede neural;

W matriz de pesos sinápticos;

$w_{i,j}$ peso sináptico;

$x_\alpha b$ distância do eixo elástico ao centro de gravidade do perfil aerodinâmico;

y vetor de saídas da rede neural;

z autovetor;

α ângulo de ataque;

ζ, θ constantes reais;

λ autovalor;

μ parâmetro ajustável do algoritmo de Levenberg-Marquardt;

φ função contínua;

ρ densidade;

ϱ taxa de aprendizagem.

Subscritos e superescritos:

h, l elemento da matriz jacobiana;

i, j neurônios da rede neural;

M número total de camadas da rede neural;

m camada da rede neural em análise;

o, p elementos da matriz hessiana;

p dimensão do espaço;

Q número total de pares de treinamento para a rede neural;

q par de treinamento em análise;

R número total de neurônios de determinada camada;

T transposto.

ACRÔNIMOS

DFC Dinâmica de Fluidos Computacional;

MOR Modelo de Ordem Reduzida;

NARX Nonlinear Autoregressive Model With Exogenous Variables.

1 INTRODUÇÃO

1.1 Introdução.

Aeroelasticidade trata do estudo da interação mútua entre as forças aerodinâmicas, elásticas e inerciais que atuam em um corpo em movimento e imerso em um escoamento (DOWELL, 1978). Na Engenharia Aeronáutica, os fenômenos provocados por essa interação podem prejudicar o desempenho em vôo de uma aeronave ou até mesmo causar falhas catastróficas na estrutura de um avião. Sendo assim, torna-se evidente a importância de se estudar fenômenos aeroelásticos tais como *flutter*, *buffeting*, divergência, resposta dinâmica, etc.

Resposta aeroelástica é a resposta dos componentes estruturais de uma aeronave, produzida por cargas aerodinâmicas não estacionárias e subsequente interação mútua com a dinâmica estrutural (BISPLINGHOFF; ASHLEY; HALFMAN, 1996).

Os experimentos em aeroelasticidade podem ser feitos com aeronaves em escala real ou reduzida, e também com componentes tal como uma asa isolada. Os ensaios em escala real representam um grande investimento de tempo e dinheiro, o que os tornam inviáveis para aqueles testes onde há riscos de danos. Os experimentos com modelos em escala reduzida são mais usados, principalmente os realizados em túneis de vento. No entanto, alguns problemas comuns nesses testes podem fazer com que o modelo não seja completamente fiel ao sistema físico real. Entre esses problemas estão a interferência da parede do túnel ou do suporte do modelo nos resultados e a necessidade da aplicação de uma força vertical no modelo para que seu peso e massa sejam escalados corretamente, o que nem sempre é possível, pois a massa do aparelho que aplica a força pode ser maior que a massa do modelo (BISPLINGHOFF; ASHLEY; HALFMAN, 1996).

Muitas vezes o uso de modelos matemáticos, ao invés de experimentos, representa uma economia de tempo e recursos financeiros. Para a definição de um modelo aeroelástico adequado é necessário o conhecimento das características estruturais e aerodinâmicas do sistema a ser estudado. Através da análise dessas características, é possível definir se o sistema é estacionário ou não estacionário e linear ou não linear.

Um escoamento estacionário é aquele no qual os parâmetros do escoamento, por exemplo, velocidade e pressão, podem variar de um ponto para outro, mas são constantes em relação ao tempo em qualquer ponto. Isso significa que medidas dos parâmetros do escoamento tomadas em um dado ponto em vários intervalos de tempo permanecem as mesmas (HOUGHTON; CARRUTHERS, 1982). Sendo assim, conclui-se que não há interação mútua entre fluido e estrutura, pois o movimento da estrutura não altera os parâmetros do escoamento e vice-versa. Nesse caso as equações de movimento e de mecânica de fluidos, que modelam respectivamente a estrutura e o escoamento, podem ser estudadas de maneira isolada, ou seja, não é necessário trocar informações entre as equações. Além disso, algumas hipóteses simplificadoras podem ser adotadas, tais como fluido irrotacional, incompressível e não viscoso.

Um escoamento não estacionário é aquele em que os parâmetros do escoamento variam com o tempo em qualquer ponto. Pode haver interação entre fluido e estrutura e, dessa forma, não é permitido estudar o problema de maneira desacoplada. Esta consideração torna o problema mais próximo da realidade física, mas requer uma solução mais complexa, pois além da necessidade de troca de informações entre as equações de movimento e de mecânica de fluidos, a dependência temporal não pode mais ser eliminada das equações.

Quanto à questão de um sistema aeroelástico ser tratado como linear ou não linear pode-se dizer que devido ao avanço tecnológico das aeronaves modernas, os efeitos não lineares não devem mais ser desconsiderados na modelagem aerodinâmica e estrutural. A aproximação não linear do problema aeroelástico pode fornecer informações importantes, tais como: a influência das não linearidades na resposta aeroelástica subcrítica e a classificação das instabilidades aeroelásticas em benignas ou catastróficas. Em outras palavras, tal aproximação fornece a possibilidade de

determinar, por exemplo, em quais condições a velocidade de *flutter* pode ser excedida sem a ocorrência de uma falha catastrófica da asa (MARZOCCA; LIBRESCU; SILVA, 2002).

Os sistemas aeroelásticos possuem fontes de não linearidades estruturais e aerodinâmicas típicas. Entre as fontes físicas de não linearidades em estruturas elásticas estão: (i) folgas que surgem a partir de componentes estruturais conectados frouxamente; (ii) não linearidade geométrica, a qual compreende a rigidez não linear que tem origem nos grandes gradientes de deslocamento; e (iii) atrito seco ou amortecimento não linear que surge a partir de componentes estruturais em contato deslizante (DOWELL; TANG, 2002). Marques (1997) e Dowell e Tang (2002) mostram que as fontes físicas de não linearidade no escoamento aerodinâmico são o escoamento separando-se, que pode ser induzido por uma onda de choque ou por um alto ângulo de ataque, e os efeitos de compressibilidade que levam ao surgimento e movimento de ondas de choque.

Pelos fatos descritos nessa Seção e também segundo Marques (1997), para aplicações aeroelásticas o modelo aero-estrutural ideal e, talvez, mais geral seria aquele baseado em soluções das equações de mecânica dos fluidos, que considerassem escoamento não estacionário, compressível e viscoso, simultaneamente com a solução das equações de movimento. Os estados instantâneos que fossem gerados pelas equações de mecânica dos fluidos e de movimento seriam permutados e a solução simultânea global produziria tanto a resposta aerodinâmica, quanto o histórico de movimento estrutural, o qual depende de condições iniciais fornecidas. Tal modelo encontra-se representado na Figura 1. No entanto, para aplicações práticas a solução das equações gerais de mecânica dos fluidos pode ser obtida somente por meio de técnicas numéricas ou métodos de dinâmica de fluidos computacional que normalmente demandam um esforço computacional significativo (MARQUES, 1997). Isso motiva a aplicação dos chamados Modelos de Ordem Reduzida (MOR).

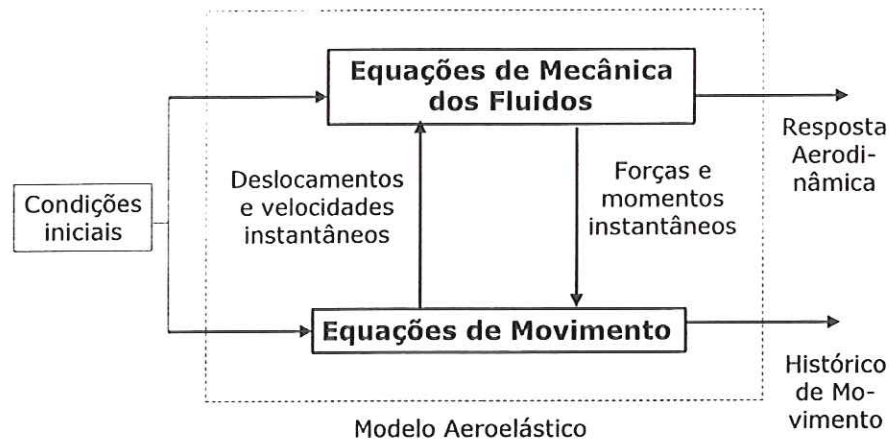


Figura 1: Modelo aero-estrutural ideal (MARQUES, 1997).

Um Modelo de Ordem Reduzida (SILVA, 1999) é um modelo matemático desenvolvido para ser uma representação matemática simplificada de um sistema dinâmico. Seu tamanho e complexidade reduzidos permitem a utilização de algoritmos computacionais mais eficientes com um custo computacional menor (MARQUES, 2002). Essa classe de modelos engloba, por exemplo, os métodos semi-empíricos, os métodos baseados na teoria de funcionais e as redes neurais artificiais.

1.2 Modelagem da resposta aeroelástica.

Um sistema físico útil para o estudo de muitos problemas aeroelásticos é o chamado de seção típica, o qual é um sistema aeroelástico bidimensional que consiste de um perfil aerodinâmico apoiado por uma suspensão elástica em flexão e torção. Figura 2 representa esse sistema.

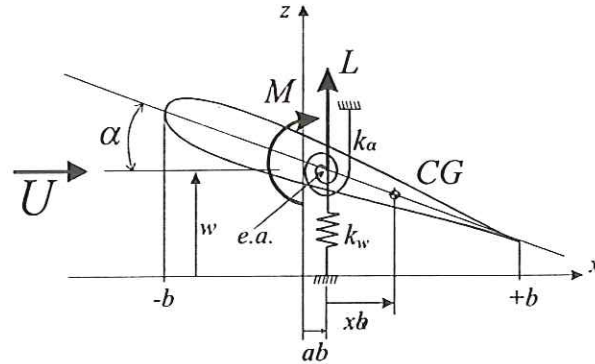


Figura 2: Sistema aeroelástico bidimensional (Seção típica).

Nesse modelo x e z são as coordenadas de deslocamento horizontal e vertical, respectivamente. O ângulo de rotação do perfil aerodinâmico é dado por $\alpha(t)$ e corresponde ao ângulo de ataque. O deslocamento na vertical é dado por $w(t)$. O ponto onde a suspensão elástica é fixada ao aerofólio é chamado eixo elástico ($e.a.$). A distância da origem do sistema de coordenadas ao eixo elástico é dada por ab ; e a distância do eixo elástico ao centro de gravidade (CG) do perfil aerodinâmico é dada por $x_\alpha b$. A velocidade do escoamento é dada por U , enquanto que o momento de arfagem e a força de sustentação são dados, respectivamente, por M e L . As constantes de rigidez para deslocamento vertical e torsional são dadas, respectivamente, por k_w e k_α .

As equações de movimento para a seção típica são (BISPLINGHOFF; ASHLEY; HALFMAN, 1996):

$$\frac{d}{dt}(m\dot{w}(t) + S_\alpha\dot{\alpha}) + k_w w(t) + L(t) = 0 \quad (1.1a)$$

$$\frac{d}{dt}(S_\alpha\dot{w}(t) + I_\alpha\dot{\alpha}(t)) + k_\alpha(t) - M(t) = 0 \quad (1.1b)$$

onde: $m \equiv \int \rho dx$ é a massa por unidade de comprimento, $S_\alpha \equiv \int \rho x dx$ é o momento estático, $I_\alpha \equiv \int \rho x^2 dx$ é o momento de inércia, $L \equiv \int P dx$ e $M \equiv \int P x dx$ são os esforços de sustentação e momento de arfagem, respectivamente, P é a pressão aerodinâmica e ρ é a densidade. Um conjunto de coordenadas generalizadas é um conjunto de coordenadas de deslocamento linearmente independentes as quais são suficientes para descrever uma configuração arbitrária de um sistema (CRAIG, 1981). Neste caso, as coordenadas generalizadas são dadas por $w(t)$ e $\alpha(t)$.

A maior dificuldade na solução das Equações (1.1a) e (1.1b) é encontrar os termos L e M , pois eles conduzem às soluções das equações de mecânica dos fluidos. As equações dinâmico-estruturais modelam a estrutura, enquanto que as equações de mecânica dos fluidos modelam a resposta aerodinâmica do sistema. Esses dois grupos de equações podem ser resolvidos de diversas maneiras e devem trocar informações a cada instante de tempo.

Os modelos dinâmico-estruturais podem ser obtidos através das leis de Newton ou das equações de Lagrange e podem ser contínuos ou discretos. Os modelos discretos geralmente possuem N graus de liberdade, que resultam em N equações diferenciais ordinárias escritas na forma matricial. A partir dessa forma matricial surge uma equação característica cuja solução fornece as frequências naturais e modos de vibrar do sistema dinâmico.

Os modelos dinâmico-estruturais contínuos resultam em equações diferenciais parciais. A aplicação da equação de Lagrange para esses sistemas conduz ao método dos modos assumidos, no qual define-se um conjunto de funções e cada uma irá representar uma forma de deslocamento da estrutura em questão. O método dos modos assumidos deu origem ao método dos elementos finitos (REDDY, 1993), no qual cada função representa uma forma de deflexão de uma porção (elemento finito) da estrutura. Esse método é amplamente utilizado em diversas áreas da engenharia, sendo uma versão mais robusta do método dos modos assumidos e contornando as desvantagens do mesmo (CRAIG, 1981).

A escolha do método de modelagem dinâmica-estrutural dependerá das características e do grau de complexidade da estrutura em questão. Segundo Dowell e Tang (2002), há casos em que para a análise aeroelástica (linear ou não linear) o modelo dinâmico-estrutural pode ser expresso de uma forma mais compacta em termos dos autovetores e de suas amplitudes modais. Uma vantagem dessa abordagem reside no fato que um modelo estrutural típico de elementos finitos pode ter entre 10^2 e 10^3 graus de liberdade, enquanto um modelo de autovetores típico de acurácia comparável terá entre 10 e 10^2 graus de liberdade.

A Dinâmica de Fluidos Computacional (DFC) apresenta métodos que buscam as soluções numéricas das equações de mecânica dos fluidos. A limitação desses métodos é o alto custo computacional requerido. Um modelo DFC típico em um escoamento tridimensional pode possuir entre 10^5 e 10^6 graus de liberdade, enquanto que um modelo modal aerodinâmico de ordem reduzida tem geralmente menos que uma centena (DOWELL; TANG, 2002). Por esse motivo, para a maioria dos computadores usados atualmente, os métodos DFC ainda representam um grande esforço.

Outros métodos usados na modelagem da resposta aerodinâmica são os Métodos dos Painéis (KATZ; PLOTKIN, 1991), os quais usam as singularidades aerodinâmicas e suas combinações para simular uma fronteira, ou seja, uma superfície de resistência ao escoamento. Para modelagem aerodinâmica de um corpo em um escoamento uniforme é necessário distribuir um número finito de painéis por todo o corpo, cada painel contém uma singularidade cuja intensidade é desconhecida. Posteriormente, usando a função potencial e algumas condições de contorno, chega-se a um sistema de equações cujas incógnitas são as intensidades das singularidades que podem ser obtidas por métodos numéricos convencionais. Essa distribuição de intensidade, quando apropriada, faz com que a superfície do corpo torne-se uma linha de corrente do fluxo simulando uma superfície de resistência ao escoamento (ANDERSON, 1991).

O método dos painéis só funciona bem para hipóteses de escoamento potencial e possui um alto custo computacional. Além disso, são necessários cuidados especiais na modelagem de regiões como junção da asa da aeronave com a fuselagem, ponto de fixação de um flape, pontos onde há entrada de ar, por exemplo, devido a limitações impostas pelas hipóteses assumidas. O método dos painéis assume como hipótese que o escoamento é potencial. Comparando com métodos DFC, o seu custo computacional é baixo, porém pode ser proibitivo de acordo com a complexidade de discretização do modelo.

Uma forma computacionalmente mais econômica de se obter modelos aeroelásticos é o uso de modelos aproximados, ou modelos de ordem reduzida (MOR). Como já se mencionou anteriormente, os modelos de ordem reduzida incluem os métodos semi-empíricos, os modelos baseados nas formas funcionais e as redes neurais artificiais.

1.3 Modelos de Ordem Reduzida.

Os modelos de ordem reduzida podem ser subdivididos nos métodos semi-empíricos, nos modelos baseados nas formas funcionais, nas redes neurais artificiais, etc. Os Métodos Semi-Empíricos, ou modelos fenomenológicos, usam informações provenientes de cálculos matemáticos combinadas com dados de experimentos para confeccionar modelos de sistemas. Métodos semi-empíricos possuem a vantagem de serem rápidos computacionalmente, no entanto, precisam de dados experimentais precisos e específicos (MARQUES, 1997). Os métodos semi-empíricos são bastante utilizados na modelagem da resposta aerodinâmica não estacionária e os trabalhos de Tran e Petot (1981) e Mahajan, Kaza e Dowell (1993) são exemplos dessas aplicações.

Um funcional é uma generalização da noção de uma função, onde o domínio é um conjunto de funções, ao invés de ser um conjunto de pontos. Um dos pioneiros a usar o conceito de funcional para modelos aerodinâmicos foi Tobak (TOBAK; PEARSON, 1964). Seu trabalho propôs a obtenção de um funcional da resposta aerodinâmica não estacionária não linear através da reformulação das funções indiciais como funcionais da história de movimento e da aplicação de uma generalização do princípio de superposição. As equações integrais resultantes são bastante complexas e seu uso prático somente é permitido por simplificações.

Uma representação funcional de grande importância na identificação de sistemas dinâmicos é a série de Volterra. Volterra (1959) mostrou que expansões da definição da série de Taylor para uma função podem ser generalizadas para funcionais. A série funcional resultante é a denominada série de Volterra, que é um somatório infinito de integrais de convolução multidimensionais de ordem crescente.

A identificação de sistemas dinâmicos não lineares baseada na representação de Volterra requer que os núcleos da série de Volterra sejam medidos. O primeiro núcleo de Volterra representa a resposta linear de um sistema a uma entrada impulso unitário, enquanto que os núcleos de ordens mais elevadas são as respostas não lineares do sistema a múltiplas entradas impulso unitário. Os núcleos de ordens mais elevadas são medidas da não linearidade, ou da influência relativa

da entrada anterior na resposta atual, que caracteriza o efeito temporal para um sistema não linear. A identificação de sistemas dinâmicos não lineares baseada nas séries de Volterra requer a determinação dos núcleos de ordens mais elevadas. Esta necessidade é a maior desvantagem de usar essa representação, pois a determinação dos núcleos de Volterra para ordem acima de dois não é uma tarefa trivial. Algumas aproximações contornam esse problema simplesmente assumindo que o sistema possui uma não linearidade fraca. Exemplos de aplicações em Aeroelasticidade desse método estão nos trabalhos de Silva et al. (2001), Silva (1999) e Marzocca, Librescu e Silva (2002).

Relacionadas com as expansões do tipo Volterra de séries funcionais, outras metodologias foram desenvolvidas (BILLINGS, 1980). Wiener (1958) estendeu o uso de somatórios de polinômios ortogonais (técnica padrão em aproximação de funções) para análise funcional através de uma modificação da série de Volterra para fornecer uma expansão composta de funcionais ortogonais. Essa extensão produziu a série de Wiener na qual todo funcional é ortogonal em relação aos outros quando o sinal de entrada é um ruído branco Gaussiano (WRAY; GREEN, 1994).

O método de Wiener contorna os dois principais problemas relacionados com a aproximação da série de Taylor: (i) permite aproximação de funções não analíticas, o que não é possível com a série de Taylor, pois para calcular os coeficientes da série a função a ser aproximada precisa ser analítica; (ii) soluciona o problema de truncamento pois o somatório de polinômios até grau n fornece a melhor aproximação de grau n , o que não acontece com a série de Taylor, onde o truncamento para grau n em geral não fornece a melhor aproximação de grau n (WRAY; GREEN, 1994).

Apesar dessas vantagens, a série de Wiener apresenta consideráveis limitações. Um dos problemas mais óbvios é que o ruído branco ideal não pode ser gerado porque tem potência infinita, o que significa que outro tipo de ruído tem que ser usado em situações práticas (MARMARELIS; MARMARELIS, 1978). Esse desvio do ideal conduz a erro nos núcleos. Além disso, o número excessivo de coeficientes requerido para identificar a série funcional, mesmo para sistemas não lineares de baixa ordem, torna essa técnica pouco prática e difícil de aplicar.

Entre os trabalhos sobre as séries de Wiener e Volterra estão os livros de Rugh (1981) e Schetzen (1980) que tratam da aplicação das séries de Volterra e de Wiener na representação de sistemas não lineares. Entre as aplicações estão o trabalho de Silva (1993) que apresenta uma metodologia para o uso da teoria de Wiener-Volterra na análise e projeto aeroservoelástico; Hunter e Korenberg (1986) que revêem esquemas de identificação de sistemas que podem ser representados por um subsistema dinâmico linear precedido ou seguido por uma não linearidade estática; Schetzen (1981) que discute os conceitos básicos da teoria de Wiener e apresenta vários métodos de modelagem pelos quais um sistema não linear pode ser modelado. Além desses, pode-se citar também os artigos de Palm e Poggio (1977), Palm e Poppel (1985) que tratam da validade e problemas da representação de Volterra e Wiener.

Uma desvantagem comum às Séries de Volterra e Métodos de Wiener é que a aproximação funcional pode ser obtida de forma razoável somente para modelos de uma entrada e uma saída. Formas funcionais multivariáveis possuem uma complexidade crescente (MARQUES, 1997).

Outras formas de modelo de ordem reduzida de sistemas não lineares são as redes neurais artificiais. Estudos recentes mostraram que redes neurais artificiais (HAYKIN, 1999) possuem grande potencial para modelagem de sistemas não lineares, podendo identificar um sistema por si próprio, ou auxiliar outros tipos de aproximações. O uso de redes neurais artificiais na identificação de sistemas dinâmicos já é bastante difundido na literatura técnica, como é o caso de: Narendra e Parthasaraty (1990), Masri, Chassiakos e Caugley (1993), Chen et al. (1995), Jagannathan e Lewis (1996) e Sjoberg et al. (1995). A eficiência das redes neurais em tarefas de identificação de sistemas é explicada por algumas propriedades como: *i*) a capacidade de adaptação ou de aprendizado que as permite aprender a partir de dados de entrada-saída de sistemas através da atualização dos pesos; *ii*) o paralelismo que as torna mais rápidas computacionalmente para execução de algumas tarefas e permite que operações independentes sejam executadas simultaneamente; *iii*) robustez ou tolerância à falhas que faz com que o desempenho da rede continue satisfatório mesmo que alguma conexão tenha sido danificada.

Uma outra propriedade, e talvez a mais importante no contexto de identificação de sistemas, é que a rede neural pode ser linear ou não linear. Se os seus neurônios possuem função de ativação não linear, então a rede é não linear e pode, conseqüentemente, identificar sistemas não lineares. Como os principais desafios atuais estão na modelagem de sistemas não lineares, essa propriedade associada àquelas descritas anteriormente tornam as redes neurais uma valiosa ferramenta alternativa para modelagem.

1.4 Redes neurais artificiais.

Uma rede neural artificial é um conjunto de unidades processadoras que armazenam conhecimento obtido através de treinamento. Para o treinamento da rede neural utiliza-se um algoritmo de aprendizagem que modifica os seus parâmetros para atingir um objetivo. Desta forma o funcionamento de uma rede neural é semelhante ao de um cérebro animal em dois aspectos: *i*) o conhecimento é adquirido pela rede através de um processo de aprendizagem; *ii*) as conexões estabelecidas pelos pesos sinápticos são usadas para armazenar o conhecimento (HAYKIN, 1999).

Dentre as vantagens do uso das redes neurais artificiais pode-se citar: *i*) relativa facilidade com a qual relações entrada-saída acopladas e não lineares podem ser calculadas; *ii*) não é necessário o conhecimento da forma do sistema de equações não lineares para determinar o modelo matemático do sistema físico; *iii*) são apropriadas para interpolação multidimensional, onde os dados não são estruturados, *iv*) os dados para obtenção do modelo podem vir de uma variedade de fontes incluindo experimentos e simulações numéricas (FALLER; SCHRECK, 1996).

Apesar das vantagens, as redes neurais possuem algumas limitações que devem ser observadas para que o processo de modelagem seja eficiente. Por exemplo, elas podem não permitir extrapolação fora do espaço de parâmetros definido experimentalmente. A definição das entradas e saídas desejadas da rede neural deve levar em consideração a física do problema para que não se obtenha resultados falsos ou insatisfatórios. Além disso, para que as redes neurais estáticas possam lidar com mapeamentos entrada-saída dinâmicos é necessário que valores passados das entradas e saídas também façam parte do vetor de entrada da rede. Redes desse tipo podem ser denominadas recorrentes (FALLER; SCHRECK, 1996).

As redes neurais têm sido utilizadas com sucesso não somente em engenharia, mas em diversas áreas como medicina, finanças, segurança, etc. As aplicações incluem: associação e reconhecimento de padrão, filtragem, identificação de sistemas (Werbos (1989); Narendra e Parthasaraty (1990); Chen et al. (1995); Pal et al. (1995); Silva e Khater (2001)), controle (Levin e Narendra (1996); Nguyen e Widrow (1990); Bhat e McAvoy (1990)), etc.

Na engenharia aeronáutica o uso de redes neurais já se encontra bastante difundido. Rauch et al. (1993) exploraram a detecção de defeitos, isolamento e reconfiguração através de redes neurais para o controle de superfícies de um avião de alto desempenho; Mead (1993) utilizou as redes neurais para a aproximação de dados experimentais. Faller e Schreck (1995), Faller, Schreck e Luttgies (1995), Marques e Anderson (1996), Marques e Anderson (2001), Marques (1997) e Marques (2002) aplicaram as redes neurais em situações envolvendo aerodinâmica não estacionária. Troudet, Sanjay e Merrill (1991), Rokhsaz e Steck (1991) e Rokhsaz, Steck e Shue (1994) desenvolveram sistemas de controle baseados em redes neurais. Rai e Madavan (1998), Rai e Madavan (2000) e Huang, Miller e Steck (1994) aplicaram as redes neurais em projetos aerodinâmicos. Além dessas referências, algumas outras são apresentadas nos trabalhos de Faller e Schreck (1996) e Rai e Madavan (2000), que fornecem uma revisão das aplicações dessa ferramenta matemática nos diversos ramos da pesquisa aeronáutica.

As redes neurais artificiais também podem ser usadas na representação de Funcionais Multicamadas. Os funcionais multicamadas são basicamente uma combinação não linear de funcionais lineares. O desenvolvimento dessa formulação teve seu início com os pesquisadores Chen e Chen (1993) e Modha e Hecht-Nielsen (1993) que, partindo do princípio de que a maioria dos sistemas é melhor representado por funcionais que por funções, estudaram generalizações do teorema de aproximação universal de forma a justificar o uso de representações funcionais na identificação de sistemas dinâmicos não lineares.

O teorema da aproximação universal (Cybenko (1989) e Hornik, Stinchcombe e White (1989)) fornece uma formalização e explicação matemática apropriada das propriedades de identificação das redes neurais para sistemas não lineares. Além disso, fornece uma base conceitual apropriada

para o uso de redes neurais convencionais como modelos paramétricos entrada-saída de sistemas dinâmicos não lineares (MARQUES, 1997).

No contexto da modelagem da resposta aerodinâmica não estacionária, o uso de funcionais multicamadas obtidos por redes neurais foi aplicado primeiramente por Marques e Anderson (1996) na modelagem e identificação de cargas aerodinâmicas não estacionárias e não lineares. A equação que define um funcional multicamada é dada por:

$$FM[\mathbf{u}_t] = \sum_{i=1}^k \zeta_i \varphi(\theta_i + L_i[\mathbf{u}_t]), \quad (1.2)$$

onde, FM é o funcional multicamada que representa a resposta do sistema, k é o número de unidade de processamento, ζ_i e θ_i são constantes reais, φ é uma função contínua monotônica crescente discriminada e não constante e $L_i[\mathbf{u}_t]$ denota funcionais lineares do vetor de entradas do sistema \mathbf{u}_t .

Diferentes classes de funcionais multicamadas podem ser obtidas pela especificação do funcional linear L_i na Equação (1.2). No caso de funcionais multicamadas obtidos por redes neurais multicamadas, L_i pode ser representado como uma superposição ponderada de funcionais afim em espaços p -dimensionais de valores reais, modificado por uma função não linear sigmoidal, ou seja:

$$FM[\mathbf{u}] = \sum_{i=1}^k \zeta_i \varphi(\theta_i + \sum_{j=1}^p w_{j,i} u_j), \quad (1.3)$$

onde, FM , k , ζ_i , θ_i , φ e \mathbf{u} são como definidos na Equação (1.2) e $w_{j,i}$ denota o valor do peso sináptico utilizado como parâmetro da rede neural.

A arquitetura da rede neural que realiza esse funcional multicamada é a rede multicamada convencional (com função de ativação sigmóide). A rede neural resultante é basicamente uma rede estática. Para representar um sistema dinâmico com tal modelo a arquitetura deve incluir atrasos. Quando se usa realimentação de saídas atrasadas obtém-se uma rede neural recorrente, que pode ser treinada usando algoritmos típicos para redes estáticas (MARQUES, 2002).

1.5 Objetivos do trabalho.

O objetivo do trabalho é realizar a identificação, através de redes neurais artificiais, da resposta aeroelástica de um modelo de asa obtido de ensaios experimentais.

O modelo de asa é ensaiado em túnel de vento. O ensaio consiste em fixar a asa a uma plataforma giratória que fornece movimento de incidência e expor à mesma um escoamento aerodinâmico em diferentes ângulos de ataque. Diferentes condições de ensaio são obtidas através da variação do ângulo de ataque e da velocidade do escoamento. As respostas aeroelásticas do modelo para diferentes condições de ensaio são obtidas por extensômetros distribuídos ao longo da superfície da semi-envergadura de asa. As não linearidades são introduzidas às respostas aeroelásticas através de ensaios com altos ângulos de incidência variando no tempo, o que leva ao típico fenômeno de estol dinâmico.

A rede neural utilizada representa um funcional multicamada. Pretende-se com essa representação otimizar o processo de identificação valendo-se das vantagens apresentadas pelas redes neurais em tarefas de identificação e da melhor representação que os funcionais fornecem de sistemas não lineares em relação as funções. Utiliza-se o algoritmo de Levenberg-Marquardt para o treinamento da rede neural, pois esse engloba a garantia de convergência do método de gradiente e a velocidade do método de Newton.

O processo de identificação consiste em apresentar à rede neural as mesmas entradas (excitações) fornecidas ao modelo de asa e comparar a saída de ambas. Os parâmetros da rede neural são atualizados até que sua saída se aproxime da resposta aeroelástica da asa. Após essa etapa, denominada treinamento, os resultados são validados através da generalização, que consiste em apresentar à rede neural entradas que não foram usadas no treinamento e observar se suas saídas se aproximam das saídas do modelo de asa.

1.6 Organização da Dissertação.

A Dissertação está dividida em quatro Capítulos. O Capítulo 1 apresenta a motivação para o uso de um modelo de ordem reduzida como as redes neurais artificiais. Mostra a vantagem de se usar uma rede neural que representa um funcional multicamada e a formulação para se obter essa representação. Apresenta também algumas características das redes neurais, sua aplicação em identificação de sistemas, seu uso em aeronáutica e explica as vantagens do treinamento com o algoritmo de Levenberg-Marquardt.

Capítulo 2 apresenta um breve histórico do desenvolvimento das redes neurais. Explica o modelo matemático de neurônio artificial e sua analogia com o neurônio biológico. Apresenta uma rápida definição de redes de camada única, multicamadas e recorrentes. Finalmente apresenta uma descrição detalhada do algoritmo de retropropagação do erro e do algoritmo de Levenberg-Marquardt.

Capítulo 3 trata dos resultados obtidos. Começa descrevendo o experimento realizado, especificando as características do modelo de asa para ensaios em túnel de vento, a forma de coleta dos dados e os equipamentos utilizados. Apresenta um estudo comparativo da influência da variação de parâmetros no desempenho em treinamento da rede neural. Apresenta os resultados da identificação de resposta aeroelástica linear e não linear para diferentes tipos de excitações.

As conclusões e sugestões para trabalhos futuros são apresentadas no Capítulo 4.

2 REDES NEURAIS ARTIFICIAIS

2.1 Introdução.

Na sua forma mais convencional a modelagem da resposta aeroelástica envolve a solução das equações de mecânica de fluidos e das equações de movimento, sendo que, para casos não estacionários, é necessário uma troca de informação entre essas equações. Nesse contexto, as redes neurais artificiais surgem como uma valiosa ferramenta matemática alternativa, pois ao lidarem somente com dados de entrada-saída do sistema, não requerem a solução dessas complexas equações e, desde que bem treinadas, são capazes de representar um modelo aeroelástico satisfatoriamente. Esse Capítulo inclui um breve histórico sobre as redes neurais artificiais. O modelo matemático de um neurônio artificial fazendo uma analogia com o neurônio biológico é descrito e uma rápida definição de redes neurais de camada única, multicamadas e recorrentes é apresentada. A técnica de treinamento supervisionado pelo algoritmo de Levenberg-Marquardt é descrita detalhadamente.

2.2 Breve histórico sobre redes neurais artificiais.

Os primeiros trabalhos em redes neurais ocorreram no final do século XIX e início do século XX e foram feitos por cientistas como Hermann von Helmholtz, Ernst Mach e Ivan Pavlov. Esses trabalhos enfatizavam teorias gerais de aprendizagem, visão, condicionamento, etc., e não incluíam um modelo matemático específico de operação de um neurônio. Em 1943, Warren McCulloch e Walter Pitts uniram seus conhecimentos em neurofisiologia e matemática desenvolvendo um modelo matemático de um neurônio e mostrando que redes artificiais de neurônios poderiam computar qualquer função lógica ou aritmética (HAGAN; DEMUTH; BEALE, 1996).

McCulloch e Pitts foram seguidos por Donald Hebb, que em 1949 publicou o livro *The Organization of Behavior* (HEBB, 1949). Esse livro apresentou a primeira regra de aprendizagem para modificação de parâmetros da rede neural.

A primeira aplicação prática de redes neurais artificiais foi proposta por Frank Rosenblatt em 1958, que construiu uma rede de perceptrons. O perceptron é a forma mais simples de uma rede neural usada para classificação de padrões linearmente separáveis, possui neurônios do tipo lógicos e conexões modificáveis (HAYKIN, 1999). Rosenblatt demonstrou que uma rede perceptron poderia ser treinada para classificar certos conjuntos de padrões. Rosenblatt propôs também um algoritmo para ajuste de pesos para o perceptron e provou sua convergência quando os padrões são linearmente separáveis (HAGAN; DEMUTH; BEALE, 1996).

Pouco tempo depois, em 1960, Widrow e Hoff (1960) introduziram a rede ADALINE (ADaptive LInear NEuron) e o algoritmo LMS (Least Mean Square). A rede perceptron e a ADALINE eram similares, sendo que a primeira usava como função de transferência a função degrau e a segunda usava a função linear. A regra de aprendizagem do perceptron era mais sensível ao ruído e o algoritmo LMS encontrou mais aplicações. No entanto, o ADALINE e o perceptron sofriam a mesma limitação, sendo capazes somente de resolver problemas linearmente separáveis (HAGAN; DEMUTH; BEALE, 1996).

Marvin Minsky e Seymour Papert publicaram um livro em 1969 onde demonstraram que as redes de Widrow e Rosenblatt sofriam limitações. O impacto desse livro, associado às limitações tecnológicas da época, fizeram com que as pesquisas em redes neurais artificiais reduzissem consideravelmente. Entre 1970 e 1980 poucos pesquisadores continuaram os estudos em redes neurais e, especificamente em engenharia, pode-se dizer que foi um período de completa estagnação (HAYKIN, 1999).

A partir de 1980 novos avanços foram sendo conquistados em diversos campos, novos conceitos e novas formas de projetar redes neurais foram introduzidos. Entre os novos conceitos, pode-se dizer que dois foram responsáveis pelo renascimento das redes neurais. O primeiro foi

desenvolvido por Hopfield em 1982 e consistiu do uso de uma função de energia para formular uma nova maneira de entender a computação desenvolvida por redes recorrentes. O segundo foi o algoritmo de retropropagação (*backpropagation*), o qual tornou-se o algoritmo mais popularmente usado no treinamento de redes de perceptron multicamadas. O algoritmo foi descoberto ao mesmo tempo por três autores: Rumelhart, Hinton e Williams (1986), Parker (1985) e LeCun (1985). Posteriormente, descobriu-se que o algoritmo já havia sido descrito em 1974 na tese de doutorado de Paul Werbos (WERBOS, 1974) pela Universidade de Harvard (HAYKIN, 1999).

Os avanços das pesquisas em redes neurais são inúmeros, abrangendo desde novas estruturas de redes (Kosmatopoulos et al. (1995)) até novos algoritmos de treinamento (Lou e Perez (1996); Gupta e Sinha (1999); Pal et al. (1995); Huang e Huang (1990)). As aplicações são difundidas em diversas áreas e incluem associação e reconhecimento de padrão, filtragem, identificação de sistemas (Chen et al. (1995); Pal et al. (1995); Narendra e Parthasaraty (1990)), controle (Levin e Narendra (1996); Nguyen e Widrow (1990); Bhat e McAvoy (1990)), etc.

Em engenharia aeronáutica as redes neurais têm sido aplicadas principalmente em estudos aerodinâmicos não lineares. Entre esses estudos estão os trabalhos de Marques e Anderson (1996), Marques e Anderson (2001) que tratam da modelagem, identificação e predição de cargas aerodinâmicas transônicas não estacionárias por redes neurais. Faller, Schreck e Luttgés (1995) estudam a identificação e controle através de redes neurais de fluxos tridimensionais separados e não estacionários. Rai e Madavan (2000) tratam do projeto aerodinâmico usando redes neurais. Ku e Hajela (1998) estudam um controlador neural para um sistema aeroelástico.

2.3 Modelo matemático de um neurônio.

O neurônio biológico é uma célula com os seguintes componentes: dendrito, corpo da célula e axônio. Uma ilustração de um neurônio biológico é apresentada na Figura 3. O dendrito recebe informação de outros neurônios, o corpo da célula processa a informação e o axônio transmite a informação de um neurônio para outro. As informações que chegam ao corpo da célula são na

realidade pulsos elétricos conhecidos como impulsos nervosos ou potenciais de ação. Os dendritos de um neurônio se comunicam com os axônios de outros neurônios através de sinapses. As sinapses são regiões eletroquimicamente ativas, onde ocorre a transferência de estímulo nervoso que resulta em uma alteração no potencial elétrico (KOVACS, 1996).

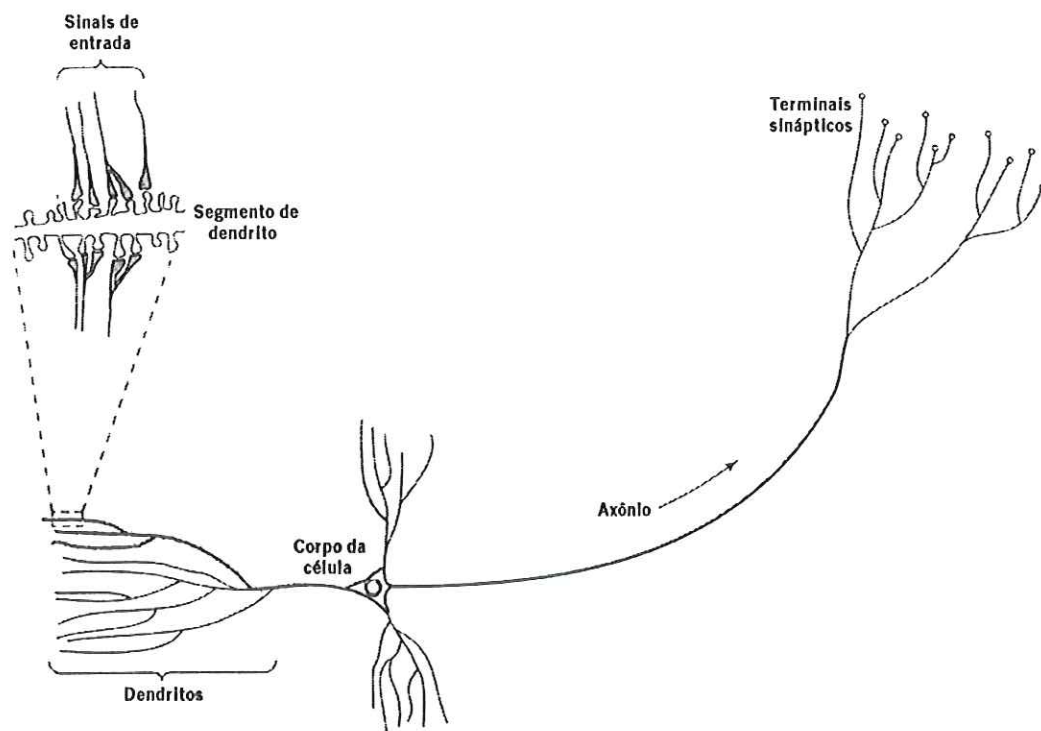


Figura 3: Neurônio biológico (HAYKIN, 1999).

Modelos matemáticos de neurônios artificiais foram inspirados no neurônio biológico. Um neurônio artificial típico é constituído de um conjunto de sinapses caracterizado por pesos $w_{i,j}$ que conectam as entradas p_j a um neurônio i . Essa conexão é feita multiplicando-se os pesos $w_{i,j}$ pelas entradas p_j . Um neurônio pode possuir também o chamado *bias* (b_i), o qual corresponde a um peso sináptico com entrada constante igual a 1. As entradas p_j multiplicadas pelos pesos

sinápticos $w_{i,j}$ são somadas à entrada 1 multiplicada pelo *bias* b_i . O valor obtido após o somatório é denominado potencial de ativação (n_i) e tem sua amplitude limitada por uma função de ativação ($f(n_i)$) (HAYKIN, 1999). A Figura 4 mostra um diagrama de blocos representativo de um neurônio artificial.

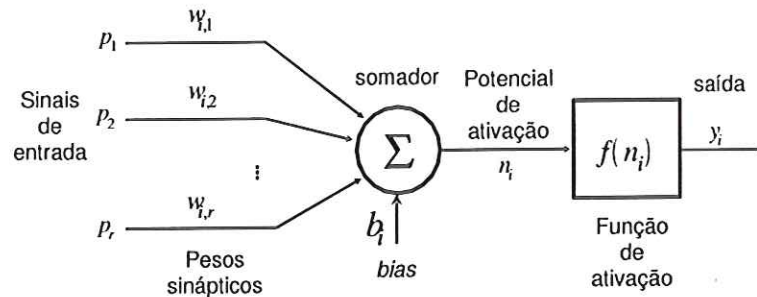


Figura 4: Diagrama de blocos de um neurônio artificial.

A equação matemática que descreve o modelo matemático do i -ésimo neurônio artificial em uma rede é a seguinte:

$$y_i = f(n_i), \quad (2.1)$$

onde

$$n_i = \sum_{j=1}^R w_{i,j} p_j + b_i, \quad (2.2)$$

onde, também, R é o número de elementos do vetor de entradas e y_i é a saída do neurônio.

Fazendo uma analogia com um neurônio biológico, pode-se dizer que os pesos $w_{i,j}$ correspondem à intensidade das sinapses, o corpo da célula é representado pelo somatório e pela função de ativação e o sinal transmitido pelo axônio corresponde à saída y_i (HAGAN; DEMUTH; BEALE, 1996).

Há vários tipos de função de ativação e sua escolha é baseada na tarefa que o neurônio irá executar. Entre as funções de ativação pode-se citar: degrau, rampa, e sigmóides. As funções lineares são geralmente usadas nos neurônios que compõem a última camada de uma rede neural e as funções sigmóides são as mais usadas nos neurônios das camadas intermediárias, mais comumente conhecidas como camadas ocultas.

2.4 Arquiteturas de redes neurais artificiais.

Uma rede neural é tipicamente formada por camadas de neurônios. Uma rede neural de camada única é aquela em que os neurônios que recebem as entradas são os mesmos que fornecem a saída. Essas redes foram muito usadas em problemas de classificação de padrões, sendo que o número de categorias em que um padrão poderia ser classificado era diretamente proporcional ao número de neurônios (HAGAN; DEMUTH; BEALE, 1996). No entanto, somente problemas linearmente separáveis poderiam ser resolvidos e isso motivou a busca de redes mais poderosas computacionalmente.

Uma rede neural multicamada é aquela que contém uma ou mais camadas de neurônios entre a entrada e a saída. Essas camadas são denominadas camadas ocultas ou intermediárias e a saída de uma camada oculta é a entrada para a camada seguinte. A Figura 5 apresenta uma ilustração de uma arquitetura típica de rede neural multicamada.

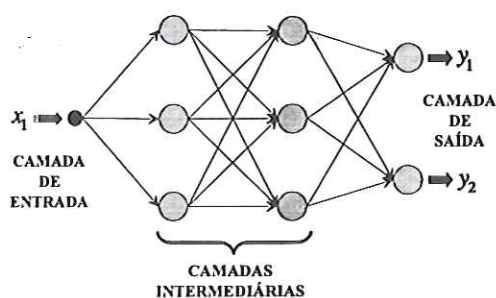


Figura 5: Arquitetura de uma rede multicamada com duas camadas ocultas ou intermediárias.

Cybenko (1989), baseado no Teorema do Aproximador Universal, demonstrou que uma rede com apenas uma camada oculta pode implementar qualquer função contínua não linear, desde que o número de neurônios na camada seja suficiente. Isso não significa que arquiteturas maiores não devam ser usadas. Há casos em que um número maior de camadas ocultas pode melhorar o desempenho da rede e sua capacidade de generalização.

As redes neurais multicamada são úteis em muitas aplicações, no entanto, algumas modificações têm sido propostas para melhorar sua aplicabilidade. Em particular, a incorporação de um elemento dinâmico é importante para a modelagem de sistemas dinâmicos. Uma aproximação

para introduzir dinâmica é adotar a filosofia dos modelos NARX e usar dados atrasados no tempo (LENNOX et al., 2001). Esses dados atrasados podem ser, tanto valores passados de entrada, como realimentações da saída. Os modelos NARX (do termo em inglês: *Nonlinear AutoRegressive model with eXogenous variables*) são modelos discretos no tempo que explicam o valor da saída $y(k)$ em função de valores prévios dos sinais de saída e de entrada, ou seja:

$$y(k) = f[y(k-1), \dots, y(k-a_y), p(k-1), \dots, p(k-a_p)], \quad (2.3)$$

onde a_y e a_p são os maiores atrasos em y e p , respectivamente (AGUIRRE, 2000).

Uma arquitetura de rede neural do tipo recorrente é aquela que possui realimentações. As realimentações podem ser globais (em toda a rede) ou locais (em um único neurônio) e podem ser dos neurônios da camada de saída para a camada de entrada ou dos neurônios da camada oculta para a camada de entrada.

A Figura 6 ilustra uma rede neural recorrente com atrasos na entrada e realimentações da saída como entradas da rede.

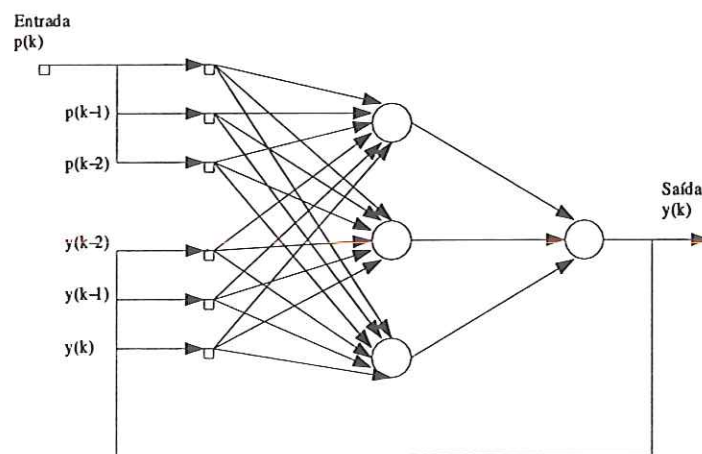


Figura 6: Rede neural recorrente.

Teoricamente, uma rede recorrente com realimentação global pode aprender a dinâmica de um ambiente não estacionário armazenando o conhecimento obtido a partir da amostra de treinamento em um conjunto fixo de pesos. Mais importante, a rede pode rastrear as variações estatísticas do ambiente se a amostra de treinamento for representativa do comportamento não estacionário do ambiente e se seu treinamento for adequado. Um treinamento pode ser considerado adequado quando ele não é insuficiente a ponto da rede ser incapaz de aproximar a saída do sistema quando a entrada aplicada é diferente da usada no treinamento, nem exagerado a ponto da rede aprender características que não pertencem ao sistema a ser modelado como, por exemplo, ruído (HAYKIN, 1999).

2.5 Treinamento de redes neurais artificiais.

A propriedade que torna as redes neurais uma ferramenta tão útil é a sua capacidade de aprender a partir de dados entradas-saídas de um sistema e melhorar seu desempenho através da aprendizagem. Esse processo de aprendizagem é feito através de iterações ou épocas de treinamento. Em cada iteração os pesos sinápticos e *bias* são ajustados até que se obtenha um desempenho desejado (HAYKIN, 1999). Existem várias regras de aprendizagem e elas se dividem basicamente na aprendizagem supervisionada e não supervisionada.

Na aprendizagem não supervisionada as saídas desejadas não são conhecidas e os pesos e *bias* são ajustados em função somente das entradas da rede. Uma vez que a rede se torna afinada com as regularidades estatísticas dos dados de entrada, ela desenvolve a habilidade de formar representações internas para codificar características de entrada e criar novas classes automaticamente. A maioria dos algoritmos que implementam esse tipo de aprendizagem desenvolvem operações de agrupamento e a rede aprende a categorizar os padrões de entrada em um número finito de classes (HAGAN; DEMUTH; BEALE, 1996).

Para desenvolver uma aprendizagem não supervisionada usa-se geralmente uma regra de aprendizagem competitiva. A rede neural para esse tipo de treinamento é constituída de duas camadas,

ou seja, uma camada que recebe entradas externas e uma camada competitiva. A camada competitiva é constituída de neurônios que competem entre si pela oportunidade de responder a características contidas nos dados de entrada. Nessa estratégia o neurônio com a maior entrada total vence a competição e é ativado, os demais neurônios são desativados (HAYKIN, 1999).

Na aprendizagem supervisionada, para cada entrada há uma saída desejada previamente conhecida. O processo de identificação de sistemas dinâmicos geralmente é feito por aprendizagem supervisionada. Nesse processo, como um primeiro passo, deve-se escolher aleatoriamente os parâmetros (pesos e *bias*) da rede neural e, em seguida, são aplicadas as mesmas entradas ao sistema dinâmico e à rede neural. Compara-se a saída de ambos, de onde se obtém o erro entre os respectivos valores de saída e a aprendizagem é feita através da correção do erro, o qual é uma função dos pesos e *bias* da rede. Em redes neurais multicamadas essa função pode ser vista como uma superfície de erro multidimensional contendo vários pontos de mínimo local e um ponto de mínimo global. O treinamento é feito por algum algoritmo que busque um ponto de mínimo da função erro através do ajuste dos pesos e *bias*, visando obter a parametrização que mais adequadamente representa a dinâmica do sistema (HAGAN; DEMUTH; BEALE, 1996). Por se tratar de um processo de minimização de função, os métodos de análise numérica e otimização podem ser aplicados. O processo de identificação de sistemas dinâmicos através de treinamento supervisionado está representado na Figura 7.

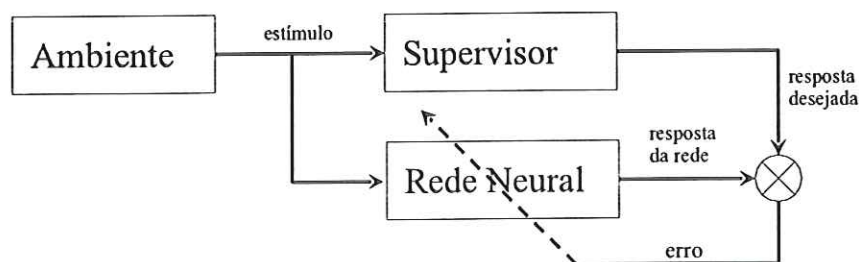


Figura 7: Representação esquemática de um treinamento supervisionado típico.

Entre as técnicas de otimização numérica aplicáveis para redes neurais, considera-se o método do gradiente e o método de Newton. O método de Newton é um método de segunda ordem baseado na hipótese de que a função a ser minimizada pode ser aproximada por uma função quadrática na região próxima a um ponto de mínimo. Uma desvantagem desse método é que ele não é confiável quando a busca pelo mínimo está ocorrendo fora da região onde a função é aproximadamente parabólica. Em contraste, o método do gradiente é apropriado para aproximar o mínimo partindo-se de um ponto fora da região onde a função é aproximadamente parabólica. Contudo, o método do gradiente não converge rapidamente quando se encontra na vizinhança imediata de um ponto de mínimo. Conseqüentemente, seria conveniente um algoritmo que se comportasse como o método do gradiente para a primeira parte da busca do mínimo e como o método de Newton depois que houvesse convergência para a região próxima ao mínimo. Essas características são encontradas no método de Levenberg-Marquardt, o qual foi sugerido por Levenberg (1944) e Marquardt (1963) e foi aplicado com sucesso para o treinamento de redes neurais por Hagan e Menhaj (1994).

As seções que se seguem descrevem o algoritmo de retropropagação do erro, o qual implementa o método do gradiente para uma rede neural multicamada. Posteriormente, será feita uma descrição do método de Levenberg-Marquardt voltada para o treinamento de redes neurais. Os métodos apresentados nesse trabalho são baseados nas descrições apresentadas em Hagan, Demuth e Beale (1996) e Haykin (1999).

2.5.1 Algoritmo de retropropagação do erro.

O algoritmo de retropropagação do erro tem sido usado como algoritmo padrão para o treinamento de redes neurais multicamadas. O nome retropropagação é proveniente do fato de que as derivadas parciais da função de custo (índice de desempenho) com respeito aos parâmetros livres (pesos sinápticos e *bias*) da rede são determinadas retropropagando-se os sinais de erro (computados pelos neurônios de saída) através da rede, camada por camada (HAYKIN, 1999). O algoritmo de retropropagação do erro foi primeiramente descrito na tese de doutorado de Paul Werbos em 1974 (WERBOS, 1974).

Em redes neurais multicamadas a saída de uma camada se torna a entrada para a camada seguinte. As equações que descrevem essa operação são:

$$\mathbf{y}^{m+1} = \mathbf{f}^{m+1}(\mathbf{W}^{m+1}\mathbf{y}^m + \mathbf{b}^{m+1}), \quad \text{para } m = 0, 1, 2, \dots, M-1, \quad (2.4)$$

onde, M é o número de camadas da rede neural, \mathbf{W}^{m+1} é a matriz de pesos sinápticos dos neurônios da camada $m + 1$, \mathbf{y}^m é o vetor das saídas dos neurônios da camada m , \mathbf{b}^{m+1} é o vetor de *bias* dos neurônios da camada $m + 1$, \mathbf{f}^{m+1} é o vetor das funções de ativação dos neurônios da camada $m + 1$. Os neurônios da primeira camada recebem entradas externas, ou seja:

$$\mathbf{y}^0 = \mathbf{p}, \quad (2.5)$$

onde, \mathbf{p} é o vetor das entradas externas aplicadas à rede neural.

As saídas dos neurônios da última camada são consideradas saídas da rede e representadas por:

$$\mathbf{y} = \mathbf{y}^M. \quad (2.6)$$

No processo de treinamento supervisionado, fornece-se ao algoritmo um conjunto de pares de treinamento da forma,

$$\{\mathbf{p}_1, \mathbf{t}_1\}, \{\mathbf{p}_2, \mathbf{t}_2\}, \dots, \{\mathbf{p}_Q, \mathbf{t}_Q\}, \quad (2.7)$$

onde, \mathbf{t}_q é o vetor de saídas desejadas para a rede neural correspondente ao vetor de entradas \mathbf{p}_q e Q é o número total de pares de treinamento.

Para cada entrada ou vetor de entradas apresentados, a rede produz uma saída, que é comparada com a saída desejada. Dessa comparação surge uma medida do erro, ou seja, um índice de desempenho. O algoritmo de retropropagação usa como índice de desempenho o erro médio quadrático e tem como objetivo minimizar essa medida do erro através do ajuste dos parâmetros da rede (pesos e *bias*).

Se cada saída desejada ocorre com igual probabilidade, o erro médio quadrático é proporcional a soma dos erros quadráticos, e pode ser aproximado por:

$$F(\mathbf{x}) = (\mathbf{t}(k) - \mathbf{y}(k))^T (\mathbf{t}(k) - \mathbf{y}(k)) = \mathbf{e}(k)^T \mathbf{e}(k), \quad (2.8)$$

onde, $F(\mathbf{x})$ é o índice de desempenho, \mathbf{x} é o vetor que engloba os pesos e *bias* da rede e $\mathbf{e}(k)$ é o vetor de erro dos neurônios da última camada na iteração k .

As equações de atualização dos parâmetros para o método do gradiente são:

$$w_{i,j}^m(k+1) = w_{i,j}^m(k) - \rho \frac{\partial F}{\partial w_{i,j}^m}, \quad (2.9)$$

$$b_i^m(k+1) = b_i^m(k) - \rho \frac{\partial F}{\partial b_i^m}, \quad (2.10)$$

onde, ρ é a taxa de aprendizagem, $w_{i,j}^m$ é o peso entre o neurônio i da camada m e a entrada y_j da camada $m-1$ e b_i^m é o *bias* para o neurônio i da camada m .

Como em redes neurais multicamadas o erro não é uma função explícita dos pesos e *bias* na camada oculta, é necessário usar a regra da cadeia para obter as derivadas:

$$\frac{\partial F}{\partial w_{i,j}^m} = \frac{\partial F}{\partial n_i^m} \frac{\partial n_i^m}{\partial w_{i,j}^m}, \quad (2.11)$$

$$\frac{\partial F}{\partial b_i^m} = \frac{\partial F}{\partial n_i^m} \frac{\partial n_i^m}{\partial b_i^m}, \quad (2.12)$$

onde, n_i^m é o potencial de ativação do neurônio i na camada m e é dado por:

$$n_i^m = \left(\sum_{j=1}^{R^{m-1}} w_{i,j}^m y_j^{m-1} \right) + b_i^m, \quad (2.13)$$

onde R^{m-1} é o número de neurônios na camada $(m-1)$.

Sendo assim, o segundo termo das Equações (2.11) e (2.12) pode ser calculado, resultando:

$$\frac{\partial n_i^m}{\partial w_{i,j}^m} = y_j^{m-1}, \quad (2.14)$$

$$\frac{\partial n_i^m}{\partial b_i^m} = 1. \quad (2.15)$$

Definindo, $S_i^m = \frac{\partial F}{\partial n_i^m}$ como a sensibilidade do índice de desempenho F à mudanças no i -ésimo elemento da entrada da rede na camada m , as Equações (2.11) e (2.12) podem ser simplificadas, resultando em:

$$\frac{\partial F}{\partial w_{i,j}^m} = S_i^m y_j^{m-1}, \quad (2.16)$$

$$\frac{\partial F}{\partial b_i^m} = S_i^m. \quad (2.17)$$

O algoritmo de retropropagação pode, então, ser expresso como:

$$w_{i,j}^m(k+1) = w_{i,j}^m(k) - \rho S_i^m y_j^{m-1}, \quad (2.18)$$

$$b_i^m(k+1) = b_i^m(k) - \rho S_i^m. \quad (2.19)$$

Na forma matricial as Equações (2.18) e (2.19) tornam-se, respectivamente:

$$\mathbf{W}^m(k+1) = \mathbf{W}^m(k) - \rho \mathbf{S}^m (\mathbf{y}^{m-1})^T, \quad (2.20)$$

$$\mathbf{b}^m(k+1) = \mathbf{b}^m(k) - \rho \mathbf{S}^m, \quad (2.21)$$

onde, para R^m sendo o número de neurônios da camada m , então:

$$\mathbf{S}^m \equiv \frac{\partial \mathbf{F}}{\partial \mathbf{n}^m} = \begin{bmatrix} \frac{\partial F}{\partial n_1^m} \\ \frac{\partial F}{\partial n_2^m} \\ \vdots \\ \frac{\partial F}{\partial n_{R^m}^m} \end{bmatrix}. \quad (2.22)$$

Resta calcular as sensibilidades \mathbf{S}^m , o que requer outra aplicação da regra da cadeia. Nesse processo ocorre uma relação de recorrência, na qual a sensibilidade na camada m é calculada pela sensibilidade na camada $m+1$. É dessa relação de recorrência que origina a denominação do algoritmo "retropropagação" e é através dela que os neurônios das camadas ocultas são responsabilizados pelo erro produzido na saída da rede.

Para determinar a relação de recorrência para as sensibilidades, usa-se a seguinte matriz jacobiana:

$$\frac{\partial \mathbf{n}^{m+1}}{\partial \mathbf{n}^m} \equiv \begin{bmatrix} \frac{\partial n_1^{m+1}}{\partial n_1^m} & \frac{\partial n_1^{m+1}}{\partial n_2^m} & \cdots & \frac{\partial n_1^{m+1}}{\partial n_{R^m}^m} \\ \frac{\partial n_2^{m+1}}{\partial n_1^m} & \frac{\partial n_2^{m+1}}{\partial n_2^m} & \cdots & \frac{\partial n_2^{m+1}}{\partial n_{R^m}^m} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial n_{R^{m+1}}^{m+1}}{\partial n_1^m} & \frac{\partial n_{R^{m+1}}^{m+1}}{\partial n_2^m} & \cdots & \frac{\partial n_{R^{m+1}}^{m+1}}{\partial n_{R^m}^m} \end{bmatrix}. \quad (2.23)$$

É necessário encontrar uma expressão para a matriz jacobiana (Equação (2.23)) e para isso considera-se, por exemplo, o elemento i, j da matriz:

$$\frac{\partial n_i^{m+1}}{\partial n_j^m} = \frac{\partial (\sum_{j=1}^{R^m} w_{i,j}^{m+1} y_j^m + b_i^{m+1})}{\partial n_j^m} = w_{i,j}^{m+1} \frac{\partial y_j^m}{\partial n_j^m} = w_{i,j}^{m+1} \frac{\partial f^m(n_j^m)}{\partial n_j^m} = w_{i,j}^{m+1} \dot{f}^m(n_j^m), \quad (2.24)$$

onde,

$$\dot{f}^m(n_j^m) = \frac{\partial f^m(n_j^m)}{\partial n_j^m}. \quad (2.25)$$

Consequentemente, a matriz jacobiana pode ser escrita como:

$$\frac{\partial \mathbf{n}^{m+1}}{\partial \mathbf{n}^m} = \mathbf{W}^{m+1} \dot{\mathbf{f}}^m(\mathbf{n}^m), \quad (2.26)$$

onde,

$$\dot{\mathbf{f}}^m(\mathbf{n}^m) = \begin{bmatrix} \dot{f}^m(n_1^m) & 0 & \cdots & 0 \\ 0 & \dot{f}^m(n_2^m) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \dot{f}^m(n_{R^m}^m) \end{bmatrix}. \quad (2.27)$$

É possível agora escrever a relação de recorrência para a sensibilidade usando a regra da cadeia na forma de matriz, resultando em:

$$\mathbf{S}^m = \frac{\partial \mathbf{F}}{\partial \mathbf{n}^m} = \left(\frac{\partial \mathbf{n}^{m+1}}{\partial \mathbf{n}^m} \right)^T \frac{\partial \mathbf{F}}{\partial \mathbf{n}^{m+1}} = \dot{\mathbf{f}}^m(\mathbf{n}^m) (\mathbf{W}^{m+1})^T \frac{\partial \mathbf{F}}{\partial \mathbf{n}^{m+1}} = \dot{\mathbf{f}}^m(\mathbf{n}^m) (\mathbf{W}^{m+1})^T \mathbf{S}^{m+1}. \quad (2.28)$$

As sensibilidades são retropropagadas pela rede da última camada para a primeira, ou seja, para obter a sensibilidade da primeira camada (\mathbf{S}^1) é necessário a sensibilidade da segunda camada (\mathbf{S}^2), conforme Equação (2.28).

Resta ainda um passo para completar o algoritmo de retropropagação. É necessário um ponto inicial \mathbf{S}^M para a relação de recorrência, o qual é obtido na última camada, ou seja:

$$S_i^M = \frac{\partial F}{\partial n_i^M} = \frac{\partial (\mathbf{t} - \mathbf{y})^T (\mathbf{t} - \mathbf{y})}{\partial n_i^M} = \frac{\partial \sum_{i=1}^{R^M} (t_i - y_i)^2}{\partial n_i^M} = -2(t_i - y_i) \frac{\partial y_i}{\partial n_i^M}. \quad (2.29)$$

Como,

$$\frac{\partial y_i}{\partial n_i^M} = \frac{\partial y_i^M}{\partial n_i^M} = \frac{\partial f^M(n_i^M)}{\partial n_i^M} = \dot{f}^M(n_i^M), \quad (2.30)$$

pode-se escrever,

$$S_i^M = -2(t_i - y_i) \dot{f}^M(n_i^M), \quad (2.31)$$

que na forma matricial, resulta:

$$\mathbf{S}^M = -2\dot{\mathbf{f}}^M(\mathbf{n}^M)(\mathbf{t} - \mathbf{y}). \quad (2.32)$$

Apesar de muito usado, o algoritmo de retropropagação apresenta consideráveis desvantagens. Além de muito lento, é possível obter resultados influenciados por um mínimo local. Por esses motivos, muitos pesquisadores vêm desenvolvendo novos algoritmos. Alguns algoritmos são baseados em modificações no algoritmo de retropropagação, as quais incluem taxa de aprendizagem variável e inclusão de um termo de momento que torna o algoritmo mais estável reduzindo as oscilações da trajetória de busca do ponto de mínimo. Outros algoritmos são baseados em métodos numéricos, tais como, o do gradiente conjugado e o algoritmo de Levenberg-Marquardt. Há também algoritmos especialmente desenvolvidos para o treinamento de redes dinâmicas, como os apresentados em Werbos (1990), Narendra e Parthasaraty (1991) e Kosmatopoulos et al. (1995).

2.5.2 Algoritmo de Levenberg-Marquardt.

O algoritmo de Levenberg-Marquardt (HAGAN; DEMUTH; BEALE, 1996) é uma variação do método de Newton, que por sua vez é baseado na expansão de segunda ordem da série de Taylor. Fazendo a expansão da função $F(\mathbf{x})$ que representa o índice de desempenho (ver Equação (2.8)):

$$F(\mathbf{x}(k+1)) = F(\mathbf{x}(k) + \Delta\mathbf{x}(k)) \approx F(\mathbf{x}(k)) + \mathbf{g}^T(k)\Delta\mathbf{x}(k) + \frac{1}{2}\Delta\mathbf{x}^T(k)\mathbf{A}(k)\Delta\mathbf{x}(k), \quad (2.33)$$

onde,

$$\mathbf{A}(k) \equiv \nabla^2 F(\mathbf{x}) |_{\mathbf{x}=\mathbf{x}(k)} \text{ é a matriz hessiana} \quad (2.34)$$

e

$$\mathbf{g}(k) \equiv \nabla F(\mathbf{x}) |_{\mathbf{x}=\mathbf{x}(k)} \text{ é o gradiente de } F(\mathbf{x}). \quad (2.35)$$

O princípio do método é localizar o ponto onde o gradiente desta aproximação quadrática de $F(\mathbf{x})$ é zero, ou seja, o ponto onde esta função tem, pelo menos, um mínimo local.

Para encontrar o gradiente desta função quadrática é necessário associar a forma geral de uma função quadrática, ou seja:

$$H(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T \mathbf{a} \mathbf{x} + \mathbf{d}^T \mathbf{x} + c. \quad (2.36)$$

O gradiente de $H(\mathbf{x})$ é dado por:

$$\nabla H(\mathbf{x}) = \mathbf{a} \mathbf{x} + \mathbf{d}. \quad (2.37)$$

Então, por analogia, o gradiente da função $F(\mathbf{x}(k+1))$ será dado por:

$$\nabla F(\mathbf{x}(k+1)) = \mathbf{A}(k)\Delta\mathbf{x}(k) + \mathbf{g}(k). \quad (2.38)$$

Igualando a Equação (2.38) a zero e isolando $\Delta\mathbf{x}(k)$, obtém-se o valor de atualização de parâmetros que tenderá a levar ao ponto onde o gradiente é zero, ou seja:

$$\Delta\mathbf{x}(k) = -\mathbf{A}^{-1}(k)\mathbf{g}(k). \quad (2.39)$$

Assim sendo, o método de Newton é definido como:

$$\mathbf{x}(k+1) = \mathbf{x}(k) - \mathbf{A}^{-1}(k)\mathbf{g}(k). \quad (2.40)$$

O índice de desempenho mais utilizado para o treinamento de uma rede neural multicamada é o erro médio quadrático. Se cada saída desejada ocorre com igual probabilidade, o erro médio quadrático é proporcional a soma dos erros quadráticos sobre os Q pares de treinamento, ou seja:

$$F(\mathbf{x}) = \sum_{q=1}^Q (\mathbf{t}_q - \mathbf{y}_q)^T (\mathbf{t}_q - \mathbf{y}_q) = \sum_{q=1}^Q \mathbf{e}_q^T \mathbf{e}_q = \sum_{q=1}^Q \sum_{j=1}^{R^M} (e_{j,q})^2 = \sum_{i=1}^N (v_i)^2 = \mathbf{v}^T(\mathbf{x})\mathbf{v}(\mathbf{x}), \quad (2.41)$$

onde, $e_{j,q}$ é o j -ésimo elemento do erro para o q -ésimo par entrada/saída e $\mathbf{v}(\mathbf{x})$ é um vetor ou matriz (se houver mais de um neurônio na camada de saída) que contém os erros de todos os neurônios para todos os pares de treinamento apresentados à rede.

O j -ésimo elemento do gradiente será, então:

$$[\nabla F(\mathbf{x})]_j = \frac{\partial F(\mathbf{x})}{\partial x_j} = 2 \sum_{i=1}^N v_i(\mathbf{x}) \frac{\partial v_i(\mathbf{x})}{\partial x_j}. \quad (2.42)$$

Na forma matricial a Equação (2.42) é escrita como:

$$\nabla F(\mathbf{x}) = 2\mathbf{J}^T(\mathbf{x})\mathbf{v}(\mathbf{x}), \quad (2.43)$$

onde, $\mathbf{J}^T(\mathbf{x})$ é a matriz jacobiana dada por:

$$\mathbf{J}(\mathbf{x}) = \begin{bmatrix} \frac{\partial v_1(\mathbf{x})}{\partial x_1} & \frac{\partial v_1(\mathbf{x})}{\partial x_2} & \cdots & \frac{\partial v_1(\mathbf{x})}{\partial x_n} \\ \frac{\partial v_2(\mathbf{x})}{\partial x_1} & \frac{\partial v_2(\mathbf{x})}{\partial x_2} & \cdots & \frac{\partial v_2(\mathbf{x})}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial v_N(\mathbf{x})}{\partial x_1} & \frac{\partial v_N(\mathbf{x})}{\partial x_2} & \cdots & \frac{\partial v_N(\mathbf{x})}{\partial x_n} \end{bmatrix}. \quad (2.44)$$

O próximo passo é encontrar a matriz hessiana, $\mathbf{A}(k)$. O elemento o, p da matriz hessiana é dado por:

$$[\nabla^2 F(\mathbf{x})]_{o,p} = \frac{\partial^2 F(\mathbf{x})}{\partial x_o \partial x_p} = 2 \sum_{i=1}^N \left\{ \frac{\partial v_i(\mathbf{x})}{\partial x_o} \frac{\partial v_i(\mathbf{x})}{\partial x_p} + v_i(\mathbf{x}) \frac{\partial^2 v_i(\mathbf{x})}{\partial x_o \partial x_p} \right\}. \quad (2.45)$$

Na forma matricial tem-se,

$$\nabla^2 F(\mathbf{x}) = 2\mathbf{J}^T(\mathbf{x})\mathbf{J}(\mathbf{x}) + 2\mathbf{P}(\mathbf{x}), \quad (2.46)$$

sendo,

$$\mathbf{P}(\mathbf{x}) = \sum_{i=1}^N v_i(\mathbf{x})\nabla^2 v_i(\mathbf{x}). \quad (2.47)$$

Adotando que $\mathbf{P}(\mathbf{x})$ seja desprezível em relação a primeira parcela da Equação (2.46), a matriz hessiana pode ser aproximada por:

$$\nabla^2 F(\mathbf{x}) \approx 2\mathbf{J}^T(\mathbf{x})\mathbf{J}(\mathbf{x}). \quad (2.48)$$

Substituindo a Equação (2.48) na equação Equação (2.40), obtém-se o método de Gauss-Newton, ou seja:

$$\mathbf{x}(k+1) = \mathbf{x}(k) - 2\mathbf{J}^T(\mathbf{x}(k))\mathbf{v}(\mathbf{x}(k))[2\mathbf{J}^T(\mathbf{x}(k))\mathbf{J}(\mathbf{x}(k))]^{-1} \quad (2.49)$$

$$= \mathbf{x}(k) - \mathbf{J}^T(\mathbf{x}(k))\mathbf{v}(\mathbf{x}(k))[\mathbf{J}^T(\mathbf{x}(k))\mathbf{J}(\mathbf{x}(k))]^{-1}. \quad (2.50)$$

A vantagem do método de Gauss-Newton sobre o de Newton está em não requerer o cálculo de derivadas de segunda ordem. Um problema com o método de Gauss-Newton é que a matriz $[\mathbf{J}^T\mathbf{J}]$ pode não ser inversível. Esse inconveniente pode ser superado usando a seguinte modificação:

$$[\mathbf{J}^T\mathbf{J}] \approx [\mathbf{J}^T\mathbf{J}] + \mu\mathbf{I}, \quad (2.51)$$

onde, \mathbf{I} é a matriz identidade e μ é um parâmetro de valor variável.

Supondo que os autovalores e autovetores de $[\mathbf{J}^T\mathbf{J}]$ sejam $\{\lambda_1, \lambda_2, \dots, \lambda_n\}$ e $\{\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_n\}$, respectivamente, então da Equação (2.51), vem que:

$$[\mathbf{J}^T\mathbf{J} + \mu\mathbf{I}]\mathbf{z}_i \approx [\mathbf{J}^T\mathbf{J}]\mathbf{z}_i + \mu\mathbf{z}_i \approx \lambda_i\mathbf{z}_i + \mu\mathbf{z}_i \approx (\lambda_i + \mu)\mathbf{z}_i. \quad (2.52)$$

Conseqüentemente, os autovetores de $[\mathbf{J}^T \mathbf{J} + \mu \mathbf{I}]$ são os mesmos que os de $[\mathbf{J}^T \mathbf{J}]$, e os autovalores de $[\mathbf{J}^T \mathbf{J} + \mu \mathbf{I}]$ são $\lambda_i + \mu$. Aumentando μ até que $\lambda_i + \mu > 0$ para todo i , $[\mathbf{J}^T \mathbf{J} + \mu \mathbf{I}]$ torna-se positiva definida e conseqüentemente inversível.

A modificação dada na Equação (2.51) é substituída na Equação (2.49) e chega-se no algoritmo de Levenberg-Marquardt, ou seja:

$$\mathbf{x}(k+1) = \mathbf{x}(k) - \mathbf{J}^T(\mathbf{x}(k))\mathbf{v}(\mathbf{x}(k))[\mathbf{J}^T(\mathbf{x}(k))\mathbf{J}(\mathbf{x}(k)) + \mu(k)\mathbf{I}]^{-1}, \quad (2.53)$$

ou

$$\Delta \mathbf{x}(k) = -\mathbf{J}^T(\mathbf{x}(k))\mathbf{v}(\mathbf{x}(k))[\mathbf{J}^T(\mathbf{x}(k))\mathbf{J}(\mathbf{x}(k)) + \mu(k)\mathbf{I}]^{-1}. \quad (2.54)$$

Assim sendo, quando aumenta-se $\mu(k)$ o algoritmo se aproxima do método do gradiente com pequena taxa de aprendizagem, ou seja:

$$\mathbf{x}(k+1) \cong \mathbf{x}(k) - \frac{1}{\mu(k)} \mathbf{J}^T(\mathbf{x}(k))\mathbf{v}(\mathbf{x}(k)) = \mathbf{x}(k) - \frac{1}{2\mu(k)} \nabla F(\mathbf{x}) \quad (2.55)$$

Para situação contrária, quando $\mu(k)$ é reduzido para zero, o algoritmo se torna o método de Gauss-Newton.

É usual que o algoritmo comece com um valor pequeno para $\mu(k)$. Se o valor de $F(\mathbf{x})$ não for reduzido, a iteração é repetida com $\mu(k)$ multiplicado por algum fator μ_{inc} maior que 1. Feito isso, espera-se que $F(\mathbf{x})$ diminua porque será dado um pequeno passo na direção do gradiente. Se $F(\mathbf{x})$ for reduzido, $\mu(k)$ é multiplicado por algum fator μ_{dec} menor que 1 e o algoritmo se aproxima do método de Gauss-Newton, o que deve prover uma convergência mais rápida. Dessa forma, conclui-se que o algoritmo de Levenberg-Marquardt engloba a velocidade do método de Newton e a garantia de convergência do método do gradiente.

Uma etapa importante do algoritmo de Levenberg-Marquardt é o cálculo da matriz jacobiana. Antes de apresentar o procedimento para o cálculo da matriz jacobiana, é conveniente observar da Equação (2.44), o vetor de erro,

$$\mathbf{v}^T = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \dots & \mathbf{v}_N \end{bmatrix} = \begin{bmatrix} e_{1,1} & e_{2,1} & \dots & e_{R^M,1} & e_{1,2} & \dots & e_{R^M,Q} \end{bmatrix}, \quad (2.56)$$

e o vetor de parâmetros, sendo:

$$\mathbf{x}^T = \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \dots & \mathbf{x}_n \end{bmatrix} = \begin{bmatrix} w_{1,1}^1 & w_{1,2}^1 & \dots & w_{R^1,R^0}^1 & b_1^1 & \dots & b_{R^1}^1 & w_{1,1}^2 & \dots & b_{R^M}^M \end{bmatrix}, \quad (2.57)$$

onde, R^0 é o número de entradas da rede, $N = QR^M$ e $n = R^1(R + 1) + R^2(R^1 + 1) + \dots + R^M(R^{M-1} + 1)$ e M é o número de camadas da rede neural.

Substituindo as Equações (2.56) e (2.57) na Equação (2.44), a matriz jacobiana para o treinamento de redes neurais multicamadas pode ser escrita como:

$$\mathbf{J}(\mathbf{x}) = \begin{bmatrix} \frac{\partial e_{1,1}}{\partial w_{1,1}^1} & \frac{\partial e_{1,1}}{\partial w_{1,2}^1} & \dots & \frac{\partial e_{1,1}}{\partial w_{R^1,R^0}^1} & \frac{\partial e_{1,1}}{\partial b_1^1} & \dots \\ \frac{\partial e_{2,1}}{\partial w_{1,1}^1} & \frac{\partial e_{2,1}}{\partial w_{1,2}^1} & \dots & \frac{\partial e_{2,1}}{\partial w_{R^1,R^0}^1} & \frac{\partial e_{2,1}}{\partial b_1^1} & \dots \\ \vdots & \vdots & & \vdots & \vdots & \\ \frac{\partial e_{R^M,1}}{\partial w_{1,1}^1} & \frac{\partial e_{R^M,1}}{\partial w_{1,2}^1} & \dots & \frac{\partial e_{R^M,1}}{\partial w_{R^1,R^0}^1} & \frac{\partial e_{R^M,1}}{\partial b_1^1} & \dots \\ \frac{\partial e_{1,2}}{\partial w_{1,1}^1} & \frac{\partial e_{1,2}}{\partial w_{1,2}^1} & \dots & \frac{\partial e_{1,2}}{\partial w_{R^1,R^0}^1} & \frac{\partial e_{1,2}}{\partial b_1^1} & \dots \\ \vdots & \vdots & & \vdots & \vdots & \end{bmatrix}. \quad (2.58)$$

Os elementos da matriz jacobiana são, então, calculados da seguinte maneira:

$$[\mathbf{J}]_{h,l} = \frac{\partial v_h}{\partial x_l} = \frac{\partial e_{k,q}}{\partial x_l}, \quad (2.59)$$

onde, $h = (q - 1)R^M + k$ é o índice referente à linha da matriz jacobiana que está sendo calculada, k é o índice referente ao neurônio da camada de saída que fornece o erro que está sendo obtido, l é o índice referente à coluna da matriz jacobiana que está sendo calculada, determinado pelo elemento de peso ou *bias* que está sendo considerado e q é o par de treinamento (entrada-saída) que está sendo considerado.

O algoritmo de retropropagação calcula as sensibilidades através de uma relação de recorrência da última camada para a primeira camada da rede neural. Pode-se usar o mesmo conceito para calcular os termos necessários para a matriz jacobiana. Seja uma sensibilidade do algoritmo de Levenberg-Marquardt da forma:

$$S_{i,h}^m \equiv \frac{\partial v_h}{\partial n_{i,q}^m} = \frac{\partial e_{k,q}}{\partial n_{i,q}^m}. \quad (2.60)$$

Os elementos da matriz jacobiana podem ser obtidos através de:

$$[\mathbf{J}]_{h,l} = \frac{\partial v_h}{\partial x_l} = \frac{\partial e_{k,q}}{\partial w_{i,j}^m} = \frac{\partial e_{k,q}}{\partial n_{i,q}^m} \frac{\partial n_{i,q}^m}{\partial w_{i,j}^m} = S_{i,h}^m y_{j,q}^{m-1}, \quad (2.61)$$

ou se x_l for um *bias*, então,

$$[\mathbf{J}]_{h,l} = \frac{\partial v_h}{\partial x_l} = \frac{\partial e_{k,q}}{\partial b_i^m} = \frac{\partial e_{k,q}}{\partial n_{i,q}^m} \frac{\partial n_{i,q}^m}{\partial b_i^m} = S_{i,h}^m. \quad (2.62)$$

Na última camada da rede neural, a sensibilidade será dada por:

$$S_{i,h}^M = \frac{\partial v_h}{\partial n_{i,q}^M} = \frac{\partial e_{k,q}}{\partial n_{i,q}^M} = \frac{\partial (t_{k,q} - y_{k,q}^M)}{\partial n_{i,q}^M} = -\frac{\partial y_{k,q}^M}{\partial n_{i,q}^M} = \begin{cases} -\dot{f}^M(n_{i,q}^M) & \text{para } i = k \\ 0 & \text{para } i \neq k \end{cases} \quad (2.63)$$

Quando uma entrada \mathbf{p}_q for aplicada à rede neural e a correspondente saída \mathbf{y}_q^M for calculada, a retropropagação do algoritmo de Levenberg-Marquardt será iniciada com,

$$\mathbf{S}_q^M = -\dot{\mathbf{f}}^M(\mathbf{n}_q^M), \quad (2.64)$$

onde, $\dot{\mathbf{f}}^M(\mathbf{n}_q^M)$ é definida na Equação (2.27). Cada coluna da matriz \mathbf{S}_q^M deve ser retropropagada através da rede usando a Equação (2.28) para produzir uma linha da matriz jacobiana. As colunas também podem ser retropropagadas usando,

$$\mathbf{S}_q^m = \dot{\mathbf{f}}^m(\mathbf{n}_q^m)(\mathbf{W}^{M+1})^T \mathbf{S}_q^{m+1}. \quad (2.65)$$

As matrizes totais de sensibilidade do algoritmo de Levenberg-Marquardt para cada camada são, então, criadas aumentando as matrizes calculadas para cada entrada, ou seja:

$$\mathbf{S}^m = [\mathbf{S}_1^m \mathbf{S}_2^m \dots \mathbf{S}_Q^m]. \quad (2.66)$$

Para cada entrada que é apresentada à rede, serão retropropagados R^M vetores de sensibilidade. Isso acontece porque está sendo calculada a derivada de cada erro individual, ao invés da derivada do somatório dos erros quadráticos. Para toda entrada aplicada na rede haverá R^M erros (um para cada elemento da saída da rede). Para cada erro haverá uma linha da matriz jacobiana.

As iterações do algoritmo de Levenberg-Marquardt podem ser resumidas como segue:

1. Apresentar todas as entradas à rede neural e calcular as saídas correspondentes (usando Equação (2.4) e Equação (2.5)) e os erros. Calcular o índice de desempenho $F(\mathbf{x})$ usando Equação (2.41).
2. Calcular a matriz jacobiana (Equação (2.58)). Calcular as sensibilidades com as relações de recorrência da Equação (2.65) iniciando com Equação (2.64). Calcular os elementos da matriz jacobiana com Equação (2.61) e Equação (2.62).
3. Resolver a Equação (2.54) para obter $\Delta \mathbf{x}_k$.
4. Recalcular as somas dos erros quadráticos usando $\mathbf{x}_k + \Delta \mathbf{x}_k$. Se esta nova soma é menor que aquela calculada no passo 1, multiplica-se μ por μ_{dec} , deixa-se $\mathbf{x}_{k+1} = \mathbf{x}_k + \Delta \mathbf{x}_k$ e retorna-se ao passo 1. Se a soma não for reduzida, multiplica-se μ por μ_{inc} e dirige-se ao passo 3.

Figura 8 apresenta um diagrama de blocos representando resumidamente o algoritmo de Levenberg-Marquardt.

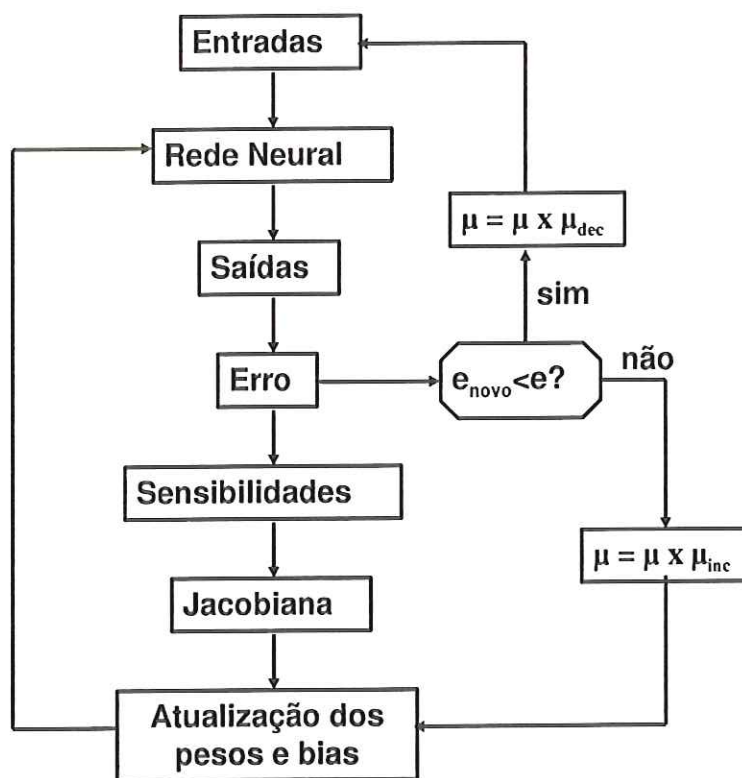


Figura 8: Representação esquemática do algoritmo de Levenberg-Marquardt.

3 IDENTIFICAÇÃO DO FUNCIONAL DA RESPOSTA AEROELÁSTICA.

3.1 Introdução.

Este Capítulo descreve o procedimento de experimento para coleta de dados aeroelásticos e como os parâmetros utilizados no treinamento da rede neural foram definidos. Posteriormente, os resultados de treinamento da rede neural como representação do funcional aeroelástico são apresentados. A identificação do funcional aeroelástico foi realizada para dados experimentais representativos de comportamento linear e não linear de um sistema aeroelástico. O desempenho da rede neural para essa tarefa é, finalmente, verificado.

3.2 Aparato experimental.

Os dados de entrada e saída desejada utilizados no processo de identificação da resposta aeroelástica via redes neurais, foram obtidos através de ensaios experimentais. Ensaiou-se em túnel de vento para uma semi-asa reta retangular composta de uma estrutura principal feita de fibra de vidro e resina epóxi revestida por espuma de alta densidade (styrofoam[®]) e folha de mogno de $2mm$ de espessura para adquirir perfil aerodinâmico. A asa possui aerofólio NACA0012 da raiz até a ponta, sendo a envergadura de $800mm$ e $290mm$ de corda. A estrutura principal internamente à asa é de uma placa plana com taxa de afilamento de 1:1,67, sendo a largura da raiz da asa $250mm$.

Para obtenção da placa afilada usada como estrutura principal laminou-se manualmente 20 camadas de fibras de vidro (0° , 90°) impregnadas com uma resina epóxi. Para minimizar o efeito da espuma e do mogno na rigidez da estrutura, fez-se a segmentação do revestimento a cada 100mm no sentido da envergadura.

O modelo de asa foi fixado a uma plataforma giratória que fornecia movimento de incidência através de um motor elétrico (servomotor TMC-1000 THOMSON sem escova e amplificador para o motor TD-50) montado abaixo da plataforma giratória. Nove extensômetros foram distribuídos ao longo de três linhas da superfície da estrutura principal para fornecer medidas da sua resposta dinâmica (ver Figura 9). A primeira e última linhas possuíam três extensômetros cada (enumerados de 1 a 3 e 7 a 9, respectivamente), todos para capturar movimentos de flexão. A linha intermediária possuía três extensômetros para capturar movimentos de torção (enumerados de 4 a 6).

A Figura 9 apresenta uma ilustração do aparato experimental com indicação das posições dos extensômetros na estrutura principal do modelo de asa. A Figura 10 mostra mais aspectos do modelo usado no experimento.

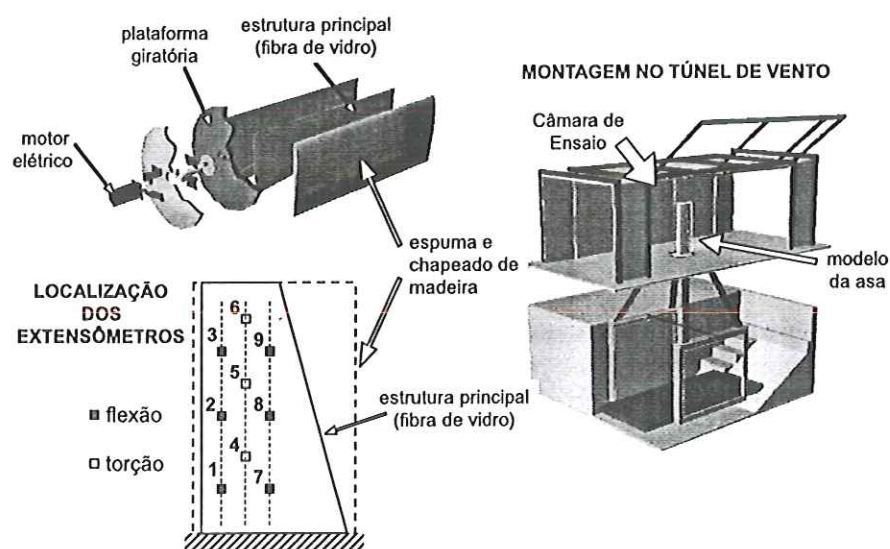


Figura 9: Esquema de montagem do modelo da asa no túnel de vento e localização dos extensômetros.

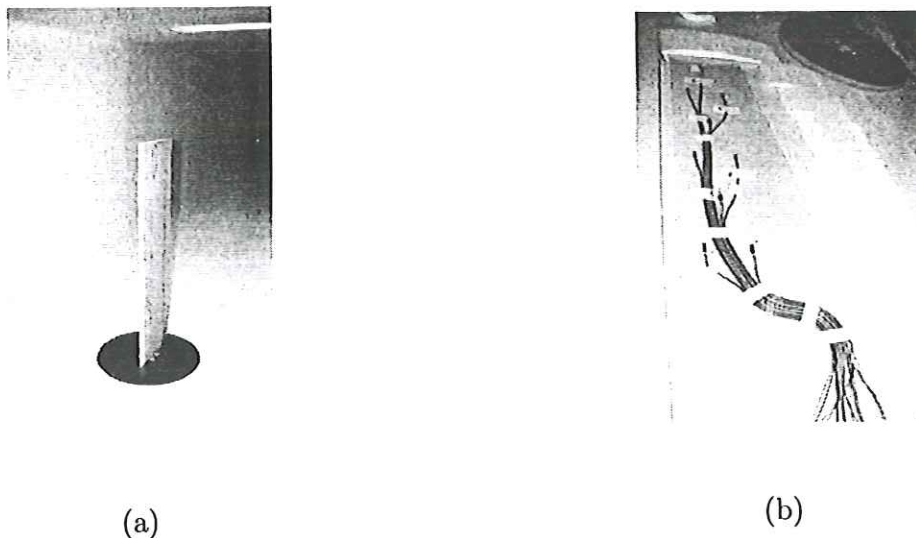


Figura 10: Modelo de asa instalado no túnel de vento (a) e detalhe da estrutura principal com os extensômetros colados (b).

O túnel de vento de circuito fechado apresenta uma câmara de teste com seção transversal de $1300 \times 1400 \text{ mm}$. A máxima velocidade de escoamento aerodinâmico obtida no túnel é 50 m/s com nível de turbulência de $0,3\%$.

Aquisição de dados e controle de movimento do servomotor foram obtidos utilizando uma placa controladora *dSPACE*[®] DS1103 PPC e uma interface em tempo real para *SIMULINK*[®]. O amplificador HBM KWS 3073 para energizar a ponte de extensômetros foi usado para adquirir e amplificar os sinais dos extensômetros. Os sinais resultantes foram adquiridos diretamente pela placa controladora *dSPACE*[®] permitindo subsequente armazenagem de dados em um microcomputador.

A idéia dos ensaios é que, através dos sinais captados pelos extensômetros, se relacione a resposta aeroelástica à variações dinâmicas na incidência do modelo de asa para uma determinada faixa de velocidades. Desta maneira, procura-se estabelecer uma relação entrada/saída na forma de excitações aerodinâmicas por variação de incidência da asa (imersa no escoamento) com a resposta aeroelástica.

Os experimentos foram realizados considerando-se uma faixa de velocidades de escoamento no túnel de vento entre 10 e 17m/s. Esta faixa foi considerada observando-se algumas respostas e respectivas frequências reduzidas, que deveriam ser mantidas entre 0,01 e 0,2, aproximadamente. Esses valores de frequência reduzida são típicos em Aeroelasticidade. Por frequência reduzida, também chamada de número de Strouhal (BISPLINGHOFF; ASHLEY; HALFMAN, 1996), entende-se $k = \frac{wc}{v}$, onde w é a frequência de oscilação, c é a corda de referência e v é a velocidade do escoamento.

A variação da incidência da asa foi baseada em sinais característicos, sendo que para este trabalho apenas os sinais oscilatórios e aleatórios foram considerados. Os movimentos oscilatórios foram considerados para pequenos e relativamente grandes ângulos de incidência, para garantir carregamento aerodinâmico não estacionário linear e não linear, respectivamente. Para o caso de respostas aeroelásticas lineares, a amplitude de oscilações máxima admitida foi de $5,5^{\circ}$ com incidência média α_m de zero grau. Outras amplitudes consideradas nas oscilações foram 2,3 e 4 graus (para $\alpha_m = 0^{\circ}$). Com relação às frequências da plataforma, foram explorados casos com 1,4,8,10 e 13rad/s. A Tabela 1 apresenta um resumo dos parâmetros que caracterizam os experimentos com movimento oscilatório da plataforma.

Tabela 1: Características de movimentos oscilatórios da plataforma (relacionada à carregamento aerodinâmico linear).

Velocidade do escoamento (m/s)	Frequência (plataforma)(rad/s)	Frequência reduzida
11,22	1; 4; 8	0,02576; 0,10338; 0,20676
13,65	1; 4; 8; 10	0,02124; 0,08497; 0,16994; 0,21242
16,93	1; 4; 8; 13	0,01713; 0,06852; 0,13683; 0,22235

O caso não linear foi realizado para a mesma amplitude máxima admitida de $5,5^\circ$, porém com incidência média, α_m , de $9,5^\circ$, o que leva a ângulos de incidência instantâneos de até 15° . Certamente, nestas condições, efeitos de escoamento separado e estol dinâmico estão presentes. A Tabela 2 apresenta um resumo dos parâmetros dos experimentos com movimentos oscilatórios da plataforma, característicos de carregamento aerodinâmico não estacionário não linear.

Tabela 2: Características dos movimentos oscilatórios da plataforma (relacionada à carregamento aerodinâmico não linear).

Velocidade do escoamento (m/s)	Frequência (plataforma) (rad/s)	Frequência reduzida
11,19	3; 8	0,0778; 0,2073
13,58	3; 8; 10	0,00641; 0,1709; 0,2136
16,89	8; 13	0,1374; 0,2233

Para efeito de ilustração as Figuras 11 e 12 mostram exemplos de respostas aeroelásticas lineares e não lineares geradas por movimentos de incidência oscilatórios.

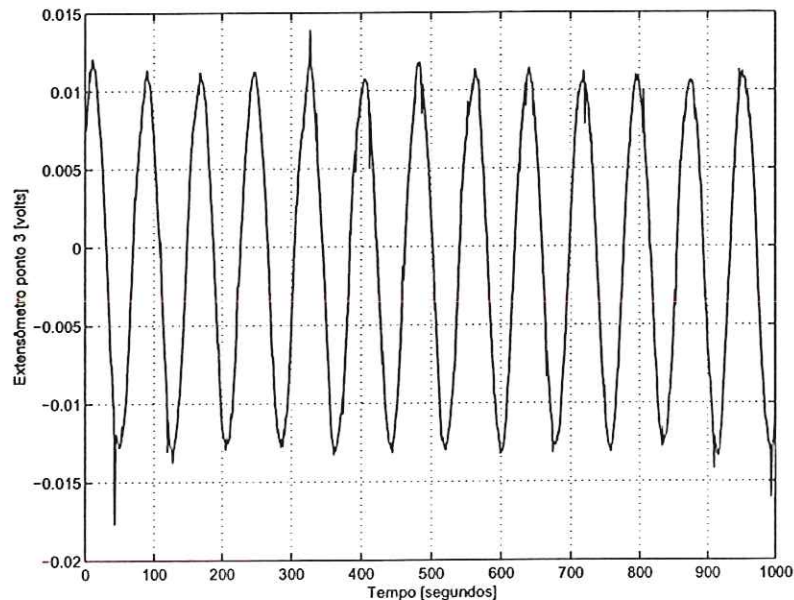


Figura 11: Resposta aeroelástica linear para ângulo de incidência oscilatório (amplitude $5,5^\circ$, $\alpha_m = 0^\circ$, frequência de $8 \frac{rad}{s}$, $k = 0,20676$ e velocidade de escoamento de $11,22 \frac{m}{s}$).

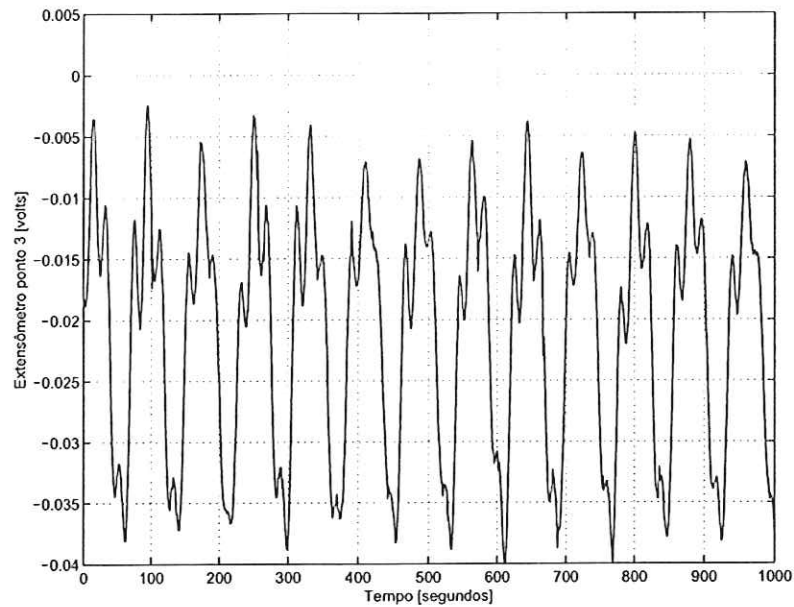


Figura 12: Resposta aeroelástica não linear para ângulo de incidência oscilatório (amplitude $5,5^{\circ}$, $\alpha_m = 9,5^{\circ}$, frequência de $8 \frac{rad}{s}$, $k = 0,2073$ e velocidade de escoamento de $11,19 \frac{m}{s}$).

Com relação aos movimentos aleatórios da plataforma, optou-se por uma forma suavizada de um sinal aleatório de distribuição uniforme. A suavização é em termos de uma filtragem baseada em um sistema de segunda ordem com expressão:

$$\frac{a}{s^2 + s + b}, \quad (3.1)$$

onde, s é a variável de Laplace e a, b são valores ajustáveis para o sinal aleatório desejado. Neste caso, foram escolhidos os valores: $a = 11$ e $b = 30$, para uso nos experimentos. Tais valores são escolhidos por tentativa e erro, tomando como base os sinais produzidos pela filtragem.

A suavização se mostrou necessária, pois um sinal puramente aleatório não era possível de ser corrigido pelo motor de acionamento da plataforma. Apesar dessa limitação, foi possível garantir sinais satisfatórios para gerar respostas aeroelásticas adequadas para o presente trabalho.

3.3 Rede neural para representar o funcional aeroelástico.

A rede neural escolhida para a identificação do funcional da resposta aeroelástica é do tipo perceptron multicamada convencional, portanto, de comportamento estático. Adotando uma arquitetura de rede recorrente, conforme Figura 6, introduz-se dinâmica ao modelo. Para o treinamento da rede neural utilizou-se o algoritmo de Levenberg-Marquardt, como descrito no Capítulo 2.

Dos dados experimentais foi estabelecido que o modelo aeroelástico a ser identificado seria àquele onde como entrada do sistema tem-se a variação dinâmica do ângulo de incidência proporcionado pela plataforma. Como saídas, são tomados os valores dos sinais medidos diretamente nos extensômetros, que são representativos da resposta aeroelástica. Chamando de α e y os vetores das histórias no tempo da entrada e saída, respectivamente, o modelo em redes neurais, ou seja, o funcional da resposta aeroelástica (MF) para uma velocidade de escoamento particular, pode ser definido como:

$$y(k) = MF(U_\infty, \alpha(k), \alpha(k-1), \dots, \alpha(k-M), y(k-1), y(k-2), \dots, y(k-N)) \quad (3.2)$$

onde, U_∞ é a velocidade do escoamento, k é o tempo discreto, M é o número de atrasos no tempo para a entrada e N é o número de atrasos no tempo para a saída.

Observa-se da Equação 3.2 que um parâmetro estático é introduzido como entrada do modelo de rede neural. Este parâmetro é a velocidade do escoamento U_∞ , que possibilita o modelo aeroelástico contabilizar os efeitos do escoamento.

3.3.1 Análise dos parâmetros de treinamento da rede neural.

É conveniente realizar uma análise preliminar de diversas possibilidades para valores iniciais de parâmetros devido à sensibilidade do processo à mínimos locais. Sendo assim, a análise dos parâmetros de treinamento da rede neural realizada nesse trabalho tem por objetivo encontrar valores iniciais de parâmetros que conferem à rede neural um melhor desempenho em treinamento e, conseqüentemente, uma boa identificação do modelo aeroelástico. Os parâmetros analisados foram: *i*) valores iniciais dos pesos e bias; *ii*) arquitetura da rede neural; *iii*) número de atrasos

na entrada; *iv*) tamanho das janelas temporais, que representam o trecho da história no tempo das variáveis envolvidas no treinamento; *v*) parâmetros μ , μ_{inc} e μ_{dec} do algoritmo Levenberg-Marquardt.

Para definição dos melhores parâmetros de treinamento serão analisados os valores de EMQ e Gradiente atingidos pelo algoritmo de Levenberg-Marquardt ao final de 50 épocas para cada valor inicial de parâmetro analisado.

O valor de gradiente apresentado é calculado de acordo com a Equação (3.3) e corresponde a uma medida da inclinação da curva do erro na última época, ou seja, é um gradiente local. Ele indica a tendência de decaimento do erro, sendo que quanto maior seu valor, maior é a expectativa de redução do erro.

$$grad = \sqrt{(\mathbf{J}^T \times \mathbf{e})^T \times (\mathbf{J}^T \times \mathbf{e})}; \quad (3.3)$$

onde, \mathbf{J} é a matriz jacobiana e \mathbf{e} é o vetor de erro médio quadrático (EMQ).

Um gradiente maior indica uma maior tendência de redução de erro mas não garante que, continuando-se o treinamento atinge-se um EMQ significativamente menor. Sendo assim, o valor principal de avaliação de desempenho da rede neural será o EMQ e o valor do gradiente será um critério de desempate entre valores de EMQ muito próximos.

3.3.2 Valores iniciais de pesos e bias.

Neste procedimento de comparação, os valores iniciais de pesos e *bias* são 0,01; 0,1; 0,5 e 1. Para que uma comparação possa ser feita mantem-se os demais parâmetros constantes com os seguintes valores: 50 épocas; arquitetura 12-8-1, sendo que os neurônios nas camadas ocultas são não lineares e na camada de saída são lineares; 4 atrasos para a variável de entrada (ângulo de incidência) e 4 atrasos para a variável de saída (resposta aeroelástica); 1000 janelas temporais, ou seja, o trecho das histórias no tempo das variáveis envolvidas no modelo; $\mu = 0,01$; $\mu_{inc} = 10$; $\mu_{dec} = 0,1$.

A Figura 13 apresenta as curvas de decaimento do erro médio quadrático (EMQ) para cada valor analisado. Tabela 3 apresenta os principais resultados obtidos para o erro final do treinamento, bem como o gradiente local.

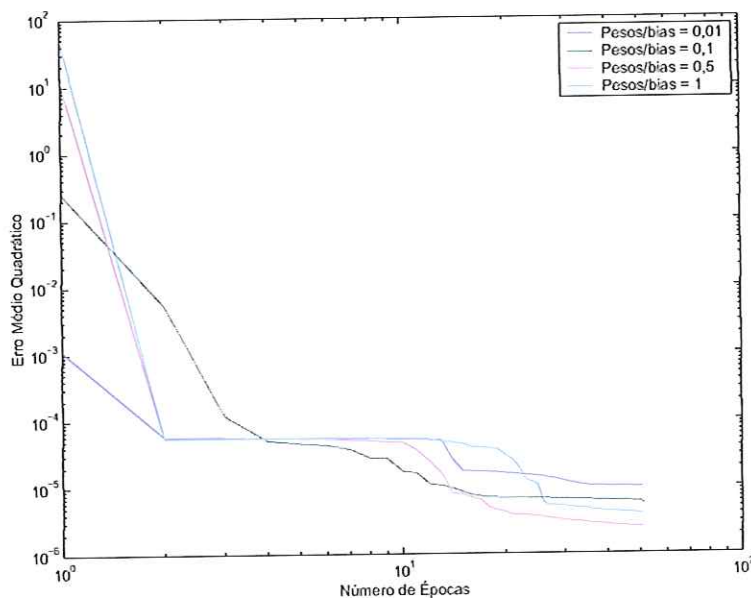


Figura 13: Curvas de decaimento do erro para comparação entre valores iniciais de pesos e *bias*.

Tabela 3: Resultados da comparação entre valores iniciais de pesos e *bias*.

Pesos e <i>bias</i> iniciais	EMQ	Gradiente
0,01	$9,634 \times 10^{-6}$	0,009
0,1	$5,297 \times 10^{-6}$	1,156
0,5	$2,376 \times 10^{-6}$	0,049
1	$3,843 \times 10^{-6}$	0,013

Analisando os dados da Tabela 3 percebe-se que o valor de pesos e *bias* inicial que apresentou menor EMQ foi 0,5, sendo o valor inicial de pesos e *bias* usado nas análises de parâmetros que serão apresentadas.

Segundo Hagan, Demuth e Beale (1996), os pesos e *bias* iniciais não devem ter valor zero pois a propriedade de simetria das rede neurais faz com que a origem do espaço de parâmetros seja um ponto de sela da superfície de desempenho. Os pesos e *bias* iniciais também não devem ter valores altos pois a superfície de desempenho tende a possuir regiões muito planas quando se está

distante do ponto ótimo. Assim, para se evitar o ponto de sela na origem sem o deslocamento para regiões muito planas da superfície de desempenho, é aconselhável que os pesos e bias iniciais tenham valores pequenos e aleatórios. Sendo assim, para o treinamento da rede que representa o funcional aeroelástico os valores iniciais de pesos e *bias* foram escolhidos aleatoriamente no intervalo ± 1 e o valor inicial de pesos e *bias* de 0,5 foi usado somente nos treinamentos com intuito de comparação de parâmetros

3.3.3 Arquiteturas para a rede.

Nesses testes variou-se apenas a arquitetura da rede neural e manteve-se os demais parâmetros constantes com os seguintes valores: 50 épocas, pesos e *bias* iniciais todos iguais a 0,5, 4 atrasos para a variável ângulo de incidência e 4 atrasos para a variável resposta aeroelástica, 1000 janelas temporais, $\mu = 0,01$, $\mu_{inc} = 10$, $\mu_{dec} = 0,1$. Arquiteturas com uma e duas camadas ocultas foram verificadas.

A Figura 14 apresenta as curvas de decaimento do erro para cada arquitetura analisada. Tabela 4 resume os principais resultados obtidos.

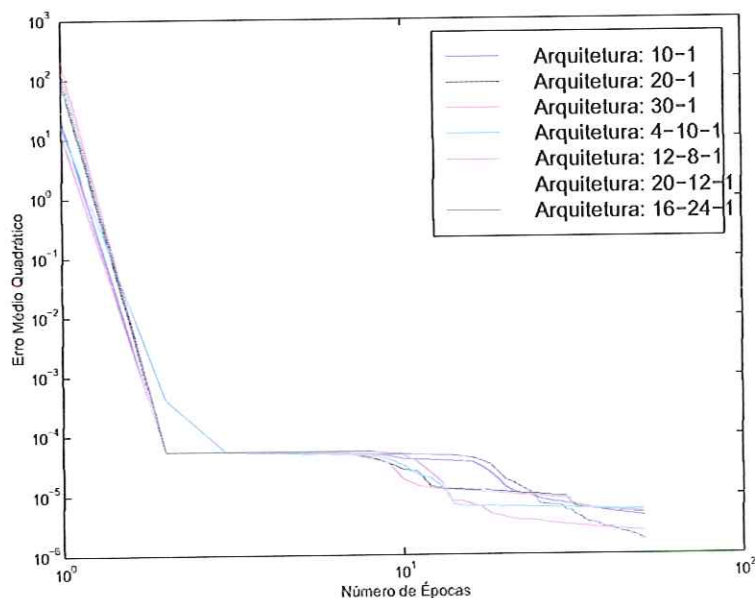


Figura 14: Curvas de decaimento do erro para comparação entre arquiteturas da rede neural.

Tabela 4: Resultados da comparação entre arquiteturas.

Arquitetura	EMQ	Gradiente
10-1	$4,214 \times 10^{-6}$	0,068
20-1	$4,830 \times 10^{-6}$	0,067
30-1	$4,972 \times 10^{-6}$	0,075
4-10-1	$5,163 \times 10^{-6}$	1,213
12-8-1	$2,376 \times 10^{-6}$	0,049
20-12-1	$2,804 \times 10^{-6}$	0,010
16-24-1	$1,654 \times 10^{-6}$	0,630

Analisando os dados da Tabela 4, percebe-se que a arquitetura 16-24-1 além de apresentar um EMQ menor, apresenta um gradiente elevado.

3.3.4 Número de atrasos usados no vetor de entrada.

Aqui a comparação é feita entre o número de atrasos nas variáveis de entrada e saída que compõem o vetor de entrada da rede neural. Nesse caso os parâmetros fixos foram os seguintes: 50 épocas, pesos e *bias* iniciais todos iguais a 0,5, arquitetura 16-24-1, 1000 janelas temporais, $\mu = 0,01$, $\mu_{inc} = 10$, $\mu_{dec} = 0,1$. A Figura 15 apresenta as curvas de decaimento do erro para cada caso analisado. Tabela 5 resume os principais resultados obtidos.

Tabela 5: Resultados da comparação entre os atrasos no vetor de entrada.

Atrasos	EMQ	Gradiente
2-2	$4,343 \times 10^{-6}$	0,118
4-4	$4,375 \times 10^{-6}$	0,967
6-6	$4,364 \times 10^{-5}$	0,142
8-8	$9,401 \times 10^{-6}$	1,6064

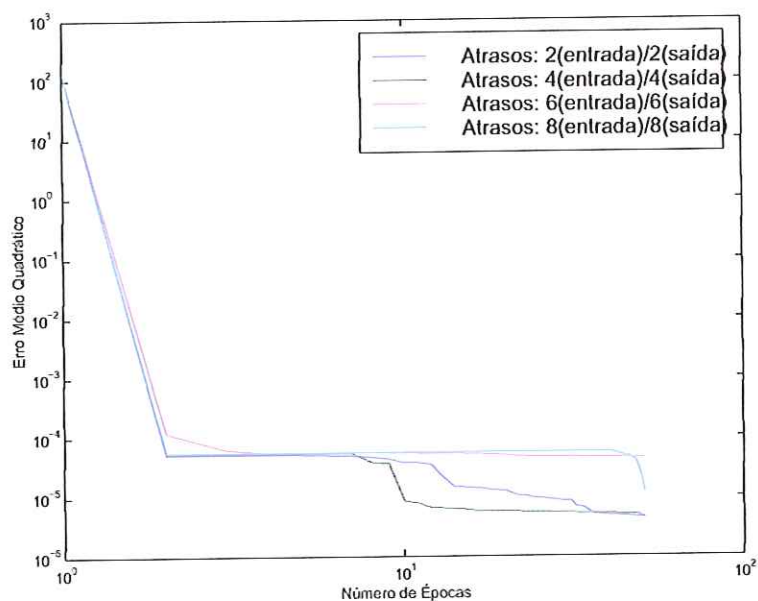


Figura 15: Curvas de decaimento do erro para comparação entre o número de atrasos na rede neural.

Analisando os resultados apresentados na Tabela 5, conclui-se que os valores de dois e quatro atrasos para a variável ângulo de incidência e para a variável resposta aeroelástica apresentaram resultados de EMQ que só se diferenciam a partir da oitava casa decimal. Sendo assim, opta-se pelo valor que apresentou maior valor de gradiente.

3.3.5 Tamanho das janelas temporais.

Uma janela temporal representa o trecho da história no tempo das variáveis envolvidas no treinamento. O número de janelas temporais representam o número de exemplos à que a rede neural está sendo treinada. Nesse caso os parâmetros fixos foram: 50 épocas, pesos e *bias* iniciais todos iguais a 0,5, arquitetura 16-24-1, 4 atrasos para a variável ângulo de incidência e 4 para a variável resposta aeroelástica, $\mu = 0,01$, $\mu_{inc} = 10$, $\mu_{dec} = 0,1$. A Figura 16 apresenta as curvas de decaimento do erro para cada caso analisado. Tabela 6 resume os principais resultados obtidos.



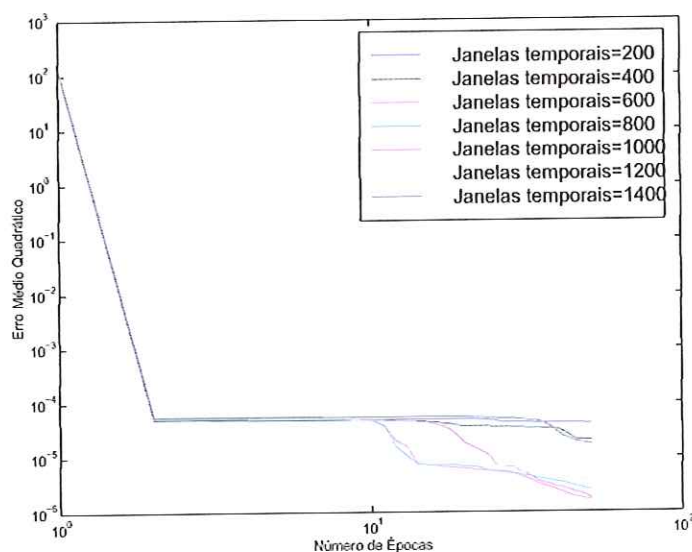


Figura 16: Curvas de decaimento do erro para tamanhos diferentes de janelas temporais usadas na entrada da rede neural.

Tabela 6: Resultados para comparação entre número de janelas temporais.

Número de janelas	EMQ	Gradiente
200	$4,111 \times 10^{-5}$	0,009
400	$2,010 \times 10^{-5}$	0,522
600	$1,617 \times 10^{-6}$	0,034
800	$2,390 \times 10^{-6}$	0,271
1000	$1,654 \times 10^{-6}$	0,630
1200	$3,739 \times 10^{-6}$	0,030
1400	$1,672 \times 10^{-5}$	0,075

Observando os dados da Tabela 6, percebe-se que o EMQ alcançado quando se tem 600 e 1000 janelas temporais só se diferencia a partir da oitava casa decimal. Opta-se então por 1000 janelas temporais por apresentar um gradiente maior.

3.3.6 Parâmetro μ

O parâmetro μ (conforme Equação (2.54)) é o parâmetro do algoritmo de Levenberg-Marquardt que ora tem seu valor aumentado fazendo com que o algoritmo se aproxime do método do gradiente, e ora tem seu valor reduzido fazendo com que o algoritmo de Levenberg-Marquardt se aproxime do método de Newton. Avaliou-se o desempenho em treinamento de alguns valores iniciais de μ . Nesse caso os parâmetros fixos foram os seguintes: 50 épocas, pesos e *bias* iniciais todos iguais a 0,5, arquitetura 16-24-1, 4 atrasos para a variável ângulo de incidência e 4 para a variável resposta aeroelástica, 1000 janelas temporais, $\mu_{inc} = 10$, $\mu_{dec} = 0,1$. Tabela 7 resume os principais resultados obtidos.

Tabela 7: Resultados para comparação entre os parâmetros μ .

μ	EMQ	Gradiente
0,00001	$8,534 \times 10^{-6}$	0,033
0,0001	$3,730 \times 10^{-5}$	0,678
0,001	$8,137 \times 10^{-6}$	0,391
0,01	$1,654 \times 10^{-6}$	0,630
0,1	$5,890 \times 10^{-6}$	0,295

μ	EMQ	Gradiente
0,2	$8,657 \times 10^{-6}$	1,135
0,3	$1,272 \times 10^{-6}$	0,165
0,4	$2,714 \times 10^{-6}$	0,092
0,5	$2,284 \times 10^{-6}$	0,048

A Figura 17 apresenta as curvas de decaimento do erro para cada valor do parâmetro μ testado. Analisando os dados da Tabela 7 percebe-se que o parâmetro $\mu = 0,3$ forneceu o menor valor de EMQ.

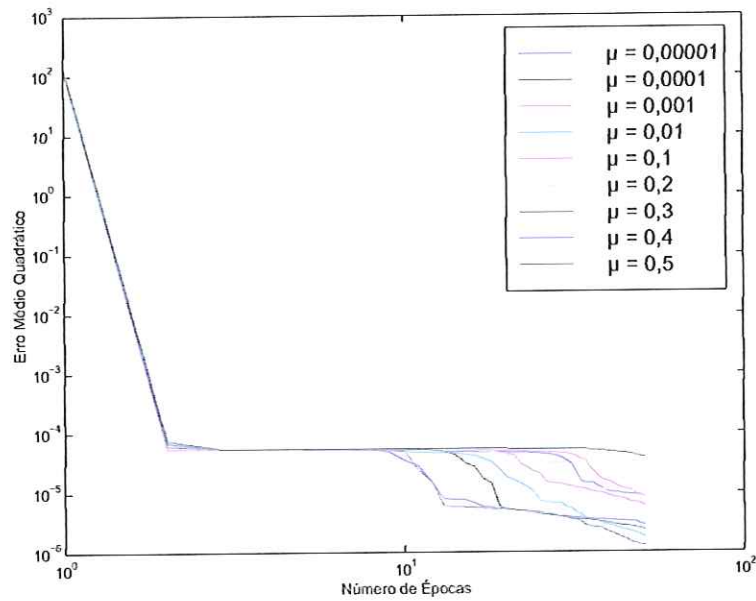


Figura 17: Curvas de decaimento do erro para comparação entre valores do parâmetro μ usados no treinamento da rede neural.

3.3.7 Parâmetro μ_{inc} .

O parâmetro μ_{inc} tem valor maior que 1 e é o responsável por aumentar o valor do parâmetro μ fazendo com que o algoritmo de Levenberg-Marquardt se aproxime do método do gradiente. Nesse caso, os parâmetros fixos foram os seguintes: 50 épocas, pesos e *bias* iniciais todos iguais a 0,5, arquitetura 16-24-1, 4 atrasos para a variável ângulo de incidência e 4 para a variável resposta aeroclástica, 1000 janelas temporais, $\mu = 0,3$, $\mu_{dec} = 0,1$. A Figura 18 apresenta as curvas de decaimento do erro para cada valor do parâmetro μ_{inc} testado. Tabela 8 resume os principais resultados obtidos. Analisando os dados da Tabela 8 percebe-se que o valor $\mu_{inc} = 10$ apresenta um bom resultado de EMQ.

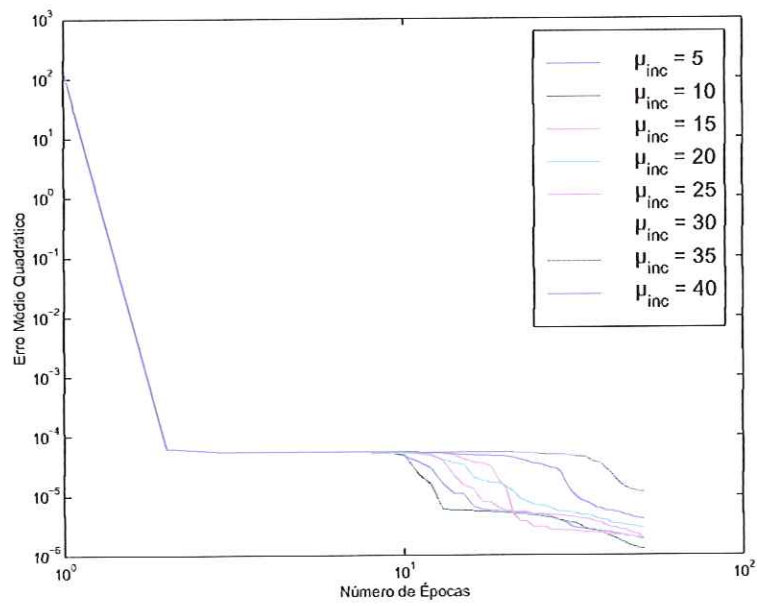


Figura 18: Curvas de decaimento do erro para comparação entre valores do parâmetro μ_{inc} usados no treinamento da rede neural.

Tabela 8: Resultados para comparação entre os parâmetros μ_{inc} .

μ_{inc}	EMQ	Gradiente
5	$1,778 \times 10^{-6}$	0,445
10	$1,272 \times 10^{-6}$	0,165
15	$1,942 \times 10^{-6}$	0,013
20	$2,842 \times 10^{-6}$	0,015
25	$1,941 \times 10^{-6}$	0,025
30	$7,867 \times 10^{-6}$	0,046
35	$1,121 \times 10^{-5}$	0,014
40	$3,955 \times 10^{-6}$	0,028

3.3.8 Parâmetro μ_{dec} .

O parâmetro μ_{dec} tem valor menor que 1 e é o responsável por reduzir o valor do parâmetro μ fazendo com que o algoritmo de Levenberg-Marquardt se aproxime do método de Newton. Nesse caso os parâmetros fixos foram os seguintes: 50 épocas, pesos e *bias* iniciais todos iguais a 0,5, arquitetura 16-24-1, 4 atrasos para a variável ângulo de incidência e 4 para a resposta aeroelástica, 1000 janelas temporais, $\mu = 0,3$, $\mu_{inc} = 10$. A Figura 19 apresenta as curvas de decaimento do erro para cada valor do parâmetro μ_{dec} testado. Tabela 9 resume os principais resultados obtidos. Analisando os dados da Tabela 9 percebe-se que o valor 0,1 apresenta um bom resultado de EMQ.

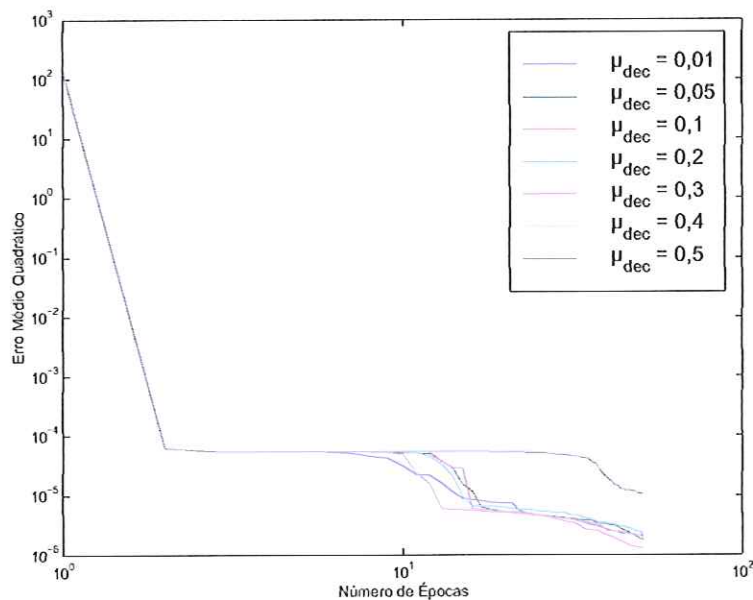


Figura 19: Curvas de decaimento do erro para comparação entre valores do parâmetro μ_{dec} usados no treinamento da rede neural.

Tabela 9: Resultados para comparação entre os parâmetros μ_{dec} .

μ_{dec}	EMQ	Gradiente
0,01	$2,049 \times 10^{-6}$	0,075
0,05	$1,790 \times 10^{-6}$	0,045
0,1	$1,272 \times 10^{-6}$	0,165
0,2	$2,228 \times 10^{-6}$	0,123
0,3	$1,983 \times 10^{-6}$	0,003
0,4	$1,154 \times 10^{-5}$	6,422
0,5	$1,005 \times 10^{-5}$	0,072

A análise apresentada nesta seção procurou mostrar um procedimento sistemático de escolha dos melhores parâmetros da rede neural, bem como os de treinamento. Um sumário dos valores encontrados é apresentado na Tabela 10.

Tabela 10: Conjunto dos parâmetros da rede neural e de treinamento determinados pela análise comparativa.

Item	Valor
Pesos e <i>bias</i> iniciais	todos iguais à 0,5
Arquitetura	16 – 24 – 1
Atrasos	4 para entrada e saída
Janelas temporais	1000
μ	0,3
μ_{inc}	10
μ_{dec}	0,1

3.4 Identificação da resposta aeroelástica linear.

Nesse experimento a resposta aeroelástica compreende a resposta da semi-envergadura de asa produzida pela variação do ângulo de incidência a qual causa variação nas cargas aerodinâmicas e subsequente interação mútua com a dinâmica estrutural da semi-envergadura. À medida que se aumenta o ângulo de incidência começam a surgir as não linearidades aerodinâmicas devido à separação do escoamento na superfície da semi-envergadura de asa. Sendo assim, para se trabalhar com respostas aeroelásticas lineares utiliza-se como entrada para o sistema uma história temporal de ângulo de incidência baixo (até $\pm 5,5^\circ$ e $\alpha_m = 0$).

3.4.1 Treinamento.

Os parâmetros tomados como base para o treinamento foram aqueles que forneceram melhor desempenho na Seção 3.3, exceto a arquitetura da rede neural. Por se tratar de comportamento linear, uma arquitetura com 16 neurônios na primeira camada oculta e 24 na segunda, seria uma arquitetura muito robusta para a tarefa em questão. Sabe-se que arquiteturas com número excessivo de neurônios em geral leva ao sobre ajuste do modelo. Sendo assim, optou-se pela arquitetura composta de duas camadas ocultas com 12 e 8 neurônios respectivamente, todos com função de ativação tangente hiperbólica, e uma camada de saída com um neurônio com função de ativação linear. Os valores de pesos e *bias* iniciais foram tomados aleatoriamente entre ± 1 , seguindo as tendências mostradas nos testes dos parâmetros de treinamento.

Para compor o vetor de entrada tem-se, quatro valores atrasados do ângulo de incidência, quatro valores atrasados da resposta aeroelástica e a velocidade do escoamento aerodinâmico no túnel de vento como parâmetro estático (fixo), somando assim, dez entradas. A Figura 20 apresenta as principais características da rede neural usada para identificação da resposta aeroelástica linear. Foram usados um total de 1000 janelas temporais escolhidas aleatoriamente na base de dados. Como saída desejada utilizou-se a resposta aeroelástica dada pelo extensômetro 3 (flexão), o qual é representado na Figura 9. Os parâmetros do algoritmo Levenberg-Marquardt foram $\mu = 0,3$,

$\mu_{inc} = 10$ e $\mu_{dec} = 0,1$, baseados nos estudos comparativos da seção anterior. O treinamento foi feito por lote e realizou um total de 100 épocas, atingindo um erro da ordem de 10^{-7} , conforme ilustrado na Figura 21.

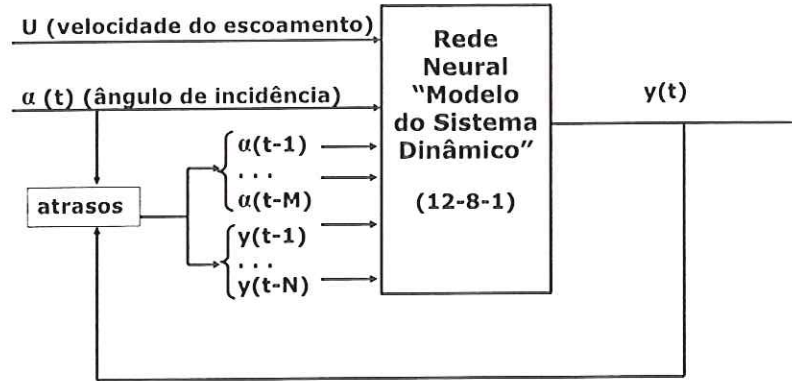


Figura 20: Características da rede neural usada para identificação da resposta aeroelástica linear.

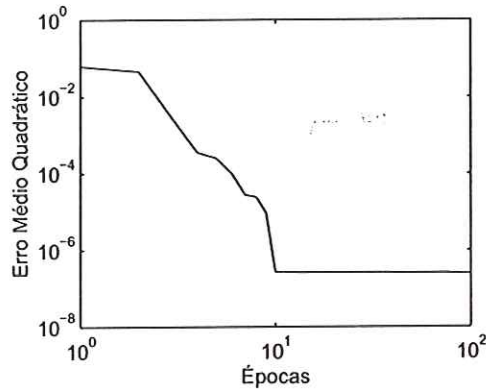


Figura 21: Decaimento do erro durante treinamento.

As Figuras 22 a 24 mostram os resultados após o treinamento da rede neural. São apresentadas respostas aeroelásticas no tempo com as respectivas saídas da rede neural. As histórias temporais do ângulo de incidência (α) usadas na entrada foram geradas através de sinais aleatórios obtidos usando ruído branco e um filtro de segunda ordem para atenuar a intensidade do sinal. Os testes foram realizados nas seguintes velocidades de escoamento: 11,27; 13,64 e 16,96m/s.

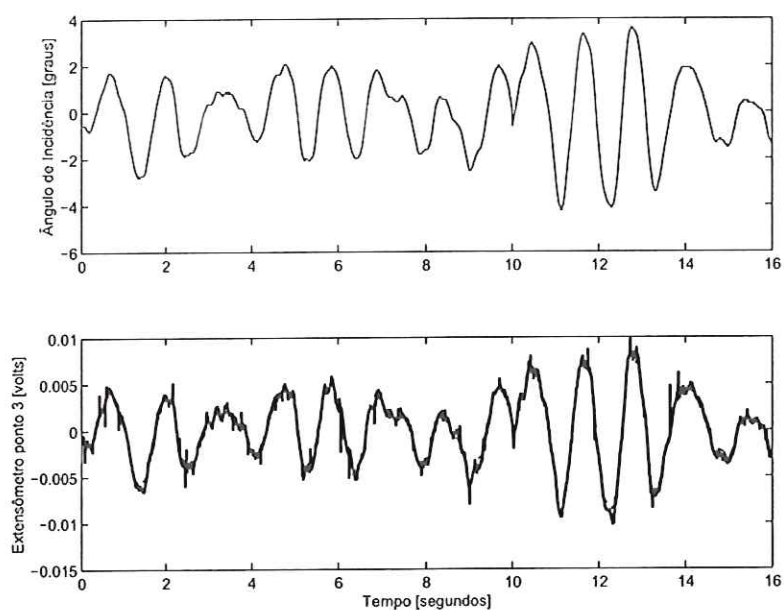


Figura 22: Resultado do treinamento da rede neural para caso de resposta aeroelástica linear; $U = 11,27 \frac{m}{s}$. Linha sólida: experimento; Linha tracejada: modelo neural.

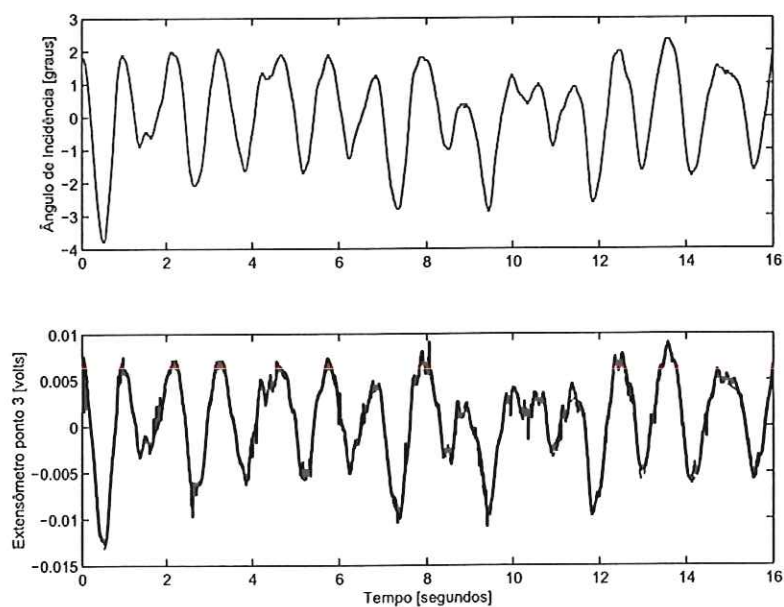


Figura 23: Resultado do treinamento da rede neural para caso de resposta aeroelástica linear; $U = 13,64 \frac{m}{s}$. Linha sólida: experimento; Linha tracejada: modelo neural.

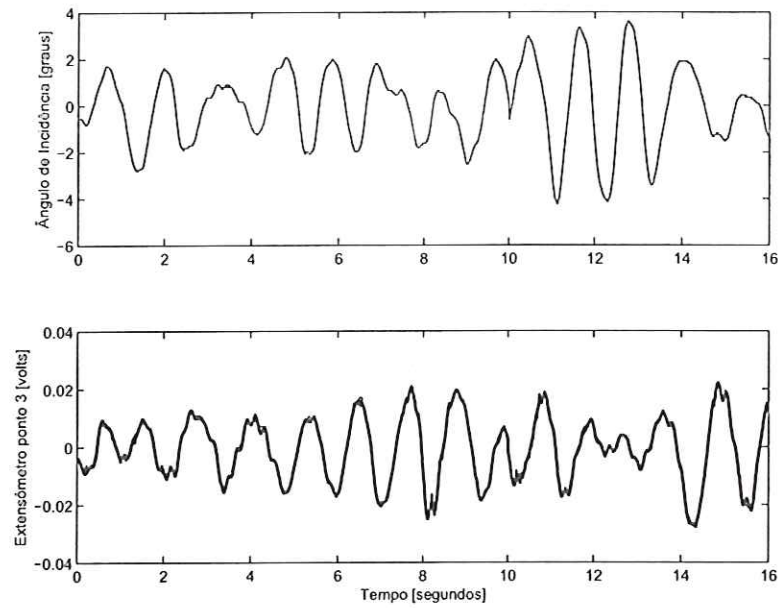


Figura 24: Resultado do treinamento da rede neural para caso de resposta aeroelástica linear; $U = 16,96 \frac{m}{s}$. Linha sólida: experimento; Linha tracejada: modelo neural.

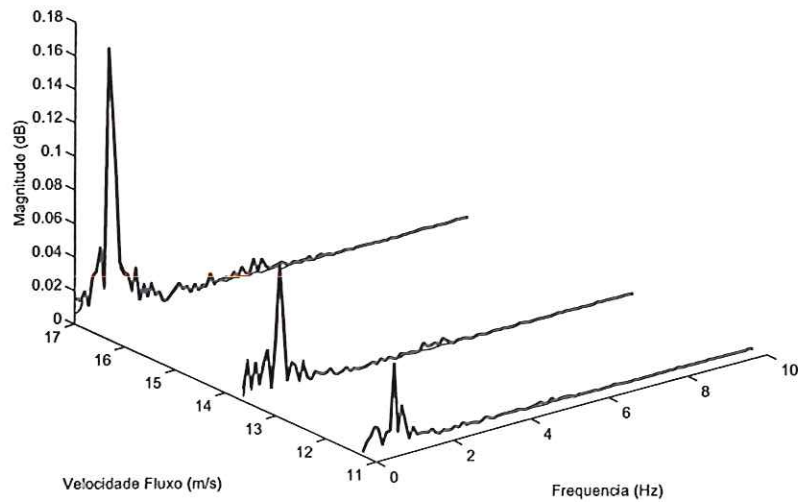


Figura 25: Resultado do treinamento da rede neural para casos de respostas aeroelásticas lineares. Linha sólida: experimento; Linha tracejada: modelo neural.

A Figura 25 apresenta as mesmas respostas aeroelásticas apresentadas nas Figuras 22 a 24 com as respectivas saídas da rede neural após treinamento, porém em frequência e para as três velocidades testadas. Analisando os resultados do treinamento é possível perceber que os sinais medidos do sistema e da rede neural estão sobrepostos, indicando que o treinamento pelo algoritmo de Levenberg-Marquardt foi eficiente. Observando a Figura 25 percebe-se que em velocidades maiores novas frequências vão surgindo e que a rede neural tende a capturá-las. Outra característica a se observar nas Figuras 22 a 25 é que a curva da saída da rede neural apresenta menos ruído que a curva da saída do sistema, o que indica que o modelo de rede neural além de conseguir capturar a dinâmica do sistema através de sinais ruidosos, ele também apresenta propriedades de filtragem.

3.4.2 Generalização.

As Figuras 26 a 33 mostram resultados de testes de generalização da rede neural. Os testes de generalização consistem em apresentar à rede neural entradas desconhecidas, ou seja, entradas que não foram usadas no treinamento. As Figuras 26 a 28 mostram testes de generalização com entradas que também usam história temporal de ângulo de incidência aleatório, porém diferentes daquelas usadas no treinamento. As velocidades de escoamento para cada caso são de 11,10; 13,63 e 16,89m/s, respectivamente. A Figura 29 apresenta as mesmas respostas aeroelásticas apresentadas nas Figuras 26 a 28 com as respectivas saídas da rede neural após testes de generalização, porém em frequência e para as três velocidades testadas.

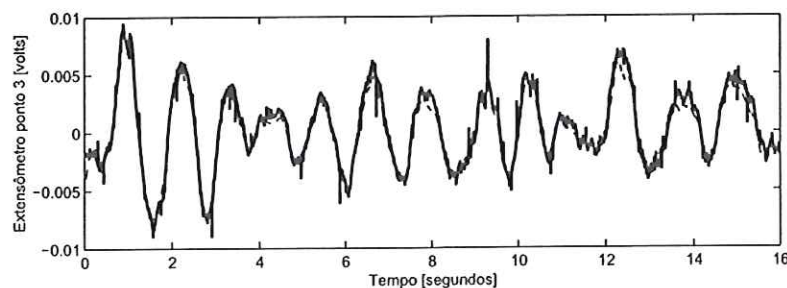


Figura 26: Resultado de teste de generalização da rede neural para caso de resposta aeroelástica linear; $U = 11,10 \frac{m}{s}$. Linha sólida: experimento; Linha tracejada: modelo neural.

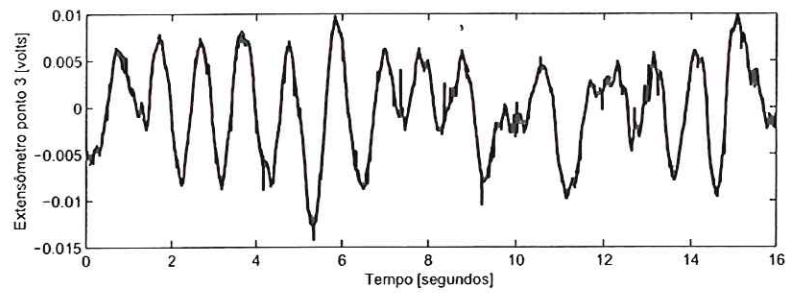


Figura 27: Resultado de teste generalização da rede neural para caso de resposta aeroelástica linear; $U = 13,63 \frac{m}{s}$. Linha sólida: experimento; Linha tracejada: modelo neural.

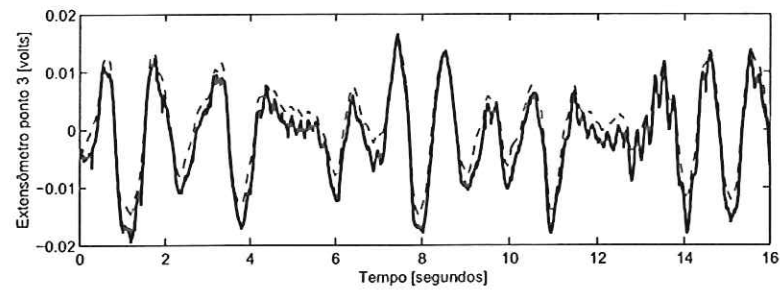


Figura 28: Resultado de teste de generalização da rede neural para caso de resposta aeroelástica linear; $U = 16,89 \frac{m}{s}$. Linha sólida: experimento; Linha tracejada: modelo neural.

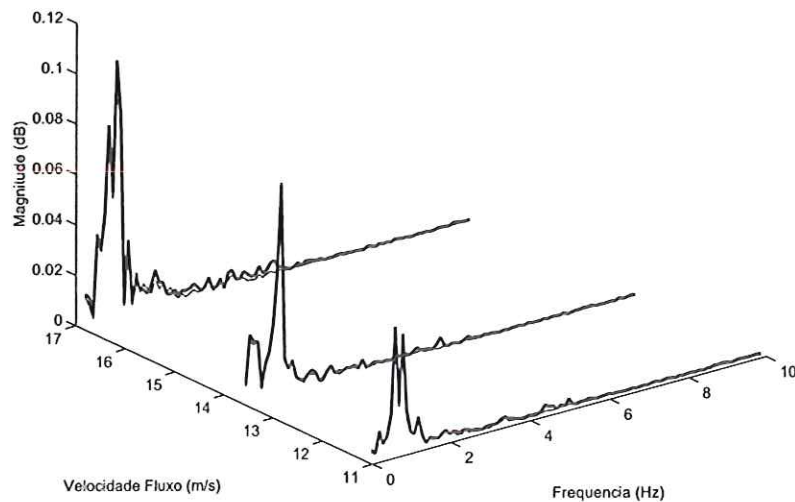


Figura 29: Resultado de testes de generalização da rede neural para casos de respostas aeroelásticas lineares. Linha sólida: experimento; Linha tracejada: modelo neural.

Observando as Figuras 26 a 29 é possível perceber que a rede neural foi capaz de identificar a resposta aeroelástica para as três velocidades de escoamento testadas.

As Figuras 30 a 32 mostram resultados de testes de generalização para entradas que usam história temporal de ângulo de incidência oscilatório com amplitude de 2 graus e frequência de oscilação de 4rad/s . Os testes foram realizados nas velocidades de escoamento de 11,22; 13,65 e $16,93\text{m/s}$ o que resulta nas seguintes frequências reduzidas: 0,10338; 0,08497 e 0,06852, respectivamente. A Figura 33 apresenta as mesmas respostas aeroelásticas apresentadas nas Figuras 30 a 32 com as respectivas saídas da rede neural após testes de generalização, porém em frequência e para as três velocidades testadas.

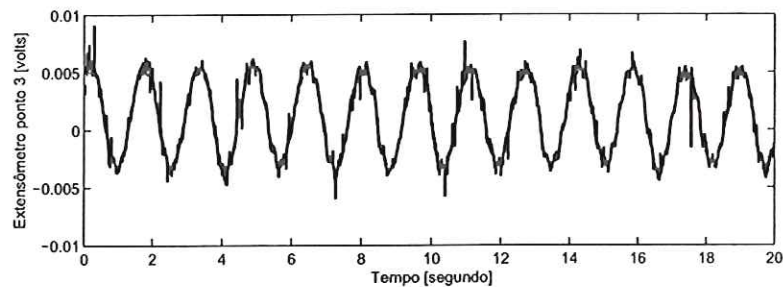


Figura 30: Resultado de teste de generalização da rede neural para caso de resposta aeroelástica linear; $U = 11,22\frac{\text{m}}{\text{s}}$ e $k = 0,10338$. Linha sólida: experimento; Linha tracejada: modelo neural.

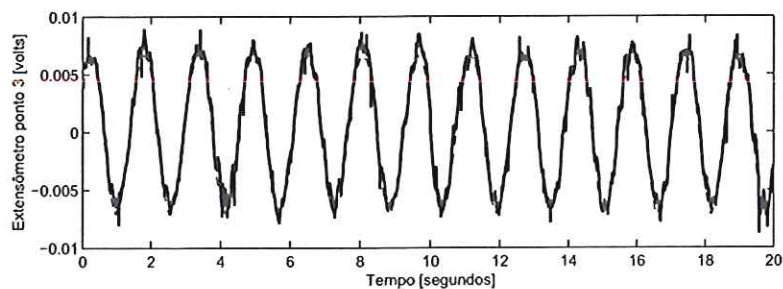


Figura 31: Resultado de teste de generalização da rede neural para caso de resposta aeroelástica linear; $U = 13,65\frac{\text{m}}{\text{s}}$ e $k = 0,08497$. Linha sólida: experimento; Linha tracejada: modelo neural.

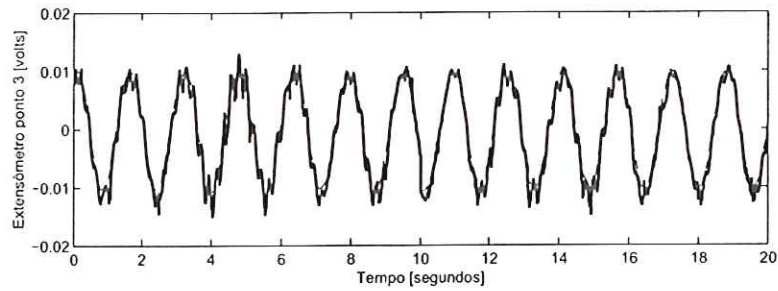


Figura 32: Resultado de teste de generalização da rede neural para caso de resposta aeroelástica linear; $U = 16,93 \frac{m}{s}$ e $k = 0,06852$. Linha sólida: experimento; Linha tracejada: modelo neural.

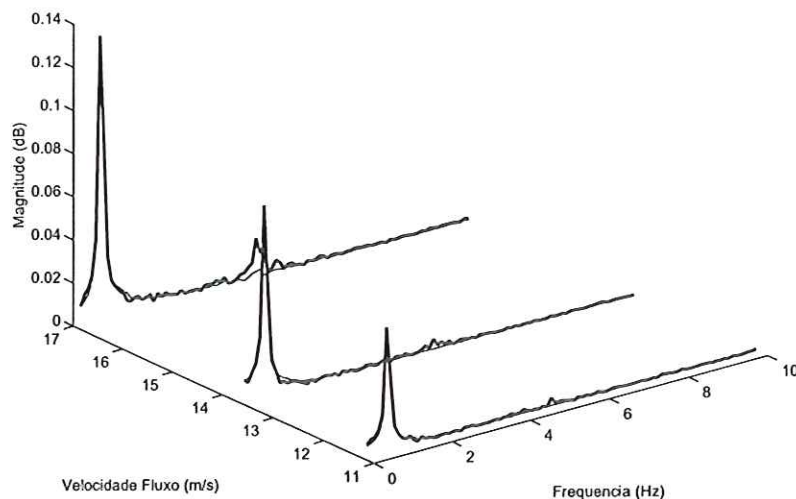


Figura 33: Resultado de testes de generalização da rede neural para casos de respostas aeroelásticas lineares. Linha sólida: experimento; Linha tracejada: modelo neural.

Os resultados mostrados nas Figuras 30 a 33 indicam que o modelo de rede neural foi capaz de identificar o sistema aeroelástico mesmo quando o tipo de entrada aplicada foi diferente. Isso indica que a rede neural é sensível a mudanças no tipo de excitação, ou seja, uma rede que foi treinada com sinais de entrada aleatórios consegue aproximar a resposta do sistema mesmo quando se usa como entrada sinais oscilatórios.

3.5 Identificação da resposta aeroelástica não linear.

As fontes físicas de não linearidades no escoamento aerodinâmico são a separação do escoamento e os efeitos de compressibilidade. Os efeitos de compressibilidade surgem à medida que se aumenta a velocidade do escoamento, no entanto, devido à limitações de velocidade do túnel de vento, não é possível obter respostas aeroelásticas não lineares devido a efeitos de compressibilidade. Sendo assim, as respostas aeroelásticas não lineares foram obtidas usando-se histórias temporais de ângulos de incidência mais altos. Para o experimento em questão, utilizou-se ângulo de incidência máximo de $5,5^\circ$, porém com $\alpha_m = 9,5^\circ$ levando a valores instantâneos de ângulos de incidência de até 15° .

3.5.1 Treinamento.

Assim como no caso linear, os parâmetros tomados como base para o treinamento foram aqueles que forneceram melhor desempenho na Seção 3.3. Sendo assim, utilizou-se a arquitetura com 16 neurônios na primeira camada oculta, 24 na segunda e 1 neurônio na camada de saída. Os neurônios das camadas ocultas têm função de ativação tangente hiperbólica e o da camada de saída tem função de ativação linear. Os pesos e *bias* iniciais tiveram valores aleatórios entre ± 1 .

Para compor o vetor de entrada tem-se quatro valores atrasados do ângulo de incidência, quatro valores atrasados da resposta aeroelástica e a velocidade do escoamento aerodinâmico no túnel de vento como parâmetro estático (fixo), somando assim, dez entradas. A Figura 20 apresenta as principais características da rede neural usada para identificação da resposta aeroelástica não linear. Foram usadas um total de 1000 janelas temporais escolhidas aleatoriamente na base de dados. Como saída desejada utilizou-se a resposta aeroelástica dada pelo extensômetro 3 (flexão), o qual é representado na Figura 9. Os parâmetros do algoritmo Levenberg-Marquardt foram $\mu = 0,3$, $\mu_{inc} = 10$ e $\mu_{dec} = 0,1$. O treinamento foi feito por lote e realizou-se um total de 100 épocas, atingindo um erro da ordem de 10^{-7} , conforme ilustrado na Figura 35.

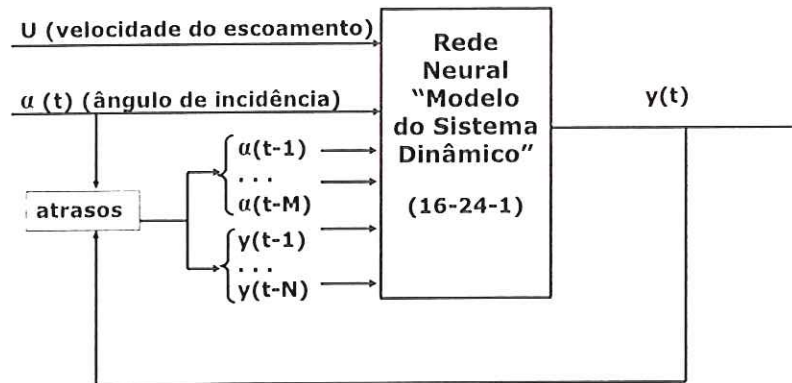


Figura 34: Características da rede neural usada para identificação da resposta aeroelástica não linear.

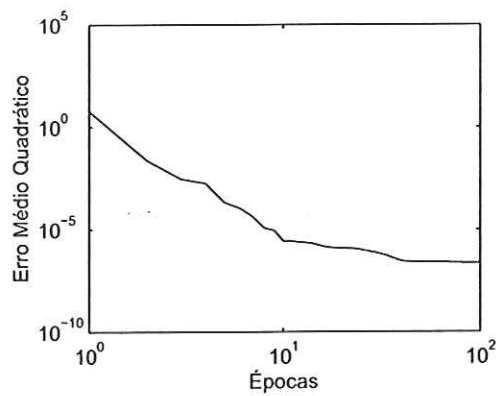


Figura 35: Redução do erro durante treinamento.

As Figuras 36 a 38 mostram os resultados após treinamento da rede neural, o qual utilizou na entrada história temporal de ângulo de incidência aleatório em velocidades de escoamento de 11, 11; 13, 58 e 16, 89 m/s . respectivamente. A Figura 39 apresenta as mesmas respostas aeroelásticas apresentadas nas Figuras 36 a 38 com as respectivas saídas da rede neural após treinamento, porém em frequência e para as três velocidades testadas.

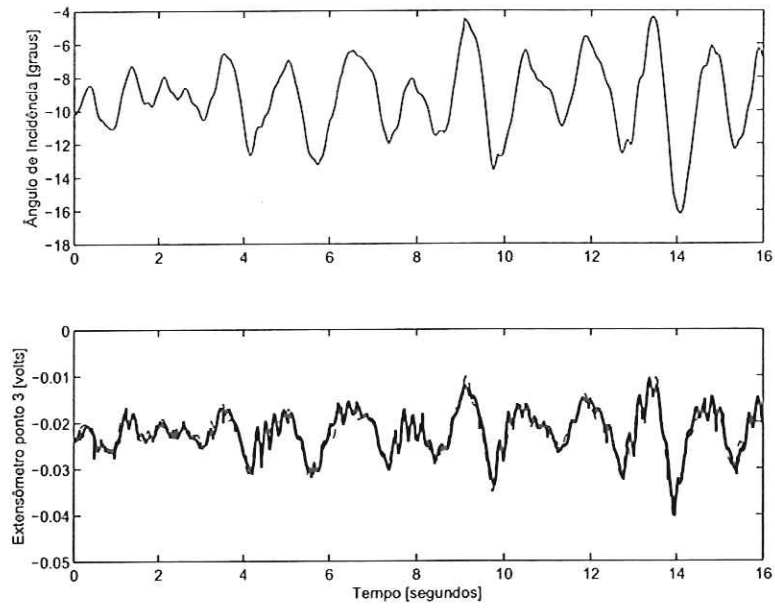


Figura 36: Resultado do treinamento da rede neural para caso de resposta aeroelástica não linear; $U = 11,11 \frac{m}{s}$. Linha sólida: experimento; Linha tracejada: modelo neural.

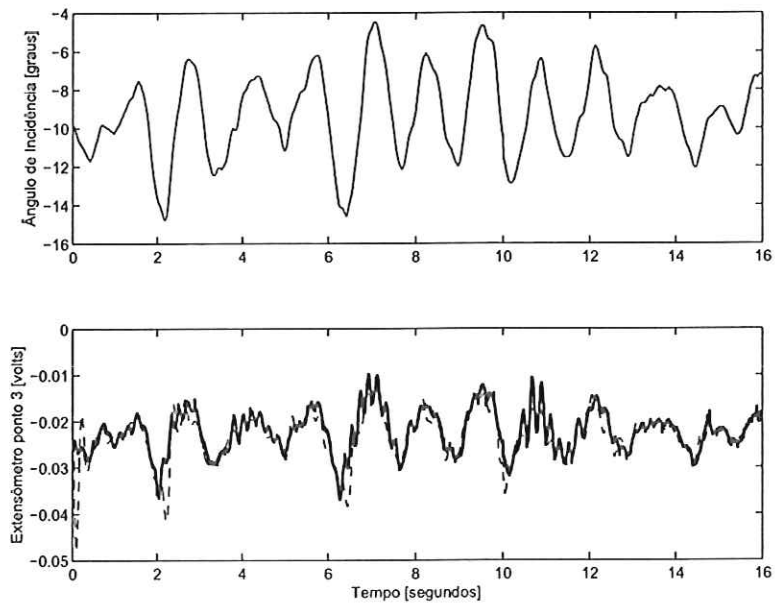


Figura 37: Resultado do treinamento da rede neural para caso de resposta aeroelástica não linear; $U = 13,58 \frac{m}{s}$. Linha sólida: experimento; Linha tracejada: modelo neural.

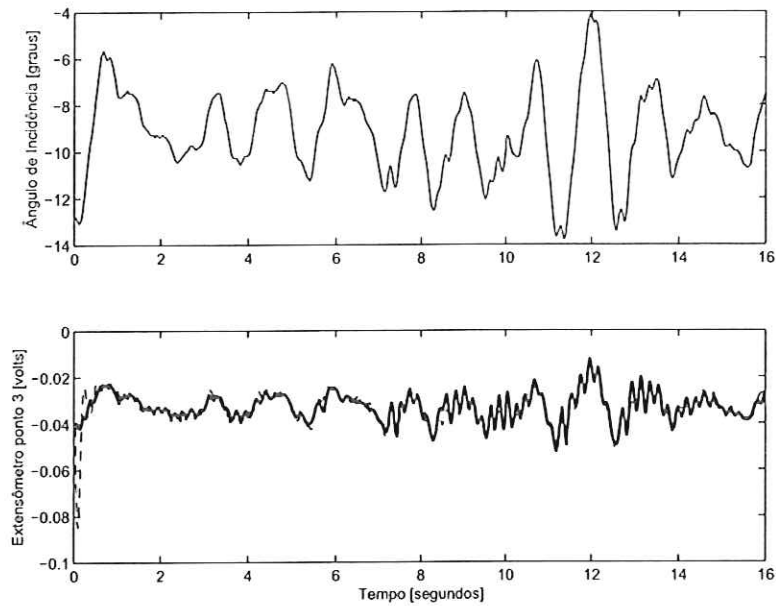


Figura 38: Resultado do treinamento da rede neural para caso de resposta aeroelástica não linear; $U = 16,89 \frac{m}{s}$. Linha sólida: experimento; Linha tracejada: modelo neural.

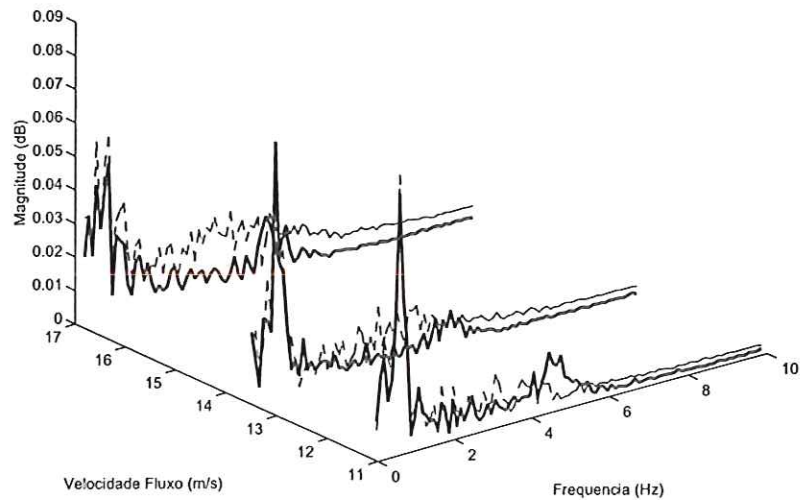


Figura 39: Resultado do treinamento da rede neural para casos de respostas aeroelásticas não lineares. Linha sólida: experimento; Linha tracejada: modelo neural.

Observando as Figuras 36 a 38 percebe-se que os sinais medidos da rede neural e do sistema estão sobrepostos, o que mostra que a rede efetuou uma boa aproximação da saída do sistema. Percebe-se uma maior dificuldade de identificação para as respostas não lineares em frequência para as velocidades mais elevadas (Figura 39).

3.5.2 Generalização.

As Figuras 40 a 42 apresentam resultados de testes de generalização da rede neural obtidos quando se utilizou na entrada da rede neural história temporal de ângulo de incidência aleatório, diferente daquele usado para o treinamento. A Figura 43 apresenta as mesmas respostas aeroelásticas apresentadas nas Figuras 40 a 42 com as respectivas saídas da rede neural após testes de generalização, porém em frequência e para as três velocidades testadas.

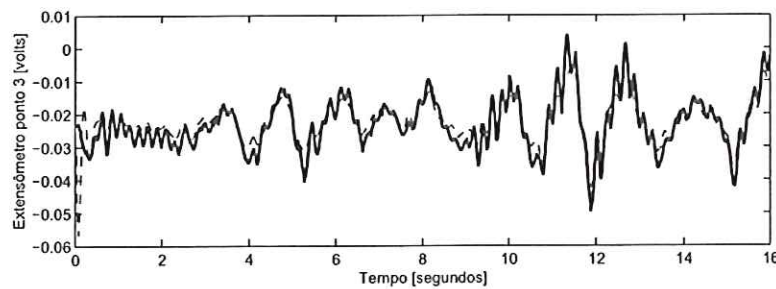


Figura 40: Resultado de teste de generalização da rede neural para caso de resposta aeroelástica não linear, $U = 11,11 \frac{m}{s}$. Linha sólida: experimento; Linha tracejada: modelo neural.

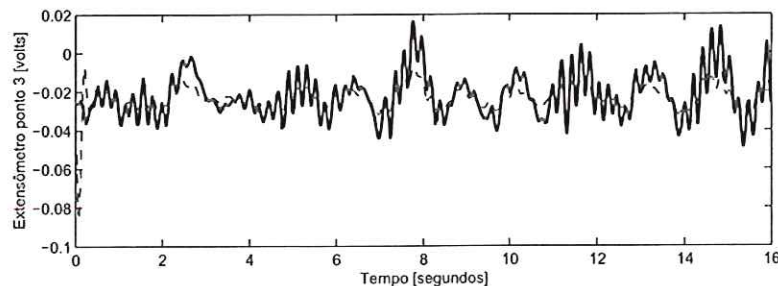


Figura 41: Resultado de teste de generalização da rede neural para caso de resposta aeroelástica não linear, $U = 13,58 \frac{m}{s}$. Linha sólida: experimento; Linha tracejada: modelo neural.

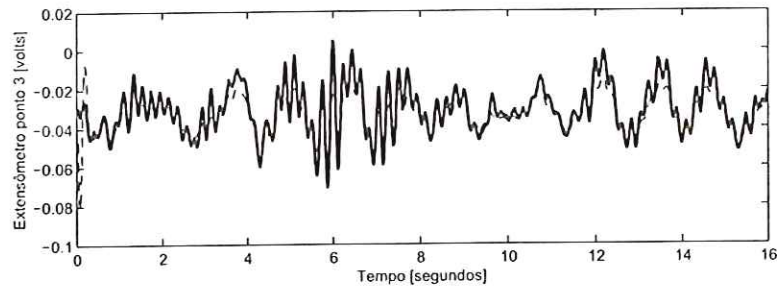


Figura 42: Resultado de teste de generalização da rede neural para caso de resposta aeroelástica não linear, $U = 16,89 \frac{m}{s}$. Linha sólida: experimento; Linha tracejada: modelo neural.

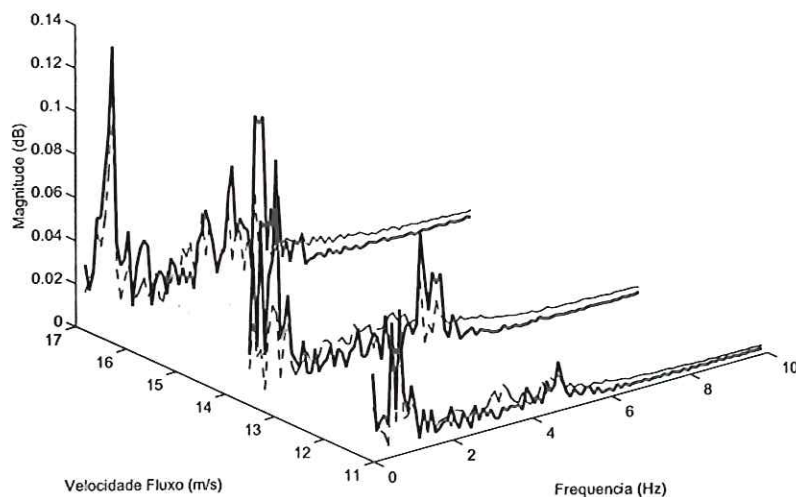


Figura 43: Resultado de testes de generalização da rede neural para casos de respostas aeroelásticas não lineares. Linha sólida: experimento; Linha tracejada: modelo neural.

Observando as Figuras 40 a 43 é possível perceber que houve uma aproximação satisfatória entre as saídas da rede neural e a do sistema quando utilizou-se na entrada da rede uma história temporal de ângulo de incidência aleatório, em especial para a velocidade de escoamento mais baixa ($U = 11,11 m/s$) onde as não linearidades são mais fracas.

Os demais testes de generalização foram feitos com entradas que usam histórias temporais de ângulos de incidência oscilatórios em diversas combinações de amplitudes, frequências e velocidade de escoamento. As Figuras 44 a 63 apresentam os resultados obtidos, com os respectivos sinais em frequência.

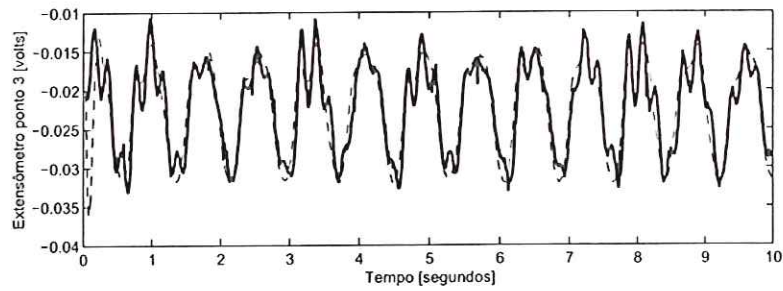


Figura 44: Resultado de teste de generalização da rede neural para caso de resposta aeroelástica não linear, amplitude 3° , frequência $8 \frac{rad}{s}$ e $U = 11,19 \frac{m}{s}$. Linha sólida: experimento; Linha tracejada: modelo neural.

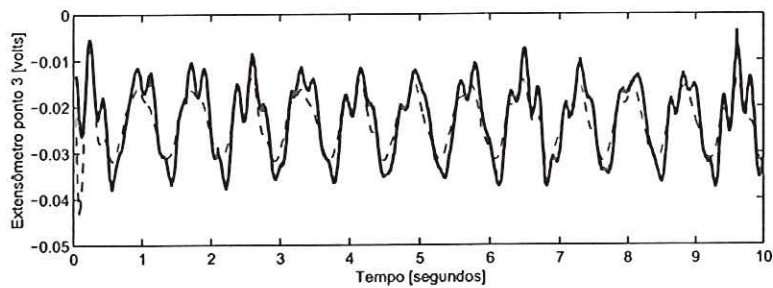


Figura 45: Resultado de teste de generalização da rede neural para caso de resposta aeroelástica não linear, amplitude 3° , frequência $8 \frac{rad}{s}$ e $U = 13,58 \frac{m}{s}$. Linha sólida: experimento; Linha tracejada: modelo neural.

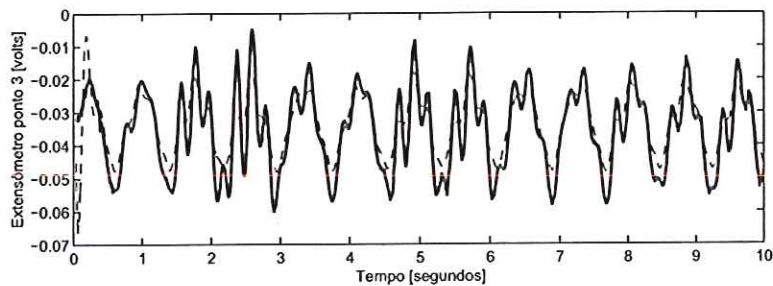


Figura 46: Resultado de testes de generalização da rede neural para casos de resposta aeroelástica não linear, amplitude 3° , frequência $8 \frac{rad}{s}$ e $U = 16,89 \frac{m}{s}$. Linha sólida: experimento; Linha tracejada: modelo neural.

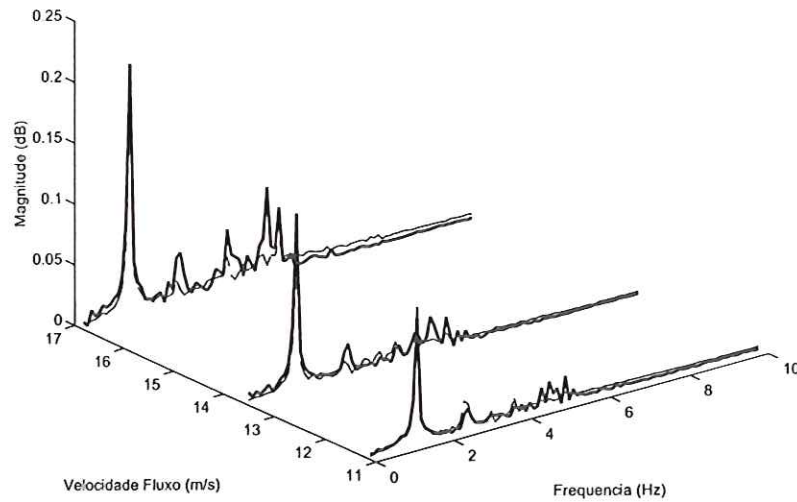


Figura 47: Resultado de testes de generalização da rede neural para casos de respostas aeroelásticas não lineares, amplitude 3° e frequência $8 \frac{rad}{s}$. Linha sólida: experimento; Linha tracejada: modelo neural.

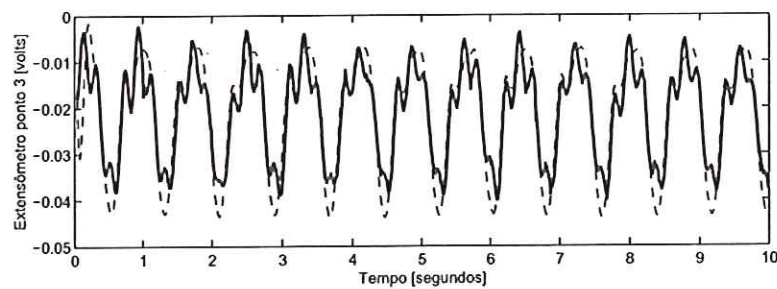


Figura 48: Resultado de teste de generalização da rede neural para caso de resposta aeroelástica não linear, amplitude $5,5^\circ$, frequência $8 \frac{rad}{s}$ e $U = 11,19 \frac{m}{s}$. Linha sólida: experimento; Linha tracejada: modelo neural.

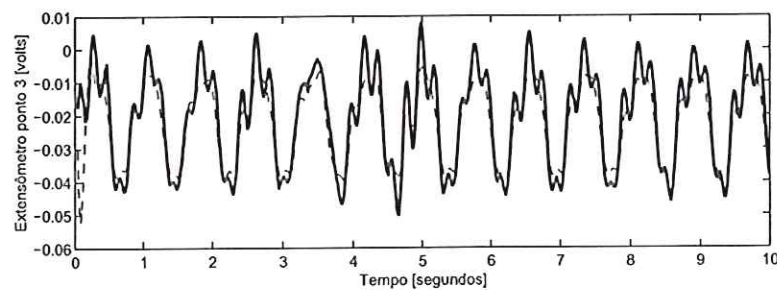


Figura 49: Resultado de teste de generalização da rede neural para caso de resposta aeroelástica não linear, amplitude $5,5^\circ$, frequência $8 \frac{rad}{s}$ e $U = 13,58 \frac{m}{s}$. Linha sólida: experimento; Linha tracejada: modelo neural.

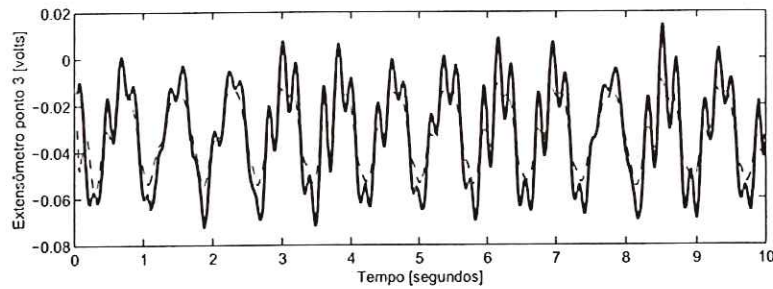


Figura 50: Resultado de teste de generalização da rede neural para caso de resposta aeroelástica não linear, amplitude $5,5^\circ$, frequência $8 \frac{rad}{s}$ e $U = 16,89 \frac{m}{s}$. Linha sólida: experimento; Linha tracejada: modelo neural.

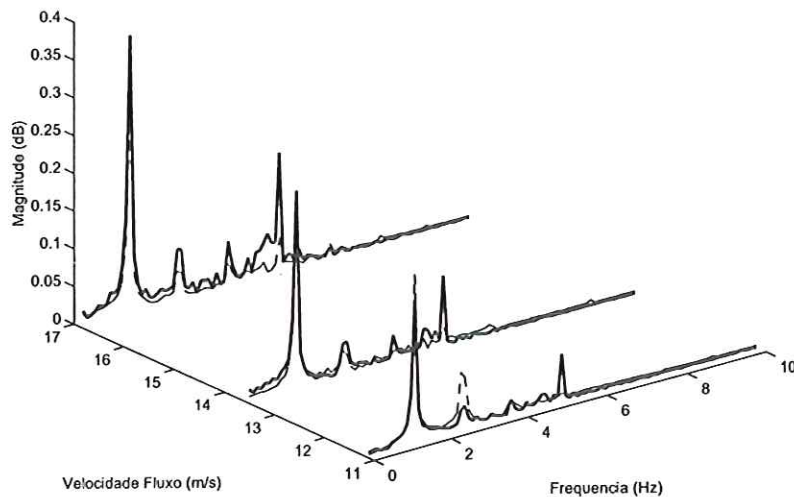


Figura 51: Resultado de testes de generalização da rede neural para casos de respostas aeroelásticas não lineares, amplitude $5,5^\circ$ e frequência $8 \frac{rad}{s}$. Linha sólida: experimento; Linha tracejada: modelo neural.

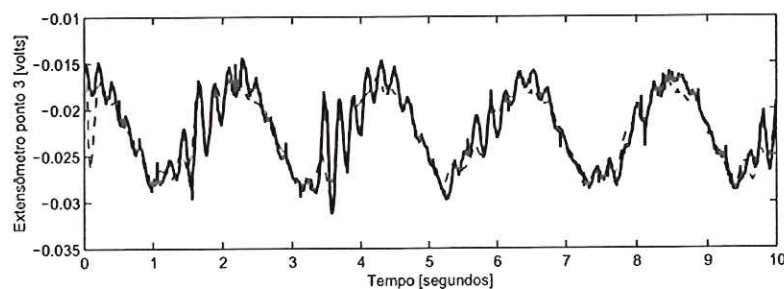


Figura 52: Resultado de teste de generalização da rede neural para caso de resposta aeroelástica não linear, amplitude 3° , frequência $3 \frac{rad}{s}$ e $U = 11,19 \frac{m}{s}$. Linha sólida: experimento; Linha tracejada: modelo neural.

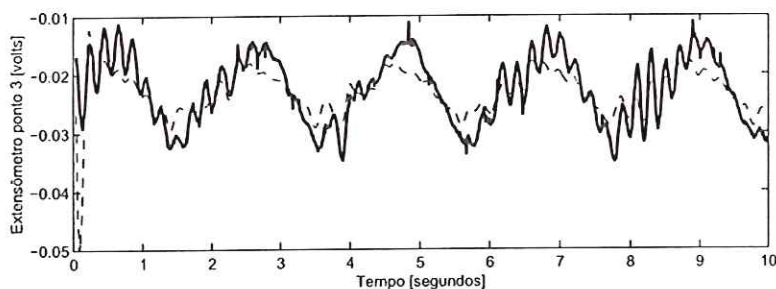


Figura 53: Resultado de teste de generalização da rede neural para caso de resposta aeroelástica não linear, amplitude 3° , frequência $3 \frac{rad}{s}$ e $U = 13,58 \frac{m}{s}$. Linha sólida: experimento; Linha tracejada: modelo neural.

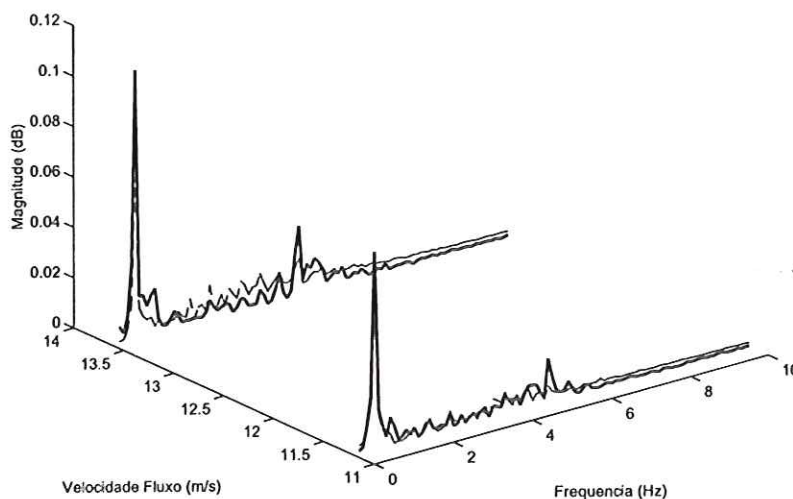


Figura 54: Resultado de testes de generalização da rede neural para casos de respostas aeroelásticas não lineares, amplitude 3° e frequência $3 \frac{rad}{s}$. Linha sólida: experimento; Linha tracejada: modelo neural.

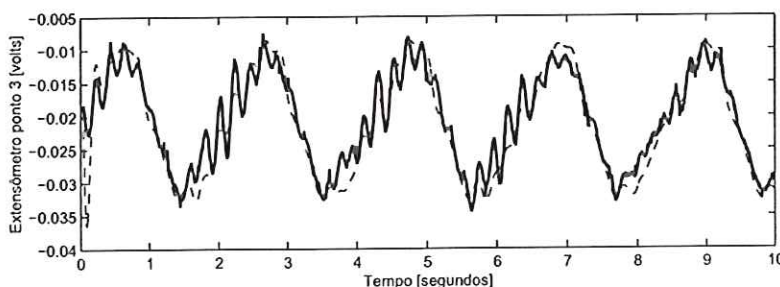


Figura 55: Resultado de teste de generalização da rede neural para caso de resposta aeroelástica não linear, amplitude $5,5^\circ$, frequência $3 \frac{rad}{s}$ e $U = 11,19 \frac{m}{s}$. Linha sólida: experimento; Linha tracejada: modelo neural.

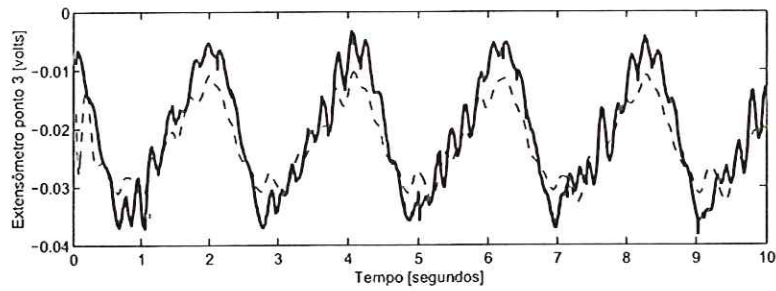


Figura 56: Resultado de teste de generalização da rede neural para caso de resposta aeroelástica não linear, amplitude $5,5^\circ$, frequência $3\frac{rad}{s}$ e $U = 13,58\frac{m}{s}$. Linha sólida: experimento; Linha tracejada: modelo neural.

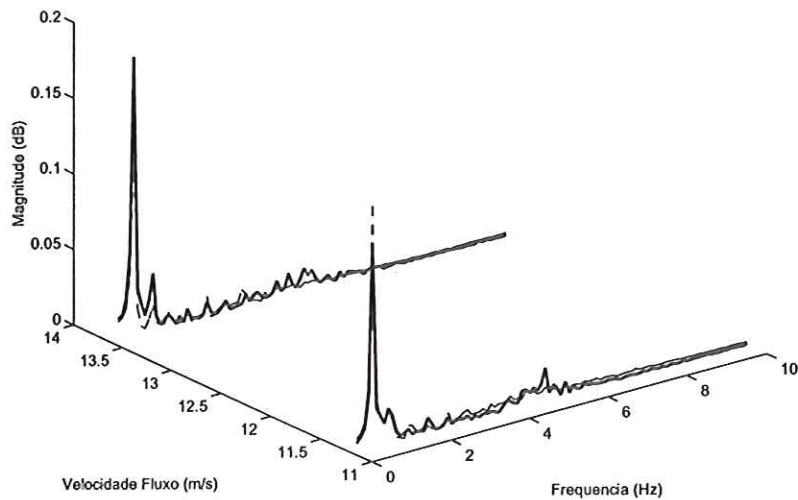


Figura 57: Resultado de testes de generalização da rede neural para casos de respostas aeroelásticas não lineares, amplitude $5,5^\circ$ e frequência $3\frac{rad}{s}$. Linha sólida: experimento; Linha tracejada: modelo neural.

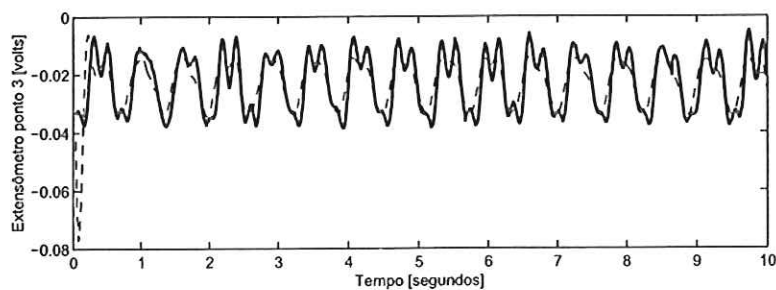


Figura 58: Resultado de teste de generalização da rede neural para caso de resposta aeroelástica não linear, amplitude 3° , frequência $10\frac{rad}{s}$ e $U = 13,58\frac{m}{s}$. Linha sólida: experimento; Linha tracejada: modelo neural.

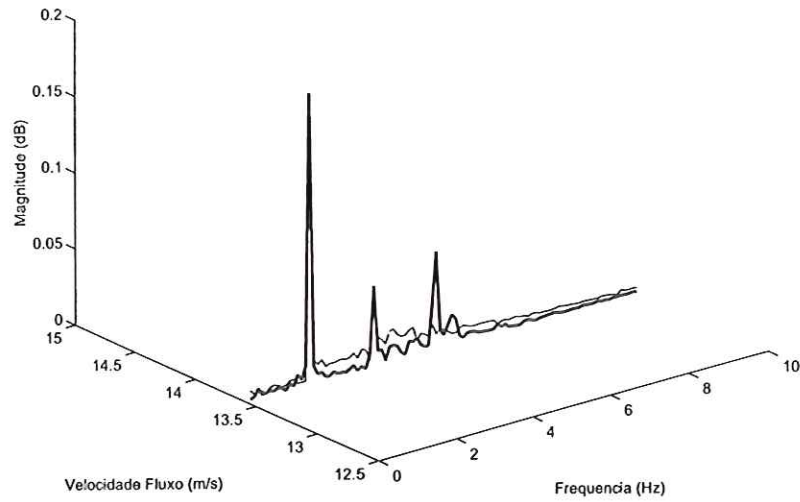


Figura 59: Resultado de teste de generalização da rede neural para caso de resposta aeroelástica não linear, amplitude 3° e frequência $10 \frac{rad}{s}$. Linha sólida: experimento; Linha tracejada: modelo neural.

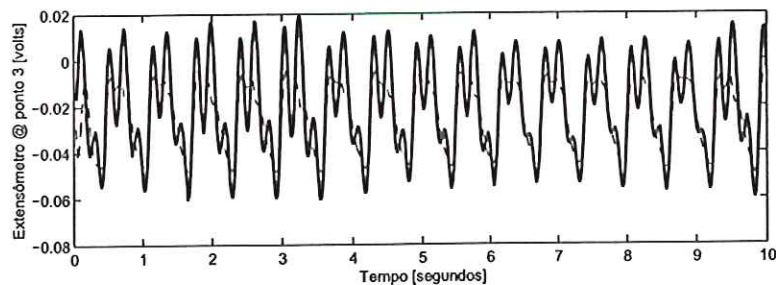


Figura 60: Resultado de teste de generalização da rede neural para caso de resposta aeroelástica não linear, amplitude $5,5^\circ$, frequência $10 \frac{rad}{s}$ e $U = 13,58 \frac{m}{s}$. Linha sólida: experimento; Linha tracejada: modelo neural.

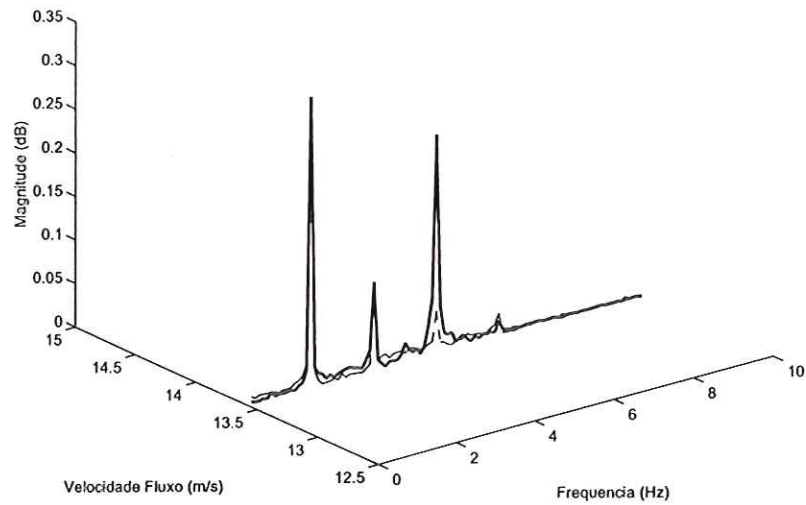


Figura 61: Resultado de teste de generalização da rede neural para caso de resposta aeroelástica não linear, amplitude $5,5^\circ$ e frequência $10 \frac{rad}{s}$. Linha sólida: experimento; Linha tracejada: modelo neural.

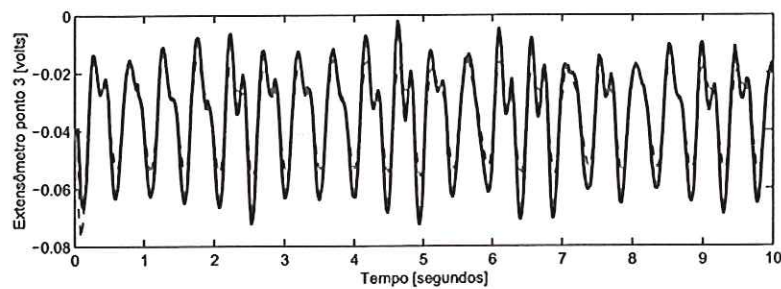


Figura 62: Resultado de teste de generalização da rede neural para caso de resposta aeroelástica não linear, amplitude 3° , frequência $13 \frac{rad}{s}$ e $U = 16,89 \frac{m}{s}$. Linha sólida: experimento; Linha tracejada: modelo neural.

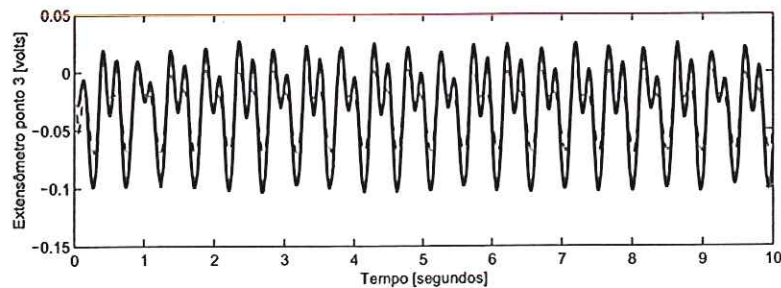


Figura 63: Resultado de teste de generalização da rede neural para caso de resposta aeroelástica não linear, amplitude $5,5^\circ$, frequência $13 \frac{rad}{s}$ e $U = 16,89 \frac{m}{s}$. Linha sólida: experimento; Linha tracejada: modelo neural.

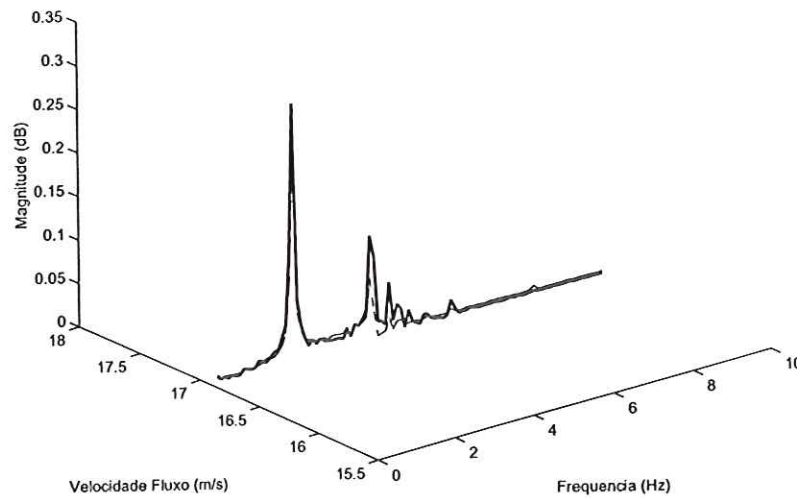


Figura 64: Resultado de teste de generalização da rede neural para caso de resposta aeroelástica não linear, amplitude 3° e frequência $13 \frac{rad}{s}$. Linha sólida: experimento; Linha tracejada: modelo neural.

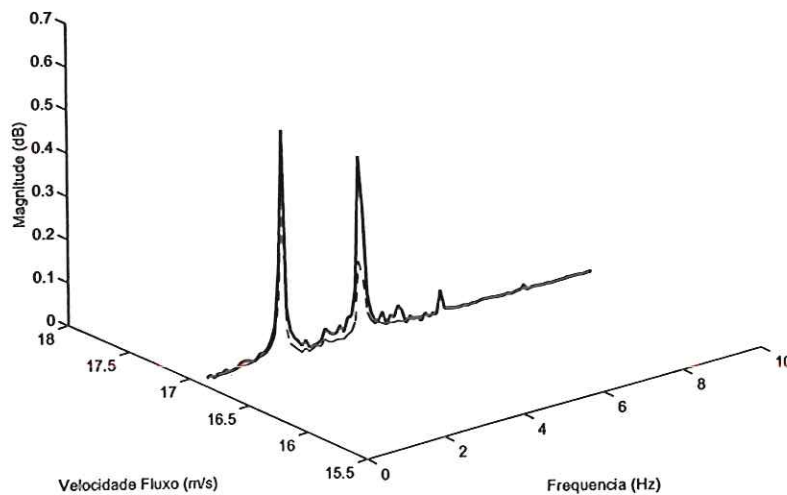


Figura 65: Resultado de teste de generalização da rede neural para caso de resposta aeroelástica não linear, amplitude $5,5^\circ$ e frequência $13 \frac{rad}{s}$. Linha sólida: experimento; Linha tracejada: modelo neural.

Analisando-se as Figuras 44 a 65 percebe-se que na maioria dos casos houve uma boa identificação. É possível perceber também que os fatores que mais negativamente influenciam na qualidade da identificação são a amplitude do sinal oscilatório e a velocidade do escoamento. Isso se

explica pelo fato de um aumento no valor desses fatores corresponder a um acréscimo na severidade da não linearidade e, conseqüentemente, a um aumento na complexidade da tarefa de identificação.

Assim como no caso linear, a rede consegue capturar a dinâmica do sistema apesar do ruído que há nos sinais. Isso é demonstrado nas respostas aeroelásticas da rede neural em frequência, onde na maior parte dos casos, a rede neural consegue identificar as principais frequências do sistema. Além disso, se a rede neural consegue reproduzir a dinâmica do sistema com menos ruído, pode-se dizer ela tem propriedades de filtragem.

Também como no caso linear, fato de a rede neural ter sido treinada com sinais de entrada aleatórios e conseguir generalizar para entradas oscilatórias indica que ela é sensível a mudanças no tipo de excitação. A rede neural consegue aproximar a saída do sistema mesmo quando excitada com um sinal significativamente diferente daquele que foi usado para treiná-la.

4 CONCLUSÕES E SUGESTÕES PARA TRABALHOS FUTUROS.

A modelagem da resposta aeroelástica não linear apresenta grandes desafios. Um modelo completo que possa calcular todos os efeitos aeroelásticos relevantes ainda não é viável em aplicações industriais. De fato, muitos dos esquemas de modelagem aeroelástica que servem tanto à pesquisa quanto à indústria, apresentam muitas simplificações e linearizações. Essas limitações inibem o projeto e análise do comportamento dinâmico de aviões que poderia considerar não linearidades, tais como em manobras severas com altos ângulos de ataque ou em regimes transônicos.

Uma alternativa para contornar tais problemas de modelagem foi apresentada nesse trabalho. Apresentou-se um modelo de ordem reduzida baseado no conceito de funcional multicamada representado por uma rede neural artificial. Os resultados obtidos mostram que o modelo neural utilizado foi capaz de representar o sistema aeroelástico de forma satisfatória, tanto para o caso linear quanto para o não linear.

Esses resultados confirmam as vantagens já conhecidas das redes neurais. A capacidade de realizar tarefas complexas de maneira simples e rápida é a vantagem que mais se evidencia. Além disso, o fato de uma rede neural funcionar como um modelo tipo caixa preta é uma desvantagem que não se aplica a esse trabalho visto que a rede neural aqui utilizada representa um funcional multicamada e, por esse motivo, tem seu funcionamento claramente explicado.

Uma outra forma funcional que pode ser combinada à uma rede neural artificial é a série de Volterra. A série funcional de Volterra ((VOLTERRA, 1959)), são expansões da definição de série de

Taylor para uma função que podem ser generalizadas para funcionais e consiste de um somatório infinito de integrais de convolução multidimensionais de ordem crescente.

Volterra (1959) provou que qualquer sistema contínuo, causal, não linear e invariante no tempo pode ser modelado por sua série. A grande desvantagem da série de Volterra, no entanto, é o cálculo de seus núcleos (Billings (1980) e Schetzen (1980)). Tentativas de desenvolvimento de uma aproximação sistemática para a identificação dos núcleos não forneceram um progresso substancial no campo de identificação não linear. No entanto, estudos recentes mostraram que as redes neurais artificiais podem ser utilizadas para obter os núcleos de Volterra. A relação entre redes neurais e série de Volterra foi previamente examinada de uma maneira rudimentar por David e Gasperi (1991), Govind e Ramamoorthy (1990), e foram sugeridos métodos para estimação indireta dos núcleos de Volterra, entre eles o trabalho de Wray e Green (1994). Nesse trabalho provou-se que uma arquitetura de rede particular pode ser equivalente a uma representação de série de Volterra de um sistema dinâmico. Além disso, os núcleos de qualquer ordem também podem ser extraídos dos parâmetros da rede.

Assim como os funcionais multicamadas representados por rede neurais, as séries de Volterra e as redes neurais são duas ferramentas matemáticas que se completam quando associadas. Por um lado um bom treinamento de uma rede neural pode levar a estimativas adequadas dos núcleos da série de Volterra. Por outro lado, a série de Volterra confere à rede neural uma interpretação científica, justificando sua utilização na identificação de sistemas. Sendo assim, o desenvolvimento de uma rede neural capaz de fornecer os núcleos de Volterra é um tema interessante para um trabalho futuro em identificação da resposta aeroelástica não linear.

REFERÊNCIAS

- AGUIRRE, L. A. *Introdução à identificação de sistemas - Técnicas lineares e não lineares aplicadas a sistemas reais*. [S.l.]: Editora UFMG, 2000.
- ANDERSON, J. D. *Fundamentals of aerodynamics*. [S.l.]: New York: McGraw-Hill, 1991.
- BHAT, N.; MCAVOY, T. J. Use of neural nets for dynamic modeling and control of chemical process systems. *Computers Chem. Engng.*, Printed in Great Britain, v. 14, p. 573-583, 1990.
- BILLINGS, S. A. Identification of nonlinear systems - A survey. *Proc. IEE, part D*, v. 6, p. 272-285, 1980.
- BISPLINGHOFF, R. L.; ASHLEY, H.; HALFMAN, R. L. *Aeroelasticity*. [S.l.]: Dover Edition, 1996.
- CHEN, H. M. et al. Neural network for structural dynamic model identification. *Journal of Engineering Mechanics*, p. 1377-1381, december 1995.
- CHEN, T.; CHEN, H. Approximation of continuous functionals by neural networks with application to dynamic systems. *IEEE Transactions on Neural Networks*, v. 4, 1993.
- CRAIG, R. R. J. *Structural Dynamics - an introduction to computer methods*. [S.l.]: Jhon Wiley & Sons, 1981.
- CYBENKO, G. Approximation by superpositions of a sigmoidal function. *Mathematics of Control, Signals, and Systems*, v. 2, 1989.
- DAVID, G. W.; GASPERI, M. L. ANN modelling of Volterra systems. In: *IJCNN'91*. [S.l.: s.n.], 1991.
- DOWELL, E. H. *A Modern Course in Aeroelasticity*. [S.l.]: Alphen aan den Rijn, 1978.
- DOWELL, E. H.; TANG, D. Nonlinear aeroelasticity and unsteady aerodynamics. *AIAA Journal*, v. 40, p. 1697-1707, 2002.
- FALLER, W. E.; SCHRECK, S. J. Unsteady fluid mechanics applications of neural networks. In: *Proc. of the 33rd Aerospace Sciences Meeting*. [S.l.]: AIAA paper 95-0529, 1995.
- FALLER, W. E.; SCHRECK, S. J. Neural networks: applications and opportunities in aeronautics. *Prog. Aerospace Sci.*, v. 32, p. 433-456, 1996.
- FALLER, W. E.; SCHRECK, S. J.; LUTTGES, M. W. Neural network prediction and control of three-dimensional unsteady separated flow fields. *J. Aircraft*, AIAA paper 94-0532, p. 1213-1220, 1995.

- GOVIND, G.; RAMAMOORTHY, P. A. Multilayered neural networks and Volterra series: The missing linking. In: *Proc. Int. Conf. Syst. Eng.* [S.l.: s.n.], 1990.
- GUPTA, P.; SINHA, N. K. An improved approach for nonlinear system identification using neural networks. *Journal of the Franklin Institute*, v. 336, p. 721–734, 1999.
- HAGAN, M. T.; MENHAJ, M. B. Training feedforward networks with the Marquardt algorithm. *IEEE Trans. on Neural Networks*, v. 15, p. 989–993, november 1994.
- HAGAN, T. M.; DEMUTH, B. H.; BEALE, M. *Neural Network Design*. [S.l.]: PWS Publishing Company, 1996.
- HAYKIN, S. *Neural networks - A comprehensive foundation*. [S.l.]: Macmillan College Publ., 1999.
- HEBB, D. O. *The Organization of Behavior*. [S.l.]: Willey, 1949.
- HORNIK, K.; STINCHCOMBE, M.; WHITE, H. Multilayer feedforward networks are universal approximators. *Neural Networks*, v. 2, 1989.
- HOUGHTON, E. L.; CARRUTHERS, N. B. *Aerodynamics for engineering students*. [S.l.]: Edward Arnold Ltd, 1982.
- HUANG, S.; HUANG, Y. Learning algorithms for perceptrons using back-propagation with selective updates. *IEEE Control Systems Magazine*, p. 56–61, April 1990.
- HUANG, S. Y.; MILLER, L. S.; STECK, J. E. An exploratory application of neural networks to airfoil design. *AIAA paper*, AIAA, 1994.
- HUNTER, I. W.; KORENBERG, M. J. The identification of nonlinear biological systems: Wiener and Hammerstein cascade models. *Biological Cybernetics*, v. 55, p. 135–144, 1986.
- JAGANNATHAN, S.; LEWIS, F. L. Identification of nonlinear dynamical systems using multilayered neural networks. *Automatica*, v. 32, 1996.
- KATZ, J.; PLOTKIN, A. *Low-speed aerodynamics: from wing theory to panel methods*. [S.l.]: New York: McGraw-Hill, 1991.
- KOSMATOPOULOS, E. B. et al. High-order neural network structures for identification of dynamical systems. *IEEE Trans. on Neural Networks*, v. 6, p. 422–431, march 1995.
- KOVACS, Z. L. *Redes neurais artificiais: fundamentos e aplicações*. [S.l.]: Editora Collegium Cognition e Edição Acadêmica, 1996.
- KU, C. S.; HAJELA, P. Neural-network-based controller for nonlinear aeroelastic system. *AIAA Journal*, v. 36, n. 2, p. 249–255, February 1998.
- LECUN, Y. Une procedure d'apprentissage pour reseau a seuil assymetrique. *Cognitive*, v. 85, 1985.
- LENNOX, B. et al. Industrial application of neural networks - an investigation. *Journal of Process Control*, v. 11, p. 497–507, 2001.

- LEVENBERG, K. A method for the solution of certain non-linear problems in least squares. *Quart. Appl. Math.*, v. 2, p. 164–168, 1944.
- LEVIN, A. U.; NARENDRA, K. S. Control of nonlinear dynamical systems using neural networks - Part II: observability, identification, and control. *IEEE Trans. on Neural Networks*, v. 7, p. 30–42, january 1996.
- LOU, K.; PEREZ, R. A. A new system identification technique using Kalman filtering and multilayer neural networks. *Artificial Intelligence in Engineering*, Elsevier Science Limited, v. 10, p. 1–8, 1996.
- MAHAJAN, A. J.; KAZA, K. R. V.; DOWELL, E. H. Semi-empirical model for prediction of unsteady forces on an airfoil with application to flutter. *Journal of Fluids and Structures*, v. 7, 1993.
- MARMARELIS, P. Z.; MARMARELIS, V. Z. *Analysis of physiological systems: the white noise approach*. [S.l.]: Plenum Press, 1978.
- MARQUARDT, D. W. An algorithm for least-square estimation of nonlinear parameters. *Journal of the Society for Industrial and Applied Mathematics*, v. 11, p. 431–441, june 1963.
- MARQUES, F. D. *Multi-layer functional approximation of non-linear unsteady aerodynamic response*. Tese (Doutorado) — University of Glasgow, 1997.
- MARQUES, F. D. Neural network based unsteady transonic loads reduced-order model. In: *43rd AIAA/ASME/ASCE/AHS/ACE Structures, Structural Dynamics, and Materials Conference*. [S.l.]: AIAA paper, 2002.
- MARQUES, F. D.; ANDERSON, J. Modeling and identification of non-linear unsteady aerodynamic loads by neural networks and genetic algorithm. In: *Proc. of the 20th Congress of the International Council of Aeronautical Sciences*. [S.l.]: ICAS 96, 1996. p. 243–251.
- MARQUES, F. D.; ANDERSON, J. Identification and prediction of unsteady transonic aerodynamics loads by multi-layer functionals. *Journal of Fluids and Structures*, v. 15, p. 83–106, 2001.
- MARZOCCA, P.; LIBRESCU, L.; SILVA, W. A. Aeroelastic response of nonlinear wing sections using a functional series technique. *AIAA*, v. 40, n. 5, p. 813–824, 2002.
- MASRI, S. F.; CHASSIAKOS, A. G.; CAUGLEY, T. K. Identification of nonlinear dynamic systems using neural networks. *Journal of Applied Mechanics*, v. 60, p. 123–133, march 1993.
- MEAD, A. J. J. An application of artificial neural networks to experimental data approximation. *AIAA paper 93-0408, 31 st Aerospace Science Meeting*, 1993.
- MODHA, D. S.; HECHT-NIELSEN, R. *Multilayers Functionals, Mathematical Approaches to Neural Networks*. [S.l.]: Elsevier Science Publishers, 1993.
- NARENDRA, S. K.; PARTHASARATY, K. Identification and control of dynamical systems using neural networks. *IEEE Trans. on Neural Networks*, v. 1, p. 4–27, march 1990.

- NARENDRA, S. K.; PARTHASARATY, K. Gradient methods for the optimization of dynamic-systems containing neural networks. *IEEE Trans. on Neural Networks*, v. 2, p. 252–262, March 1991.
- NGUYEN, D. H.; WIDROW, B. Neural networks for self-learning control systems. *IEEE Control Systems Magazine*, p. 18–23, april 1990.
- PAL, C. et al. Dynamic system identification by neural networks. *JSME International Journal Series C*, v. 38, p. 686–692, 1995.
- PALM, G.; POGGIO, T. The Volterra representation and the Wiener expansion: validity and pitfalls. *SIAM J. Appl. Math.*, v. 33, p. 195–216, 1977.
- PALM, G.; POPPEL, B. Volterra representation and Wiener-like identification of nonlinear systems: scope and limitations. *Q. Rev. Biophys.*, v. 18, 1985.
- PARKER, D. B. *Learning-logic: Casting the cortex of humain brain in silicon*. [S.l.], 1985.
- RAI, M. M.; MADAVAN, N. K. Application of artificial neural networks to the design of turbomachinery airfoils. *AIAA paper 98-1003*, 1998.
- RAI, M. M.; MADAVAN, N. K. Aerodynamic design using neural networks. *AIAA Journal*, v. 36, p. 173–182, january 2000.
- RAUCH, H. E. et al. Fault detection, isolation and reconfiguration for aircraft using neural networks. *AIAA 11th Applied Aerodynamics Conference*, 1993.
- REDDY, J. N. *An introduction to the finite element method*. [S.l.]: New York: McGraw-Hill, 1993.
- ROKHSAZ, K.; STECK, J. E. Use of neural networks in control of high alpha maneuvers. *J. Guidance, Control Dynamics*, v. 15, 1991.
- ROKHSAZ, K.; STECK, J. E.; SHUE, S. Longitudinal flight control decoupling using artificial neural networks. In: *32 nd Aerospace Sciences Meeting*. [S.l.]: AIAA paper 94-0274, 1994.
- RUGH, W. J. *Nonlinear system theory - The Volterra/Wiener approach*. [S.l.]: Johns Hopkins University Press, 1981.
- RUMELHART, D. E.; HINTON, G. E.; WILLIAMS, R. J. Learning internal representations of back-propagation errors. *Nature*, v. 323, 1986.
- SCHETZEN, M. *The Volterra and Wiener theories of nonlinear systems*. [S.l.]: John Wiley and Sons, 1980.
- SCHETZEN, M. Nonlinear system modeling based on the Wiener theory. In: *Proc. of the IEEE*. [S.l.: s.n.], 1981. v. 69, p. 1557–1573.
- SILVA, A. P. C.; KHATER, E. The Levenberg-Marquardt algorithm applied in nonlinear dynamic system. In: *Proc. of the V Brazilian Conference on Neural Networks*. [S.l.: s.n.], 2001.
- SILVA, W. Application to nonlinear systems theory to transonic unsteady aerodynamic responses. *J. Aircraft*, v. 30, p. 660–668, 1993.

- SILVA, W. A. Reduced-order models based on linear and nonlinear aerodynamic impulse responses. In: *Proc. of the 40th Structures, Struc. Dynamics, and Materials Conference*. [S.l.]: AIAA paper, 1999.
- SILVA, W. A. et al. Reduced-order modeling: cooperative research and development at the NASA Langley research center. In: *Proc. of the International Forum on Aeroelasticity and Structural Dynamics*. [S.l.]: IFASD paper, 2001.
- SJOBERG, J. et al. Non-linear black-box modeling in system identification: an unified overview. *Automatica*, v. 31, 1995.
- TOBAK, M.; PEARSON, W. E. *A study of nonlinear longitudinal dynamic stability*. [S.l.], 1964.
- TRAN, C. T.; PETOT, D. Semi-empirical model for the dynamic stall of airfoils in view of the application to the calculation of responses of a helicopter blade in forward flight. *Vertica*, v. 5, p. 35–53, 1981.
- TROUDET, T.; SANJAY, G.; MERRIL, W. C. Neural network application to aircraft control system design. *AIAA paper 91-2715-CP*, 1991.
- VOLTERRA, V. *Theory of functionals and of integral and integro-differential equations*. [S.l.]: Dover, 1959.
- WERBOS, J. P. Backpropagation through time: what it does and how to do it. *Proceedings of the IEEE*, v. 78, p. 1550–1560, October 1990.
- WERBOS, P. J. *Beyond regression: New tools for prediction and analysis in the behavioral sciences*. Tese (Doutorado) — Harvard University, 1974.
- WERBOS, P. J. Neural networks for control and system identification. In: *Proc. of the 28-th CDC*. [S.l.: s.n.], 1989.
- WIDROW, B.; HOFF, M. E. Adaptive switching circuits. *1960 IRE WESCON Convention Record*, p. 96–104, 1960.
- WIENER, N. *Nonlinear problems in random theory*. [S.l.]: MIT Press, 1958.
- WRAY, J.; GREEN, G. G. R. Calculation of the Volterra kernels of non-linear dynamic systems using an artificial neural network. *Biological Cybernetics*, v. 71, p. 187–195, 1994.