

REDES NEURAIIS PARA MODELAGEM DE SISTEMAS ESTRUTURAIS OFFSHORE  
DINÂMICOS NÃO-LINEARES COM HISTERESE

Djalene Maria Rocha

TESE SUBMETIDA AO CORPO DOCENTE DA COORDENAÇÃO DOS  
PROGRAMAS DE PÓS-GRADUAÇÃO DE ENGENHARIA DA UNIVERSIDADE  
FEDERAL DO RIO DE JANEIRO COMO PARTE DOS REQUISITOS NECESSÁRIOS  
PARA A OBTENÇÃO DO GRAU DE DOUTOR EM CIÊNCIAS EM ENGENHARIA  
CIVIL.

Aprovada por:

---

Prof. Nelson Francisco Favilla Ebecken, D.Sc.

---

Prof. Alexandre Gonçalves Evsukoff, Dr.

---

Profa. Beatriz de Souza Leite Pires de Lima, D.Sc.

---

Dr. Márcio Martins Mourelle, D.Sc.

---

Dr. Carlos Eduardo Luz Riodades Mendonça, D.Sc.

---

Prof. Luiz Pereira Calôba, Dr.Ing.

RIO DE JANEIRO, RJ – BRASIL

JUNHO DE 2007

# **Livros Grátis**

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

Milhares de livros grátis para download.

ROCHA, DJALENE MARIA

Redes Neurais para Modelagem de  
Sistemas Estruturais Offshore Dinâmicos não-  
Lineares com Histerese [Rio de Janeiro] 2007

X, 157 p. 29,7 cm (COPPE/UFRJ, D.Sc.,  
Engenharia Civil, 2007)

Tese – Universidade Federal do Rio de  
Janeiro, COPPE

1. Redes Neurais
2. Histerese
3. NARX

I. COPPE/UFRJ II. Título (série)

“Era um pacato cidadão sem documento  
Não tinha nome profissão não tinha tempo

Mas certo dia deu-se um caso  
E ele embarcou num disco  
E foi levado pra bem longe  
Do asterisco em que vivemos

Ele partiu e não voltou  
E não voltou porque não quis  
Quero dizer ficou por lá  
Já que por lá se é mais feliz

E um espacograma ele enviou  
Pra quem quisesse compreender  
Mas ninguém nunca decifrou  
O que ele nos mandou dizer  
Terra mar e ar atenção  
O futuro é hoje e cabe na palma da mão

Para azar de quem não sabe e não crê  
Que se pode sempre a sorte escolher  
E enterrar qualquer estrela no chão  
Viet vista visão viet vista visão

Terra mar e ar atenção  
Fica a morte por medida  
Fica a vida por prisão”

Torquato Neto

Dedico à minha querida mamãe, aos meus queridos irmãos e a Lindonor Mota.

## AGRADECIMENTOS

**Agradeço a:**

**Ao meu grande orientador e amigo Professor Nelson Ebecken**

**Dary Lottmar e Marcio Mourelle**

**Professor Luiz Pereira Calôba e Professor Alex Freitas**

**Álvaro Maia e Ricardo Beltrão (PETROBRAS)**

**Henrique Lima, Luciano Costa, Paulo Correia**

**Pulitha Liyanagama, Sebastien Marion ,Damjee Ediriweera e Myo**

**Remo Zauli , Meniconi, Marcio, Arleston e Dennis.**

**Eric Oliveira e Luis Manoel**

**Meus grandes amigos Mestre Amilcar, Beth Porto, Carlos Eduardo**

**(Dick), Claudia Claro, Cyranka (meu camarada), Maria Aparecida**

**(Cida), Lea Margarida e Raphael.**

**Minha queridas Carina Pires e Rita Mateus.**

**Carlos Bardana, Salvador, Erica e Pedro Mendes.**

**Dauro Noronha**

**Aos queridos Arleide Araújo e Eduardo Araújo.**

**Heitor Araújo, Beth Queija, Stael, Edmundo, Cláudio Amaral, Marcos**

**Donato , Nora, Carla, Thais , Guzzo, André Castelão, Alexandre Lins,**

**Lima e a toda a minha querida família PETROBRAS/MC.**

**Luis Henrique, Celso Ayabe, Cláudio Kenupp, Elesbão.**

**Meu paizão Milton Motta, Vô Icleia, minha afilhada de coração Renata**

**Porto, Aninha Porto ,Mari Porto e tio Regis Motta.**

**Ana Lucia Pimentel, Daniela, Patrícia Rocha e Flavia Maximo.**

**Ao meu pedreiro Carlos.**

**Minha mãe Lucia, meus irmãos Durvalina, Djenal, Daisy, Denise e**

**Daniella. E também a todos os meus sobrinhos e associados.**

**Lindonor Mota, pelo carinho de pai.**

Resumo da Tese apresentada à COPPE/UFRJ como parte dos requisitos necessários para a obtenção do grau de Doutor em Ciências (D.Sc.)

REDES NEURAIS PARA MODELAGEM DE SISTEMAS ESTRUTURAIS OFFSHORE  
DINÂMICOS NÃO LINEARES COM HISTERESE

Djalene Maria Rocha

Junho/2007

Orientador: Nelson Francisco Favilla Ebecken

Programa: Engenharia Civil

Este trabalho propõe um modelo empírico para o sistema com histerese formado por um tipo de conexão elastomérica riser-plataforma, em águas profundas. A conexão tem o objetivo de minimizar os momentos fletores atuantes, possui grande complexidade e alta criticidade devido às conseqüências econômicas e ambientais decorrentes de sua falha. O elemento principal da junta é composto de material elastomérico com comportamento histerético não-linear. Além disso, o sistema composto por essa conexão também apresenta não-linearidades devido à ação das cargas dinâmicas e devido aos grandes deslocamentos. Redes Neurais TDNN e Recorrentes são investigadas por possuírem a habilidade de modelar comportamentos histeréticos não-lineares, bem como de sistemas dinâmicos. Verificou-se que o modelo de Redes Neurais Recorrentes apresenta os melhores resultados para o sistema estudado. Devem ser ressaltadas ainda as dificuldades adicionais ao processo de modelagem deste sistema relacionadas ao emprego dos dados reais.

Abstract of Thesis presented to COPPE/UFRJ as a partial fulfillment of the requirements for the degree of Doctor of Science (D.Sc.)

NEURAL NETWORK MODELLING OF THE NONLINEAR DYNAMIC  
STRUCTUTAL OFFSHORE SYSTEM WITH HISTERESIS

Djalene Maria Rocha

June/2007

Advisor: Nelson Francisco Favilla Ebecken

Department: Civil Engineering

This work proposes an empirical modeling of the system formed by the riser-platform connection, in deep water. This connection has the objective of minimizing the acting bending moment, possesses high complexity and high-criticality due to economic and environmental consequences from its fault. The main element in the joint is made of elastomeric material, which reveals nonlinear hysteresis. In addition, this whole connection system presents nonlinearities due to the action of dynamic loading and large motions. TDNN and Recurrent Neural Networks are being investigated because they possess the ability to model nonlinear hysteretic behaviors and also dynamic systems. Simulation results have confirmed that Recurrent Neural Network is the one that presents the best representation of the system studied. Emphasis shall be given to the additional difficulties, which arise from the utilization of real data in the modeling process for this system.

## SUMÁRIO

<b>CAPÍTULO I - INTRODUÇÃO</b>	<b>01</b>
1.1 - Objetivo	02
1.2 – Motivação	04
1.3 – Importância do trabalho	06
1.4 – Organização do trabalho	07
<b>CAPITULO II - DESCRIÇÃO DO SISTEMA FLEXJOINT</b>	<b>09</b>
2.1 - Sistemas flutuantes de produção de petróleo e gás	10
2.2 - Tipos de risers de produção	12
2.2.1 - Risers rígidos verticais	13
2.2.2 - Risers flexíveis	14
2.2.3 - Risers rígidos em catenária (SCRs)	15
2.3 - Flexjoint	19
2.3.1 - Características da flexjoint	20
2.3.1.1 - Conceito de rigidez	20
2.3.1.2 – Resposta dinâmica da junta elástica	22
2.3.2 - Motivação para o trabalho	
<b>2.4 - Modelos da flexjoint</b>	<b>23</b>
2.4.1 - Modelo empírico experimental	23
2.4.2 - Modelo por elementos finitos	25
<b>CAPÍTULO III - FUNDAMENTOS TEÓRICOS</b>	<b>26</b>
3.1 - Modelagem de Sistemas Dinâmicos	27
3.1.1 - Introdução	27
3.1.2 - Conceitos básicos	27
3.1.3 - Finalidade da modelagem de Sistemas	30
3.1.4 - Classificação dos modelos matemáticos	31
3.1.5 - Representação de sistemas dinâmicos	34
3.1.5.1 - Representação de sistemas dinâmicos em entrada-saída.	34
3.1.5.2 - Representação de sistemas dinâmicos em espaço de estados.	35



3.1.6 - Identificação de Sistemas Dinâmicos	36
3.1.7 - Redes Neurais na Identificação de Sistemas Dinâmicos não lineares	38
3.2 - Redes neurais artificiais	41
3.2.1 - Introdução	41
3.2.2 – Histórico	42
3.2.3 - O neurônio biológico e artificial	43
3.2.4 – Modelo genérico de um neurônio	44
3.2.5 - Arquitetura de uma rede neural	46
3.2.5.1 - Redes neurais de uma camada	46
3.2.5.2 - Redes neurais multi-camadas	48
3.2.6 - Aprendizagem	50
3.2.6.1 - Algoritmo backpropagation	51
3.2.6.2 - Overfitting e early-stopping	52
3.2.6.3 - Conjunto de teste	54
3.2.7 - Redes Neurais dinâmicas	54
3.2.8 - TDNN (Time Delay Neural Network)	56
3.2.9 - Recorrentes	58
3.2.9.1 - Modelo de espaço de estados	58
3.2.9.2 Jordan	60
3.2.9.3 Modelo recorrente de entrada/saída	61
3.3 Sinais Aleatórios	63
3.3.1 - Descrição dos parâmetros	63
3.4 - Histerese	66
<b>CAPÍTULO IV - DESCRIÇÃO DOS DADOS</b>	<b>68</b>
4.1 – Introdução	68
4.2 - Sub-sistema de monitoração de esforços e vibrações	70
4.3 - Variáveis do problema	72
4.4 - Características dos dados e período de análise	76
4.5 - Análise exploratória dos dados no domínio do tempo e da frequência	77
4.5.1 - Análise no domínio do tempo	77
4.5.1.1 Correlação	85

4.5.2 - No domínio da frequência	86
4.5.2.1 - Análise de cluster	92
4.6 - Dados de ensaios de laboratório	95
<b>CAPÍTULO V - SELEÇÃO DE METODOLOGIA PARA MODELAGEM DA HISTERESE COM REDES NEURAI</b>	<b>97</b>
5.1- Introdução	97
5.2 – Modelagem da histerese	98
5.3 – Preparação dos dados	99
5.4 – Seleção da estrutura e algoritmo de aprendizado	102
5.5 – Os modelos	104
5.5.1 – Backpropagation convencional	104
5.5.1 – Rede neural de atrasos de tempo (TDNN)	106
5.5.2.1 - Determinação do tamanho da janela temporal	109
5.5.3 – Redes Recorrentes NARX	111
5.5.3.1 – Arquitetura NARX paralelo	112
5.5.3.2 - Arquitetura NARX série - paralelo	115
5.6 – Comparação dos modelos propostos	118
<b>CAPÍTULO VI - ANÁLISES E RESULTADOS</b>	<b>120</b>
6.1 – Introdução	120
6.2 – O modelo de identificação	120
6.2.1 – As variáveis dos modelos	121
6.2.2- Definição dos atrasos de tempo e parâmetro dos modelos	123
6.3 – Conjuntos de treinamento, validação e teste	125
6.4 – Comparação dos modelos série-paralelo e paralelo	131
6.5 – Influência das variáveis externas nos modelos	136
6.5.1 – Amostragem e filtragem	137
6.5.2 – Análise da influência das variáveis externas com os dados de laboratório	144
6.6 - Resumo	147
<b>CAPÍTULO VII – CONCLUSÕES E RECOMENDAÇÕES</b>	<b>148</b>
7.1 – Conclusões	148
7.2 - Recomendações para trabalhos futuros	150



## **CAPÍTULO I - INTRODUÇÃO**

O objetivo da modelagem de sistemas é identificar e representar as leis que governam os fenômenos analisados, de uma maneira matemática e precisa. Entretanto, reconhece-se que essas descrições, conhecidas como perfeitas, nem sempre são as mais adequadas. A evolução da teoria probabilística e da estatística clássica fornece meios teóricos para a descrição de incertezas, tomando como base eventos repetitivos e com longos prazos.

Recentemente, os modelos desenvolvidos através de observações têm recebido especial importância. Principalmente com o desenvolvimento de tecnologias que se valem da investigação do conhecimento armazenado em Banco de Dados. Esses modelos têm como grande vantagem sua robustez pois são tolerantes a ruídos e ausência de dados, além de poderem ser acoplados a outros modelos, fornecendo uma solução híbrida.

Os fenômenos da natureza, por envolverem grande heterogeneidade de propriedades, contornos indefinidos, condições iniciais desconhecidas e complexas interações são ideais para aplicação dessas tecnologias. Surge aí a importância dos Modelos

Baseados em Dados. Para construir esses modelos não são necessárias complexa matemática e hipóteses simplificadoras, bastam dados em número suficiente para serem representativos do fenômeno.

A extração de conhecimento explícito e compreensivo a partir de dados conhecidos, através da técnica conhecida como Mineração de Dados (DATA MINING), gera informações preciosas. É possível ainda complementar estas informações com o conhecimento obtido de especialistas. A tecnologia de DATA MINING abriu um novo campo de oportunidades, modificando profundamente as estratégias de solução de problemas. Os resultados obtidos englobam:

1. Explicitação do conhecimento através de regras e correlações;
2. Sensibilidade do comportamento em relação aos parâmetros relevantes;
3. Avaliação do conhecimento obtido através de comparação com o conhecimento existente, para entendimento do interesse da análise - construção de um modelo preditivo para utilização em futuras análises e previsões de comportamentos.

## **1.1 - O objetivo**

A PETROBRAS foi pioneira ao instalar um *riser* rígido em catenária (Steel Catenary Riser – SCR) de 10” na plataforma semi-submersível P-XVIII. A instalação desse protótipo teve como motivação a preocupação quanto a viabilidade técnica e econômica no uso de linhas flexíveis, solução predominante, para grandes diâmetros em águas ultraprofundas [MACHADO *et al*,1999]. Esta unidade de produção de petróleo está ancorada em 910 m de profundidade d’água, no Campo de Marlim, na Bacia de Campos, Rio de Janeiro( (ver Figura 1.1). Este SCR foi instalado com a finalidade de transporte de gás entre as unidades P-18 e P-26. Dado seu pioneirismo, este SCR foi monitorado, em tempo real, para aquisição de dados referentes a:

1. Posicionamento da Plataforma;

2. Condições meteorológicas e oceanográficas;
3. Posicionamento do SCR;
4. Esforços e vibrações no SCR.

O SCR foi instalado em 1996 e continua em operação com sucesso até a presente data.



**Figura 1.1 – Campo de Marlim, na Bacia de Campos**

O objetivo desta tese é a modelagem de sistemas não lineares dinâmicos que exibem histerese através de redes neurais, utilizando dados obtidos através de processos de monitoração de estruturas offshore, em escala real, analisando e comparando vantagens e desvantagens das opções disponíveis. O sistema específico em estudo é aquele que compõe a conexão entre o riser e a plataforma, que é realizada por um equipamento denominado *Flexjoint*. A *Flexjoint* (ver Figura 1.2) é um equipamento instalado no topo do SCR e que, através de uma junta elastomérica, permite que este flexione quando sujeito a

carregamentos dinâmicos ambientais de onda, vento e correnteza. A finalidade da aplicação é de ampliar o entendimento do comportamento histerético da referida junta.

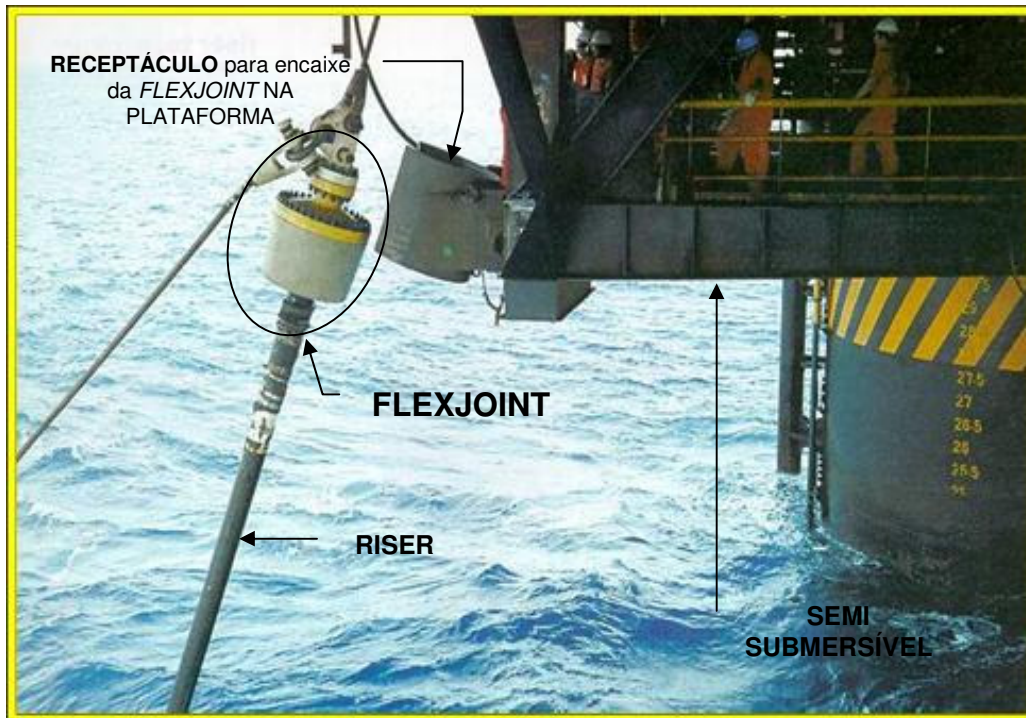


Figura 1.2 – *Flexjoint* da plataforma P-XVIII

## 1.2 - Motivação

Uma eventual degeneração do elastômero pode comprometer o desempenho do equipamento, embora ainda não seja possível concluir se no sentido do aumento ou da redução da rigidez. Contudo, pode-se prever que, na hipótese de enrijecimento da conexão do *riser* com a plataforma, os danos de fadiga seriam agravados na região próxima ao topo, reduzindo sua vida útil.

Através de inspeção local, verificou-se que o elastômero da *flexjoint* da P-XVIII apresenta algumas feições que podem estar associadas a um processo de deterioração

preocupante, possivelmente com reflexo na rigidez do elastômero, o que pode reduzir consideravelmente a vida útil do *riser* devido à fadiga na região próxima ao topo.

Existe alguma incerteza na formulação da rigidez fornecida pelo fabricante devido às premissas admitidas nas análises de laboratório. Por sua vez, a grande complexidade do sistema e a carência de informações sobre seu mecanismo dificultam a modelagem matemática através do conhecimento das leis que governam o processo do sistema.

O monitoramento de variáveis fundamentais pode nos conduzir a resultados mais amplos e mais acurados do que os obtidos em laboratórios. Isso possibilita ampliar o conhecimento sobre como funciona a *flexjoint* e, conseqüentemente, estabelecer estratégias para novas aquisições e medidas eficientes de controle das unidades em operação. E, além disso, permitir o aprimoramento de modelos matemáticos convencionais.

Dentro desse contexto, a identificação do sistema *flexjoint* pode ser bastante útil para se explorar suas características e propriedades, principalmente no tocante à rigidez rotacional.

Devido a complexidade inerente à formulação de um modelo matemático baseado em leis e princípios físicos para representação do sistema proposto e da disponibilidade de dados, técnicas de identificação de sistemas foram utilizadas para a modelagem do sistema. Neste trabalho a técnica de identificação empregada foi a de Redes Neurais.

Como alternativa a soluções matemáticas clássicas, modelos baseados em dados, como Redes Neurais, têm sido atualmente bastante empregados em diversas áreas, inclusive em problemas de Engenharia. Uma das mais importantes vantagens das Redes Neurais é a de resolver problemas do mundo real com considerável complexidade, cuja solução algorítmica ainda não existe ou é bastante complexa e dispendiosa para ser encontrada. Resumidamente, as Redes Neurais mapeiam entradas em saídas, em função de uma coleção de exemplos representativos do mapeamento entrada-saída desejado.

Para representar o comportamento dinâmico de um sistema sob histerese é necessário que o modelo, através de algum mecanismo, saiba em que estágio do carregamento se está num dado momento, ou seja, a rede precisa ter memória dinâmica. Existem muitos modelos que implementam memória dinâmica na arquitetura de rede neural. Por exemplo, a rede recorrente simples (RRS) e redes Jordan descrita em



[Elman,1990]; a arquitetura de rede recorrente treinada com retropropagação ao longo do tempo (backpropagation through time -BPTT) de (Werbos, P. J,1990); o modelo de rede com atraso de tempo de (Lang et al.,1988), entre tantas outras. É importante comentar que a maioria dos trabalhos em identificação de sistemas empregam redes estáticas *feedforward* multicamadas com algoritmo de aprendizado *backpropagation* ou variações mais eficientes deste algoritmo, que capturam a dinâmica do sistema real através de linhas de atrasos das entradas e saídas [Narendra& Parthasarathy,1990] e [Haykin,2004]. A memória de linha de atraso é então o mecanismo empregado para armazenar o conhecimento presente e histórico do processo real. Essa será a forma implementada neste trabalho. Os parâmetros levados em consideração para esse estudo são:

- O ângulo de inclinação da FlexJoint, como variável de saída.
- Esforço axial, como variável de entrada;
- Momento Fletor ,como variável de entrada;

### **1.3 - Importância do trabalho**

O desenvolvimento de novas tecnologias para a exploração e produção de petróleo em águas ultraprofundas é um desafio técnico e econômico. Esforços conjuntos dos fabricantes de tecnologias e empresas petrolíferas têm sido feitos a fim de viabilizar e aprimorar sistemas, equipamentos e unidades de produção. A modelagem proposta neste trabalho visa promover um melhor entendimento do sistema real formado pela conexão entre *riser* e plataforma através da flexjoint, e aperfeiçoar a tecnologia, conduzindo a melhorias que permitam seu uso em larga escala de forma segura e economicamente viável. A utilização de equipamentos mais seguros é imprescindível para evitar os impactos ambientais causados pela atividade da indústria petrolífera. O controle de acidentes, além de evitar o prejuízo ao meio ambiente, impede que ocorram gastos extravagantes relacionados à interrupção da produção. Além disso, a introdução de novas ferramentas de modelagem, tais como as redes neurais, possibilita o aperfeiçoamento de modelos

convencionais já existentes e abre espaço para modelagens antes nunca realizadas, pois as redes neurais não exigem conhecimento prévio sobre o comportamento interno dos sistemas a serem modelados.

## 1.4 - Organização do trabalho

Além deste capítulo introdutório, o trabalho é estruturado nos seguintes capítulos:

O **capítulo 2** descreve os sistemas de produção para águas profundas e ultraprofundas destacando a importância de investimento em novas tecnologias tais como os SCRs. Dentro desses sistemas é de particular interesse estudar a *Flexjoint*, seus principais componentes e propriedades, porque é neste componente que se encontra a motivação para este trabalho. Adicionalmente descreve-se sucintamente alguns dos modelos já existentes.

O **capítulo 3** descreve os conceitos teóricos fundamentais utilizados no presente estudo:

Modelagem de Sistemas dinâmicos

- Redes Neurais
  - *Backpropagation*
  - Redes TDNN (*Time Delay Neural Network*)
  - Redes Recorrentes
- Histerese
- Sinais aleatórios

O **capítulo 4** descreve os dados aqusitados para o estudo, a forma como foram coletados, bem como algumas análises estatísticas descritivas, correlações e análises no domínio do tempo e da frequência.

No **capítulo 5** trata basicamente de sistemas não lineares histeréticos usando dados obtidos com ensaios no elastômero da *flexjoint*. São discutidos três tipos de redes neurais

quanto a modelagem do sistema: a *feedforward MLP*, TDNN e recorrentes. O objetivo deste capítulo é estudar o comportamento desses tipos de rede a partir de dados mais bem comportados, o que servirá de base para análise da flexjoint a partir de dados obtidos em produção.

No **capítulo 6** a partir dos modelos implementados no capítulo 5, estudou-se os modelos com base nos dados reais, analisando vantagens e desvantagens, comparando-se o desempenho. Uma dessas análises foi a determinação da influência das variáveis externas do modelo.

Finalmente o **capítulo 7** é reservado para as conclusões e algumas propostas para a sua continuação.

## **CAPÍTULO II - DESCRIÇÃO DO SISTEMA *FLEXJOINT***

Neste capítulo, serão abordados os principais tópicos relativos à tecnologia de exploração de petróleo em águas profundas e ultraprofundas, tipos de flutuantes e *risers*. O foco do capítulo é a *flexjoint*, equipamento que conecta o *riser* ao flutuante (plataforma) e minimiza a transferência dos movimentos de rotação desta para o topo do *riser*. Serão abordadas suas principais características, propriedades e dificuldades encontradas nos modelos existentes que a representam.

## 2.1 - Sistemas flutuantes de produção de petróleo e gás

Inicialmente a extração de petróleo *offshore* no Brasil era efetuada em lâmina d'água denominadas rasas, com profundidades variando de 100m a 500m. Em função da possibilidade de esgotamento dos recursos petrolíferos e com o avanço de novas tecnologias, novas fontes de petróleo foram descobertas em laminas d'água cada vez mais profundas, conforme ilustra a Figura 2.1.

A indústria de petróleo tem investido cada vez mais no desenvolvimento de sistemas flutuantes de produção mais versáteis e projetados para condições ambientais diversas e em águas profundas e ultra-profundas. Regiões como Brasil, golfo do México, noroeste da Europa e oeste da África são líderes na utilização de sistemas s unidades de produção, além daquelas que já estão sendo projetadas para futuras instalações.

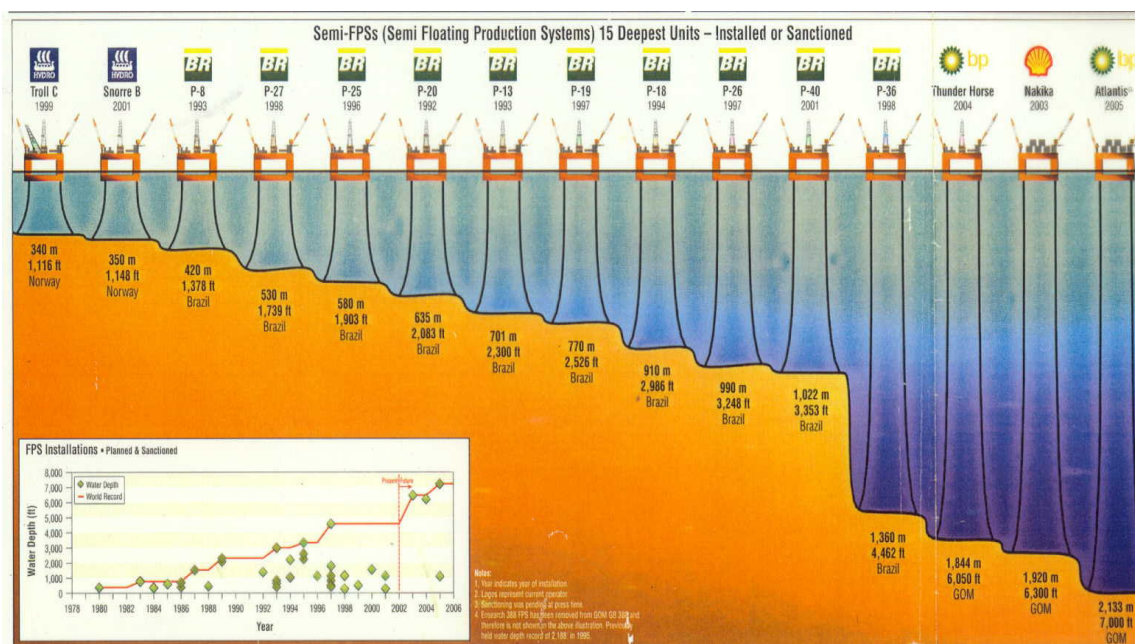
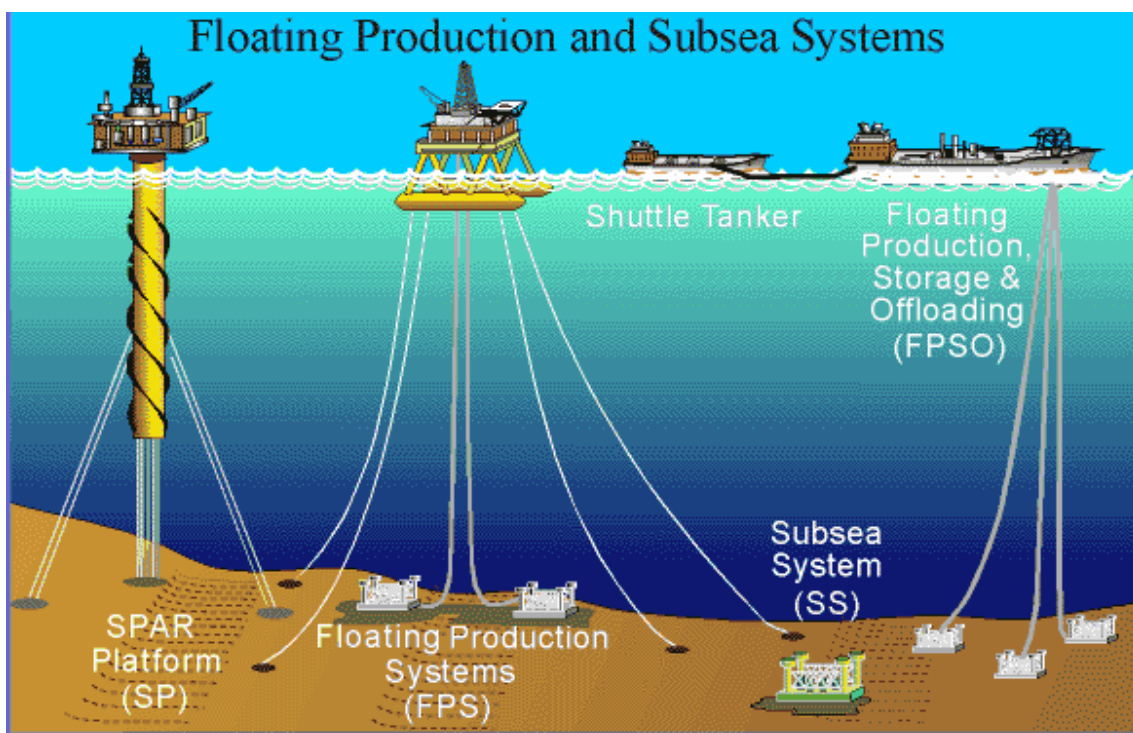


Figura 2.1 – Evolução das profundidades da lâmina d'água (Fonte – *OffshoreMagazine*).

Os tipos mais comuns de sistemas de produção mostrados na Figura 2.2 são: SPAR, TLP (Tension-Leg Platforms), FPSO (Floating Production, Storage and Offloading System), FSO (Floating Storage and Offloading System) e semi-submersíveis.



**Figura 2.2 – Modelos simbólicos de sistemas de produção de petróleo**

No Brasil, até 2006, encontram-se instaladas 22 plataformas semi-submersíveis, 8 FPSO e 1 FSO. Os sistemas flutuantes de produção são definidos da seguinte forma:

A SPAR tem um casco em forma de cilindro alongado, que flutua na posição vertical, o que lhe proporciona uma grande capacidade de estocagem de óleo, além de respostas de movimentos verticais (*heave*) minimizadas. Esta característica permite a possibilidade de utilização de *risers* rígidos acoplados à unidade.

A TLP (*Tension Leg Platform*) é constituída de estrutura com várias colunas, e caracteriza-se por um sistema de ancoragem vertical, através de cabos tensionados, que eliminam a resposta vertical da embarcação, possibilitando, também, a utilização de *risers* rígidos, porém tornando a embarcação bastante sensível à variação de peso. Este tipo de embarcação é totalmente diferente dos outros, uma vez que seus movimentos são modificados pela ancoragem.

O FPSO, ou monocasco, é constituído de um casco convencional em forma de navio e caracteriza-se pela grande capacidade de estocagem e pelas respostas aos movimentos convencionais de embarcações desta natureza. Este tipo de unidade de produção requer a utilização de *risers* flexíveis.

O casco da semi-submersível é formado por três elementos distintos: convés, colunas e *pontoons*. A característica principal deste tipo de embarcação é a pequena área de linha d'água, que faz

com que os seus movimentos tenham períodos naturais bem maiores do que os de outras embarcações (monocasco, por exemplo), gerando movimentos verticais (*heave*) bem menores. Ainda assim, este tipo de embarcação requer o uso de *risers* flexíveis, na maioria dos casos.

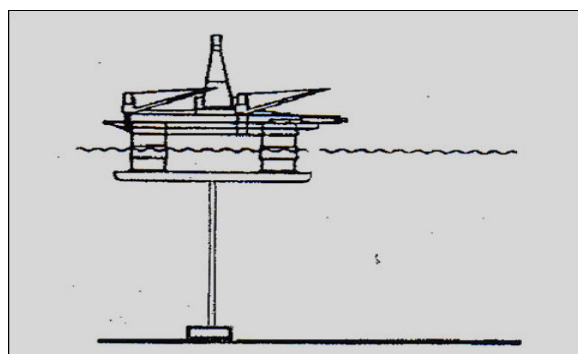
Os sistemas flutuantes de produção têm como componentes principais o casco, os *risers* e o poço. A seguir, é feita uma breve descrição dos principais tipos de *risers* empregados.

## 2.2 - Tipos de *risers* de produção

O *riser* de produção é um elemento tubular que interliga a cabeça do poço petrolífero ao flutuante na superfície do mar. É um duto esbelto que, além do movimento induzido pelo próprio flutuante, está sujeito à ação de carregamento devido ao vento, onda e corrente marítima. Pode desempenhar funções específicas como: permitir a passagem de ferramentas e conduzir fluidos dos poços até a plataforma durante a exploração. Atualmente, os tipos de *riser* mais empregados são: *riser* rígido vertical, linhas ou *risers* flexíveis e SCRs (*Steel Catenary Riser*).

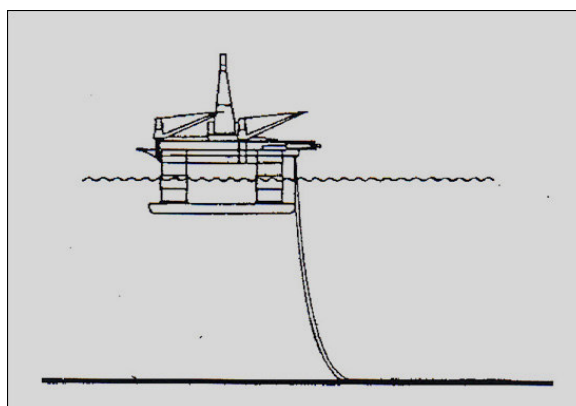
Existem vários modelos de configuração de linhas possíveis para a utilização em águas profundas:

1. Vertical: é aplicada uma força de tração no topo com a finalidade de manter o *riser* tracionado, evitando a flambagem (Figura 2.3)



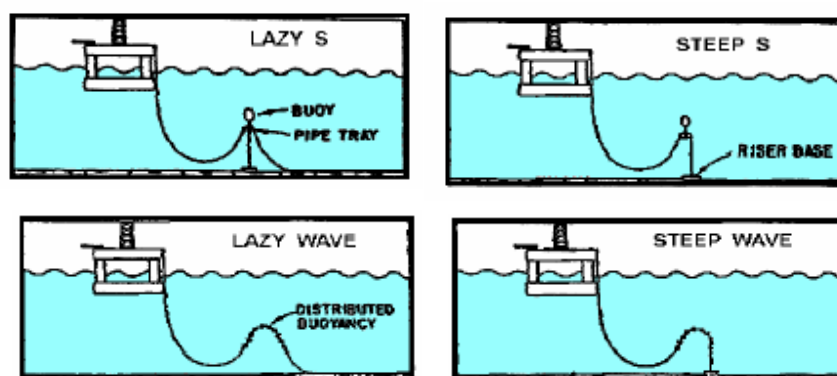
**Figura 2.3 – Riser rígido vertical**

2. Catenária: na maioria dos casos, não é aplicada força de tração no topo. As extremidades (topo e fundo) não estão no mesmo alinhamento vertical (Figura 2.4).



**Figura 2.4 – Riser rígido em catenária**

3. Complexa: derivada da configuração em catenária, o *riser* assume uma geometria em forma de catenária dupla através da instalação de flutuadores ou bóias mantidas submersas com poitas (Figura 2.5).



**Figura 2.5 – Vários esquemas de suspensão do *riser* de produção**

O critério de escolha do tipo de *riser* dependerá principalmente do tipo do flutuante, lâmina d'água, ancoragem, tipo de operação e custo. A seguir, serão descritos os tipos de *risers* mais conhecidos.

### 2.2.1 - *Risers* rígidos verticais

*Risers* rígidos são tubos de aço de aproximadamente 12 metros de comprimento, acoplados uns aos outros por processo de soldagem ou por conectores mecânicos. As limitações no comprimento



contínuo máximo dos *risers* rígidos dependem do processo de fabricação, transporte, manuseio e instalação. Podem estar envolvidos por flutuadores para diminuir o peso, quando em lâminas d'água profundas.

Este tipo de *riser* é usualmente empregado nas atividades de perfuração, completação e *workover*, que podem ser exercidas por unidades DP, capazes de manter sua posição limitada a pequenos deslocamentos. Também é utilizado em unidades de produção baseadas em completação seca (*Spar buoy* e TLP), ou seja, em que o controle do poço está na unidade e não no fundo do mar.

### 2.2.2 - *Risers* flexíveis

Estes *risers* são compostos por camadas de diversos materiais, resultando em uma estrutura que lhes confere a capacidade de permitir grandes deslocamentos e rotações sob a ação de carregamentos ambientais. São similares aos *risers* rígidos, no que se refere à capacidade de resistir aos esforços de tração, pressões externa e interna, e torção. A Figura 2.6 apresenta um esquema básico da composição de camadas de um *riser* flexível.

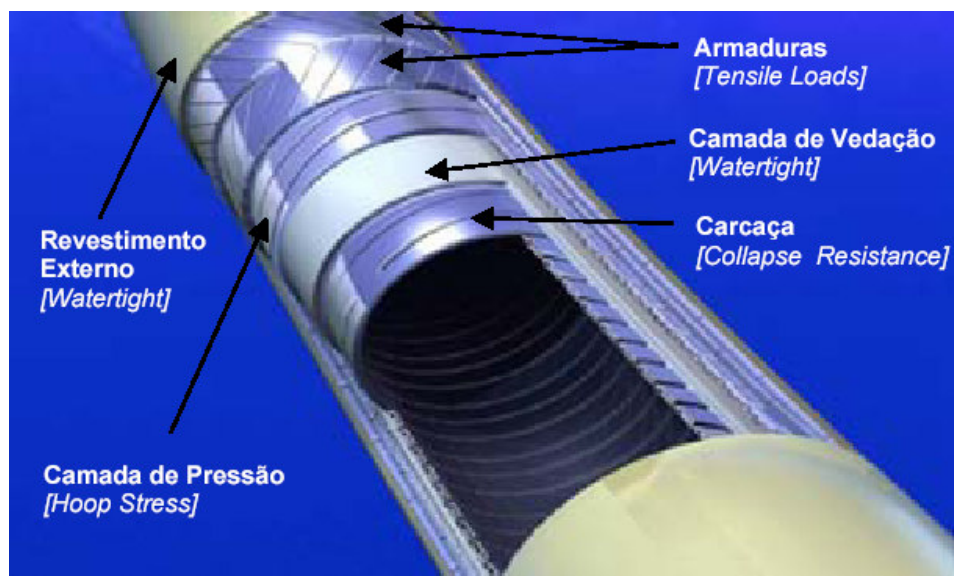


Figura 2.6 – Camadas dos *risers* flexíveis

### 2.2.3 - Risers rígidos em catenária (SCRs)

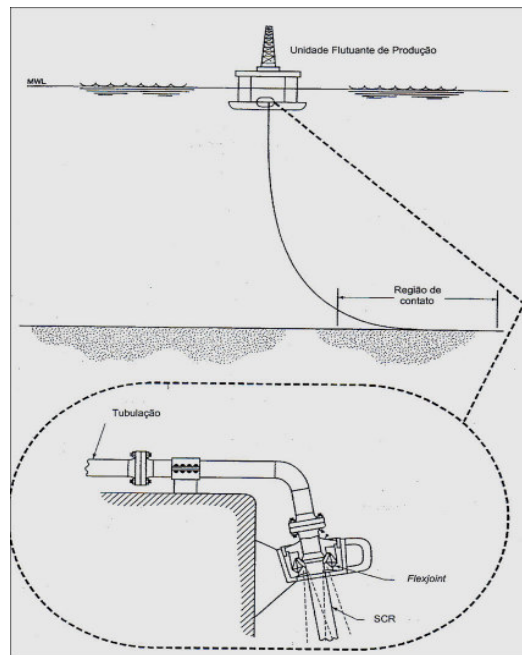
Grande parte do desafio em águas profundas e ultraprofundas reside no desenvolvimento dos *risers*. *Risers* flexíveis têm sido extensivamente empregados em sistemas flutuantes de produção. Entretanto, devido às suas limitações técnicas em altas pressões e alto custo, soluções como os SCRs têm surgido como alternativas mais versáteis, apesar de apresentarem complexidade de projeto e de instalação maiores que os flexíveis (Hatton et al, 1998).

Em 1994, no campo de Auger, no golfo do México, a Shell instalou o primeiro SCR. Desde então, os SCRs têm sido vitais para o desenvolvimento de campo em águas profundas (Mekha B.B, 2001).

No Brasil, os SCRs têm sido bastante estudados pela Petrobras, como alternativa aos *risers* flexíveis. Em 1998, na plataforma semi-submersível P-XVIII, foi instalado um SCR de 10'', com a finalidade de transportar o entre a P-26 e P-18, localizadas no campo de Marlim, na bacia de Campos, Rio de Janeiro. Além deste, dois outros de 10'' foram instalados em 2000 na plataforma P-36, campo de Roncador, com operação em lâmina d'água de 1360 metros.

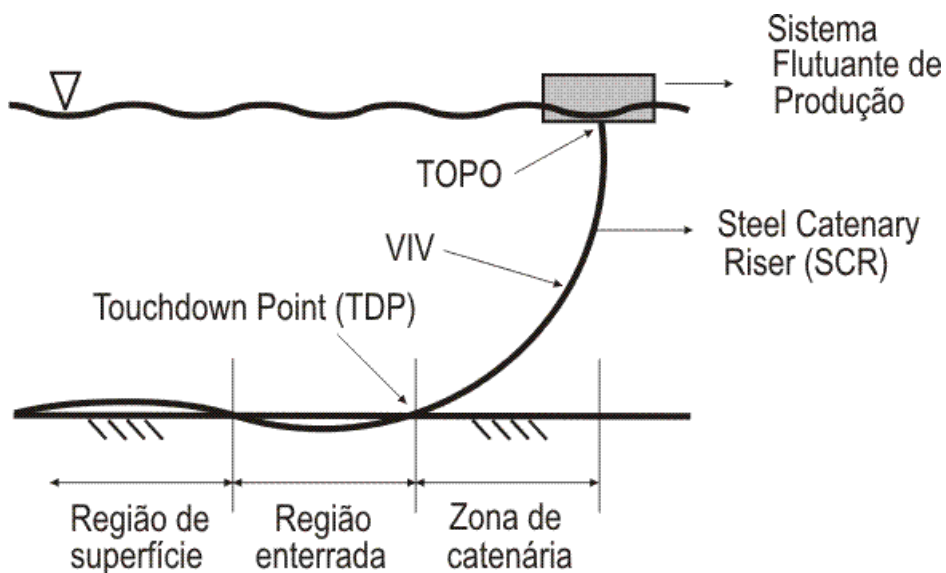
Em (Gonzalez *et al* ,2005), foi apresentado o projeto conceitual de 43 *risers* para um FPSO em 1800 m de lâmina d'água para o campo de Roncador, Rio de Janeiro. (Francis *et al*, 2004), (Tanaka *et al* ,2005) e (Roveri *et al*, 2005) apresentaram extensivos estudos para configurações com esquema *lazy wave* para as condições ambientais brasileiras, sustentação intermediária, como alternativa à configuração tradicional em catenária livre.

Os SCRs (*Steel Catenary Riser*) são dutos rígidos com configuração em catenária que se movimentam livremente entre o flutuante e o solo marinho. No topo, podem ser conectados à plataforma através de uma junta do tipo *flexjoint* (Figura 2.7) sob um ângulo de topo prescrito, que lhe confere grande flexibilidade. Estão sujeitos à mesma natureza de carregamento que os demais, porém podem ser dimensionados para suportar grandes pressões e grandes diâmetros, características que os têm tornado alternativa ao uso de flexíveis.



**Figura 2.7 – Riser rígido em catenária e sua conexão à plataforma**

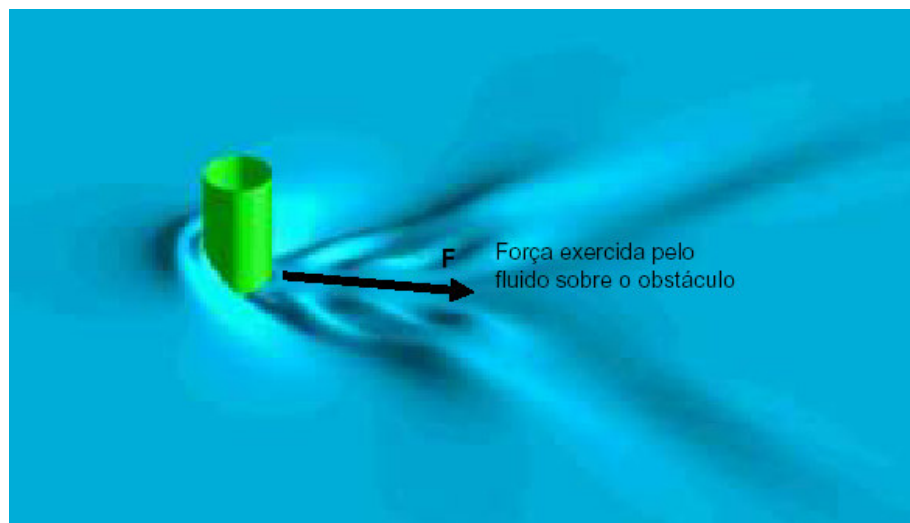
Extensas pesquisas, com investimento em monitoração e protótipos, têm sido valiosas para o entendimento do comportamento dos SCRs. Três regiões do SCR devem ser consideradas na sua descrição, conforme mostra a Figura 2.8. São identificados: o topo, a região de ocorrência de vibrações devido a vórtice (VIV) e o TDP (*Touch Down Point*).



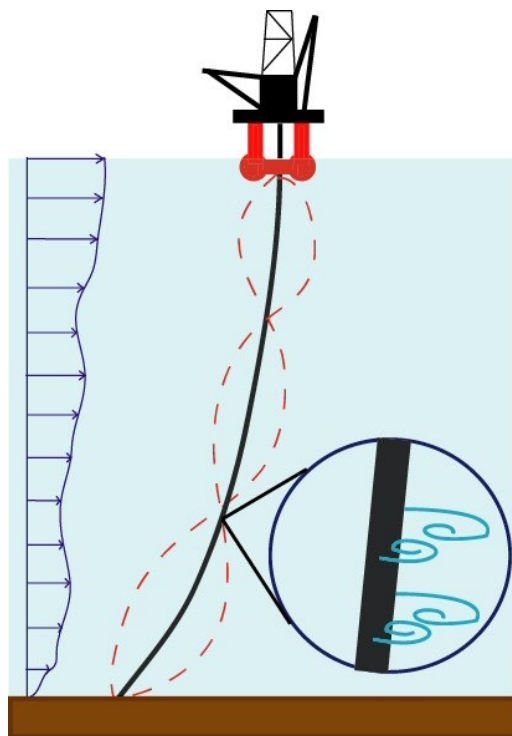
**Figura 2.8 – Regiões críticas de um SCR**

O TDP é a região de maior flutuação de tensões. Estas flutuações são decorrentes do movimento dinâmico do flutuante, que é transferido para o topo do *riser* e age diretamente no TDP, fazendo com que este se desloque ao longo do *riser*, em atrito com o solo marinho.

Vibrações induzidas pelo desprendimento de vórtices (VIV – *Vortex-Induced Vibration*) é um fenômeno que ocorre devido à ação da correnteza no *riser* (ver Figura 2.9), desprendendo vórtices deste, devido um complexo mecanismo ligado ao escoamento de fluido ao redor de um corpo rígido. Quando a frequência de desprendimento dos vórtices se aproxima da frequência natural de um dos modos de vibração do *riser*, este pode oscilar significativamente na direção transversal, isto é, perpendicular à passagem da corrente (ver Figura 2.10). Essa vibração depende de vários fatores, tais como a velocidade e perfil da correnteza, os modos naturais de vibração do *riser*, e o esforço axial ao qual está submetido o *riser*. Estas vibrações induzem a esforços cíclicos, tais como momentos fletores e esforços axiais, que podem causar redução significativa da vida útil do *riser* devido à fadiga. Supressores de vórtices, tais como *strakes* são bastante eficientes na redução dessas vibrações.



**Figura 2.9 – Ilustração da reação de um fluido em movimento a um obstáculo**



**Figura 2.10 – Formação de vórtices devido à passagem da correnteza no riser**

O topo do SCR é conectado à plataforma através de uma *flexjoint* (ver Figura 2.7), equipamento que minimiza os momentos fletores nesta região. Estes esforços são causados principalmente pelos movimentos resultantes da ação de ondas na superfície, e devido à ação da correnteza ao VIV.

As *flexjoints* são frequentemente usadas na conexão inferior do *riser* de perfuração para articular com o mínimo de esforço de flexão. Entretanto, conexões de topo e intermediárias também são utilizadas. A principal propriedade da *flexjoint* é a sua rigidez rotacional. A rigidez é função da magnitude da deflexão (ISO TC 67/SC 4N,2005) e esta propriedade pode ser importante para a fadiga. Para os SCRs elas são capazes de fornecer uma liberdade de rotação angular de  $\pm 20^\circ$  (Oilstates Site).

A deflexão do *riser* pode ser causada:

- Pelo movimento do flutuante
- Pela formação de vórtices que são desprendidos do *riser* quando a correnteza flui em torno de sua superfície externa. As deflexões causadas por estes vórtices são frequentemente difíceis de prever.
- Pela ação direta da correnteza que deflete o *riser* em sua direção.



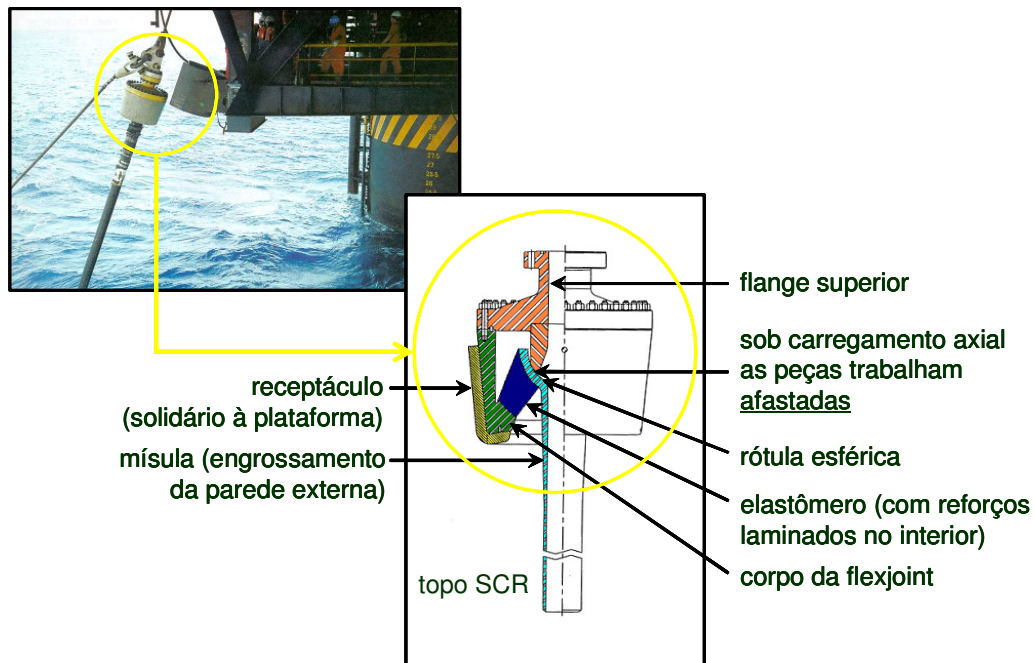
**Figura 2.11 – Flexjoint do SCR da PRINCE**  
**Fonte : OTC 14153**

A Figura 2.11 mostra a flexjoint da PRINCE. Na próxima seção, o comportamento da *flexjoint* será descrito com mais detalhes.

### **2.3 - Flexjoint**

A *flexjoint*, equipamento utilizado para a conexão de *risers* rígidos em catenária – *Steel Catenary Risers* (SCR) – com unidades marítimas de produção de petróleo (plataformas semi-submersíveis, por exemplo), minimiza a transferência dos movimentos de rotação destas unidades para o topo (*Upper Flexjoint*) dos *risers* e, por conseguinte, reduz a intensidade dos momentos fletores nesta região. O topo do *riser* é suportado sobre um apoio constituído por finas camadas de elastômero intercaladas por lâminas de aço, montado no interior do corpo da *flexjoint*. O berço assim constituído é capaz de deformar livremente para compensar, em parte, os movimentos de rotação da plataforma que, de outra maneira, seriam transmitidos integralmente ao topo do *riser*. Assim, surge uma rotação relativa entre *riser* e plataforma que tem origem em excitações a que todo o sistema está submetido, sejam dinâmicas ou quasi-estáticas. São exemplos de excitações dinâmicas as forças hidrodinâmicas produzidas pela ação de ondas, tanto sobre o casco da plataforma, quanto sobre o próprio *riser*, bem como as vibrações induzidas por vórtices que se desprendem ao redor do *riser* em virtude de determinados fluxos de corrente. Como excitações quasi-estáticas, pode-se mencionar a deriva da plataforma, resposta à ação do vento sobre o convés e da correnteza sobre o casco, bem como aos efeitos de segunda ordem produzidos pela incidência das ondas também sobre o casco. Com relação ao carregamento induzido pela corrente, cabe ainda mencionar o efeito que produz sobre o *riser*,

modificando sua configuração geométrica. A Figura 2.12 mostra a *Flexjoint* do SCR da plataforma P-XVIII instalada e um desenho esquemático de seus componentes.



**Figura 2.12 – *Flexjoint* do SCR da plataforma P-XVIII e um desenho esquemático de seus componentes**

### 2.3.1 - Características da *flexjoint*

#### 2.3.1.1 - Conceito de rigidez

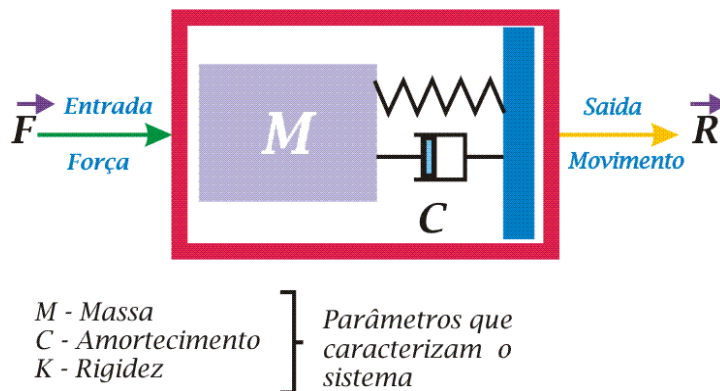
A rigidez é definida genericamente como a resistência do corpo ou sistema à deflexão ou vibração quando uma força é aplicada sobre ele. Ela depende de múltiplos fatores, entre eles a natureza das condições de carregamento, podendo ser classificada em estática ou dinâmica (Chew,2003). Considere a equação geral de equilíbrio da dinâmica:

$$M\ddot{X}(t) + C\dot{X}(t) + KX(t) = f(t) \quad \text{(Eq. 2.1)}$$

onde:

$M\ddot{X}$ = forças de inércia [N,Nm]	$\ddot{X}$ = aceleração [m/s <sup>2</sup> , rad/s <sup>2</sup> ]
$C\dot{X}$ = forças de amortecimento [N,Nm]	$\dot{X}$ = velocidade [m/s,rad/s]
$KX$ = forças elásticas [N,Nm]	$X$ = deslocamento [m,rad]
$K$ = Rigidez [N/m, Nm/rad]	

Entende-se por rigidez estática a parcela  $K$  da equação (2.1). A deflexão (ou resposta)  $(X(t))$  depende somente da rigidez  $K$  e pode ou não variar ao longo do tempo. Por rigidez dinâmica entende-se a contribuição das parcelas referentes à inércia e amortecimento. Nesse caso dinâmico, a força muda de direção e magnitude com o tempo. O movimento dinâmico (vibração) é a resposta a essa força dinâmica e depende das parcelas de inércia e amortecimento do sistema. Dessa forma, a rigidez estática resiste à deflexão, e a rigidez dinâmica à vibração. Apesar da representação linear, modelo representado pela Figura 2.13 ilustra o conceito da rigidez dinâmica.



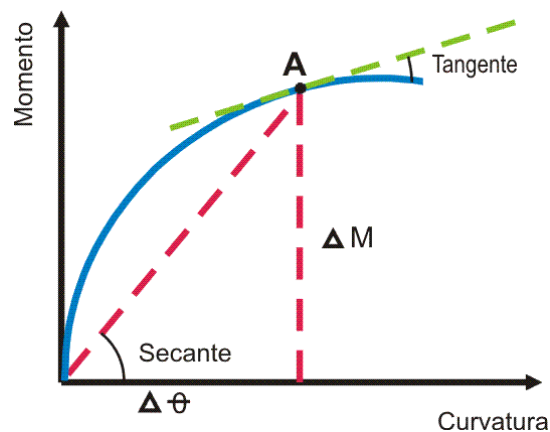
**Figure 2.13 - Rigidez dinâmica é parcela estática complementada pelos efeitos de inércia e de amortecimento de um corpo ou sistema.**

Além dos conceitos de rigidez estática e dinâmica, a rigidez pode ser obtida como uma tangente ou secante. A rigidez é denominada tangente quando corresponde ao estado de deformação ou de tensão em que o sistema encontra-se naquele estado (ponto de interesse). Rigidez secante pode ser



obtida traçando-se uma reta ligando a origem do diagrama momento-curvatura a ponto de interesse neste diagrama. A Figura 2.14 ilustra os conceitos de rigidez secante e tangente.

A rigidez secante é uma aproximação empregada no lugar do cálculo da rigidez tangente. É adotada quando são encontradas dificuldades relacionadas à convergência de modelos numéricos ou quando não há necessidade de um cálculo mais rigoroso, como ocorre no caso da obtenção através das tangentes.



**Figura 2.14 - Rigidez é a inclinação da curva força x deslocamento, e pode ser obtida como tangente e secante.**

### 2.3.1.2 – Resposta dinâmica da junta elástica

O *riser* em condições de trabalho está sujeito a movimentos causados por carregamentos aleatórios e dinâmicos de onda, vento e correnteza, que incidem direta ou indiretamente através do flutuante ao qual ele se encontra conectado. A resposta dinâmica da estrutura em função desses carregamentos é um dos mais importantes fatores de projeto, e está diretamente relacionada com a rigidez dinâmica do elemento estrutural. A rigidez dinâmica exprime a resposta da estrutura quando sujeita a cargas dinâmicas, e varia com a frequência e amplitude destas cargas. Essa rigidez pode ser expressa matematicamente como:

$$\text{Resposta Observada\_vibração} = \frac{\text{Força}}{\text{Rigidez Dinâmica}} \quad (\text{Eq. 2.2})$$

A equação 2.2 expressa a relação entre a resposta, a força e a rigidez.

## 2.4 - Modelos da *flexjoint*

Antes de começar a modelagem da *flexjoint*, um primeiro passo foi investigar alguns modelos encontrados na literatura. As seções seguintes descrevem alguns desses modelos.

### 2.4.1 - Modelo empírico experimental

Para a simulação global do comportamento estrutural de um *riser* e seus acessórios, uma das propriedades fundamentais de entrada é a sua rigidez. No sistema computacional ANFLEX (Mourelle et al, 1995), o modelo do *riser* é subdividido em zonas com propriedades constantes. O tamanho do elemento finito, as propriedades do *riser* e os coeficientes hidrodinâmicos podem mudar de zona para zona (ANFLEX). A *flexjoint* é modelada no ANFLEX como um elemento escalar com uma rigidez estática localizada em uma posição desejada do *riser*.

Atualmente, para a *flexjoint*, a rigidez rotacional empregada nos modelos é obtida através de uma formulação empírica fornecida pelo fabricante, em função de ensaios de laboratório.

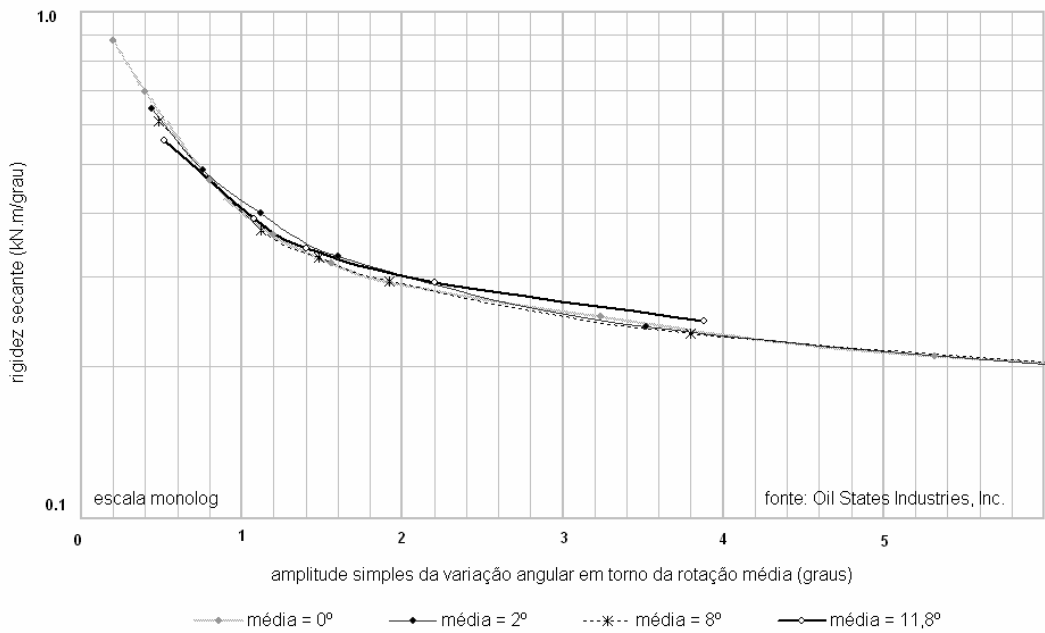
A rigidez da *flexjoint* oferecida aos movimentos de rotação não é uma propriedade linear; varia com a rotação experimentada pelo próprio equipamento, além de também ser influenciada pela pressão e pela temperatura do fluido interno. Atualmente, a Petrobras dispõe de um único exemplar de SCR: um gasoduto com diâmetro nominal de 10 polegadas, instalado na plataforma P-XVIII .

A partir de ensaios realizados pelo fabricante da *flexjoint* instalada para conexão de topo deste SCR (Oil States Industries,1996), ele pode concluir que a rigidez tangente à rotação depende preponderantemente da amplitude com que o ângulo de rotação varia em torno do valor médio. Por

exemplo, para uma mesma condição de carregamento, quanto maior a rigidez, menor será o ângulo de rotação relativa.

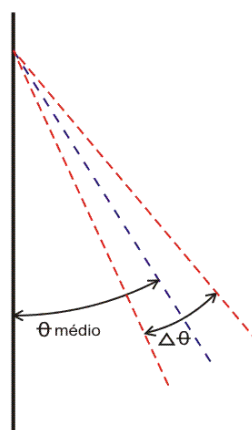
$$Fator\ de\ Enr\ ijecimento = f(\Delta\theta)$$

$$K_{secante} = f(\Delta\theta) \cdot K_{nominal}$$



**Figura 2.15: Sumário dos ensaios para determinação da rigidez à rotação da flexjoint 10” do SCR da P-XVIII**

A Figura 2.15 (Oil States Industries, 1996) mostra os resultados dos ensaios para valores médios da rotação entre 0° e 12°. Cada uma das curvas representa um comportamento da rigidez secante para uma configuração média da flexjoint, conforme ilustrado na Figura 2.16.



**Figura 2.16 – Esquema da posição da *flexjoint* em relação ao  $\theta_{\text{méd}}</math> e sua oscilação em torno do  $\Delta\theta$$**

#### **2.4.2 - Modelo por Elementos Finitos**

Os modelos numéricos disponíveis para flexjoint da PXVIII são o da OilStates e da MERL (HARRIS et al,2006) e empregam as técnicas de Elementos Finitos.

O modelo da OilStates como descrito no artigo (Sertã et al,1996) baseou-se em uma flexjoint com configuração similar para especificações de projetos e materiais e foi basicamente empregado para cálculo do dano a fadiga. A rigidez é empregada como um requisito a ser atendido na análise.

O objetivo principal do modelo da MERL é verificar a flexjoint nas condições reais de funcionamento, devido ao surgimento de fissuras no elastômero.

Embora os modelos contribuam para o entendimento do comportamento da flexjoint, é importante dizer que as análises são estáticas e não dinâmicas, diferente da abordagem do modelo experimental, simulado em ambiente com a dinâmica presente, e do modelo que será proposto neste trabalho, cujos dados monitorados são obtidos do processo dinâmico real da plataforma.

### **CAPÍTULO III - FUNDAMENTOS TEÓRICOS**

Neste capítulo são apresentados fundamentos teóricos que embasam o trabalho da tese. A idéia geral é a modelagem (ou identificação) de um sistema dinâmico não linear com comportamento histerético, em que somente os dados de entrada e saída do sistema estão disponíveis. É uma modelagem do tipo caixa preta. Assim, para entender um melhor entendimento da modelagem é importante algum conhecimento prévio sobre sistemas dinâmicos, redes neurais empregadas para a modelagem “caixa preta” do sistema, sinais aleatórios, forma como são apresentados os dados do sistema em estudo e histerese, fenômeno exibido pelo elastômero, principal componente da *flexjoint*.

Na seção 3.1 é feita uma breve introdução à teoria de modelagem de sistemas dinâmicos e Redes Neurais Artificiais são estudadas na seção 3.2 como ferramenta para modelagem de sistemas dinâmicos não lineares. São discutidas principalmente propriedades das redes neurais recorrentes e TDNN, reconhecidamente adequadas para modelagem de sistemas dinâmicos e previsão de séries temporais. Na seção 3.3 tratará da definição de alguns parâmetros que caracterizam sinais aleatórios. Finalmente, na seção 3.4 será feita

uma breve explanação sobre o fenômeno da histerese.

### **3.1 - Modelagem de Sistemas Dinâmicos**

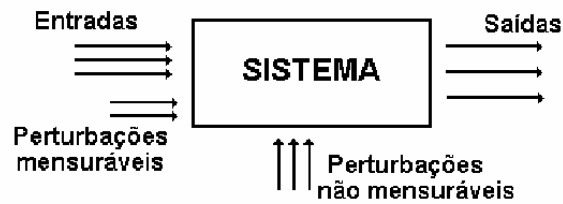
#### **3.1.1 - Introdução**

A identificação de sistemas é a área que permite obter modelos matemáticos para sistemas dinâmicos reais, a partir de dados observados de entrada e saída. Nesta seção será abordada a Identificação de Sistemas de forma introdutória, sendo focado essencialmente a terminologia e conceitos que foram empregadas neste trabalho bem como o emprego de Redes Neurais para formulação de Sistemas Dinâmicos não lineares. A área de modelagem de Sistemas Dinâmicos é extremamente ampla, com extensa bibliografia voltada para a teoria, mecanismos e ferramentas para modelagem e aplicações desses sistemas. Para estudo mais detalhado, livros como (LJUNG,1987) e (AGUIRRE,2004) são usados como referências em trabalhos na área de Identificação de Sistemas.

#### **3.1.2 - Conceitos básicos**

Primeiramente, um sistema pode ser definido como um conjunto de objetos que atuam em conjunto e realizam um certo objetivo. De acordo com esta definição, para o problema em estudo, a plataforma, o *riser* e a flexjoint formam um sistema. Mais especificamente, a conexão realizada pela flexjoint entre *riser* e plataforma é parte desse grande sistema e será tratado como um subsistema deste e é o objeto de estudo deste trabalho.

Os sistemas interagem com o mundo externo através de entradas, saídas e perturbações, como é mostrado na Figura 3.1. As perturbações podem ser de ruídos dos instrumentos de medidas ou de fatores externos não mensuráveis.



**Figura 3.1 - Descrição conceitual de sistema**

A característica básica dos sistemas conceituados como dinâmicos é traduzida como a de sistemas cujo estado assumido no tempo é condicionado pelas suas variáveis de entrada e pelo seu histórico, ou seja, pelo seu estado anterior, assumindo-se que estas variáveis são dependentes do tempo, ou seja, são funções onde:

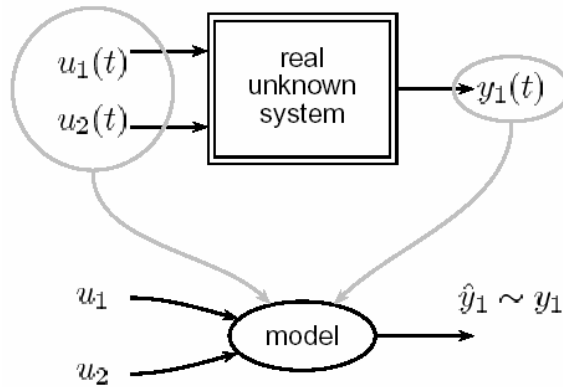
$$\frac{d\bar{x}(t)}{dt} \neq 0 \quad (\text{Eq. 3.1})$$

Em outras palavras, o estado do sistema assumido em  $t$  depende dos valores das variáveis de entrada neste instante bem como do estado no qual o sistema se encontrava em um número de instantes imediatamente anteriores, ou seja,  $t-1$ ,  $t-2$  e assim por diante.

Para modelos dinâmicos as variáveis referidas acima são sinais reais que evoluem como uma função do tempo, discreto ( $k$ ) ou contínuo ( $t$ ).

Para o desenvolvimento das áreas a exemplo da Engenharia, Física, Economia, Medicina ou mesmo em tarefas do cotidiano é fundamental a construção de modelos matemáticos para representar os sistemas. As entradas e saídas do sistema representam as variáveis do modelo, conforme ilustra a Figura 3.2. A idéia é que os modelos dos sistemas estabeleçam relações entre essas variáveis e na medida do possível, representem o sistema o mais próximo da realidade em suas características essenciais.

(1) black-box (BB) modeling



**Figura 3.2 - Esquema de modelagem de um sistema real**

Essencialmente todos os modelos são obtidos a partir de dados observados, entretanto alguns são construídos a partir de experiência ou conhecimento prévio do sistema e das variáveis envolvidas e não da observação direta dos dados observados. Modelos assim construídos são obtidos através da **modelagem física do processo**, utilizando-se das leis físicas, a exemplo das leis de Newton, da Teoria da Relatividade de Einstein entre outras e de relacionamentos bem estabelecidos que têm raízes em trabalhos empíricos anteriores (LJUNG, *op. cit.*). Este tipo de modelagem é conhecida como “caixa branca”, modelagem pela natureza do processo, modelagem fenomenológica ou conceitual.

Usualmente são modelos descritos por equações de diferenças (Eq. 3.2) ou equações diferenciais (algébricas) (Eq. 3.3).

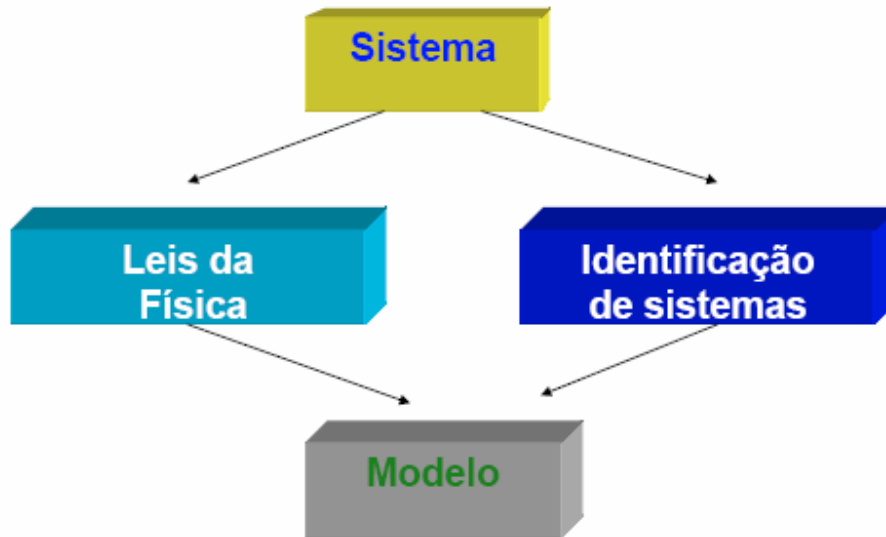
$$y_{n+1} = f(y_n) \quad (\text{Eq. 3.2})$$

$$\frac{d\bar{x}}{dt} = \bar{f}(\bar{x}, t) \quad (\text{Eq. 3.3})$$

A outra classe de modelos são chamados **modelos empíricos** e são construídos diretamente a partir dos dados observados sem que seja necessário conhecimento prévio do



comportamento interno do sistema e por isso conhecidos como modelos “caixa preta”. Identificação de Sistema é área que lida com o problema de modelagem empírica.



**Figura 3.3 – Esquema de modelagem de Sistemas**

### **3.1.3 Finalidade da modelagem de Sistemas**

Em síntese, os modelos matemáticos são representações simplificadas de um sistema real. É natural, portanto que o modelo apresente somente algumas características análogas ao do sistema, sendo que a definição de quais características são relevantes para serem representadas, é fundamental saber o objetivo do modelo.

Classicamente são conhecidos como objetivo da modelagem de sistemas:

- Predição
- Projeto de sistema de controle
- Simulação
- Diagnóstico e

- Otimização.

Embora os objetivos relacionados abaixo estejam presentes implicitamente no próprio processo de modelagem, quando se trata de novos desafios , de novas tecnologias eles podem se tornar, a priori, o objetivo principal da modelagem:

- Permitir um conhecimento mais profundo das características do sistema;
- Descobrir relações existentes entre as variáveis do problema, medidas ou não e detectar a importância delas;
- Permitir tratar o problema em seu conjunto e com todas as variáveis simultaneamente.

### **3.1.4 - Classificação dos modelos matemáticos**

Para o desenvolvimento de modelos matemáticos é importante analisar algumas propriedades fundamentais dos sistemas. Estas propriedades permitem definir que tipos de ferramentas teóricas serão necessárias para modelar o problema. Em função destas propriedades os modelos matemáticos podem ser basicamente classificados da seguinte forma:

- **Linear ou não linear**

Um modelo é dito linear quando obedece ao princípio da superposição. Caso o princípio não se aplique é chamado de não-linear. Os modelos lineares desprezam efeitos de segunda ordem. Na prática, as variáveis que os definem estão elevadas todas à primeira potência e não há produto entre variáveis. Considere os exemplos de modelos :

- $2 \ddot{y} x(t) + x(t) = u(t)$  é linear (Eq. 3.4)

- $2 \ddot{y} x(t) + x^2(t) = u(t)$  e (Eq. 3.5)

- $2 \ddot{y} x(t) + x(t)u(t) = 0$  não são lineares. (Eq. 3.6)

- **Discreto (amostrado) ou contínuo**

Quando as variáveis do modelo assumem valores em qualquer tempo  $t$  ( $t \in \mathfrak{R}^+$ ) ele é dito modelo de tempo contínuo e em geral é representado por equações diferenciais. Modelos em tempo discreto ou amostrado são aqueles cujas variáveis são seqüências de valores definidos em intervalos de tempo  $n$ , sendo  $n$  um número inteiro. As equações 3.7 e 3.8 são exemplos de modelos contínuo e discreto respectivamente.

$$5 \frac{dy}{dt} + y(t) = u(t) \quad , (t \in \mathfrak{R}^+) \quad \text{(Eq. 3.7)}$$

$$y(n) = 0.9y(n-1) + 0.1u(n-1), \quad \forall n \in \mathbb{Z} \quad \text{(Eq. 3.8)}$$

- **Variante ou invariante no tempo**

Um sistema é dito invariante quando sua resposta independe do instante de tempo em que uma entrada é aplicada, ou seja, o comportamento do sistema que está sendo modelado não varia com o tempo. Segundo (AGUIRRE, *op. cit*) isto não significa que as variáveis do sistema tem seus valores constantes, normalmente esses valores flutuam com o tempo. Caso haja ao menos uma entrada (parâmetro) variando ao longo do tempo o sistema é variante no tempo, ou seja, a resposta desse sistema depende do instante de tempo em que uma entrada é aplicada. Por exemplo, um painel solar é um sistema variante no tempo, pois a acumulação de pó diminui o seu rendimento. As equações 3.9 e 3.10 são exemplos de modelos invariantes e variantes no tempo respectivamente.

$$y(n) = x(n) + n^3 x(n-1), \forall n \in \mathbf{Z} \quad (\text{Eq. 3.9})$$

$$\ddot{x} + \frac{4}{t} \dot{x} + \frac{2}{t^2} x = u(t), t > 0 \quad (\text{Eq. 3.10})$$

De acordo com (AGUIRRE, *op. cit*) determinar se o sistema é ou não invariante no período em que se considera o sistema, é uma das premissas fundamentais para a modelagem matemática. O fato de ser invariante simplifica muito o problema da modelagem e implica em que apenas um modelo é requerido para representar o sistema.

Quanto ao número de entradas os modelos recebem as seguintes denominações:

- **SISO** — *Single Input, Single Output*: O sistema possui uma única entrada e uma única saída.
- **MISO** — *Multiple Input, Single Output*
- **SIMO** — *Single Input, Multiple Output*
- **MIMO** — *Multiple Input, Multiple Output*

A combinação de tais atributos dá origem a diversos tipos de modelos: modelos de sistemas lineares invariantes no tempo, modelos de sistemas não lineares discretos variantes no tempo, modelos não lineares contínuos invariantes no tempo.

O foco deste estudo serão os sistemas dinâmicos, não lineares, discretos e invariantes no tempo.

### 3.1.5 - Representação de sistemas dinâmicos

As duas formas de representações matemáticas de sistemas citadas por (HAYKIN, 2004) são: representação de modelos em espaços de estados e representação entrada-saída. A representação por espaço de estados é reconhecidamente a mais completa e, portanto mais vantajosa, porque além das variáveis de entrada e saída, as informações internas do sistema estão disponíveis, ou seja, variáveis de estado. Entretanto, nem sempre é possível se obter as variáveis de estado, sendo para problemas não lineares uma tarefa de alta complexidade. Para os casos em que se pode assumir que as variáveis de entrada e saída observadas contém informação suficiente para representar os estados do sistema, modelos representados por entrada-saída são mais comumente empregados (PAIVA, 1999).

#### 3.1.5.1 - Representação de sistemas dinâmicos em entrada-saída.

Um sistema é descrito por um modelo de entrada-saída quando é modelado através de um mapeamento de um conjunto de sinais de entrada  $u(t)$  para um conjunto de sinais de saída  $y(t)$ . Para sistemas discretos representados por equações a diferenças o mapeamento pode ser escrito como:

$$y(k) = f[u(k-1), u(k-2), \dots, u(k-N), y(k-1), y(k-2), \dots, y(k-N)] \quad (\text{Eq. 3.11})$$

$$y(k) \in \mathfrak{R}^p \text{ - saídas do sistema}$$

$$u(k) \in \mathfrak{R}^s \text{ - entradas do sistema}$$

onde  $N$  é um valor inteiro que descreve o número de exemplos anteriores de sinais  $u$  e  $y$  que são necessários para prever a próxima saída.

### 3.1.5.2 - Representação de sistemas dinâmicos em espaço de estados.

A representação de sistema entrada-saída não utiliza qualquer conhecimento da estrutura interna do sistema. Por sua vez, teorias mais modernas incorporam aos modelos uma descrição mais rica da dinâmica do sistema através do espaço de estados do sistema, que são definidos em função das variáveis internas deste, conhecidas como variáveis de estado. Logo, o conceito de estado é fundamental na formulação de sistemas dinâmicos.

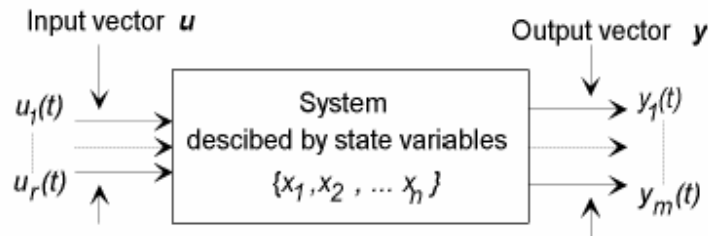
O estado de um sistema dinâmico é formalmente definido como um conjunto de quantidades que resumem toda a informação sobre o comportamento passado que é necessária para descrever unicamente o seu comportamento futuro, exceto pelos efeitos puramente externos que surgem devido à entrada (excitação) aplicada (HAYKIN, *op. cit.*). Um particular sistema determinado por seu estado (*state-determined*) tem a característica que (ROWELL *et. al.*, 2002):

“Uma descrição matemática do sistema em termos mínimo de um conjunto de variáveis  $x_i(t), i = 1, \dots, n$ , juntamente com o conhecimento destas variáveis em um instante inicial de tempo  $t_0$  e as entradas do sistema para um tempo  $t \geq t_0$ , são suficientes para prever o estado futuro do sistema para todos os tempos  $t > t_0$ ”.

Ou seja, o comportamento dinâmico do sistema determinado por seu estado é completamente caracterizado pela resposta do conjunto de  $n$  variáveis  $x_i(t)$ , onde o número  $n$  é definido como sendo a ordem do sistema.

O sistema mostrado na Figura 3.4 tem duas entradas  $u_1(t)$  e  $u_2(t)$ , e quatro variáveis de saída  $y_1(t), \dots, y_4(t)$ . Se o sistema é *state-determined*, o conhecimento de suas variáveis de estado  $(x_1(t_0), x_2(t_0), \dots, x_n(t_0))$  em algum tempo inicial  $t_0$ , e as entradas  $u_1(t)$  e  $u_2(t)$  para  $t \geq t_0$ , é suficiente para determinar todo o comportamento futuro do

sistema.



**Figura 3.4 – Representação em espaço de estado de um sistema dinâmico.**

A representação matemática de um sistema de tempo discreto ( $k$ ) não linear com uma entrada e uma saída no espaço dos estados e feita a partir das seguintes equações:

$$x(k+1) = a(x(k), u(k), k) \text{ Equação de Estado} \quad (\text{Eq. 3.11})$$

$$y(k) = c(x(k), u(k), k) \text{ Equação de Saída} \quad (\text{Eq. 3.12})$$

As funções  $a(\cdot)$  e  $c(\cdot)$  são não lineares para o exemplo considerado. Um modelo equivalente para a representação entrada-saída é dado pela equação 3.13

$$y(k) = h(y(k-1), \dots, y(k-r), u(k-1), \dots, u(k-r)) \quad (\text{Eq. 3.13})$$

### 3.1.6 - Identificação de Sistemas Dinâmicos

Como foi mencionado na seção 4.2, a identificação de sistemas é a área da teoria de matemática sistemas que lida com o problema da representação matemática de modelos baseados em dados experimentais, medidos ou observados da própria operação do sistema real. A idéia básica é permitir a construção de modelos matemáticos de um sistema dinâmico baseado em dados medidos, fazendo-se ajustes de parâmetros para um dado

modelo de forma que as saídas se aproximem o mais próximo possível das saídas medidas (LJUNG, *op. cit.*).

Essencialmente, o procedimento de construção de modelos a partir de dados é composto das seguintes etapas (LJUNG, *op. cit.*):

### **1- Obtenção de dados**

Os dados para modelagem tanto podem ser obtidos através de um projeto experimental, de um processo de monitoração por exemplo, quanto da operação normal do sistema. A vantagem da monitoração é que o usuário pode determinar quais as variáveis que devem ser observados e por quanto tempo.

### **2- Escolha da representação e estrutura dos modelos**

A escolha do tipo de representação dos modelos depende basicamente da finalidade destes, do conhecimento que se tem do sistema real, da sua complexidade e também da qualidade e quantidade de dados medidos. Se a representação matemática for uma Rede Neural, pode-se representá-la através de espaço de estado ou entrada-saída, por exemplo, e a escolha do número de neurônios e número de camadas de neurônios definem a estrutura do modelo.

### **3- Validação dos modelos**

A validação consiste em determinar se os modelos escolhidos representam características essenciais do sistema real em estudo, ou seja, como os modelos se comportam quando tentam reproduzir os dados medidos do sistema real.

Em função da complexidade dos modelos, do fato de serem lineares ou não lineares, do conhecimento da distribuição estatísticas das variáveis, os modelos empíricos podem ser paramétricos e não paramétricos.

Métodos de identificação de sistemas paramétricos são aqueles cuja estrutura do relacionamento funcional entre as variáveis dependentes e independentes é conhecida (LIMA,2000). Estes métodos se baseiam no conhecimento de expressões analíticas da



distribuição de probabilidades. Isto significa dizer que para a construção do modelo são assumidas hipóteses da distribuição que aproxima os dados do sistema, uma distribuição gaussiana, por exemplo. São também conhecidos como modelos “caixa branca”.

Para os métodos não paramétricos não existe conhecimento (ou pouco) *a priori* sobre a estrutura do modelo, que será determinada pelos dados medidos do sistema. São conhecidos como métodos de distribuição livre ou “**caixa preta**”.

Vale ressaltar que existe uma escala entre os modelos paramétricos e não paramétricos. Há modelos puramente não paramétricos numa extremidade, nenhuma suposição estatística sobre as variáveis é feita, sendo que para construí-los é necessário uma grande quantidade de dados para o treinamento. Numa outra extremidade encontram-se os modelos puramente paramétricos, para os quais a quantidade de dados necessária é mínima e conseqüentemente é exigido um conhecimento grande sobre a estrutura do modelo. É comum na literatura o termo semiparamétrico ou “caixa-cinza” para expressar modelos que se utilizam tanto da abordagem paramétrica quanto da não paramétrica.

Diversos métodos paramétricos e não-paramétricos são amplamente empregados para identificação de sistemas a exemplo de Redes Neurais, Análise Multivariada e Métodos Box-Jenkins. As redes neurais são consideradas ferramentas não paramétricas se a arquitetura da rede neural puder ser definida em função do problema de aproximação. Se a estrutura (estrutura de conexão e número de neurônios) da rede neural for definida previamente, independente do problema de aproximação, ela representa um modelo paramétrico (VON ZUBEN, 1996).

As redes neurais serão empregadas para identificação do sistema em estudo neste trabalho.

### **3.1.7 - Redes Neurais na Identificação de Sistemas Dinâmicos não lineares**

Teorias, novas arquiteturas e aplicações de redes neurais para o problema de identificação e controle de sistemas dinâmicos lineares e não lineares têm sido reportadas por diversos autores (NARENDRA & PARTHASARATHY, 1990; CHEN AND

BILLINGS, 1992; NERRAND *et al.*, 1994; RIVALS *et al.*, 1996). São citadas como poderosas ferramentas para identificação e controle principalmente quando:

- há pouco conhecimento sobre o sistema;
- a estrutura do sistema apresenta não-linearidades significativas e
- somente dados de entrada-saída estão disponíveis.

Ou seja, as redes neurais são usualmente empregadas como ferramentas do tipo “caixa-preta” e estatisticamente falando são métodos não paramétricos e conseqüentemente não têm uma distribuição de probabilidade associada.

São muitas as arquiteturas de redes neurais possíveis para identificação. A maioria se baseia nas redes neurais recorrentes (NARENDRA & PARTHASARATHY, *op. cit*) e nas redes TDNN (Time Delay Neural Network). Estas redes podem ser treinadas através do algoritmo padrão backpropagation, como foi proposto por (NARENDRA & PARTHASARATHY, *op. cit*) para diferentes modelos de sistemas dinâmicos.

De acordo com a variável de entrada, também chamada regressor, os modelos recebem denominações diferentes. É importante lembrar que a dinâmica das redes é realizada quase sempre através de valores atrasados das variáveis de entradas e saídas do sistema. Sendo assim, os modelos recebem as denominações a seguir e podem ser expressos analiticamente pelas seguintes equações:

$$y(k) = f(\Theta, \varphi(k)),$$

$$\varphi(k)$$

- NFIR, *Nonlinear Finite Impulse Response models*; neste caso o regressor é um vetor ,

$\varphi(k) = [u(k-1), u(k-2), \dots, u(k-N)]$  e a saída

$$y(k) = f[u(k-1), u(k-2), \dots, u(k-N)]$$

- NARX, *Nonlinear AutoRegressive with eXogenous input models*; o regressor é um vetor composto pelo passado da entrada e saída observadas :

$$\varphi(k) = [u(k-1), u(k-2), \dots, u(k-N), y(k-1), y(k-2), \dots, y(k-N)]$$

- NOE, *Nonlinear Output Error Models*, o vetor regressor é composto do passado da entrada e do passado da saída estimada do modelo,

$$\varphi(k) = [u(k-1), u(k-2), \dots, u(k-N), \hat{y}(k-1), \hat{y}(k-2), \dots, \hat{y}(k-N)]$$

- NARMAX, *Nonlinear AutoRegressive Moving Average with eXogenous input models*,

$$\varphi(k) = [u(k-1), \dots, u(k-N), y(k-1), \dots, y(k-N), \varepsilon(k-1), \dots, \varepsilon(k-N)]$$

- NBJ, *Nonlinear Box-Jenkins models*; onde os regressores são o passado das entradas, o passado estimado das saídas, a estimação dos erros usando o passado das saídas observadas e a estimação dos erros usando o passado estimado das saídas.

$$\varphi(k) = [u(k-1), \dots, u(k-N), \hat{y}(k-1), \dots, \hat{y}(k-N), \varepsilon(k-1), \dots, \varepsilon(k-N), \hat{\varepsilon}(k-1), \dots, \hat{\varepsilon}(k-N)]$$

A seguinte seção apresenta uma breve introdução aos conceitos de Redes Neurais artificiais e especial foco é dado às redes recorrentes e TDNN que serão avaliadas e empregadas neste trabalho. É importante deixar claro que outros tipos de redes neurais, p.ex. RBF (*Radial Basis Function*), são empregados para modelagem do comportamento histerético e dinâmico de sistemas, ver (SUDHEER *et al.*,2003) e (SUN & FUJIWARA, 2002). Entretanto, não serão explorados aqui.

### **3.2 - Redes neurais artificiais**

O objetivo desta seção é descrever a teoria de redes neurais, assim como sua origem, histórico, arquiteturas e aplicações mais comuns. Entretanto, por se tratar de um problema amplamente conhecido da comunidade acadêmica, será apresentada aqui apenas uma visão global da teoria das redes neurais.

#### **3.2.1 - Introdução**

A principal característica de uma rede neural artificial é a sua semelhança com o funcionamento do cérebro humano, tanto no processo de aprendizado, como no modo de armazenamento das informações e interligações entre elas. Assim, a base de uma rede neural é o neurônio artificial, que inspirado no neurônio biológico, também pode ser treinado através de exemplos.

Normalmente implementadas através de componentes eletrônicos ou através de programas computacionais, as redes tiveram origem na década 40, com o desenvolvimento do primeiro modelo neural por McCulloch e Pitts.

Desde então, a aplicação das redes neurais artificiais tem sido vasta, incluindo desde a síntese de discursos e o diagnóstico de problemas, até disciplinas como a medicina, a economia, processamento de sinais ou qualquer outra área que possa ser enquadrada na categoria de reconhecimento de padrões.

### 3.2.2 – Histórico

O primeiro trabalho em redes neurais artificiais surgiu em 1943, quando o psiquiatra e neuroanatomista McCulloch e o matemático Pitts modelaram matematicamente o funcionamento simplificado de um neurônio biológico, criando então o primeiro neurônio artificial.

Em 1949, com a publicação do livro “The organization of behavior” escrito pelo neuropsicólogo Hebb, houve um significativo avanço com a apresentação da primeira regra de aprendizado para o neurônio, segundo a qual a eficiência de uma sinapse variável entre dois neurônios é aumentada através da ativação repetida de um neurônio causada pelo outro neurônio participante.

No final dos anos 50, Rosenblatt apresentou um novo modelo, de uma camada de neuronios, capaz de aprender algumas classificações por meio da atribuição de pesos às conexões: o Perceptron. Também durante esse período, em 1960, Widrow e Hoff deram significativas contribuições ao estudo das redes neurais, ao desenvolverem uma regra de aprendizado conhecida como Regra de Widrow e Hoff ou Regra Delta. Em 1962 ROSENBLATT apresentou o Teorema de Convergência do Perceptron.

A pesquisa em redes neurais artificiais seguia seu rumo normalmente, quando em 1969, Minsky e Papert observaram algumas limitações no Perceptron e concluíram que ele não seria capaz de resolver o problema clássico do XOR (exclusive or). Esta conclusão contribuiu muito para o declínio das redes neurais durante a década de 70.

O interesse pelas redes neurais, contudo, foi retomado na década de 80, com a publicação de novos trabalhos, tais como:

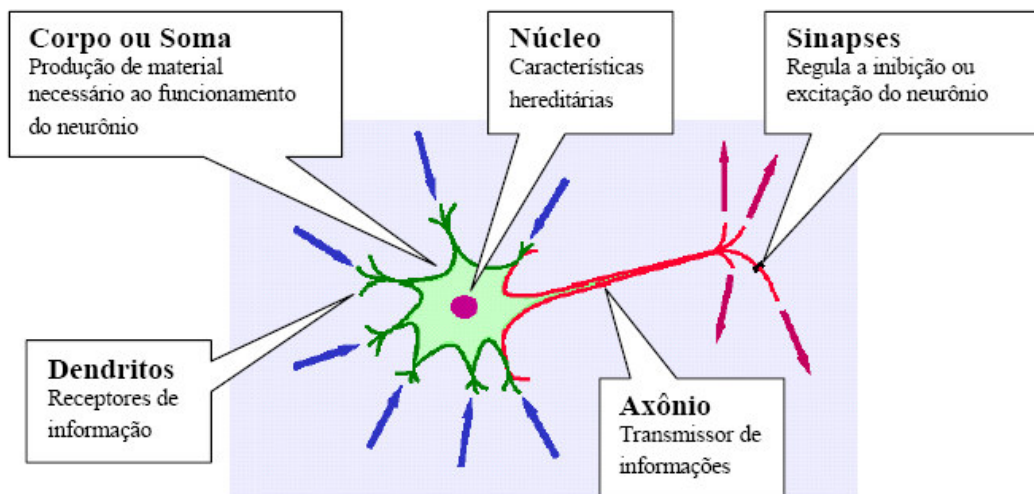
- Função de Energia – Hopfield;
- Self-organizing maps – Kohonen;
- Reinforcement learning – Barto e Anderson;

Finalmente, o estudo das redes neurais ressurgiu ao final dos anos 80, quando em 1986, Rumelhart, Hinton e Williams apresentaram uma descrição do algoritmo de retropropagação de erro, mostrando que a visão de Misnky e Papert sobre o Perceptron foi demasiadamente pessimista. Hoje, reconhece-se que o Perceptron foi uma das principais contribuições para o desenvolvimento das redes neurais.

### 3.2.3 - O neurônio biológico e artificial

A unidade básica do sistema nervoso é a célula nervosa ou neurônio, que foi descrito por RAMÓN Y CAJAL em 1911. O neurônio divide-se em três partes, conforme descrito abaixo:

- Dendritos (entrada)
- Corpo celular
- Axônio (saída)



**Figura 3.5 – Modelo simplificado de um neurônio biológico**

O neurônio recebe informações através dos dendritos, enviado-as para o corpo

celular, que as processa e emite uma resposta através do axônio para outro neurônio, estabelecendo-se assim, o tratamento biológico das informações no cérebro. Um neurônio é capaz de trocar informações com milhares de outros, mas à medida que as informações recebidas vão se acumulando e excedem o seu nível crítico, ele descarrega um pulso elétrico que caminha através do axônio e é liberado para o próximo neurônio pelos dendritos. Este fenômeno de liberação, também conhecido por despolarização, é seguido de um período durante o qual o neurônio fica impossibilitado de funcionar, chamado de período refratário. A comunicação dentro dos neurônios se dá através de sinais elétricos. Entre neurônios, a comunicação é feita por neurotransmissores e é comumente chamada de sinapse. O alcance da passagem de um sinal entre neurônios depende de muitos fatores, tais como a quantidade de neurotransmissores disponível, o número e arranjo de receptores, a quantidade de neurotransmissores reabsorvidos, etc.

Ainda há muito para se descobrir a respeito do funcionamento do cérebro no processo das informações. Sabe-se, contudo, que no cérebro humano a passagem dos sinais de um neurônio para outro pode ser excitante ou inibidora, ou seja, a sinapse tem a capacidade de converter a atividade de um axônio em efeitos elétricos que podem inibir ou excitar os neurônios seguintes. Quando um neurônio recebe maior quantidade de impulsos excitantes do que inibidores, ele libera um pulso elétrico para o seu axônio. O aprendizado ocorre através do sucesso da sinapse, de modo que a influência de um neurônio sobre o outro aconteça.

### **3.2.4 - Modelo genérico de um neurônio**

O primeiro modelo do funcionamento de um neurônio foi desenvolvido pelo psiquiatra e neuro-anatomista McCulloch juntamente com o matemático Pitts, em 1947. O neurônio artificial de McCulloch possui entradas (dendritos) e apenas uma saída (axônio). Para modelar o acúmulo de informações e se chegar à sinapse, cada entrada possui um peso, que dependendo do tipo de passagem de sinal, pode ser positivo (excitante) ou negativo (inibidor). Os pesos armazenam a quantidade de informação e estabelecem a intensidade com que cada entrada contribuirá para o resultado final do neurônio. O corpo

celular é representado pela soma dos produtos entre entradas e pesos correspondentes ( $\sum_{i=1}^m x_i \cdot \omega_i$ ). Sempre que o somatório atingir o limiar (nível crítico), o corpo celular é emulado com o valor 1, caso contrário, ele recebe o valor 0.

$x_1 \dots x_m$  = entradas

$y$  = saída

$\omega_1 \dots \omega_m$  = pesos

$\Theta$  = limiar

Assim, o neurônio artificial de McCulloch e Pitts, pode ser representado por:

$$\sum_{i=1}^m x_i \cdot \omega_i > \Theta \quad (\text{Eq. 3.14})$$

Esse comportamento pode ser representado por uma função a qual se dá o nome de função de ativação ou função de transferência. A função de ativação é uma função matemática que, aplicada à combinação linear entre as variáveis de entrada e pesos que chegam a determinado neurônio, retorna ao seu valor de saída. Existem diversas funções matemáticas que são utilizadas como função de ativação. As funções de ativação mais comumente usadas são: função logística e a função tangente hiperbólica.

Uma das principais diferenças entre o neurônio artificial de McCulloch e Pitts e o biológico é a simultaneidade. Os neurônios artificiais trabalham simultaneamente, enquanto os biológicos não, uma vez que suas saídas dependem dos neuro-transmissores e das ativações anteriores, devido ao período de inatividade após a sinapse.

Devido à sua grande complexidade, o funcionamento do neurônio biológico ainda não é completamente conhecido. Além disso, a modelagem do neurônio está limitada à capacidade de nossos computadores. Assim, o neurônio artificial, embora inspirado no biológico, pode ser descrito como uma aproximação grosseira deste.



### 3.2.5 - Arquitetura de uma rede neural

De posse do modelo de neurônio descrito na seção anterior, pode-se criar redes de neurônios. Uma rede neural é um conjunto de unidades computacionais interconectadas e organizadas em camadas. A interligação entre camadas é unidirecional partindo da entrada e atravessando cada camada até a última (camada de saída). Esse é o esquema básico de uma rede neural ao qual se dá o nome de redes *feedforward* ou redes progressivas.

A seguir descrevemos uma rede *feedforward* de uma ou múltiplas camadas.

#### 3.2.5.1 - Redes neurais de uma camada

Uma camada de neurônios com  $R$  elementos de entrada e  $S$  neurônios pode ser descrito da seguinte forma (acompanhe na Figura 3.6):

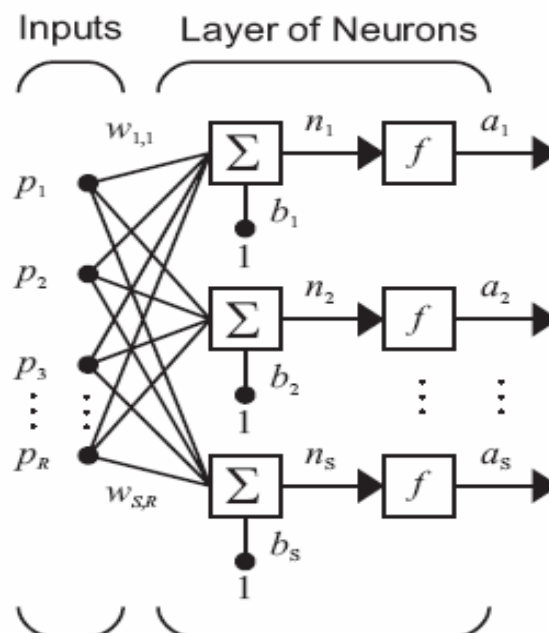


Figura 3.6 - Uma rede neural com uma camada de neurônios(Fonte: Neural Networks MATLAB – help)

Cada elemento de entrada está conectado a cada um dos neurônios de entrada de forma que cada neurônio perceba todos os elementos de entrada (vetor  $\mathbf{p}_R$ ). Esta conexão se dá pelo vetor de pesos  $\mathbf{W}_i$  e pelo escalar de polarização do neurônio  $\mathbf{b}_i$  (armazenado no vetor  $\mathbf{b}_S$ ). Assim, podemos armazenar os pesos de toda a camada de neurônios numa matriz  $\mathbf{W}_{SR}$ . Note que o número de neurônios não é, necessariamente, igual ao de elementos da entrada. Cada neurônio possui seu somatório que reúne o vetor  $\mathbf{p}$  de entrada e os seu vetor  $\mathbf{W}_i$  de pesos para formar seu escalar  $\mathbf{n}_i$ . Estes escalares podem ser armazenados num vetor  $\mathbf{n}_S$ . Finalmente, cada neurônio fornece, como saída,  $\mathbf{a}_i$ : o resultado da aplicação de sua função de transição  $\mathbf{f}$ . Mais uma vez podemos representar todos os resultado no vetor  $\mathbf{a}_S$ .

Assim sendo, o vetor  $\mathbf{a}$  possui a seguinte expressão:  $\mathbf{a} = \mathbf{f}(\mathbf{W}\mathbf{p} + \mathbf{b})$ .

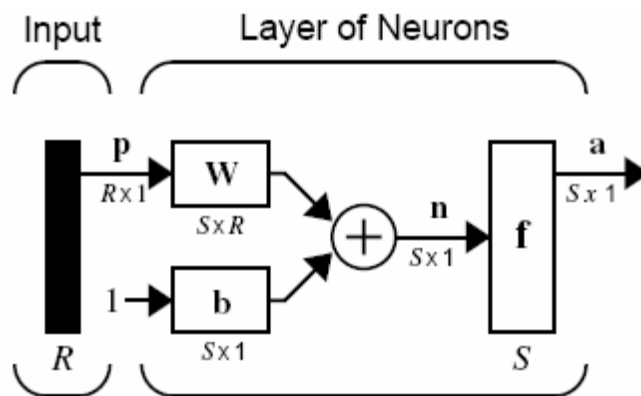


Figura 3.7 – Esquema de uma RN com uma camada de neurônios

A Figura 3.7 ilustra a notação abreviada para o esquema de uma camada de neurônios.

Pode-se criar uma camada de neurônios com mais de um tipo de função de transição, bastando, para tanto, empilhar várias camadas como as aqui descritas.

### 3.2.5.2 Redes neurais multi-camadas

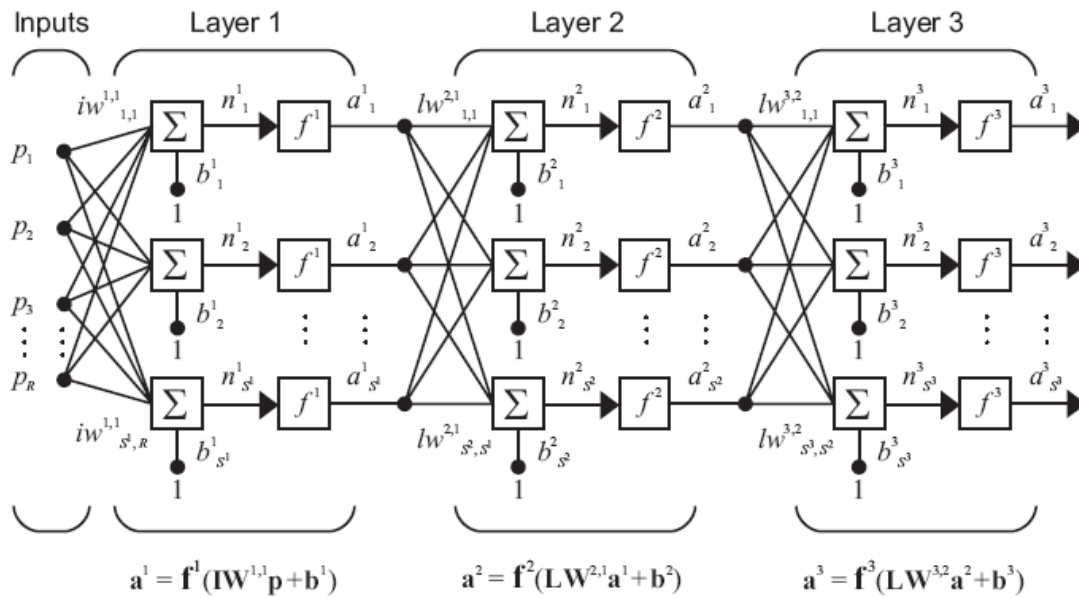
Para criar uma rede neural com mais de uma camada, antes precisamos introduzir alguns conceitos novos e mudar um pouco a notação.

A última camada de neurônios da rede é dita a *camada de saída*. Todas as outras camadas são chamadas de *camadas intermediárias* (alguns autores chamam a camada de elementos de entrada de *camada de entrada* apesar de não serem neurônios). As matrizes de pesos conectados à entrada serão chamados de *pesos de entrada*. As matrizes de pesos na saída de uma camada são chamadas de *pesos de saída* (desta camada).

Para nos referirmos os elementos de cada camada (vetor de pesos, função de transição, etc) adicionaremos um índice acima. Portanto:  $S^1$  é o número de neurônios na primeira camada intermediária,  $f^1$  é a função de transição da primeira camada,  $\mathbf{W}^{1,2}$  indicando que os pesos têm como origem a camada 2 (segundo índice) e destino a camada 1 (primeiro índice).

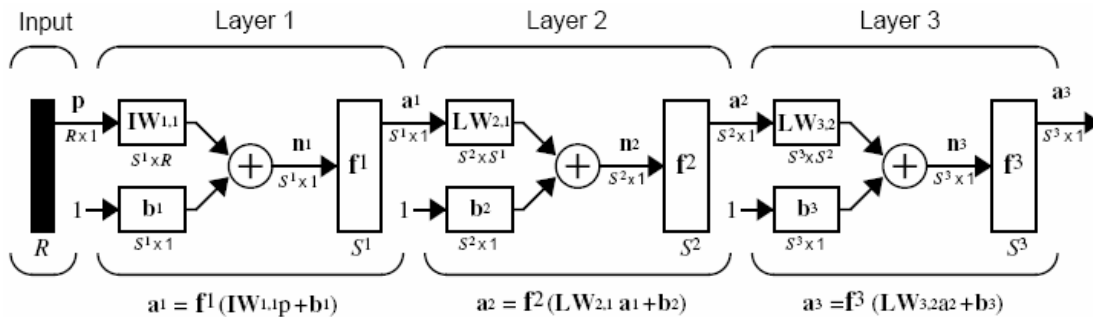
A seguir, a Figura 3.8 ilustra uma rede neural de 3 camadas e o uso dos índices superiores na indicação da camada. Note que a saída de cada camada constitui a entrada da camada subsequente ( $\mathbf{p}^{i+1} = \mathbf{a}^i$ ). A expressão da saída da rede neural seria então:

$$\mathbf{a}^3 = f^3(\mathbf{LW}^{3,2} f^2(\mathbf{LW}^{2,1} f^1(\mathbf{IW}^{1,1} \mathbf{p} + \mathbf{b}^1) + \mathbf{b}^2) + \mathbf{b}^3)$$



**Figura 3.8 – Esquema de uma RN com 3 camada de neurônios**

E a sua abreviação (Figura 3.9):



**Figura 3.9 – Esquema de uma RN com 3 camada de neurônios(abreviação)**

Redes de múltiplas camadas são muito poderosas. Por exemplo, uma rede neural de duas camadas (a primeira sigmóide e a segunda linear) é capaz de representar muito bem qualquer função com um número finito de descontinuidades.

### 3.2.6 - Aprendizagem

A característica mais importante de uma rede neural artificial é a sua capacidade de aprender um problema. No processo de aprendizado a rede otimiza o seu desempenho, até chegar a uma solução generalizada para o problema.

O processo de aprendizado de uma rede neural artificial se dá através da modificação das sinapses dos neurônios, sendo que tais alterações são diretamente conectadas à ativação destes. Todo o conhecimento de uma rede está armazenado nas sinapses, através dos pesos atribuídos a cada uma delas. Assim, as sinapses mais ativadas da rede são fortalecidas, enquanto que as menos utilizadas são enfraquecidas ao longo do processo de aprendizado.

O conjunto de regras definidas para a solução de um problema de aprendizado é chamado de algoritmo de aprendizado. Existem vários tipos de algoritmos, que diferem entre si principalmente pelo modo como os pesos são modificados.

O processo de aprendizado de uma rede neural artificial deve ser rigoroso e verdadeiro, sendo necessário tempo suficiente para se treinar uma rede neural artificial, a fim de se evitar modelos espúrios. Assim, durante o processo de aprendizado, é importante que uma parte da base de dados seja separada para o treinamento da rede, a fim de que a rede “aprenda” as regras em vez de “decorar” os exemplos. O restante dos dados é armazenado para uma fase posterior de testes, a fim de que a rede chegue corretamente aos resultados esperados.

De modo geral, classificam-se os processos de aprendizado em três tipos distintos: o supervisionado, o não-supervisionado e o híbrido.

O processo supervisionado ocorre quando a rede recebe um conjunto de entradas e seus padrões correspondentes de saída, ocorrendo ajustes sinápticos até que o erro encontrado seja aceitável.

No processo não supervisionado a rede trata os dados de forma a perceber algumas propriedades do conjunto, sendo a partir daí constituído o aprendizado.

No processo híbrido ocorre uma mistura do supervisionado com o não-

supervisionado, podendo uma camada utilizar-se de um tipo, enquanto outra camada utiliza-se de outro tipo.

### 3.2.6.1 - Algoritmo backpropagation

É o principal algoritmo de aprendizado empregado em redes neurais pela sua eficiência (RUMELHART *et al.*, *apud* HIROSE *et al.*, 1991), bem como por apresentar excelentes resultados na análise e na resolução de problemas de confiabilidade estrutural (Saraiva, 1997).

Este algoritmo foi criado como uma generalização do algoritmo Widrow-Hoff e baseia-se no princípio de aprendizado de correção de erros onde o erro é propagado no sentido contrário em que os dados são processados (da camada de saída para a primeira). Sendo assim, uma rede neural possuiria dois passos: (1) o *feedforward* onde os dados de entrada são processados e os pesos sinápticos são atualizados para refletir a exposição do neurônio à informação e (2) o *backpropagation* onde o resultado da rede é comparado com a solução real do problema apresentado e a diferença entre o valor calculado e o real é usada para corrigir os pesos sinápticos, desta vez, na direção contrária.

Ao conjunto de dados de entrada e sua respectiva saída, dá-se o nome de conjunto de treinamento. É o conjunto de treinamento que é usado para treinar a rede durante a fase de em que esta está aprendendo um novo problema. Através do conjunto de treinamento, mostrado em (8), composto por um vetor  $\mathbf{x}$  com  $n$  entradas e um vetor de saída exato  $\mathbf{y}_e$ :

$$Y = \{x_i, y_{ei}\} \quad i = 1, \dots, n \quad (\text{Eq. 3.15})$$

pode-se definir o erro quadrático médio sobre este em (Eq. 3.16):

$$E_R = \frac{1}{n} \sum_{i=1}^n (y_i - y_{ei})^2 \quad (\text{Eq. 3.16})$$

onde  $y_i$  é a saída fornecida pela rede e  $y_{ei}$  é o valor exato correspondente à saída da rede.

A expressão (9) pode ser reescrita em função dos pesos, como apresentado em (10):

$$E_R = \frac{1}{n} \sum_{i=1}^n (f(\mathbf{W}^T \cdot \mathbf{x}) - y_{ei})^2 \quad (\text{Eq. 3.17})$$

A equação 3.17 é da forma de uma função objetivo de um problema de otimização onde se deseja minimizar  $E_R$  e cujas variáveis são os pesos  $W_{ij}$ . Para resolver esse problema, há vários métodos. Dentre eles há os que usam gradiente (mais fácil de calcular mas de convergência lenta), a hessiana (altos custos computacionais, mas converge mais rápido) e, em particular o método de Levenberg-Marquardt que usa a matriz jacobiana como aproximação da hessiana no intuito de balancear custo e eficiência. Este último sugere que os pesos sejam ajustados segundo a expressão:

$$\mathbf{W}(k+1) = \mathbf{W}(k) - [\mathbf{J}^T(\mathbf{W}) \cdot \mathbf{J}(\mathbf{W}) + \mu_k \mathbf{I}]^{-1} \cdot \mathbf{J}^T(\mathbf{W}) \cdot \mathbf{e}(\mathbf{W}) \quad (\text{Eq. 3.18})$$

onde:  $\mathbf{I}$  é a matriz identidade e  $\mu_k$  é a constante do método de Levenberg - Marquardt

### 3.2.6.2 - Overfitting e Early-stopping

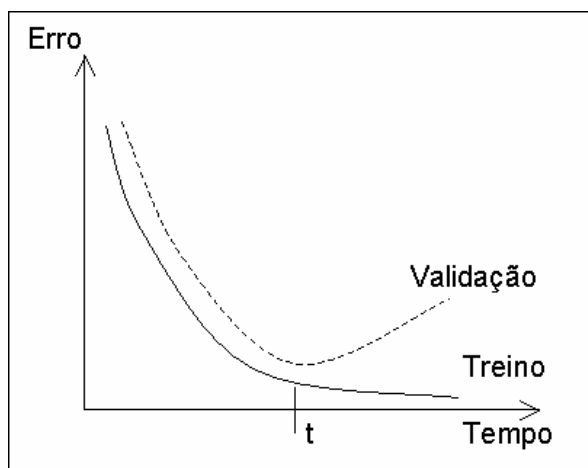
Quando se aplicam algoritmos de minimização do erro entre os valores do conjunto de treinamento e os valores calculados pela rede durante o aprendizado, insere-se uma nova complicação no método. Ao tentar minimizar o erro se está, de certa forma, fazendo a rede “decorar” o comportamento do conjunto de treino ao invés de “aprender” o comportamento do problema. Ou seja, pode-se chegar ao ponto de minimizar o erro, a rede estará bem ajustada para o conjunto de treino e qualquer outro conjunto que não possua grandes

divergências deste. Esta situação foi batizada de *overfitting*, uma vez que a rede se ajusta aos pontos do conjunto de treino (e só a eles). Porém, o que se pretende é que mesmo o conjunto de treino não sendo representativo de todo o espaço do problema, que a rede possa dar respostas boas para qualquer caso quer ele pertença ao conjunto de treino ou não, uma vez que ela tenha aprendido o comportamento do problema.

Há várias formas de evitar (ou contornar) o *overfitting*. Uma delas é chamada técnica de parada antecipada (ou *early-stopping*). Esta técnica se apóia no fato de se usar algoritmos de aprendizado baseados em derivadas. O uso desses algoritmos propicia que os parâmetros que têm maior influência no erro da saída da rede serão modificados nas primeiras iterações, ficando o ajuste fino para as iterações posteriores. A idéia é terminar o treinamento antes que a fase de ajuste fino tenha início e, por conseguinte, o *overfitting*.

Para detectar este momento faz-se uso de um segundo conjunto de dados chamado conjunto de validação. O conjunto de validação, idealmente não deve diferir estatisticamente do conjunto de treino, mas também não pode ser muito similar. O conjunto de validação não será usado para treinar a rede, mas apenas para comparar o erro produzido pelo conjunto de treino com o erro produzido pelo conjunto de teste. Como fazem parte do mesmo problema, a tendência é que os erros de ambos os conjuntos diminuam ao longo do treinamento, porém, a partir de certo ponto, o erro do conjunto de treino continuará a diminuir (pois o algoritmo de aprendizado forçará esse comportamento) enquanto que o erro do conjunto de validação aumentará. Isso acontece porque o conjunto de validação não influencia o aprendizado da rede. Sendo assim, no momento em que a rede começar a se adequar melhor ao conjunto de treino do que ao conjunto de validação (um subconjunto qualquer dos casos do problema), isso significa que a rede começou a “decorar” o conjunto de treino. A Figura 3.10 ilustra esse comportamento.





**Figura 3.10 - Comportamento do erro nos conjuntos de treino e de validação**

Amari *et al.* (1996) determina que o uso de *early stopping* é necessário se o número de pesos for 30 vezes inferior ao número de exemplos, e que o conjunto de validação deve conter  $1/\sqrt{2X \text{ números pesos}}$  X número de exemplos disponíveis.

Uma vez detectado este descolamento o processo de treinamento pode ser terminado e o estado dos pesos e *bias* (vetores  $\mathbf{b}^S$ ) que obtiveram menor erro podem ser restaurados para a rede.

### 3.2.6.3 - Conjunto de teste

O conjunto de teste também não deve diferir dos outros dois conjuntos já mencionados anteriormente. Porém o conjunto de teste tem pouca ou nenhuma influência no aprendizado da rede servindo mais para comparação de modelos diferentes ou de arquiteturas diferentes de redes que pretendem resolver o mesmo problema.

### 3.2.7 - Redes neurais dinâmicas

As redes neurais descritas até aqui são ditas estáticas pois não são capazes de representar situações dinâmicas, temporais ou que apresentem alguma ordem no conjunto

de dados. As redes dinâmicas recebem esta denominação por apresentarem estruturas como atraso de tempo e retroalimentação que as tornam mais poderosas que as redes estáticas, porém mais difíceis de treinar.

O tempo é uma entidade importante no aprendizado. É através de sua inclusão na rede neural que a capacita a seguir variações estatísticas de processos não-estacionários. O tempo pode ser discreto ou contínuo e pode ser incorporado às redes neurais de duas formas:

- **Representação implícita:** “o tempo é representado pelo efeito que tem sobre o processamento de sinais de uma maneira implícita” (HAYKIN, *op. cit*), isto é, o tempo está representado na ordem em que os dados são inseridos.
- **Representação explícita:** “o tempo recebe sua própria representação particular” (HAYKIN, *op. cit*), isto é, a introdução do tempo altera a arquitetura da rede.

Na representação implícita, redes estáticas ganham propriedades dinâmicas que as tornam sensíveis à estrutura temporal dos sinais portadores de informação. Já para uma rede neural ser considerada dinâmica, ela deve representar memória. A memória das redes dinâmicas pode ser de curto ou de longo prazo. Para inserir memória de longo prazo, faz-se uso do aprendizado supervisionado, onde os dados de conjunto de treinamento são armazenados (total ou parcialmente) nos pesos sinápticos. Já a memória de curto prazo é necessária quando o problema apresenta algum tipo dimensão temporal. Uma forma simples é o uso de atrasos de tempo (*time delay*) que podem ser implementados na entrada da rede ou no nível das sinapses. A idéia dos atrasos de tempo possui fundamentação biológica.

Sendo assim, a saída das redes neurais dinâmicas depende tanto da entrada atual quanto das entradas e saídas ou estados da rede anteriores. As redes neurais podem ser divididas em duas categorias: as que possuem conexões apenas adiante e as que possuem conexões recorrentes (*feedback*).

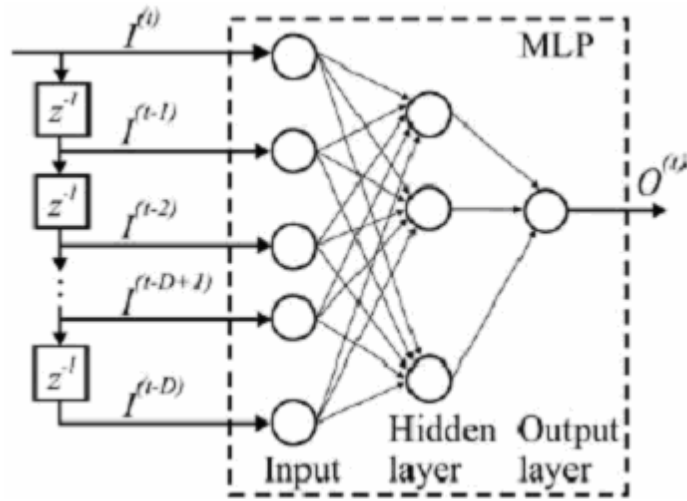
O treinamento de redes dinâmicas pode ser feito usando o algoritmo *back propagation* (descrito na seção 3.2.6.1), porém espera-se que essas redes sejam mais difíceis de treinar. Em parte por causa do duplo efeito da propagação do erro. Há um efeito direto (mudanças nos pesos numa iteração causam mudanças na saída desta mesma iteração) que o algoritmo *backpropagation* ordinário é capaz de calcular e um outro indireto (a saída alterada numa iteração faz parte da entrada da rede na iteração seguinte) que é bem mais difícil de calcular exigindo um algoritmo computacionalmente mais oneroso (c.f. (DE JESÚS *et al.*, 2001a) e (DE JESÚS *et al.*, 2001b)).

Nas seções a seguir são descritas duas principais categorias de redes neurais dinâmicas.

### 3.2.8 - TDNN (*Time Delay Neural Network*)

As redes neurais de atraso de tempo, primeiramente descritas em (LANG AND HINTON, 1988; WAIBEL *et al.*, 1989) consistem em apresentar como entrada para a rede não apenas a entrada atual, mas um vetor com a entrada atual e as últimas **D** entradas anteriores. Isso geralmente é representado por um operador  $\mathbf{z}^{-1}$  entre cada uma das entradas. A esse conjunto da entrada atual e dos operadores  $\mathbf{z}^{-1}$  produzindo entradas atrasadas, dá-se o nome de **linha de atraso derivada**. Em resumo, as TDNNs não passam de redes estáticas que aprendem não com uma entrada, mas com uma seqüência de entrada (presente e passada). As redes TDNN podem apresentar diferentes arquiteturas em função de que camada(s) apresenta(m) ou não atrasos de tempo.

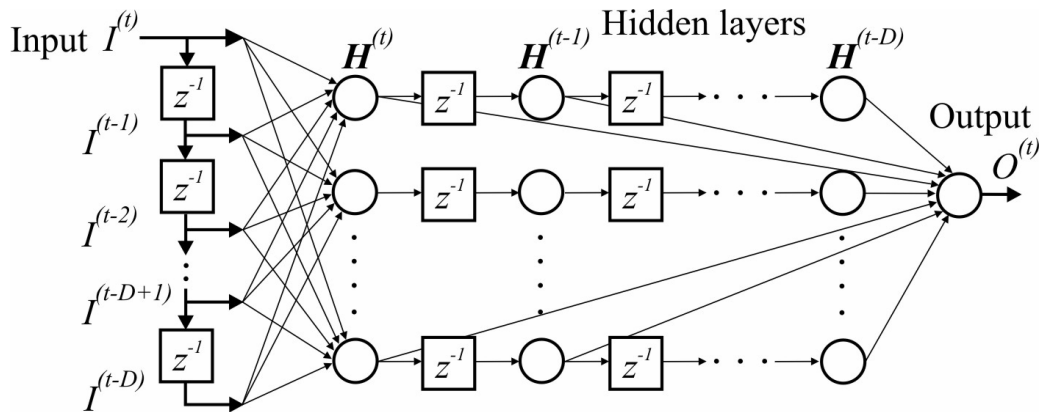
Uma abordagem consiste da primeira camada atrasada, sendo que a saída da camada de entrada é um vetor de dimensão **D**, constituído do sinal de entrada no tempo atual  $I(t)$  e os sinais passados  $I(t-1)$ ,  $I(t-2)$ , ...,  $I(t-D)$ . Para alguns autores (HAYKIN, *op. cit*) é também chamada TLFN focada (*focused Time Lagged Feedforward Network*). A Figura 3.11 ilustra esse modelo.



**Figura 3.11 – Time delay neural Network – TDNN**

O segundo tipo inclui somente atrasos de tempo nas camadas intermediárias sem atrasos na camada de entrada. A saída da estrutura que compõe as camadas intermediárias é composta de  $D$  vetores, constituída de uma primeira camada no tempo atual e camadas intermediárias  $H(t-1)$ ,  $H(t-2)$  ...,  $H(t-D)$  no tempo passado.

O terceiro tipo consiste de camadas de entrada e intermediárias atrasadas no tempo, conforme ilustrado na Figura 3.12.



**Figura 3.12 – Time Delay Neural Network – TDNN**

### 3.2.9 - Recorrentes

Para representar memória, uma alternativa ao atraso de tempo é a **realimentação** (ELMAN, 1990). A realimentação consiste em fornecer como entrada numa iteração a saída de uma interação anterior. Ela pode ser:

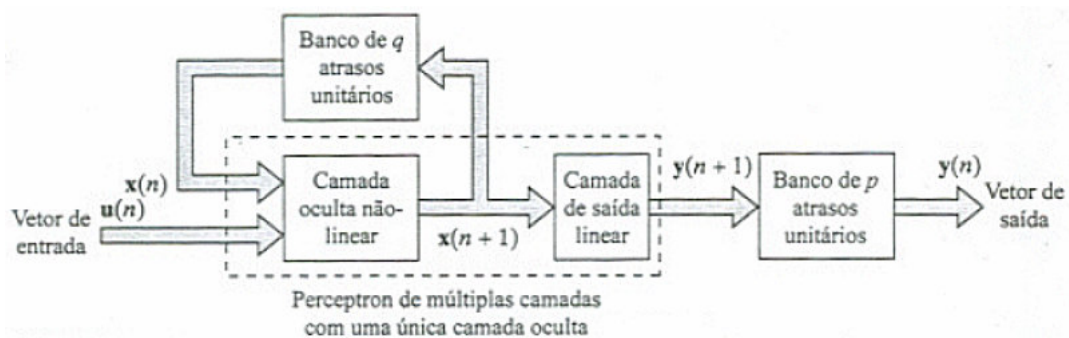
- **Local:** quando ocorrer no escopo de cada neurônio, *i.e.* quando a saída de um neurônio também participar da sua entrada ou
- **Global:** quando ocorrer no escopo de uma camada ou da rede toda, *p.ex.* a saída da rede é fornecida como entrada da rede.

Redes *feedforward* com um ou mais laços de realimentação são ditas **recorrentes**. Assim como as *TDNNs*, as redes recorrentes possuem uma variedade de arquiteturas. E é essa variedade, aliada à possibilidade de representação de memória não só da entrada, mas também da saída o que torna as redes recorrentes um caso ainda mais generalizado das redes *feedforward*.

Dentre as várias arquiteturas possíveis para as redes recorrentes, duas das redes recorrentes mais utilizadas são: o **modelo de espaço de estados** e o **modelo recorrente de entrada/saída** (HAYKIN, *op. cit.*). Tanto estes quanto os outros compartilham de duas características básicas: baseiam-se numa rede estática *feedforward* de múltiplas camadas e exploram a capacidade de mapeamento não-linear da rede estática multi-camadas.

### 3.2.9.1 - Modelo de espaço de estados

Neste modelo genérico, os neurônios da camada oculta constituem o estado da rede. A saída da camada oculta é realimentada, concatenada à entrada exógena, através de um banco de atrasos unitários. A Figura 3.13 mostra a arquitetura deste modelo. Note que a saída da rede também sofre um atraso unitário, pois a saída da rede está sempre uma unidade adiantada.



**Figura 3.13 – Arquitetura do modelo de espaço de estados**

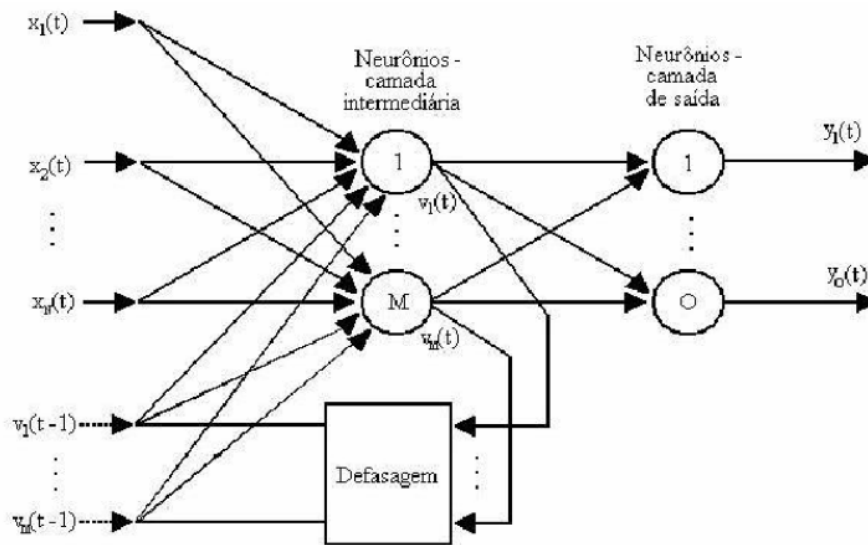
Sendo  $\mathbf{u}(\mathbf{n})$  m-por-1 o vetor de entrada exógena,  $\mathbf{x}(\mathbf{n})$  q-por-1 a saída da camada oculta,  $\mathbf{y}(\mathbf{n})$  a saída da rede todos no tempo  $\mathbf{n}$ ,  $\mathbf{f}()$  a função não-linear que caracteriza a camada oculta e  $\mathbf{C}$  a matriz de pesos, o modelo comporta-se segundo:

$$\mathbf{x}(n+1) = \mathbf{f}(\mathbf{x}(n), \mathbf{u}(n)) \quad (\text{Eq. 3.19})$$

$$\mathbf{y}(n) = \mathbf{C}\mathbf{x}(n) \quad (\text{Eq. 3.20})$$

Note que a saída da camada oculta é concatenada uma a cada intervalo de tempo graças ao banco de atrasos unitários e que este fornecimento possui uma defasagem de q períodos, i.e. a saída de uma dada iteração leva q iterações para chegar à entrada da camada novamente. Ainda digno de nota é que a camada de saída possui função característica linear.

ELMAN (1990) descreve um caso particular deste modelo chamado Rede recorrente simples onde o atraso da camada de saída não é utilizado e a função de saída pode ser não-linear. A seguir um esquema da arquitetura da rede RRS de Elman:

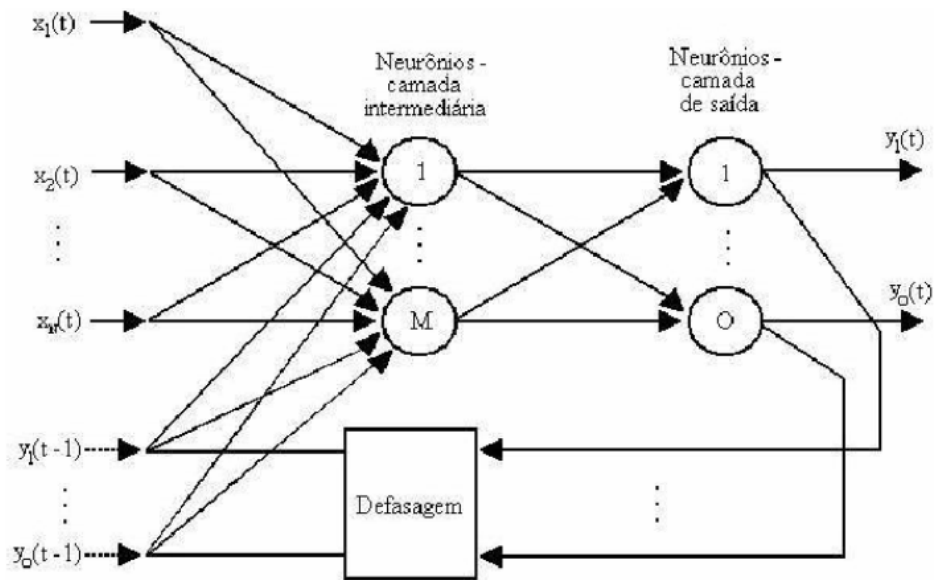


**Figura 3.14 - Rede recorrente de Elman**

Devido à existência do banco de atrasos, a rede neural de Elman pode continuar gerando informação mesmo após o término da estimulação externa da rede. Isso significa que estas redes não são apenas um registro seqüencial dos dados do passado.

### 3.9.2 - Jordan

Uma variação da rede de Elman é a rede recorrente de Jordan (MANDIC & CHAMBERS, 2001) onde a entrada não é realimentada da saída da camada oculta, mas da camada de saída como pode ser visto na Figura 3.15:

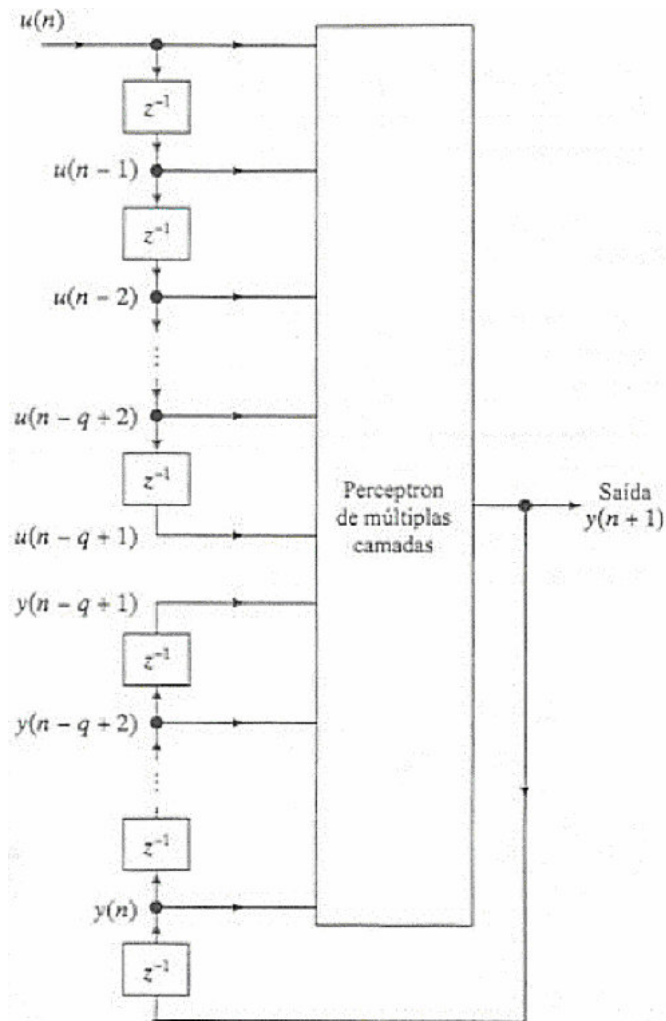


**Figura 3.15 - Rede recorrente de Jordan**

### 3.9.3 Modelo recorrente de entrada/saída

Já neste modelo, a entrada é constituída da entrada atual, uma memória de linha de atraso com derivação para esta entrada e outra para a saída da rede. O tamanho dessas memórias é independente um do outro. Aqui também, a saída apresenta-se defasada (entrada:  $u(n)$  e saída  $y(n+1)$ ).





**Figura 3.16 – Modelo auto-regressivo não linear com entradas exógenas (NARX)**

Esse modelo também é chamada *modelo auto-regressivo não-linear com entradas exógenas* (NARX, *Nonlinear AutoRegressive with eXogenous inputs*) e podem ser representados por:

$$\mathbf{y}(n+1) = \mathbf{F}(\mathbf{y}(n), \dots, \mathbf{y}(n-q+1), \mathbf{u}(n), \dots, \mathbf{u}(n-q+1)) \quad (\text{Eq. 3.21})$$

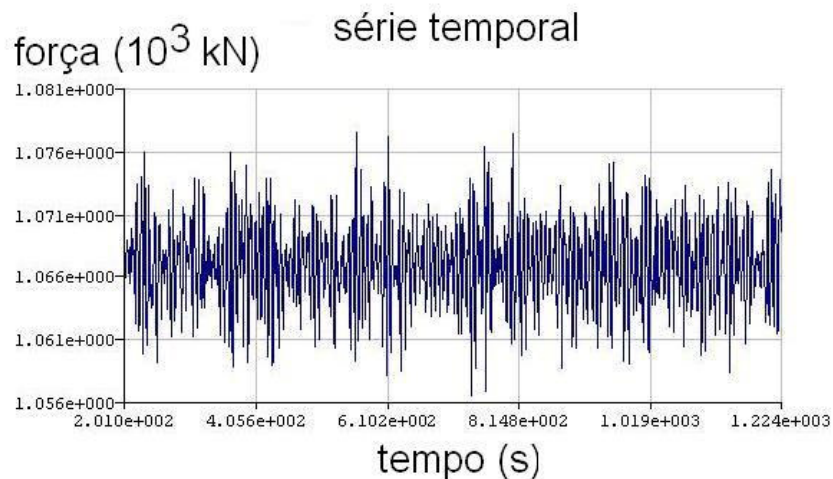
onde  $\mathbf{F}$  é a função não-linear característica da rede neural. Note mais uma vez que o valor de  $q$  para a memória dos dados exógenos e para a dos realimentados podem diferir.

### 3.3 - Sinais Aleatórios

No capítulo 4 serão descritos os dados que serão utilizados para gerar o modelo do comportamento da *Flexjoint*. Esses dados são sinais aleatórios no tempo discretos oriundos do processo de monitoração da plataforma PXVIII, sujeita à ação de cargas aleatórias de onda, vento e correnteza. Estes sinais são estatisticamente descritos por parâmetros tais como altura significativa do sinal ( $H_s$ ), período de pico do sinal ( $T_p$ ) e nesta seção será visto uma breve descrição destes.

#### 3.3.1 Descrição dos parâmetros

Seja a série temporal de um sinal aleatório conforme é ilustrado na Figura 3.17. Neste exemplo, o sinal corresponde à força axial no topo do SCR da PXVIII, avaliada para um dado carregamento através de simulação numérica.

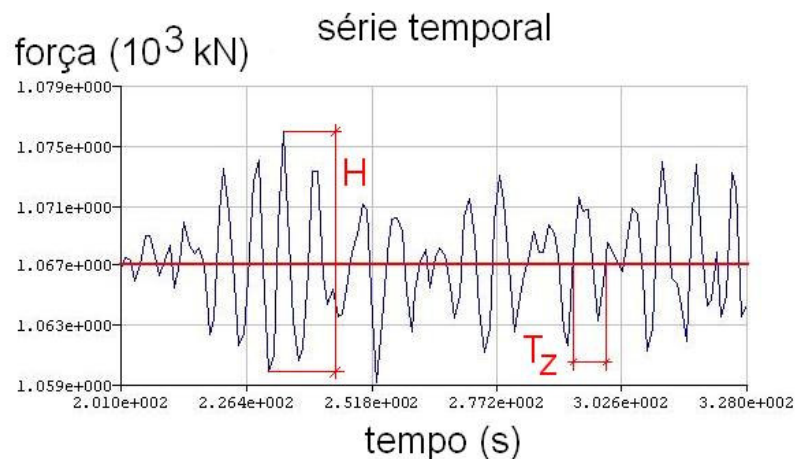


**Figura 3.17 – Série temporal de um sinal aleatório**

O tempo total  $T_s$  desta série temporal corresponde a 1024 s. Na Figura 3.17, convém observar que estão sendo considerados os registros entre os instantes 201 s e 1224 s. Do tempo total da simulação, foram descartados os registros correspondentes aos 200 s

iniciais, a fim de expurgar efeitos transientes inerentes do processo (artificialmente) gradual de aplicação das cargas no modelo numérico.

A fim de observar melhor o comportamento do sinal, seja um intervalo de 128 s tomado do início da série.



**Figura 3.18 – Detalhe da série temporal de um sinal aleatório**

Tomando por referência o valor médio do sinal, correspondente, neste exemplo, a 1067 kN, o período de cruzamento zero ( $T_z$ ) de um dado ciclo corresponde ao intervalo de tempo que decorre entre duas elevações consecutivas do valor do sinal acima do valor médio, conforme mostrado na Figura 3.18. O **período de cruzamento zero do sinal** corresponde à média de todos os  $T_z$  observados ao longo do tempo total da série ( $T_s$ ).

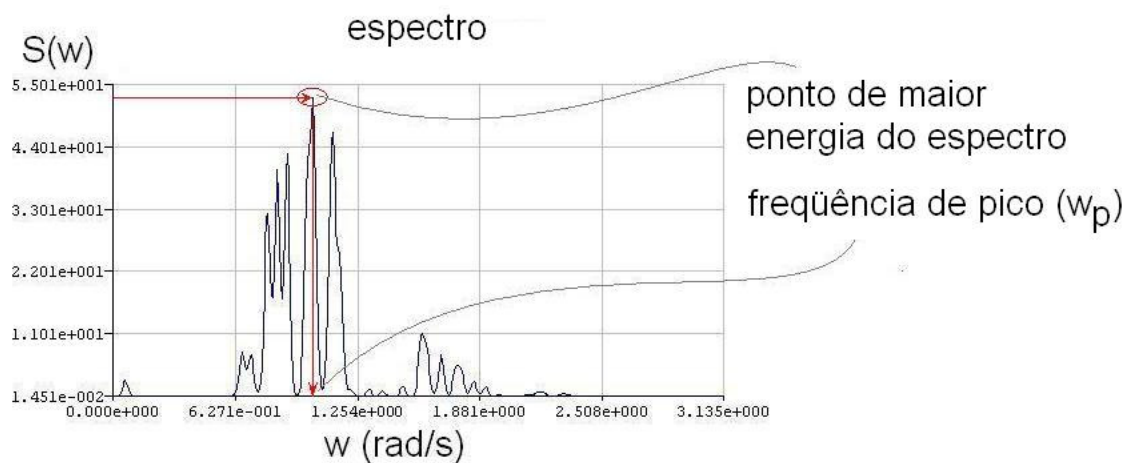
A altura ( $H$ ) de um dado ciclo corresponde à diferença entre os valores (do sinal) correspondentes a um cavado e uma crista consecutivos, conforme mostrado na Figura 3.1.8.

A **altura máxima do sinal** ( $H_{m\acute{a}x}$ ) corresponde ao maior valor  $H$  observado ao longo do tempo total da série ( $T_s$ ). A **altura significativa do sinal** ( $H_s$ ) – também chamada por  $H_{1/3}$  – corresponde à média do terço superior de todas as alturas ( $H$ ) observadas durante  $T_s$ . Para obtenção deste valor a partir da série temporal, basta ordenar de forma crescente todos os valores de altura ( $H$ ) obtidos ao longo do tempo total da série, dividir está série de

valores assim ordenada em três seqüências de igual tamanho, e calcular a média da última delas.

No caso específico de sinais correspondentes à elevação da superfície do mar devida à ação de ondas aleatórias, a altura significativa do sinal, ou melhor, da onda é muito próxima da altura de onda que é informada pelos capitães a partir de sua observação visual no mar (CHAKRABARTI, 1987).

Para entender o significado do período de pico de um sinal, é necessário considerar o espectro de energia deste sinal, o qual é obtido a partir da transformação de Fourier da função de auto-correlação calculada sobre a série temporal (CHAKRABARTI, 1987; NEWLAND, 1980). O período de pico do sinal ( $T_p$ ) é o inverso da frequência de pico do sinal ( $\omega_p$ ), que, por sua vez, é a frequência de maior energia do sinal, ou seja, ordenada correspondente ao valor máximo da função espectro  $S(\omega)$ . O espectro correspondente à série temporal ilustrada na figura 3.17 é exibido na figura 3.19 a seguir.



**Figura 3.19 – Espectro de energia de um sinal aleatório**

A título de exemplo, a frequência de pico ( $\omega_p$ ) do espectro mostrado na Figura 3.19 vale 1,13 rad/s. O período de pico ( $T_p$ ) correspondente é 5,56 s.

### 3.4 – Histerese

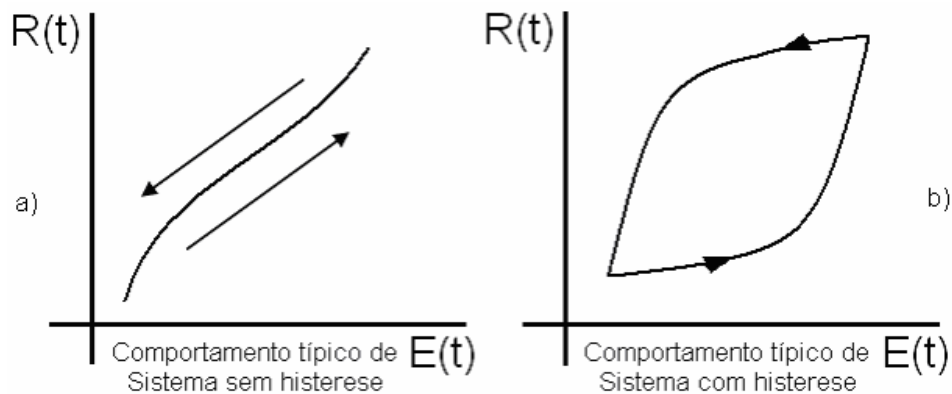
O termo histerese pode ser encontrado na literatura relacionado aos mais diversos tópicos (física, matemática, economia, biologia, computação, etc). A rigor, Histerese é um fenômeno do campo da termodinâmica que pode apresentar-se de três maneiras independente da taxa, dependente da taxa e relaxação térmica (Falabella,1984). Por ser, não completamente dominado, há muitas definições e caracterizações diferentes dependendo da orientação que se queria dar ao assunto.

Boa parte da literatura pesquisada (WEI and C.T. SUN, 2000), (TAWFIK *et al.*, 1997), (ALMEIDA, 2003) não define, especificamente, o fenômeno mas tenta passar seu significado através de propriedades matemáticas como por exemplo: “sistema com efeito de memória independente da taxa”.

Antes de conceituar histerese, faz-se necessário definir o que entendemos por sistema. Todo sistema ou material pode ser pensado como apresentando resposta  $R(t)$  a estímulos  $E(t)$  de acordo com suas características internas. E e R foram descritas como função do tempo pois pensamos em fornecer estímulos variáveis e cíclicos (carregamento e descarregamento). Geralmente, mas nem sempre, conhece-se o comportamento dos estímulos e se está interessado em conhecer as respostas do sistema e esses estímulos.

Para o trabalho, pode-se pensar em histerese como um comportamento de um sistema ou material em que a resposta ao carregamento é diferente da resposta ao descarregamento. Em outras palavras, histerese é uma propriedade de sistemas (físicos) de não seguir instantaneamente as forças aplicadas a eles mas, em vez disso, reagem mais lentamente ou não retornam completamente ao estado original, tornando o estado atual dependente do histórico dos estados recentes.

Como pode ser observado na figura 3.20.b, a relação E/R deixa de ser uma função pois existe pelo menos um valor de estímulo para o qual existem duas respostas possíveis. É de se notar também que o comportamento histerético pode dar-se com um gráfico no sentido horário ou anti-horário, o que representa situações físicas diferentes.



**Figura 3.20 – Esquema de sistema com histerese (b) e sem histerese (a)**

Mas só isso não basta para diferenciar a histerese de outros fenômenos dinâmicos. Há duas outras propriedades / características particulares: independência da taxa e memória.

A rigor, independência de taxa é uma definição matemática. Em geral, observa-se que a dependência de taxa existe, mas só ocorre / se manifesta em casos de carregamento mais extremos, podendo ser desprezada nos outros casos, levando à caracterização da histerese, principalmente na matemática, como independente da taxa. Quando a velocidade de variação de  $E(t)$  tende a zero, o sistema tende a não depender mais desta taxa. Para prever o comportamento de um sistema independente da taxa é necessário apenas os valores extremos recentes e sua a posição atual.

Além da independência de taxa, histerese também é conhecida por apresentar memória. O estado de um sistema sem memória pode ser determinado simplesmente a partir do seu estado atual e do estímulo que recebe. Sistemas com memória possuem comportamento dependente do estado atual e não só do estímulo recebido, mas também do histórico de estímulos anteriormente recebidos. É exatamente a independência da taxa que diferencia o efeito de memória histerético dos efeitos de memória de curto prazo (Pare,1998).

## **CAPÍTULO IV - DESCRIÇÃO DOS DADOS**

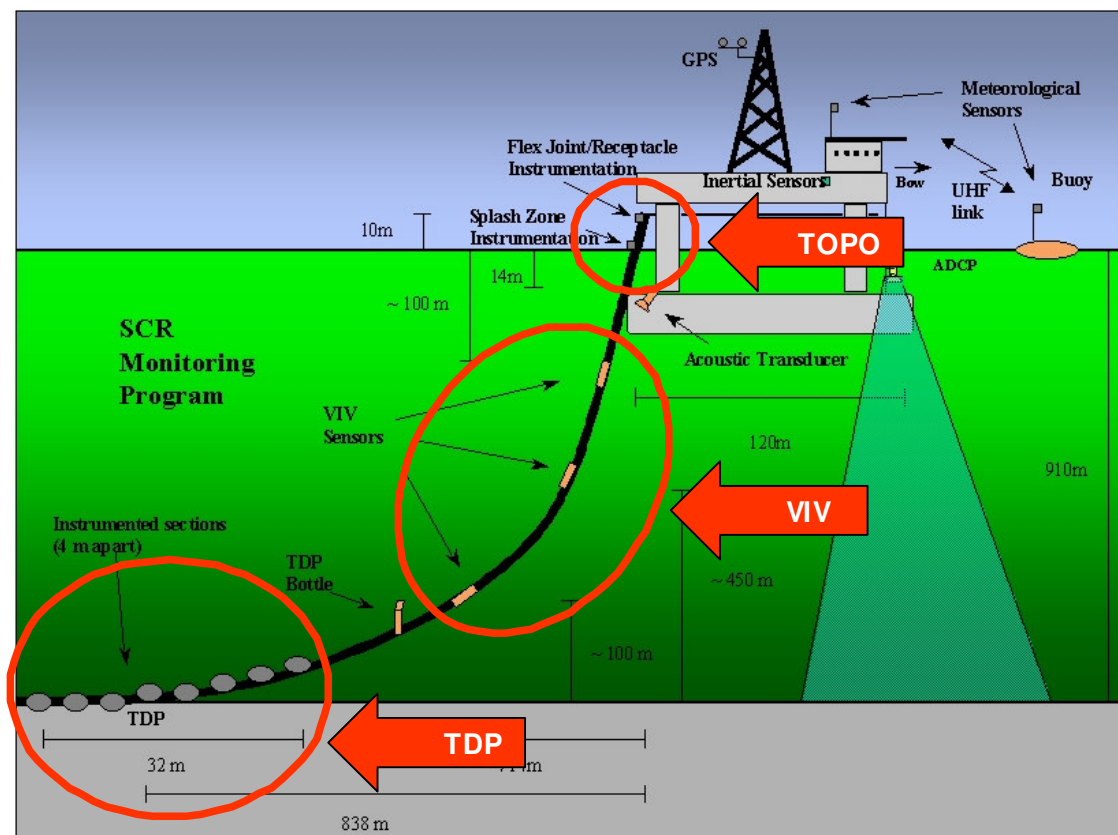
### **4.1 – Introdução**

O programa de monitoração do *riser* SCR da plataforma PXVIII da PETROBRAS foi projetado para permitir a calibração de modelos computacionais usados em vários estágios de análises da plataforma, do sistema de ancoragem, do projeto de *riser* e da modelagem de vibrações induzidas por vórtices (VIV), para sistemas de produção de petróleo em águas profundas (MACHADO et al,1999).

Em agosto de 1997, usando o projeto protótipo do SCR como guia, parâmetros foram selecionados e sensores e equipamentos foram requeridos e funcionalidades definidas. O programa também tinha que lidar com a estratégia de aquisição, requisitos de processamento de dados para a grande quantidade de dados gerados, sua implementação, coordenação e interfaces com diversos fornecedores, logística e aspectos operacionais.

O *riser* é composto de tubos de 8 m com 273,10 mm de diâmetro externo e a *flexjoint* no topo, com um comprimento suspenso na configuração média de 1319 m, e foi instalado

com um ângulo de partida no topo de 20 graus com a vertical. Os parâmetros de interesse identificados são parâmetros ambientais: perfil de corrente na coluna do mar, espectro de onda direcional, vento; movimentos e posição da plataforma; tensões nas linhas de ancoragem; carregamentos no topo do *riser*, zona de ondas e do Touch Down Point (TDP); movimentos do *riser*, incluindo aqueles devido ao VIV. Um esquema gráfico dos principais pontos de monitoração é mostrado na Figura 4.1.



**Figura 4.1 – Riser SCR e instrumentação**

**Fonte: REMO Z. Machado, *et al.*,1999.**

Inicialmente, os dados monitorados do sistema real são transferidos para arquivos e submetidos a procedimentos de pré-processamento (verificação de tamanhos e formatos de arquivos), processamento (verificação, filtragem e transformação para parâmetros de interesse) e só então armazenados em tabelas no banco de dados ORACLE v.8.0, ocupando



mais de 40 Gb de memória. Os arquivos de dados são convertidos para texto ou Excel que são mais facilmente acessíveis pelo MATLAB.

A aquisição dos dados foi realizada por quatro sub-sistemas:

- Sub-sistema de Monitoração Meteo-Oceanográfica
- Sub-sistema de Posicionamento de Superfície
- Sub-sistema de Posicionamento do SCR
- Sub-sistema de Monitoração de Esforços e Vibrações

No item seguinte encontram-se relacionados os parâmetros monitorados para o Sub-sistema de Monitoração de Esforços e Vibrações, no qual estão os dados referentes a *flexjoint*.

#### **4.2 - Sub-sistema de monitoração de esforços e vibrações**

Neste sistema, os principais pontos considerados foram: TOPO, TDP e pontos críticos sujeitos aos movimentos induzidos por ondas, ventos e correntes marinhas (VIV). Foram então monitorados as inclinações, esforços axiais e de flexão e movimentos do SCR.

##### **Principais variáveis monitoradas**

- esforços:
  - tração
  - momento fletor
- inclinações da plataforma e do topo do SCR
- vibrações induzidas por vórtices (VIV)
- pressão interna do *riser*

- posicionamento e movimentos da plataforma
  - tração nas amarras

Em particular, no topo do *riser* foram medidos esforços no seu receptáculo de sustentação, em diversos pontos abaixo da *flexjoint* (ver Figura 4.2), as inclinações nos dois eixos acima e abaixo da *flexjoint*, e a pressão interna na tubulação. Esta monitoração foi realizada por cerca de três anos.

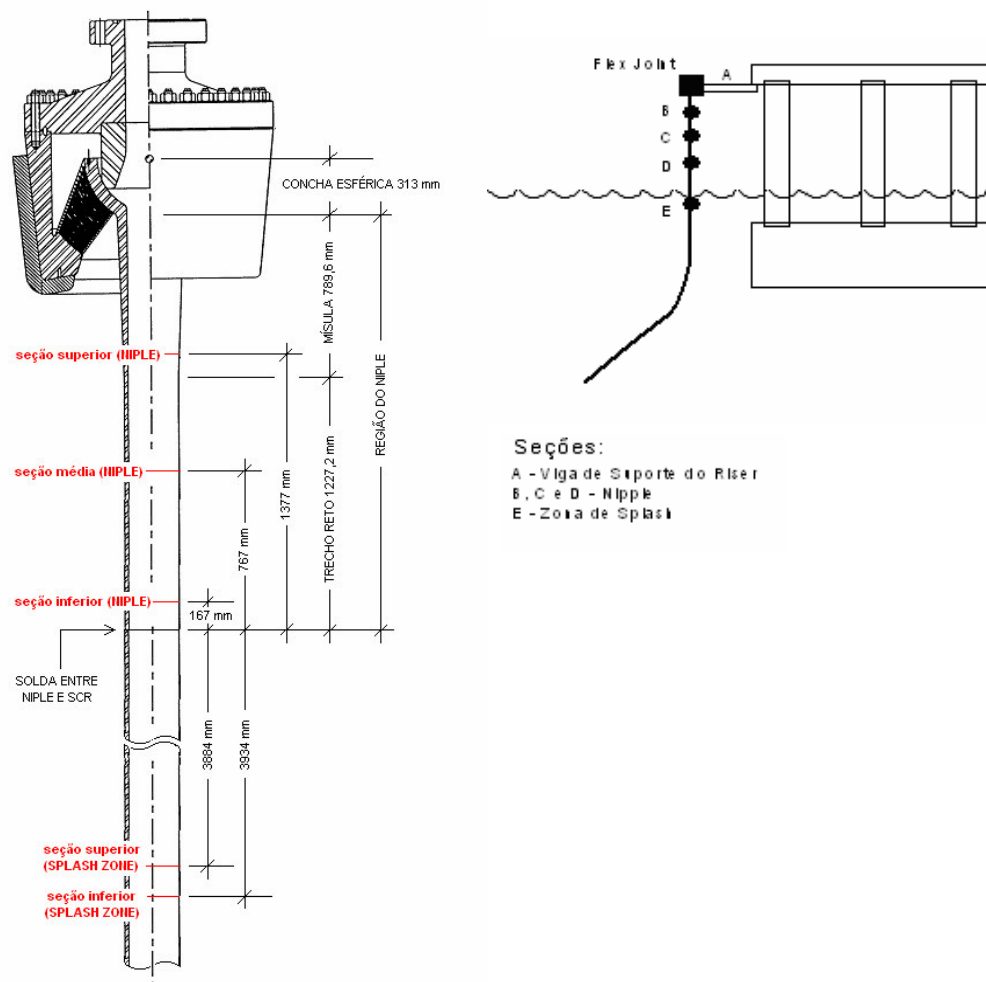


Figura 4.2 – Indicação dos pontos de instalação dos sensores na *flexjoint*

Os três pontos principais onde se localizam os sensores são descritos a seguir:

- 1) *niple*: trecho tubular da *flexjoint*, cuja extremidade inferior está soldada ao SCR enquanto a extremidade superior se abre em forma de concha esférica no interior do corpo principal da peça, onde se apóia no suporte elastomérico. Imediatamente abaixo da concha, apresenta 789,6 mm de trecho de tubo com espessura de parede variável, seguido por trecho retilíneo subjacente, com mesma seção transversal do SCR, com 1227,2 mm de comprimento. Na prática, o *niple* é uma extensão do SCR; portanto, sensores instalados nesta região fazem medições de esforços, próximo ao topo do SCR, e inclinações do SCR.
- 2) *splash zone*: região do SCR onde a lâmina d'água tem contato.
- 3) viga suporte: sensores instalados na viga suporte estão localizados, na prática, no receptáculo da *flexjoint*, ou seja, no seu suporte, o qual é solidário à plataforma. Portanto, inclinômetros instalados neste local medem as inclinações da plataforma.

### 4.3 - Variáveis do problema

Nem todos os dados provenientes do Sub-sistema de Monitoração de Esforços e Vibrações interessam para o problema em estudo, visto que o subsistema é formado por outros componentes além da *flexjoint*. As variáveis monitoradas relevantes para esse estudo são:

- Esforço de flexão (momento fletor),
- Esforço axial (tração) e
- Inclinações em relação à vertical, abaixo da *flexjoint*.

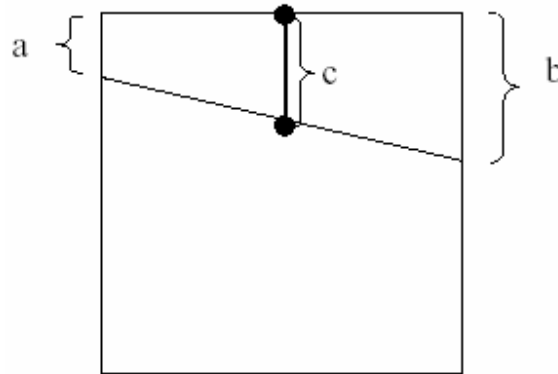
A tabela 4.1 abaixo mostra as variáveis primitivas de interesse obtidas dos bancos de dados da PXVIII.

**Tabela 4.1 – Descrição das variáveis de interesse**

nome do sensor	canal	local da instalação do sensor	posição do sensor (relativa ao local da instalação)	tipo de leitura	plano em que foi feita a medição	unidade
NUM0	19	niple	seção superior	momento fletor	plano da catenária	kN.m
NUM90	20	niple	seção superior	momento fletor	plano perpendicular à catenária	kN.m
NUT	21	niple	seção superior	força axial	plano paralelo à seção transversal do SCR	kN
IR0	34	niple		inclinação do riser	plano da catenária	graus
IR90	35	niple		inclinação do riser	plano perpendicular à catenária	graus
IP0	36	viga suporte		inclinação da plataforma	plano da catenária	graus
IP90	37	viga suporte		inclinação da plataforma	plano perpendicular à catenária	graus
APR0	57	variável calculada		ângulo relativo formado entre a plataforma e o riser	plano da catenária	graus
APR90	58	variável calculada		ângulo relativo formado entre a plataforma e o riser	plano perpendicular à catenária	graus

As variáveis NUM0 e NUM90 (NUM – *Niple Upper Momento*) referem-se aos momentos fletores na direção da catenária e na direção perpendicular à catenária do SCR medidos na posição do *niple* superior (Niple Upper). O esforço axial (tração) é representado pela variável NUT, medida no plano paralelo à seção transversal do SCR. O *strain gage* foi o sensor adotado para medição desses esforços.

O *strain gage* mede a deformação axial de um elemento metálico através de variação de sua resistência elétrica, com o auxílio de uma ponte de Wheatstone. Sendo assim, a partir das deformações axiais medidas, são determinados os esforços axiais e de flexão no plano que contém a catenária, e no perpendicular a este. Para isto, foram instalados 12 sensores por seção, ligados em ponte completa de Wheatstone, sendo dois conjuntos para medir os esforços de flexão e um para a medição do esforço axial. Então, supondo que houve uma deformação total conforme o esquema da Figura 4.3 abaixo:



**Figura 4.3 – Esquema representando o tubo com deformação axial e flexional**

A deformação devida à tração seria :  $c = \frac{a+b}{2}$  **(Eq. 4.1)**

E para o momento tem-se :  $d = \frac{a-b}{2}$  **(Eq. 4.2)**

O esforço devido à tração é dado por:

$$\text{Tração} = EA * \frac{a+b}{2} \quad \textbf{(Eq. 4.3)}$$

E o momento fletor é obtido da seguinte forma:

$$\text{Momento} = EZ * \frac{a-b}{2} \quad \textbf{(Eq. 4.4)}$$

onde :

E = módulo de elasticidade (kN/m<sup>2</sup>)

Z= módulo da seção(m<sup>4</sup>) dado por :  $Z = 2 * \frac{I}{De}$

I = momento de inércia da seção (m<sup>3</sup>)

De = diâmetro externo da seção (m)

A= área da seção(m<sup>2</sup>)

Para o SCR em estudo, essas constantes têm os valores apresentados na tabela 4.2 abaixo:

**Tabela 4.2 – Tabela com as constantes utilizadas para o cálculo dos esforços**

Descrição	Símbolo	Unid	Valores
módulo de elasticidade	E	kN/m <sup>2</sup>	2.07E+08
diâmetro externo da seção	D	m	0.3198
diâmetro interno da seção	d	m	0.2318
área da seção	A	m <sup>2</sup>	0.0381239
momento de inércia da seção	I	m <sup>3</sup>	0.0003717
módulo da seção	Z	m <sup>4</sup>	0.0023247
EA		kN	7.89E+06
EZ		kN.m	4.81E+05

Os ângulos relativos formados entre as direções 0° e 90°, entre a plataforma e o *riser* (APR0 e APR90), não são medidos diretamente, mas calculados a partir das inclinações de cada um. Assim:

$$APR0 = (IP0 - IR0) - 20,5^\circ \text{ e}$$

$$APR90 = IP90 - IR90.$$

Sendo que IP é o inclinômetro instalado na viga suporte, ou melhor, no receptáculo (que é solidário à plataforma), e IR é o inclinômetro instalado no *niple*, ou seja, no trecho tubular da *flexjoint* que, estando soldado ao topo do SCR, acompanha os movimentos do *riser*.

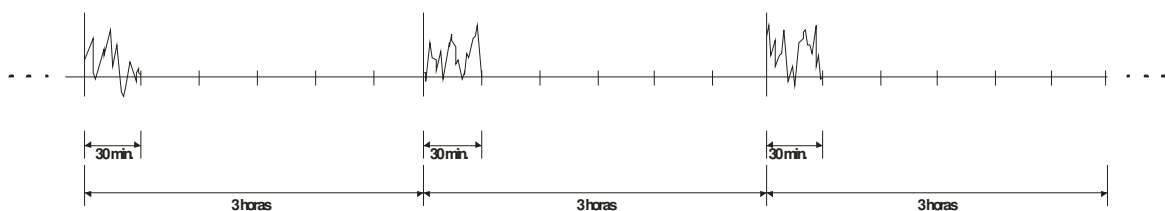
No cálculo do APR0 é descontado o ângulo de instalação de topo do SCR, que é de 20,5°. A unidade de medida dos ângulos é graus.

#### 4.4 - Características dos dados e período de análise

O período considerado neste estudo foi de **18 de julho de 2001 a 20 de novembro de 2001**, por se constituir um período de aquisição com menor incidência de erros de aquisição, ruídos e falta de dados. Por exemplo, em um dado período aconteceu de haver o descolamento do *strain gage*, que é colado na parede do tubo, devido à variação de temperatura. Esse fato gerou um acréscimo desse deslocamento nas deformações axiais.

A partir desse período bem comportado, foram obtidos do banco de dados as séries temporais de cada variável. O arquivo é composto de 948 séries discretas, cada série representando um estado de mar. Resumidamente, pelo menos três informações básicas são necessárias para especificar as condições de um certo estado de mar: altura, período e direção de onda. Durante o período que define a série e conseqüentemente o estado de mar, o mar é descrito como um processo aleatório, ergódico e estacionário. Ou seja, para um estado de mar aleatório são garantidas que as propriedades estatísticas são invariantes no decorrer do tempo. Esse fato é relevante para as análises dinâmicas aleatórias de *riser* (CAMPELLO, 2007).

O período definido para satisfazer as condições acima (ergodicidade e estacionariedade) foi de 30 minutos. Então, para cada série, os registros foram feitos por cerca de 30 minutos a cada 3 horas, 24 horas por dia. A frequência de amostragem foi de 1Hz (1s), perfazendo-se um total de cerca de 1800 pontos por série (MACHADO et al,1999). Este período foi determinado pela capacidade do sistema de aquisição de dados. A Figura 4.4 abaixo ilustra a composição das séries ao longo do período.



**Figura 4.4 – Composição das séries temporais ao longo do período**

## 4.5 - Análise exploratória dos dados no domínio do tempo e da frequência

Embora os dados da campanha de monitoração já tenham sido previamente tratados , algumas análises estatísticas e gráficas foram realizadas com a finalidade de se ter mais conhecimento quanto à adequabilidade dos dados para o estudo proposto neste trabalho.

A análise estatística foi realizada tanto no domínio do tempo quanto no domínio da frequência. A análise no tempo permitiu avaliar os limites permissíveis de cada variável descrita no memorial descritivo MD-3010.18-5520-850-PPC- 001 (MACHADO et al,1999), com os valores efetivamente realizados, e corrigir esses valores em função de condições iniciais já existentes antes da instalação dos sensores.

A análise do sinal no domínio da frequência oferece muitas informações que escapam à análise no domínio do tempo. Por exemplo, através da análise espectral, é possível identificar se as causas da deflexão são causadas pela ação das ondas ou pelo VIV, assim como identificar a participação individual de cada esforço em termos de energia na vibração do *riser*.

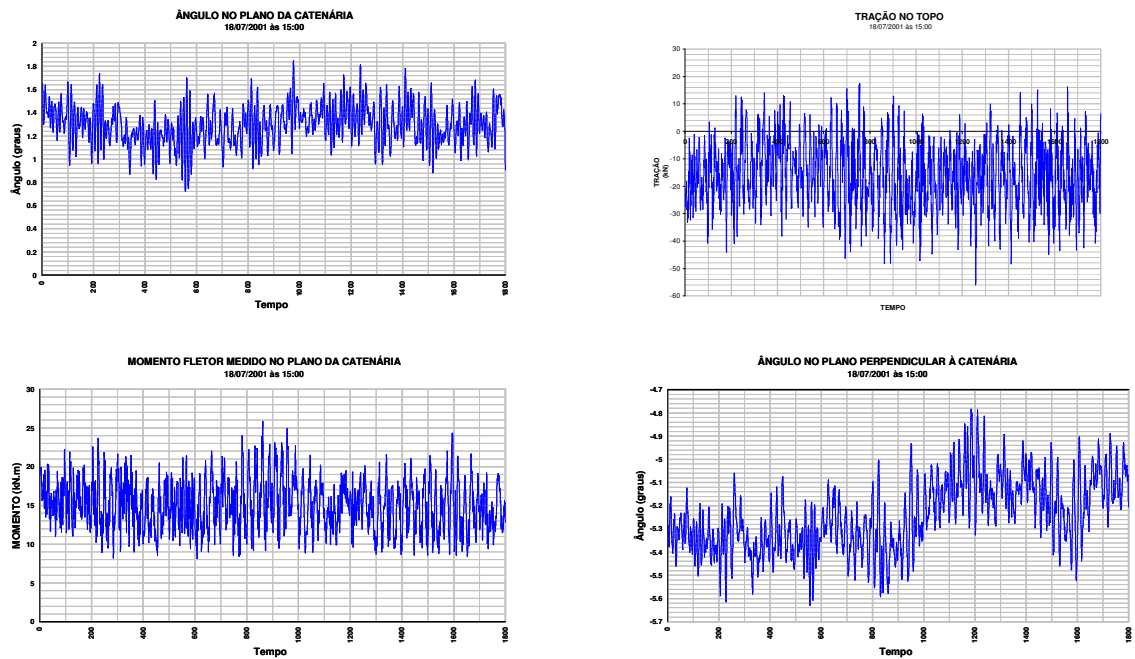
A seguir, é feita uma descrição dos procedimentos adotados para análise dos dados no domínio do tempo e da frequência.

### 4.5.1 Análise no domínio do tempo

Um exemplo das séries temporais das variáveis estudadas é mostrado na Figura 4.5.

Na Figura 4.5.b, podem-se observar valores negativos de tração, entretanto isso não significa dizer que o *riser* esteja comprimido. Os sensores (*strain gage*) foram instalados com esforços iniciais (residuais) na estrutura que não foram medidos. Conseqüentemente, os valores apresentados nas séries temporais partem de valores iniciais não observados. Uma opção para obter esses valores iniciais seria fazer uma simulação em função do estado de mar atuante no período de estudo considerado. Entretanto, para o período escolhido não foram registrados os estados de mar correspondentes.





**Figura 4.5 – Exemplo dos dados observados**

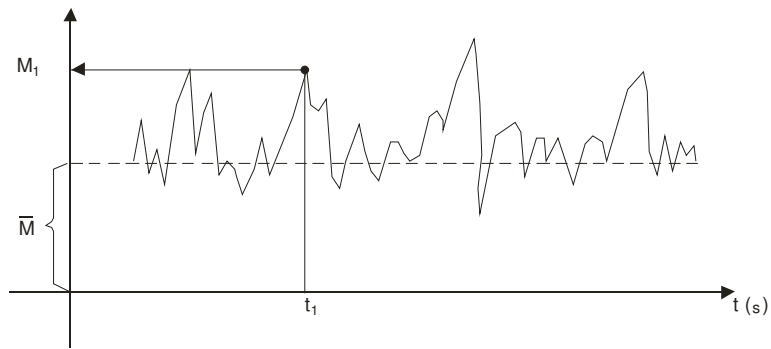
Uma alternativa seria trabalhar com valores teóricos, baseados em 97 casos de carregamento que foram considerados representativos do primeiro ano de campanha de monitoramento (MOURELLE *et al.*, 2000), fazendo-se simulações no domínio do tempo para todos eles, e construir intervalos de confiança para NUM0, NUM90 e NUT. Esses intervalos estão apresentados tabela 4.3.

**Tabela 4.3: Intervalos dos valores residuais das variáveis NUM0, NUM90 e NUT**

Variável	Unidade	Intervalo
NUM0	(kN.m)	+6.53 a +9.53
NUM90	kN.m	-1.66 a -0.94
NUT	kN	+1066.9 a +1079.4

Os intervalos foram construídos para um nível de confiança de 99.5%.

Então, os momentos e tração instantâneos serão acrescidos do valor inicial obtido da forma descrita acima. Foi utilizada a média do intervalo para cada variável considerada. Por exemplo, seja a série da Figura 4.6:



**Figura 4.6 - Série exemplo**

O momento corrigido é dado por:  $M(t_1)_{\text{corrigido}} = M_1 - \bar{M} + \hat{M}$  **(Eq. 4.5)**

onde :

$M(t_1)_{\text{corrigido}}$  = Esforço corrigido num dado tempo  $t_1$

$M_1$  = Esforço instantâneo medido

$\bar{M}$  = Esforço médio da série observada

$\hat{M}$  = Esforço médio estimado através de simulação

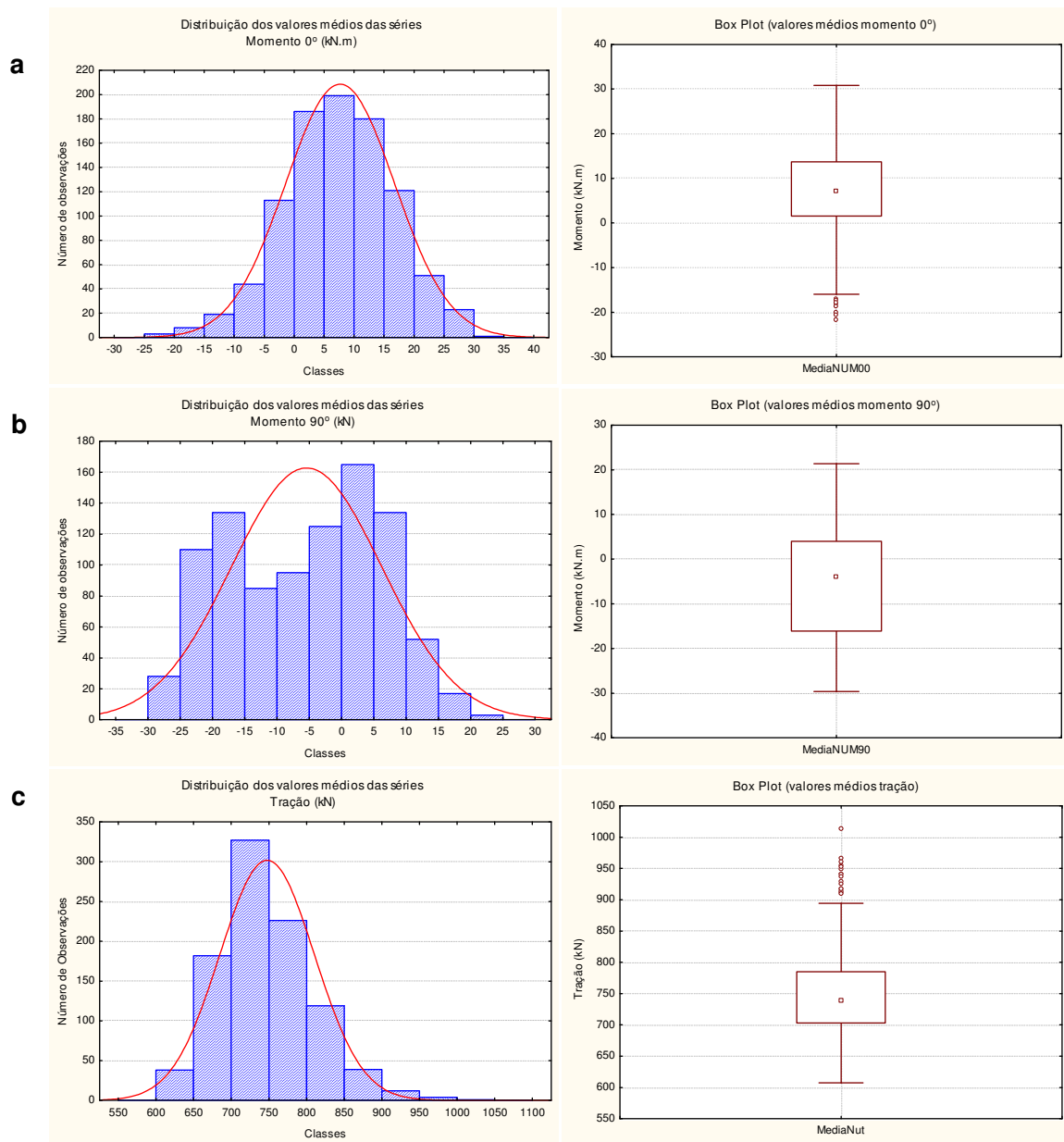
logo, para NUM0, tem-se

$$NUM0_{\text{corrigido}}(t) = NUM0(t) - \overline{NUM0} + \frac{(6.53 + 9.53)}{2} \quad \text{(Eq. 4.6)}$$

Para o ângulo, não é necessário esse ajuste porque os inclinômetros medem o ângulo realmente existente entre o *riser* e a plataforma.

Uma outra análise foi feita para avaliar se o *riser* em operação atende aos níveis de solicitação especificados no Memorial Descritivo MD-3010.18-5520-850-PPC- 001 (MACHADO et al,1999). e (OIL STATES INDUSTRIES, 1995) para detecção de possíveis valores discrepantes. A estimativa de valores típicos dos esforços foi feita em função dos estados de mar de um ano típico das condições de oceanográficas da região da PXVIII. Foram selecionados 3 estados de mar, um com menor excitação, outro com limite superior de excitação e por ultimo um estado de mar com excitação intermediária. Em condições extremas, o esforço axial pode chegar a 2000 kN e os momentos fletores a 340 kN.m (MACHADO et al,1999)..

A fim de realizar uma análise comparativa com os valores típicos (médias) e esforços extremos fornecidos, foram calculadas as estatísticas descritivas (média, desvio padrão, máximo e mínimo) das séries que compõem a janela de observação. Os histogramas e box plot da Figura 4.7 abaixo mostram a distribuição dos valores médios dos esforços .



**Figura 4.7 – Histogramas e Box plot dos valores médios dos esforços (tração e momento fletor).**

Para comparar os valores médios teóricos com os valores médios observados, adotou-se a construção de um intervalo de confiança para o desvio padrão dos valores médios observados e avaliar se o desvio padrão dos valores teóricos médios se encontram fora ou dentro do intervalo. A escolha dessa estatística se deve ao fato de que ela não reflete a influência da ausência dos valores residuais iniciais que não foram medidos, apesar de terem

sido corrigidos com valores de simulações.

Assim, baseado no Teorema do Limite Central em que, para grandes amostras, a distribuição das médias das amostras (cada serie) é descrita aproximadamente por uma normal, conforme pode ser constatado na Figura 4.7, os limites do intervalo de confiança para o parâmetro desvio padrão é dado por:

$$\sqrt{\frac{(n-1)s^2}{\chi_{n-1,\alpha/2}^2}} \leq \sigma \leq \sqrt{\frac{(n-1)s^2}{\chi_{n-1,1-\alpha/2}^2}} \quad (\text{Eq. 4.7})$$

onde:            n            = tamanho da amostra

                  s<sup>2</sup>            = variância amostral

$\chi_{n-1,\alpha/2}^2, \chi_{n-1,1-\alpha/2}^2$  = valores críticos da distribuição Qui-Quadrado

$\sigma$             = desvio padrão populacional

Para grandes amostras, o autor (COSTA NETO,1977) sugere construir o intervalo de confiança para o desvio com a seguinte expressão (equação 4.8):

$$s \pm Z_{\alpha/2} \frac{s}{\sqrt{2(n-1)}} \quad (\text{Eq. 4.8})$$

Então, com base nos dados observados da variável NUM0 (momento fletor no plano da catenária), a média dos valores médios é  $\overline{NUM0} = 15,82$ , com desvio padrão de  $s = 9,026$ . Adotando-se um grau de confiança de 95% ( $\alpha = 0,05$ ) e com n=948, os valores críticos da distribuição amostras  $\chi^2$  são  $\chi_{n-1,\alpha/2}^2 = 896,58$  e  $\chi_{n-1,1-\alpha/2}^2 = 1070,26$ . Substituindo-se os valores acima na expressão 4.7, tem-se:

$$8,68 < \sigma < 9,50$$

O desvio padrão dos valores médios obtidos através de simulação para o momento 0° (MY) é obtido através dos valores contidos Memorial Descritivo MD-3010.18-5520-850-PPC-001 (MACHADO et al,1999):

$$\mu = (5,32 + 1,99 + 0,71) / 3 = 2,67 \quad \text{(Eq. 4.9)}$$

$$\sigma = \sqrt{\frac{(5,32 - 2,67)^2 + (1,99 - 2,67)^2 + (0,71 - 2,67)^2}{3 - 1}} = 2,38 \quad \text{(Eq. 4.10)}$$

Com base no intervalo de confiança de 95% o desvio padrão dos valores médios típicos está muito além dos limites do intervalo construído baseado na amostra.

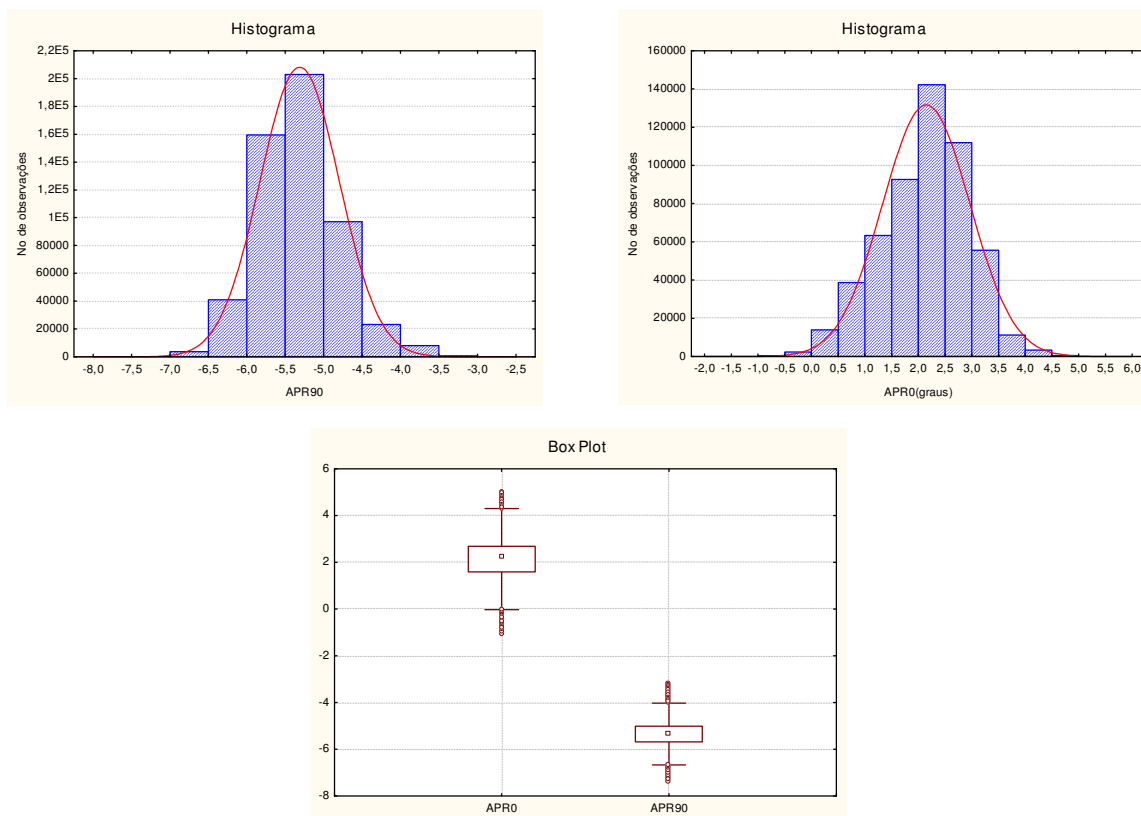
Infelizmente, os valores típicos só contemplam 3 estados de mar, ficando difícil uma comparação. Alguns fatores também fazem com que esses valores típicos se distanciem dos valores medidos. São resultados de simulações com excitação harmônica e não aleatória como é a realidade. Outro aspecto é a questão dos valores residuais iniciais não medidos. Foram adotados aqui valores que não mudaram ao longo da janela de observação e também foram obtidos a partir de simulações. Entretanto, fica aqui registrado uma possível metodologia para análise estatística considerando valores médios das series temporais.

Analisando as séries em termos de máximos, conforme pode ser visto na tabela 4.6, tanto o esforço axial quanto os momentos nas duas direções têm máximos inferiores (1451,06 kN e 64,06 kN.m) ao especificados em condições extremas (2000 kN e 340 kN.m).

Para os ângulos (APR0 e APR90) os valores limites especificados em (*Oil States Industries*, 1995) são de  $\pm 18,2^\circ$ . Como pode ser constatado nos histogramas e Box plot da Figura 4.8 e na Tabela 4.6, as variações angulares estão bem aquém do limite permitido, variando entre  $-1,48^\circ$  e  $+5,87^\circ$  para APR0 e entre  $-7,89^\circ$  e  $-3,16^\circ$  para APR90.

**Tabela 4.6 – Estatística descritiva das variáveis estudadas**

Variável	Média	Minimo	Máximo	Desvio padrão
<b>NUM0</b>	8.0297	-36.8371	64.0584	10.2938
<b>NUM90</b>	-6.4771	-52.9178	42.4167	12.5382
<b>Tração</b>	747.6484	-37.1994	1451.0581	63.9236
<b>APR0</b>	2.3945	-1.4810	5.8683	0.7158
<b>APR90</b>	-5.5315	-7.8857	-3.1620	0.5245



**Figura 4.8 – Histogramas e box plot das variáveis APR0 e APR90**

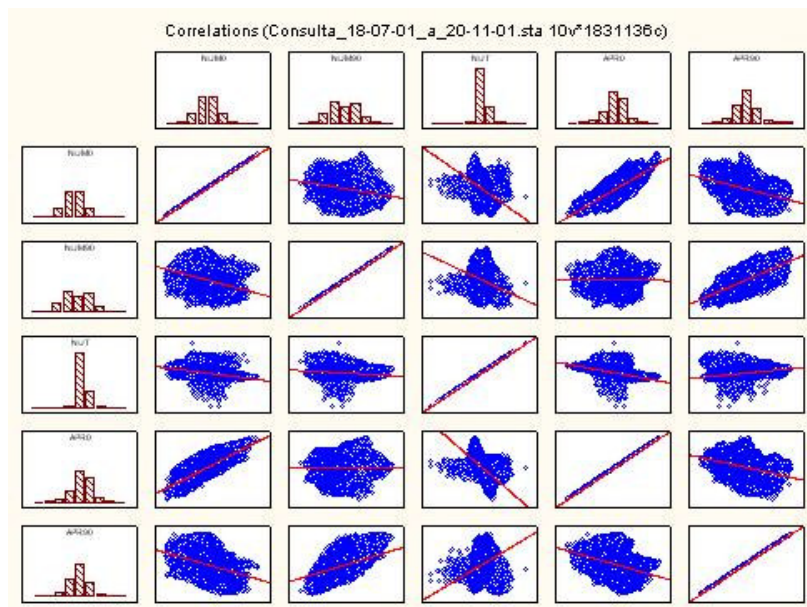
#### 4.5.1.1 Correlação

A correlação entre as variáveis medidas foi feita através da utilização da ferramenta estatística STATISTICA 7.1. A tabela 4.7 e a Figura 4.9 mostram as correlações e gráficos de dispersão entre todas as variáveis. Pode-se observar uma forte correlação positiva de 0,82 entre Num0 (momento fletor 0°) e Apr0 (Ângulo 0°), o que é coerente. A correlação entre NUM90 e APR90 é menor que entre NUM0 e APR0, porém ainda bastante significativa (0,54). A tração (NUT) mostra ter considerável influência nos ângulos e momentos. NUT e APR0 apresentam uma correlação negativa de 0,60 e entre NUT e NUM0 a correlação é de -0,45.

**Tabela 4.7 – Valores das correlações entre as variáveis**

	<b>NUM0</b>	<b>NUM90</b>	<b>NUT</b>	<b>APR0</b>	<b>APR90</b>
<b>NUM0</b>	1,00	-0,32	-0,45	0,82	-0,41
<b>NUM90</b>	-0,32	1,00	-0,24	-0,02	0,54
<b>NUT</b>	-0,45	-0,24	1,00	-0,60	0,37
<b>APR0</b>	0,82	-0,02	-0,60	1,00	-0,38
<b>APR90</b>	-0,41	0,54	0,37	-0,38	1,00





**Figura 4.9 – Gráfico de dispersão das variáveis**

#### 4.5.2 No domínio da frequência

Inicialmente, foi feita uma interpretação visual dos espectros de energia de um conjunto de 190 datas-hora, selecionadas aleatoriamente dentro do período estudado, com 950 séries. Os espectros foram obtidos através das funções disponíveis no MATLAB e também disponíveis no banco de dados da PXVIII. Regiões relevantes dos espectros obtidos são apresentadas nas Figuras abaixo. A partir delas, várias observações podem ser feitas:

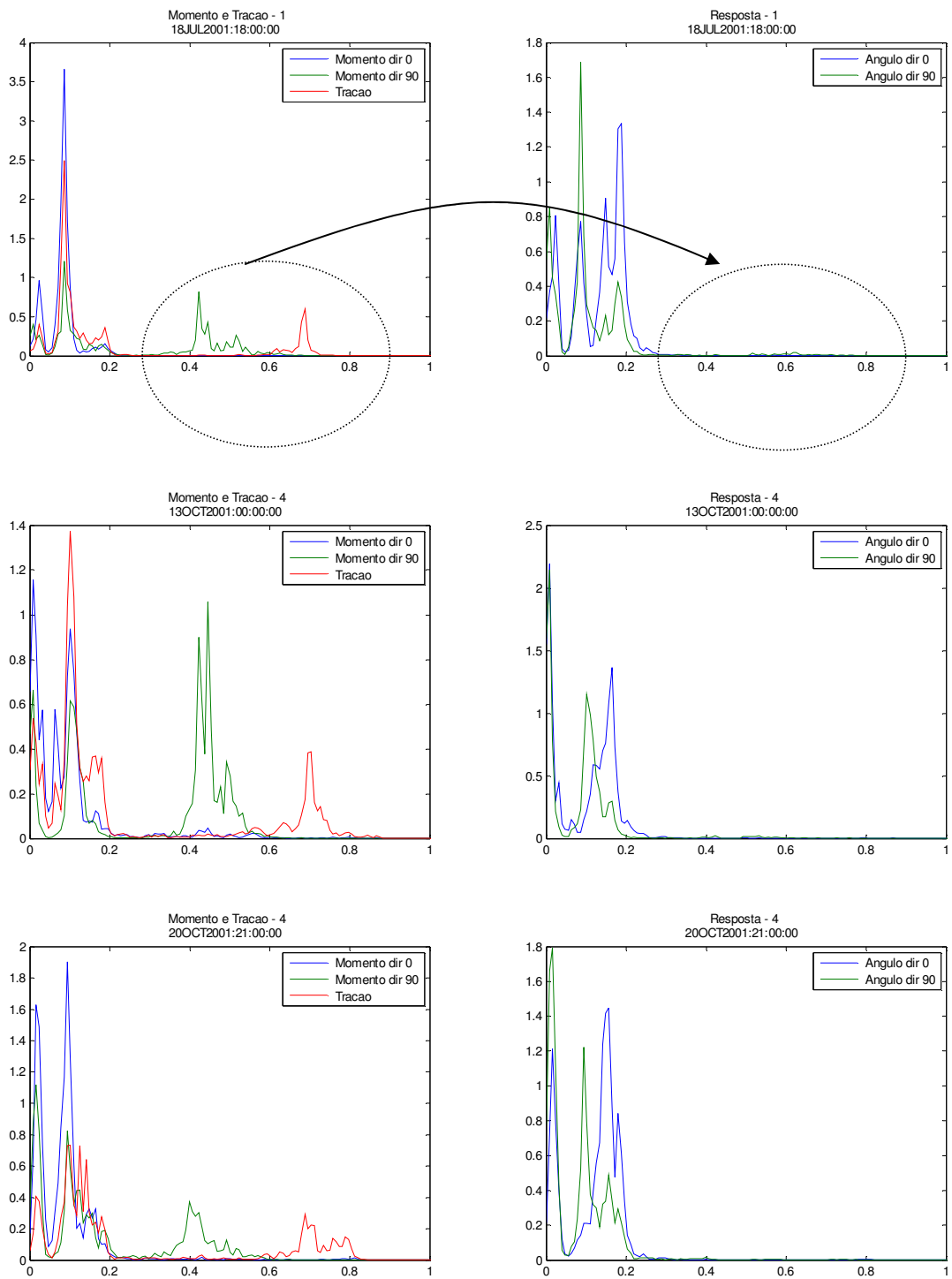
1. Pode-se observar na Figura 4.10 que nem sempre há uma correspondência entre os níveis de energia na faixa de alta frequência, devido ao VIV, entre os esforços e os ângulos medidos. Diferentemente do que ocorre na faixa de frequência de atuação da onda, em que se pode notar uma equivalência de amplitudes de energia entre esforço e ângulo.

O instrumento de medição de ângulo empregado na monitoração é um inclinômetro de pêndulo. Através de ensaios de laboratório foi constatado que o inclinômetro de pêndulo não é adequado para medições de ângulo em estruturas submetidas a carregamento dinâmico. Sendo assim, esta seria uma possível explicação para a

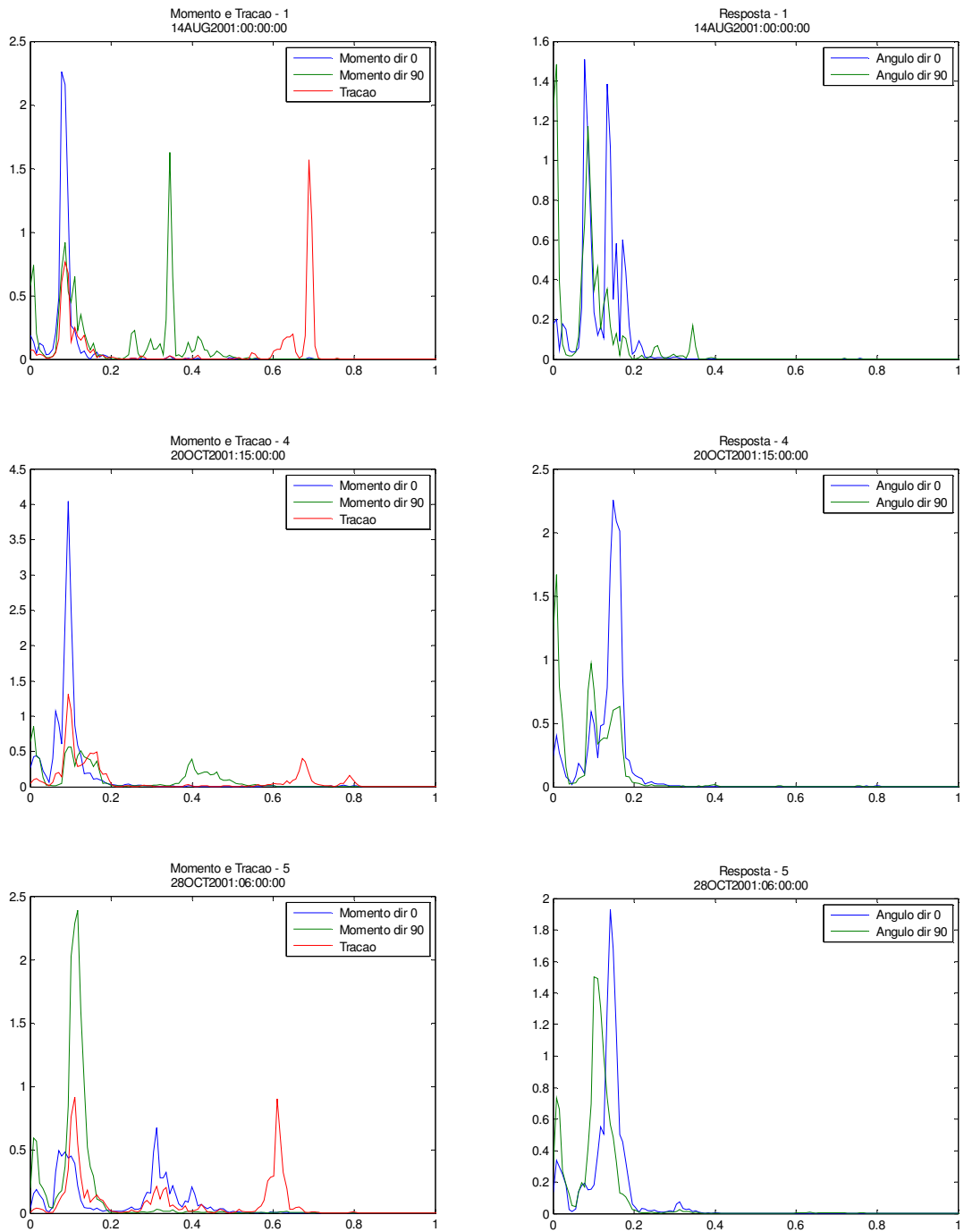
ausência de energia na faixa de frequência de VIV em alguns espectros de ângulo, quando existe a vibração. Embora o comportamento da estrutura seja dinâmico em presença das ondas, estas ocorrem em mais baixas frequências e, por isso, o inclinômetro tem melhor desempenho.

2. Na maioria dos espectros observados, há uma predominância de ocorrência de VIV na direção 90°, para o momento. Esse fato pode sugerir a predominância da correnteza na direção da catenária do *riser* neste período. Os espectros da Figura 4.11 exemplificam essa afirmativa.
3. Outro aspecto observado foi a possível ocorrência de outros fenômenos que causam deslocamentos e tensões no *riser*, além da ação da onda e correnteza (VIV). Nos espectros de tração da Figura 4.12, observa-se que há energia na faixa característica de VIV, sem que haja energia correspondente no espectro de momento. Sempre que o *riser* vibra transversalmente à ação da corrente, é natural a existência de esforço axial e de momento concomitantemente. Uma hipótese levantada para tais vibrações na direção axial é a de serem decorrentes da interação do *riser* com o solo, região do TDP (*Touch Down Point*). Essa interação causa impactos que se propagam na direção axial ao longo do tubo.
4. Nos casos em que os esforços de momento e axial apresentam energia na faixa de frequência de VIV, a frequência de vibração do último é próxima à do dobro do primeiro, como pode ser observado nos espectros da Figura 4.13.

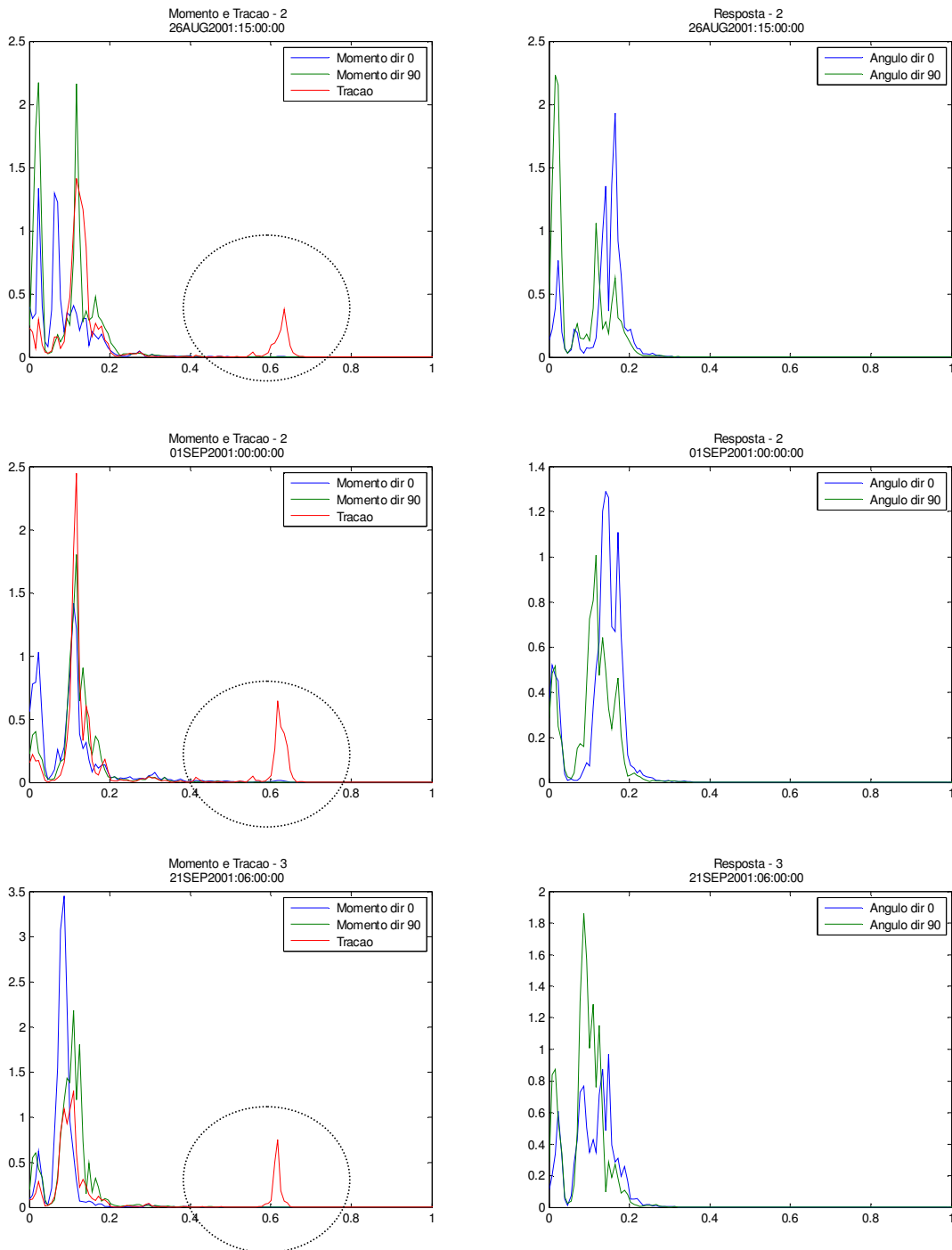
No capítulo 6, em função das análises realizadas e das observações feitas acima, sob a ótica da análise espectral, se fará uma filtragem do tipo passa-baixa, para separar a frequência de VIV da frequência característica de onda. Um outro aspecto que será também abordado no capítulo 6 é quanto ao período de amostragem. Este parâmetro corresponde ao intervalo de tempo decorrido entre duas amostras, e a sua escolha tem bastante impacto na qualidade dos modelos identificados (DALLAGNOL, 2005). Maiores explicações serão fornecidas no capítulo citado.



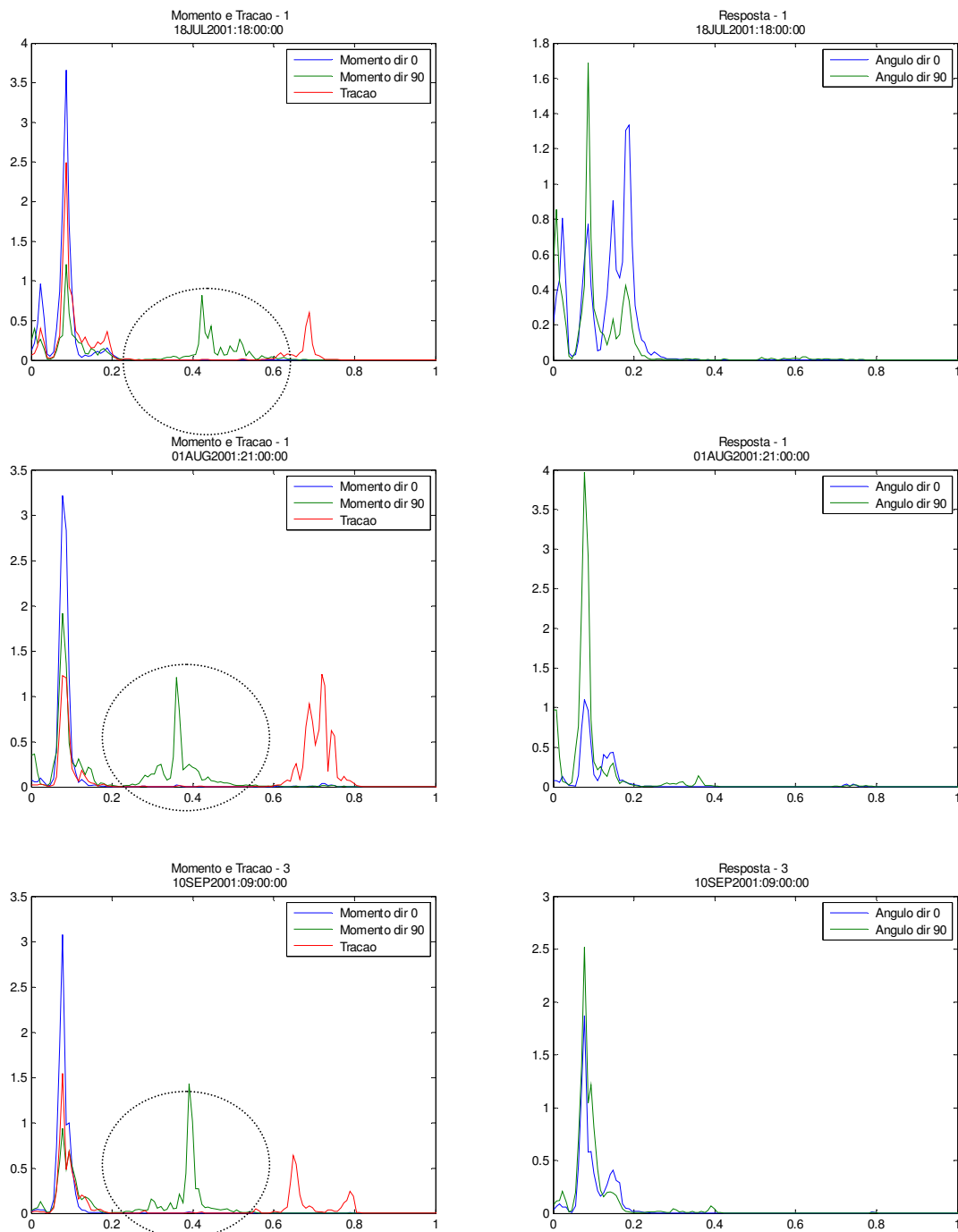
**Figura 4.10 – Exemplos de espectros de esforços e ângulos nos quais pode-se observar que há energia na faixa de frequência de VIV para os esforços, e não há para o ângulo.**



**Figura 4.11 – Exemplos de espectros de esforços e ângulos nos quais pode-se observar que a frequência de excitação da tração é próxima do dobro da frequência do momento**



**Figura 4.12 – Exemplos de espectros de esforços e ângulos, nos quais pode-se observar excitação em alta frequência da tração, na faixa de VIV, sem excitação para o momento.**



**Figura 4.13 – Exemplos de espectros de esforços e ângulos nos quais pode-se observar a predominância em alta frequência do esforço momento na direção de 90° .**

#### 4.5.2.1 Análise de cluster

Além da análise visual dos espectros, alguns parâmetros estatísticos foram obtidos para avaliar a homogeneidade das séries temporais. Com estes parâmetros, pode-se avaliar quão diferentes são umas séries das outras em termos de um comportamento médio através de uma análise de cluster. Os parâmetros do espectro calculados foram: altura significativa, período de pico e largura de banda. É importante ressaltar que estes parâmetros são calculados a partir dos espectros de esforços e guardam relação direta com os espectros de energia do estado de mar, mas seus valores, a exemplo da altura significativa, não representam a altura da onda. Estes parâmetros foram descritos no capítulo 3 sobre fundamentos teóricos, e são obtidos das seguintes formas:

- Altura significativa ( $H_s$ )

$$H_{\frac{1}{3}} = 4\sqrt{m_0} \quad [x1] \quad (\text{Eq. 4.11})$$

- Período de pico

$$T_p = \frac{1}{2\pi} \sqrt{\frac{m_4}{m_2}} \quad [x2] \quad (\text{Eq. 4.12})$$

- Largura de banda

$$\varepsilon = \sqrt{1 - \frac{m_2^2}{m_0 m_4}} \quad [x3] \quad (\text{Eq. 4.13})$$

onde :

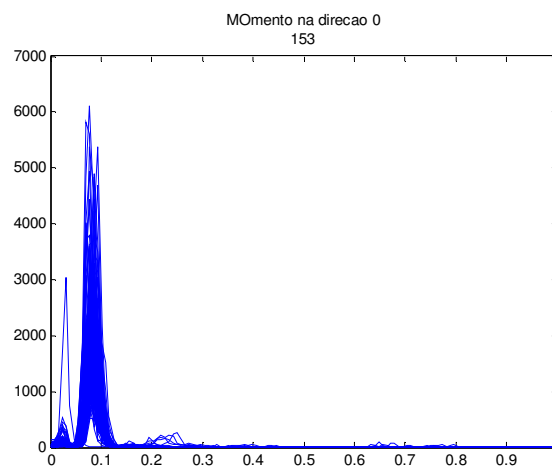
$m_0$ : momento de ordem zero

$m_2$ : momento de segunda ordem

$m_4$ : momento de quarta ordem do espectro

Os momentos  $m_0$ ,  $m_2$  e  $m_4$  estão disponíveis no banco de dados de monitoração da PXVIII, para cada faixa de frequência em que foi dividido o espectro. A divisão do espectro em faixas de frequências menores permite a avaliação dos fenômenos de forma mais isolada, principalmente aqueles referentes à atuação da onda e do VIV. Logo, será feito o cálculo dos parâmetros estatísticos propostos para cada faixa.

Na análise de agrupamento realizada nos dados, utilizando-se os parâmetros  $H_s$ ,  $T_p$  e  $\epsilon$ , revelou-se uma relevância muito grande da altura significativa ( $H_s$ ), baixa importância para  $T_p$ , e nenhuma capacidade discriminatória para a largura de banda. Dois grandes grupos foram identificados na faixa de frequência de onda (de 0 a 0,25Hz). Na Figura 4.14 encontra-se o grupo com maior energia, e, na Figura 4.15, com menor energia.



**Figura 4.14 – Grupo de espectros referentes ao momento na direção 0° com maior energia**



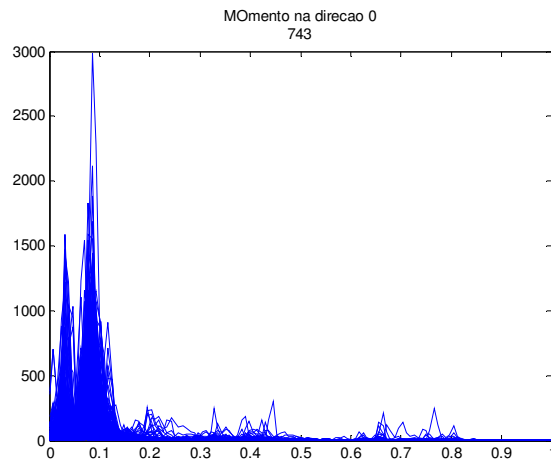


Figura 4.15 – Grupo de espectros referentes ao momento na direção 0° com menor energia

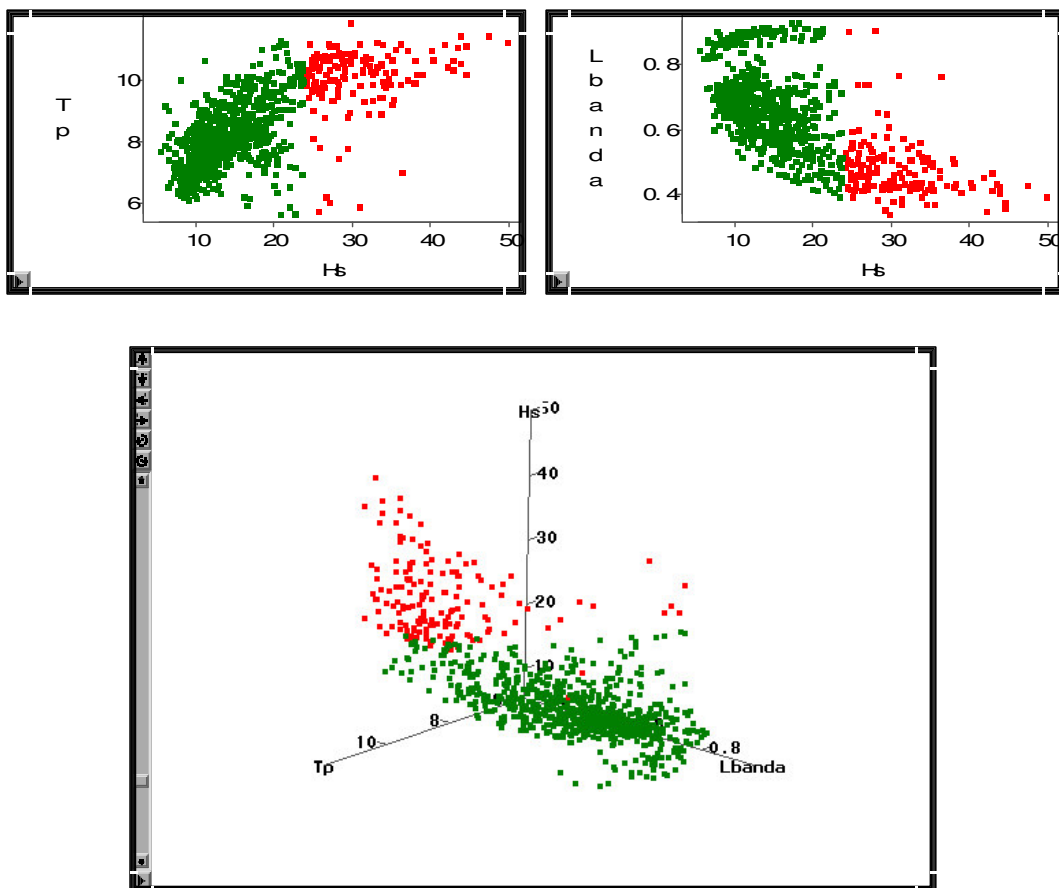


Figura 4.16 – Agrupamento dos dados em dois grupos

Como o período (julho a novembro/2001) estudado corresponde a duas estações do ano, inverno e primavera, parece razoável o agrupamento em dois níveis de energia (ver Figura 4.16) . É importante observar que a análise foi realizada em cima dos esforços, ou seja, das respostas do *riser* e não das ações de onda, vento e correnteza. Estes dados não estavam disponíveis na época. Com eles, poder-se-ia fazer uma avaliação mais quantitativa e com mais qualidade em uma análise como esta.

Logo, para a janela de observação e com as informações referentes aos esforços, pode-se ter uma visão qualitativa através da resposta, o momento, por exemplo. Pode-se ver picos distintos de energia, mares unimodais e bimodais (ver Figura 4.14 e Figura 4.15); entretanto, essas conclusões são óbvias para o período, e considerando-se o mar da bacia de Campos.

Para o processo de clusterização, adotou-se o método *WARD* disponível no software SAS 1999-2001. O método *Ward* forma grupos de dados buscando minimizar a soma das diferenças entre os elementos de cada grupo e o valor médio do grupo, minimizando o desvio padrão entre os dados de cada grupo formado (COELHO *et al.*, 2004).

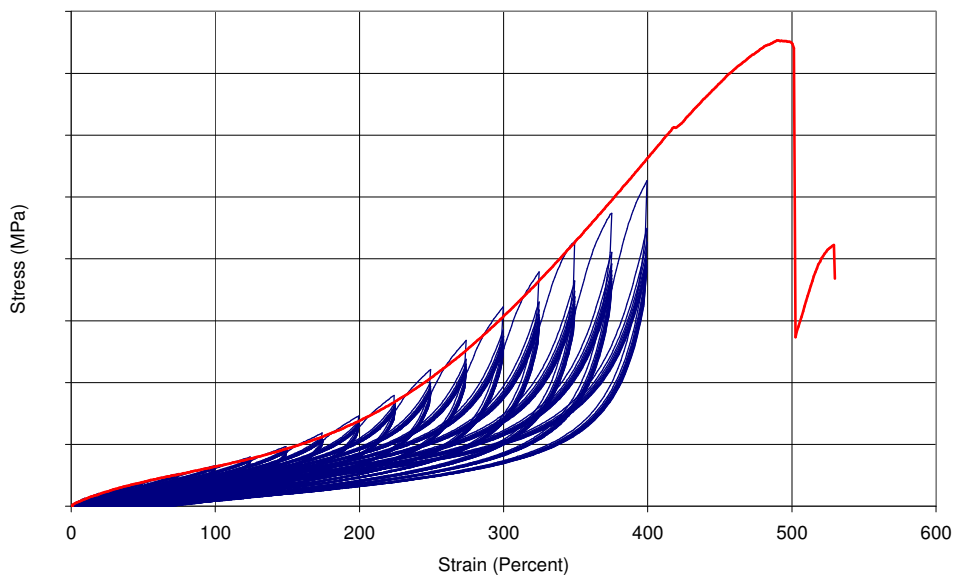
#### **4.6 Dados de ensaios de laboratório**

Paralelamente a esta pesquisa, está sendo realizada uma série de ensaios experimentais e desenvolvimento de modelos baseados em métodos matemáticos convencionais para simulação do comportamento da *flexjoint* em condições de carregamento reais, para avaliação da rigidez dinâmica e da influência dos danos causados ao elastômero pela exposição ao ozônio. Resumidamente, foram realizados até agora as seguintes etapas no trabalho da Petrobras/MERL (HARRIS *et al.*, 2006):

- Ensaios realizados em elementos de elastômero para avaliação do módulo de elasticidade através da relação tensão/deformação.
- Com os dados obtidos na etapa anterior e dados de geometria fornecidos pelo fabricante da *flexjoint*, foi possível construir um modelo de elementos finitos (MEF) da *flexjoint*.

- Com o modelo MEF, foram realizadas análises para avaliação das deformações devidas à ação da pressão interna e tração axial, e análises para avaliação da ação combinada da pressão interna, tração axial e ângulo de rotação variando até 5°.

No relatório MERL estão descritos detalhadamente todo o estudo e resultados. Resultados importantes foram obtidos e utilizados nesse estudo. O ensaio com o elastômero mostrou claramente a natureza histerética do material elastomérico da *flexjoint*, conforme mostrado na Figura 4.17. É interessante observar que os ciclos de histerese aparecem mesmo sob cargas estáticas cíclicas (quasi-estáticas), indicando que a curva tensão-deformação é insensível à velocidade de deformação. Entretanto, alguns autores (J.YI *et al.*, 2006) (BERGSTROM *et al.*, 1998) (BANKS *et al.*, 1999) afirmam que o elastomero é dependente da taxa de variação da deformação para grandes deformações. Então para os limites ensaiados, conclui-se que não foram atingidos como grandes deformações, conforme dito pelos autores citados.



**Figura 4.17 – Laços de histerese do elastômero em ensaio quasi-estático em laboratório**

Estes dados servirão de base para o estudo dos tipos de redes neurais que modelam histerese, apresentados no capítulo 5, a seguir.

## **CAPÍTULO V – SELEÇÃO DE METODOLOGIA PARA MODELAGEM DA HISTERESE COM REDES NEURAIS**

### **5.1 - Introdução**

No capítulo 3 foi visto que a Identificação de Sistemas é o processo empregado para se estimar os parâmetros de um modelo a partir de informações observadas do sistema. Uma das etapas principais do processo de identificação é a escolha da representação matemática, estrutura e estimação de parâmetros do modelo. Na abordagem utilizando Redes Neurais como representação matemática, dois dos fatores fundamentais nessa formulação são a definição da arquitetura da rede (que corresponde à etapa de escolha da estrutura) e do algoritmo de treinamento utilizado para ajustar os pesos da rede (que corresponde à etapa de estimação de parâmetros). A escolha da arquitetura e do algoritmo de treinamento dependerá da aplicação e dos objetivos a serem alcançados. Em função da disponibilidade de dados de laboratório e dos motivos expostos acima, este capítulo

propõe-se a estudar as principais características do sistema proposto com o objetivo de avaliar as potencialidades da rede neural de modelar o sistema considerando o fato de que o sistema em estudo é dinâmico, não linear e que exhibe histerese.

## 5.2 - Modelagem da histerese

A modelagem do comportamento histerético em componentes estruturais é um pré-requisito para o entendimento global da resposta da estrutura sujeita a cargas dinâmicas.

Vários modelos analíticos têm sido empregados para simulação da histerese, a exemplo de (BANKS *et al.*, 1999) que apresenta modelos dinâmicos para Histerese não linear em elastômero. Entretanto, a histerese em certos sistemas é bastante complexa para ser descrita por tais modelos. A falta de conhecimento a respeito do comportamento complexo do elastômero, componente principal da *flexjoint*, conduziu este trabalho para a modelagem de sistemas através de uma das técnicas de identificação, redes neurais, conforme discutido no capítulo 3, Fundamentos Teóricos.

Como foi dito na seção 3.4 o capítulo 3, a histerese é um fenômeno que apresenta memória. Duas das abordagens de redes neurais capazes de modelar memória são o emprego de atraso de tempo (TDNN) ou realimentação (Redes recorrentes) (HAYKIN, *op. cit*), também denominadas redes dinâmicas.

Neste capítulo, será feito um estudo da aplicação dessas abordagens para modelagem de sistemas que apresentam histerese do material. Os dados usados para este estudo são os dados de um ensaio de laboratório (ver seção 4.6) de tensão-deformação de um elastômero que compõe a *flexjoint* (ver capítulo 2).

Com a finalidade de ilustrar a influência da modelagem ou não do componente temporal, incluiu-se a rede neural convencional MLP (BRAGA *et al.*, 2000).

Para efetuar os estudos propostos, faz-se necessário definir uma topologia adequada, especificar o número de camadas e número de neurônios, selecionar o algoritmo de aprendizado e selecionar os parâmetros de aprendizado. Nas seções a seguir, discute-se o estabelecimento desses parâmetros para cada abordagem de rede neural sugerida acima.

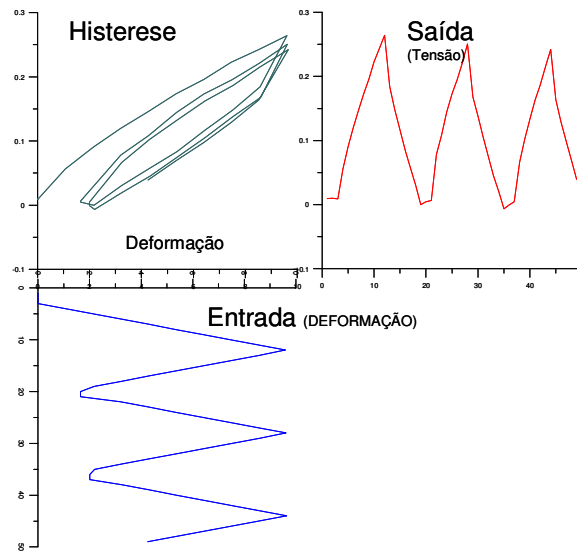
A principal característica da histerese é a sua dependência com o passado. Fundamentalmente o que uma rede neural precisa para modelar esse tipo de fenômeno é alguma forma de memória. O conceito de memória é introduzido através de atrasos de tempo e/ou através de realimentação (HAYKIN, *op. cit.*). Na literatura, dois tipos de redes são mais comumente empregados para incorporação da memória, as chamadas redes dinâmicas (HAYKIN, *op. cit.*):

- as redes neurais com atrasos de tempo, *p.ex.* TDNN (*Time Delay Neural Network*) e MLP (*multi-layer perceptron*) e
- as redes recorrentes, *p.ex.* Elman (ELMAN, 1990), Jordan e NARX.

### **5.3 – Preparação dos dados**

Os dados de laboratório apresentam curvas de histerese bem comportadas uma vez que foram produzidos a partir da aplicação de carregamentos quasi-estático. Diferentemente dessa situação, o sistema plataforma-*riser-flexjoint* está submetido a carregamentos aleatórios dinâmicos de onda, vento e correnteza.

O gráfico da Figura 5.1 mostra a composição da entrada (deformação) e saída (tensão) para formação dos laços de histerese. Os dados da tensão e deformação foram coletados em um intervalo de amostragem de 1s.



**Figura 5.1- Formação dos laços de histerese**

O conjunto de dados foi particionado em três subconjuntos: um para treinamento com 60% das amostras, um para teste com 20% e o terceiro para validação com os 20% restante. O conjunto de treinamento é utilizado para treinar a rede e conseqüentemente ajustar os seus pesos. O conjunto de validação é utilizado para implementação do *early stopping*. E o conjunto de teste para verificar a capacidade de generalização da rede. Todos os dados de entrada e saída (treinamento, validação e teste) foram normalizadas para o intervalo  $[-1; 1]$ . Este intervalo foi escolhido, dado o emprego da tangente hiperbólica como função de ativação, por ser o intervalo em que a função tem seus valores mais significativos, contribuindo melhor para o treinamento. Os dados de treinamento, teste e validação das redes são os mesmos para todos os modelos.

O conjunto de dados é formado por 180 ciclos de histerese, o que corresponde a 52194 registros e está armazenado em seqüência, ou seja, para um dado ponto arbitrário do estado tensão-deformação, os pontos que o precedem são estados que aconteceram em um tempo anterior.

Para realizar as análises, os dados foram organizados em 4 casos. Cada caso é composto pelos subconjuntos: treinamento, validação e teste. Os casos são definidos com

um número diferente de ciclos, ordenados de acordo com os valores de tensão-deformação como disposto a seguir:

Caso I – Compõe-se dos 4 primeiros ciclos do conjunto total de dados

Treinamento: 2 ciclos

Validação: 1 ciclo

Teste: 1 ciclo

Caso II – Compõe-se dos 10 primeiros ciclos do conjunto total de dados

Treinamento: 6 ciclos

Validação: 2 ciclos

Teste: 2 ciclos

Caso III – Este terceiro conjunto é composto de 15% do conjunto de dados.

Treinamento: 21 ciclos

Validação: 3 ciclos

Teste: 3 ciclos

Caso IV – Este terceiro conjunto é composto de todo o conjunto de dados.

Treinamento: 108 ciclos

Validação: 36 ciclos

Teste: 36 ciclos

A escolha de casos para análise foi arbitrária, mas de certa forma procurando-se mostrar a capacidade dos modelos de responder desde poucos ciclos até todo o conjunto de dados do ensaio de laboratório.

A Figura 5.2 mostra um esquema gráfico da separação dos dados em treinamento, validação e teste dos casos I, II, III e IV citados acima.



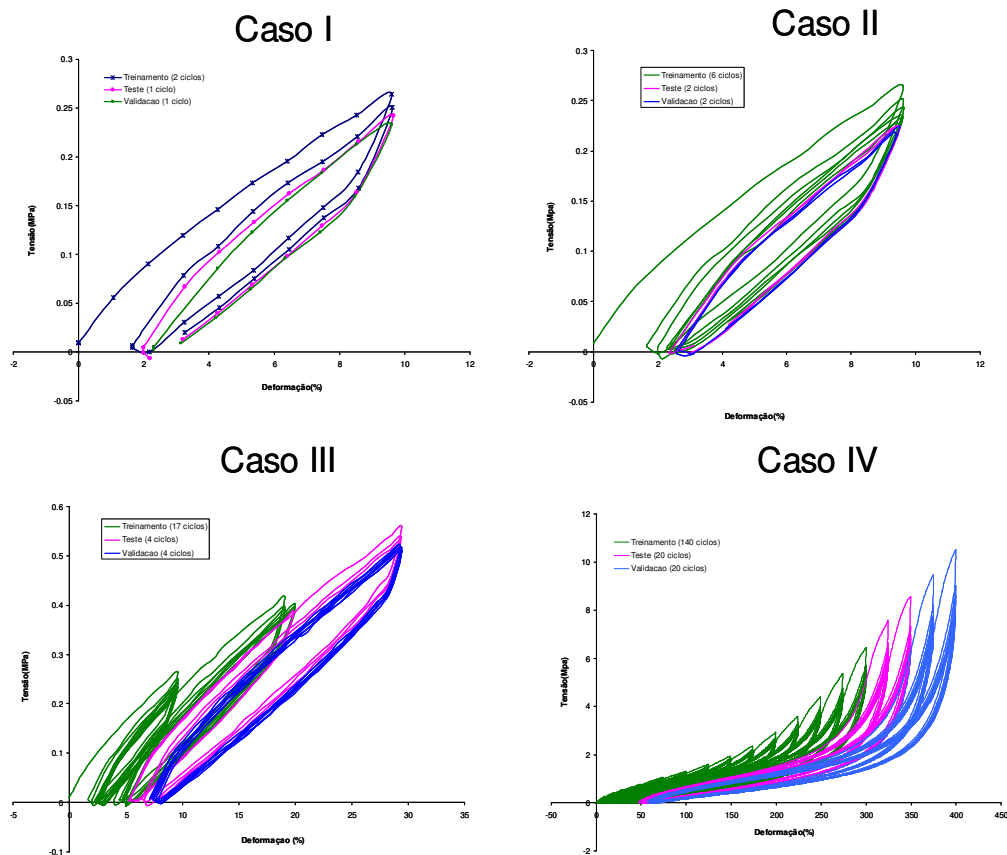


Figura 5.2 - Ciclos de histerese do elastômero para os casos I, II, III e IV.

#### 5.4 - Seleção da Estrutura e algoritmo de aprendizado

Três tipos de arquiteturas de redes neurais foram empregados para avaliar a habilidade da rede de modelar o comportamento histerético do elastômero na flexjoint. Além das redes TDNN e recorrentes do tipo NARX (paralelo e série-paralelo), também foi modelado uma rede *backpropagation* (MLP) convencional para o problema. A retropropagação convencional é um processo estático e não produz bons resultados em processos cuja dependência do tempo é fator primordial. Entretanto, foi utilizada nesta pesquisa apenas para comparar os resultados e avaliar o ganho quando se negligencia o tempo em sistemas e materiais que exibem histerese.

Os algoritmos de treinamento tradicionais baseados no gradiente (HAYKIN, *op. cit*), bastante empregado nas redes convencionais *backpropagation*, podem também ser utilizados para treinar redes dinâmicas. Porém, devido ao potencial das redes dinâmicas de processar informações ordenadas no tempo, o treinamento com esses algoritmos podem ficar significativamente lentos (MATLAB HELP, 2006). Também segundo (MATLAB HELP, 2006), a utilização do algoritmo baseado no método de otimização de Levenberg-Marquard propicia um treinamento mais adequado ao problema: "Este método minimiza a combinação dos erros quadráticos e os pesos, e então determina a combinação correta de tal forma a produzir uma rede que generalize bem".

O treinamento é supervisionado e para pelo critério *early-stopping* descrito na seção 3.2.6.2 ou se o número máximo de 600 épocas (iterações) for atingido. Basicamente, o que se pretende com o critério *early stopping* é evitar o treinamento excessivo (*overtraining*) da rede, o que pode fazer com que ela perca sua capacidade de generalização quando apresentada a novos padrões. A interrupção do processo de treinamento se dá, seguindo esse critério, quando o SSE (*Sum squared Error*) ou o MSE (*Mean Squared Error*) do conjunto de validação começa a crescer.

Um grande número de medidas de desempenho pode ser encontrado na literatura, e cada uma possui suas vantagens e limitações [MAKRIDAKIS *et al.*, 1982]. Neste estudo, a eficácia dos modelos desenvolvidos para as curvas  $\sigma$ - $\varepsilon$  foi medida através da raiz quadrada dos erros médios quadráticos (RMSE-*Root Mean Squared Error*). Quanto maior o erro, maior o erro de previsão do modelo. Sua equação (5.1) é dada por:

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^N (\hat{Y}_t - Y_t)^2}{N}} \quad (5.1) \quad \text{(Eq. 5.1)}$$

onde  $\hat{Y}_t$  é o valor de previsão no instante  $t$ ,  $Y_t$  é a saída desejada no instante  $t$  e  $N$  é o número de amostras para as quais os termos de erro são calculados.

Os modelos foram implementados com duas camadas (a intermediária e a de saída). Por simplicidade, decidiu-se adotar um mesmo número de neurônios na camada intermediária para todos os modelos. Uma estratégia empírica foi usada para escolher este número. Dado um conjunto qualquer de dados, variou-se o número de neurônios de cada modelo, anotando-se o número mínimo de neurônios necessário para a estabilização da raiz do erro médio quadrático (RMSE). Adotou-se o maior dentre os quatro valores por ser suficiente para estabilizar o RMSE dos quatro modelos. Após várias simulações (ver 5.5.2), um número de 10 neurônios na camada escondida teve bons resultados durante o treinamento.

As análises foram realizadas utilizando-se o as *funções* de redes neurais do MATLAB versão R2006 e o treinamento da rede foi realizado usando a opção *trainbr*, com momento e taxa de aprendizado adaptável. A função de ativação escolhida para os neurônios da camada escondida e da saída foi a tangente hiperbólica,  $\tanh(n) = 2/(1+\exp(-2*n))-1$  (tansig no MATLAB).

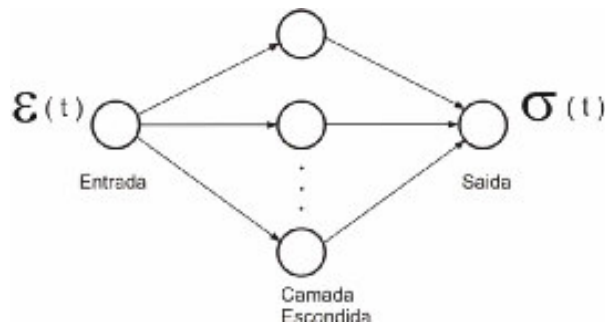
Nas seções a seguir são descritas as particularidades de cada modelo neural proposto neste trabalho e mostrado os resultados das análises realizadas com o conjunto de dados de laboratório.

## **5.5 - Os modelos**

### **5.5.1 - Backpropagation convencional**

O interesse na aplicação da rede MLP na modelagem da histerese surgiu apenas da curiosidade de avaliar o desempenho de uma rede estática, sem a consideração temporal, requerida para este tipo de fenômeno.

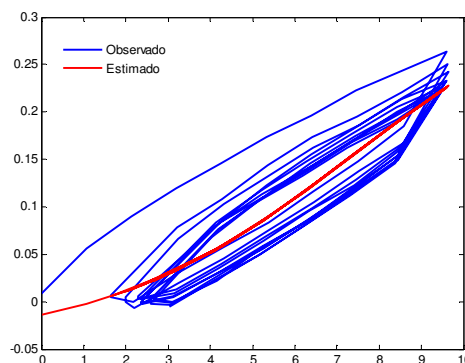
Para as redes *feedforward* a topologia adotada foi a mostrada na Figura 5.3. A entrada da rede consiste de uma única variável, a deformação no tempo  $t$  ( $\varepsilon$ ) e uma saída que é a tensão ( $\sigma$ ) correspondente, no tempo  $t$ .



**Figura 5.3 - Arquitetura da rede MLP**

O desempenho da rede MLP foi observado através de uma amostra do conjunto de dados disponível, caso II (ver 5.3). O treinamento parou após 48 épocas pelo critério de *early stopping* com um RMSE = 0,036. Foram também realizados testes sem o conjunto de validação e não houve *overtraining* para o número de épocas estabelecido, entretanto, o treinamento foi mais longo.

Como era de se esperar, o modelo com o *backpropagation* convencional, sem a introdução das variáveis defasadas no tempo, se comporta como um modelo de regressão não linear, aproximando os pontos que compõem os caminhos de carga e descarga, por uma curva que passa entre os pontos minimizando os erros, conforme pode ser visto na Figura 5.4:

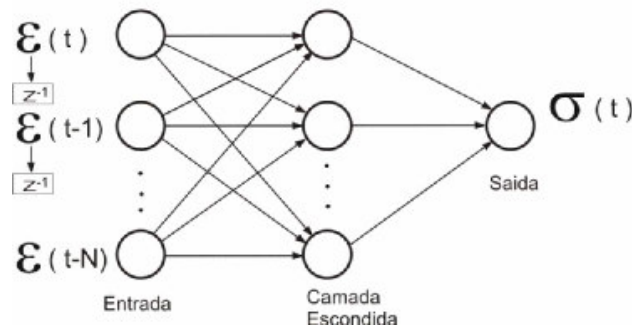


**Figura 5.4 - Curvas tensão-deformação – observados e estimados pelo modelo *backpropagation* convencional**

### 5.5.2 – Rede neural de atrasos de tempo (TDNN)

Na literatura, as redes TDNN (*Time Delay Neural Network*) são especialmente utilizadas na modelagem de séries temporais. Poucos trabalhos foram publicados para modelagem da histerese utilizando-se TDNN. Em [BASHEER *et al.*,2002], é proposto uma solução híbrida, denominada pelo autor de TDNN parcialmente recorrente, para modelagem do comportamento tensão-deformação de geomateriais submetidos a carregamentos cíclicos. Em função de tais carregamentos, os geomateriais exibem histerese.

A camada de entrada da rede TDNN é composta da variável de entrada (deformação no tempo  $t$ ) e por seus atrasos, espaçadas temporalmente em 1, isto é, em  $t-1$ ,  $t-2$ , ...,  $t-N$ . A arquitetura proposta é ilustrada pela Figura 5.5.



**Figura 5.5 - Arquitetura da rede TDNN**

A rede TDNN aproxima uma função  $f$  dada pela equação 5.2 :

$$\sigma(t) = f(\varepsilon(t), \varepsilon(t-1), \dots, \varepsilon(t-N)) \quad \text{(Eq. 5.2)}$$

onde,

$\varepsilon(t)$  : deformação no tempo  $t$

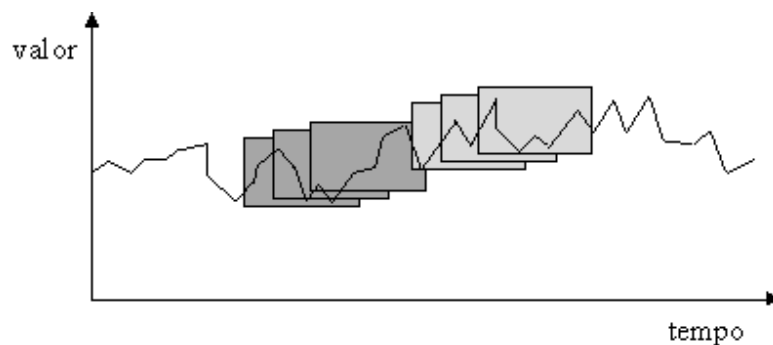
$\varepsilon(t-1)$  : deformação no tempo  $t-1$

$\varepsilon(t-N)$  : deformação no tempo  $t-N$

$\sigma(t)$  : tração no tempo  $t$

$N$  : tamanho da janela de tempo (*sliding window*)

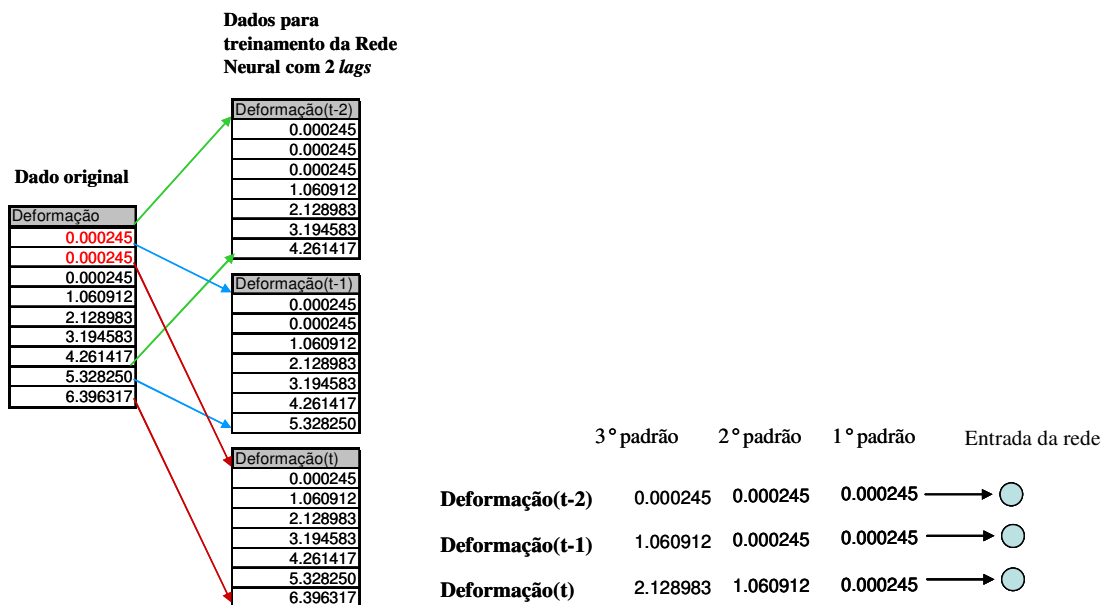
O treinamento de uma TDNN, que atua em dados em seqüência [VOJTKO,2004], é obtido através da alimentação da rede com uma janela de tamanho fixo das deformações para a rede, obtendo uma saída e fazendo o ajuste adequado dos pesos da rede. Este processo é repetido para mais uma janela, posicionada um passo adiante, até que todos os dados disponíveis tenham sido usados. A Figura 5.6, ilustra o processo:



**Figura 5.6 - ilustração das janelas de tempo**

Embora alguns softwares façam automaticamente a defasagem das variáveis, cabe aqui uma observação de como é realizada esta operação através de um exemplo com o próprio conjunto de dados estudado neste capítulo.

Considerando-se dois atrasos (*lags*) da variável de entrada (deformação), a rede vai receber, além desta variável no intervalo de tempo  $t$ , mais duas outras entradas, representando a deformação no tempo  $t-1$  e  $t-2$ . A Figura 5.7 mostra uma amostra dos dados originais e as novas variáveis decorrentes do processo de defasagem que vão alimentar a rede.



**Figura 5.7 – Ilustração do processo de defasagem de 2 atrasos da variável deformação.**

É importante ressaltar que o conjunto total é reduzido do número de atrasos estabelecido. No caso, o conjunto de dados originais continha 10 observações, e passa a ter somente 8. Se o conjunto for pequeno, a adoção de *lags* grandes pode diminuir significativamente a quantidade de informação disponível para o aprendizado.

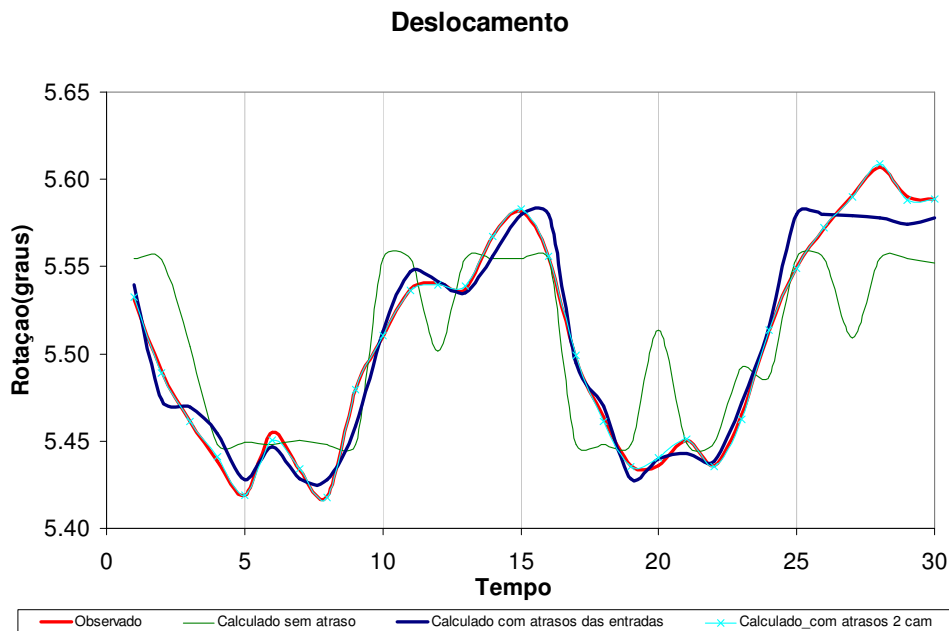
Em geral, a escolha do tamanho da janela influencia o desempenho das redes (HAYKIN, *op. cit*). A metodologia adotada para descobrir o número de atrasos (*lags*) ótimo foi empírica, através de simulações, comparado o desempenho de cada modelo e escolhendo aquele com RMSE menor. A seguir, é descrito como foram realizadas as simulações e comparações.

### 5.5.2.1 - Determinação do tamanho da janela temporal

Foram então realizadas várias simulações para definição de uma janela adequada. O tamanho de janela para essa amostra foi de 5 *lags* (atrasos).

Foi empregada uma rede cuja arquitetura está representada na Figura 3.11, variando D (tamanho do *lag*) de 0 a 5. O número de neurônios foi sendo adaptado em função do número de entradas. Foi também obtida a performance da rede com duas camadas, com e sem atraso.

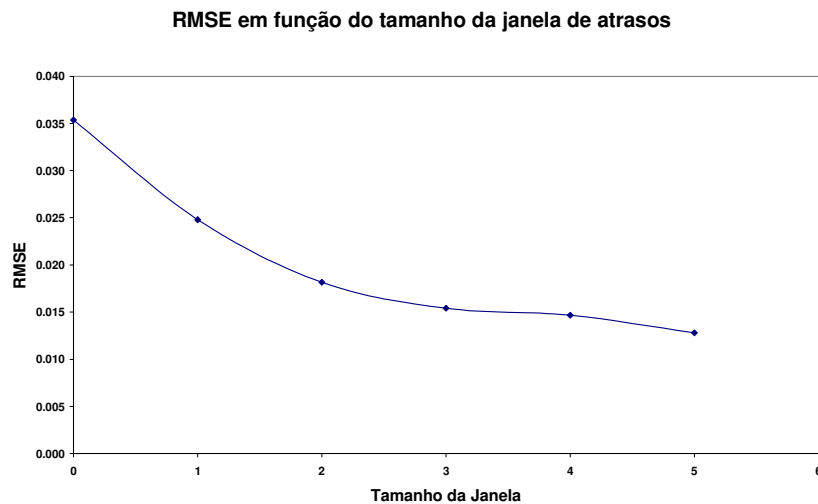
Na Figura 5.8, a linha vermelha representa os dados observados, a linha verde a resposta de uma rede estática MLP, a linha azul escuro a variável de entrada (deformação) com 5 *lags* e a linha azul claro a rede com duas camadas escondidas e os atrasos das variáveis de entrada (também com 5 *lags*). O desempenho da rede com duas camadas com e sem os atrasos foi praticamente o mesmo. Ou seja, com duas camadas, para o número de pontos da série considerado, não há influência dos atrasos. Note-se que a linha azul claro acompanha a linha vermelha (valor observado) durante praticamente todo o tempo.



**Figura 5.8 - Representação do deslocamento no tempo para vários modelos**



Diferentes janelas de tempo foram usadas para melhorar o desempenho do modelo, dado que não há uma forma sistemática de determinar quão longe um sistema deve olhar para trás a fim de prever o valor atual (T. Z. TAN et al, 2005). A Figura 5.9 mostra que o desempenho da rede, em termos de RMSE, melhora significativamente com o aumento do tamanho da janela, até um ponto de estabilização.



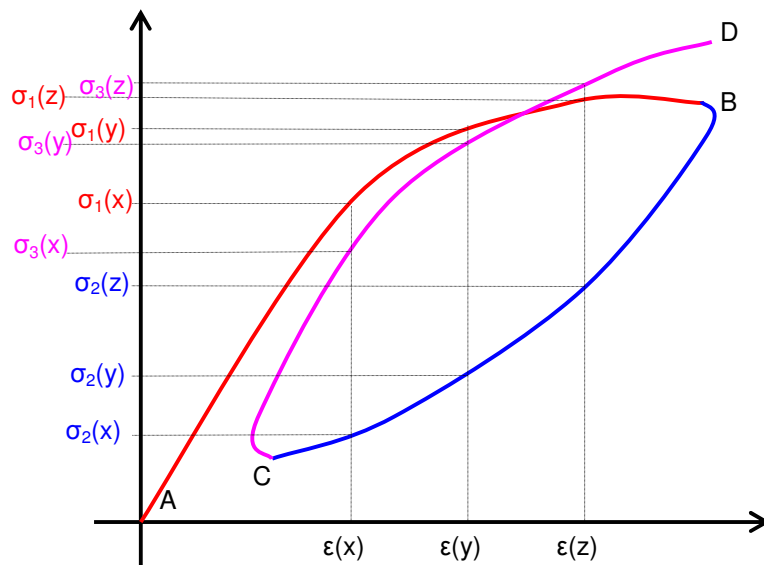
**Figura 5.9: RMSE em função do tamanho da janela de atraso**

Resumindo, a rede TDNN é composta das seguintes características:

- Rede *Feedforward* com realimentação.
- Número de camadas intermediárias = 1.
- Número de neurônios na camada intermediária = 5.
- Taxa e de aprendizagem momento adaptativos.
- Tamanho da janela de tempo da camada de entrada = 5.
- Função de ativação para todas as camadas é tangente hiperbólica.

A rede foi treinada com o caso II (ver 5.3). O treinamento parou pelo critério *early stopping* com 68 épocas e RMSE do treinamento = 0.1163 e o RMSE do teste = 0.2355 .

Sistemas cíclicos, como os que apresentam histerese, não são bem representados apenas pelo passado de suas entradas, mas sim pelo passado do par ordenado entrada-saída (no caso  $\epsilon$ - $\sigma$ ). Seja, na Figura 5.10, (1) o trecho AB (em vermelho), (2) o trecho BC (azul) e (3) o trecho CD (lilás). Nos trechos (1) e (3), o sistema deverá passar pelos mesmos pontos (na mesma ordem), em particular, pelos pontos x, y e z (considere x e y passados de z). É fácil ver que os trechos (1) e (3) são diferentes se for levado em conta os valores de  $\epsilon$  e  $\sigma$  (caso das redes recorrentes). Entretanto, se forem considerados apenas os valores de  $\epsilon$ , o modelo (no caso, TDNN) suporá, erroneamente, que se trata do mesmo ponto (z) e com o mesmo passado (x e y).



**Figura 5.10 - Representação da carga e descarga em um processo de histerese**

### 5.5.3 – Redes Recorrentes NARX

Como foi visto na seção anterior, somente as entradas são defasadas nas redes TDNN. Entretanto, para modelagem da histerese é necessário usar o par (tensão e da

deformação) que compõe o estado do carregamento passado como parâmetro de entrada do modelo. Sendo assim, a rede não linear recorrente auto-regressiva com entradas exógenas, satisfazem esse requisito e será aqui investigada a sua aplicabilidade.

As redes NARX (*Nonlinear Autoregressive with Exogenous Inputs*) são redes não lineares autoregressivas, *i.e.*, a variável de saída depende de seu passado e, por isso, também deve ser usada como entrada do sistema. Além disso, variáveis externas (exógenas) também fazem parte dos dados de entrada do modelo. Há dois tipos de arquitetura: paralelo e série-paralelo. Alguns autores denominam NARX como série-paralelo e NOE (Non-linear Output Error model) para o modelo paralelo.

Para determinar o número de atrasos de tempo e o número de neurônios da camada escondida foram adotados os mesmos procedimentos empregados para as redes TDNN. As análises concluíram o mesmo tamanho de memória, 5 atrasos, e o mesmo número de neurônios, 10.

Como o capítulo 3 descreveu em detalhes essas arquiteturas, a seguir serão analisadas cada uma dessas configurações e suas particularidades relativas ao problema, bem como suas implementações serão discutidas.

### 5.5.3.1 Arquitetura NARX Paralelo

As redes NARX tipo **paralelo** são aquelas cujas saídas do modelo são realimentadas como variáveis de entrada defasadas no tempo, juntamente com as variáveis de entrada externas. Seja  $\hat{\sigma}(t)$  a aproximação de  $\sigma(t)$ , a rede NARX paralelo é descrita pela equação 5.3:

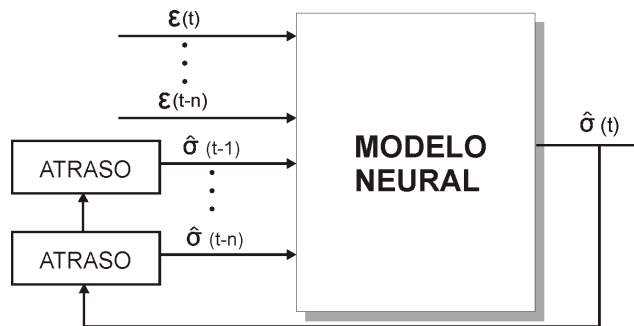
$$\hat{\sigma}(t) = f(\varepsilon(t), [\varepsilon(t-1), \hat{\sigma}(t-1)], \dots, [\varepsilon(t-N), \hat{\sigma}(t-N)]) \quad (\text{Eq. 5.3})$$

onde,

$\varepsilon(t)$  : deformação observada no tempo t

- $\varepsilon(t-1)$  : deformação observada no tempo t-1
- $\varepsilon(t-N)$  : deformação observada no tempo t-N
- $\hat{\sigma}(t)$  : tensão estimada no tempo t
- $\hat{\sigma}(t-1)$  : tensão estimada no tempo t-1
- $\hat{\sigma}(t-N)$  : tensão estimada no tempo t-N
- N : tamanho da janela de tempo (*sliding window*)

A Figura 5.11 ilustra uma rede neural tipo paralelo:



**Figura 5.11- Modelo neural paralelo**

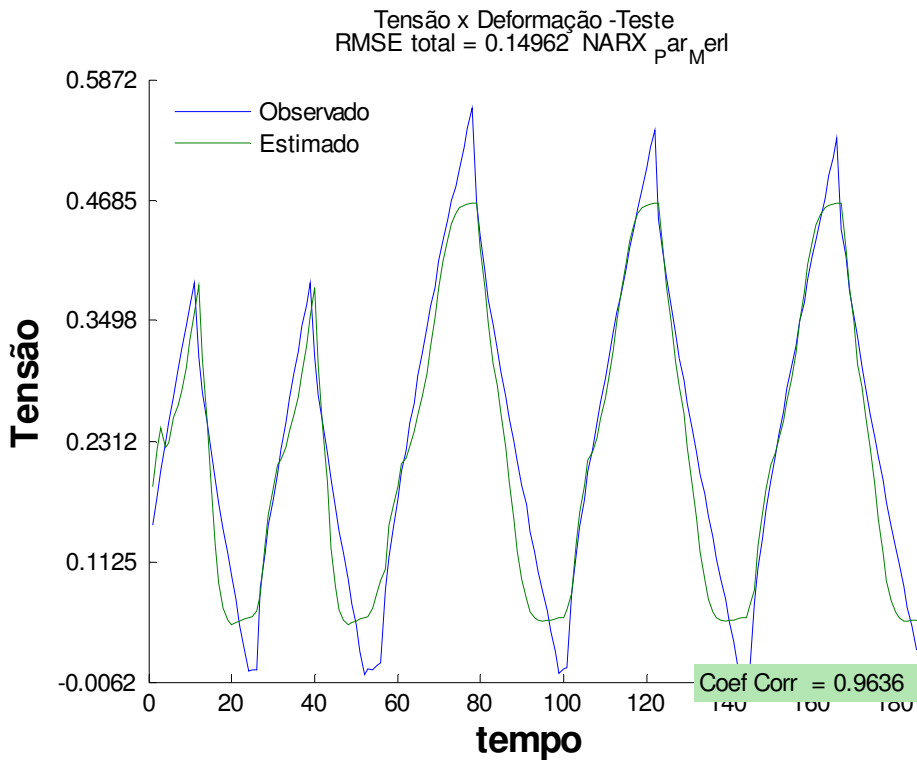
Como nos modelos anteriores, as análises foram conduzidas em função da separação do conjunto de dados em quatro partes, constituindo os casos I, II, III e IV. O desempenho dos modelos é apresentado na Tabela 5.1.

Um fator preocupante para os modelos paralelos é o tempo de treinamento da rede. Geralmente o tempo para este modelo é muito superior àqueles obtidos com o treinamento da rede série-paralelo (MATLBA HELP, 2006). Os casos I, II e III têm poucos dados e os tempos de processamento não foram significativos. Contudo, foram necessários em torno de 6 horas para o treinamento do caso IV.

**Tabela 5.1 - Desempenho dos modelos paralelo**

Modelo	Casos	RMSE			
		Treinamento		Teste	
		Com Validacao	Sem Validação	Com Validacao	Sem Validacao
Paralelo	I	0.1763	0.0191	0.1384	0.2253
	II	0.0724	0.0262	0.2180	0.3899
	III	0.0850	0.0684	0.1496	0.3165
	IV	0.0865		0.1154	

A Figura 5.12 ilustra a resposta do modelo versus o observado quando em teste. A linha azul representa a tensão real atuante no elastômero, e a linha verde é a tensão estimada pelo modelo NARX paralelo. Embora haja discrepâncias da resposta real para a estimada, o comportamento geral do modelo é muito similar ao comportamento do sistema real. Este exemplo refere-se ao caso III.



**Figura 5.12 - Resposta do modelo versus valor real do sistema**

### 5.5.3.2 - Arquitetura NARX Série-Paralelo

Em contraste às redes paralelo, as redes **série-paralelo** usam a história das saídas, ou seja, os valores observados reais, para predizer os valores futuros, ao invés dos valores estimados pelo próprio modelo, além claro das variáveis de entrada externas.

A rede recorrente aproxima uma função  $f$  dada pela equação 5.4:

$$\hat{\sigma}(t) = f\{\varepsilon(t), [\varepsilon(t-1), \dots, \varepsilon(t-N), \sigma(t-1), \dots, \sigma(t-N)]\} \quad (\text{Eq. 5.4})$$

onde,

$\varepsilon(t)$  : deformação observada no tempo t

$\varepsilon(t-1)$  : deformação observada no tempo t-1

$\varepsilon(t-N)$  : deformação observada no tempo t-N

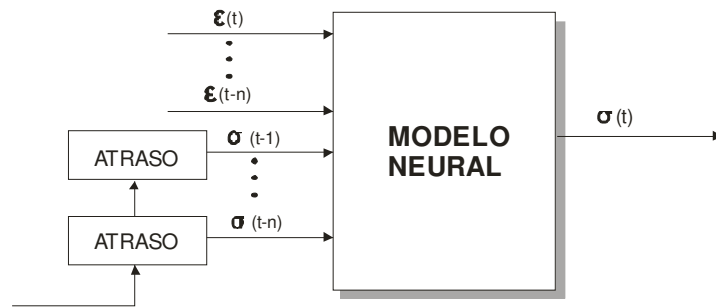
$\hat{\sigma}(t)$  : tensão estimada no tempo t

$\sigma(t-1)$  : tensão observada no tempo t-1

$\sigma(t-N)$  : tensão observada no tempo t-N

N : tamanho da janela de tempo (*sliding window*)

A Figura 5.13 ilustra uma rede neural tipo série-paralelo:



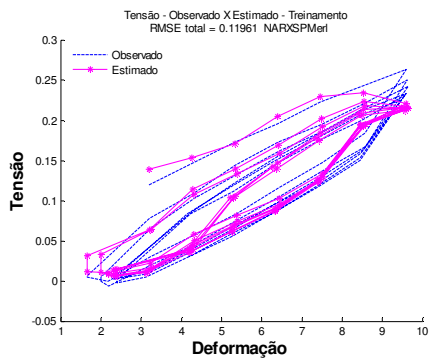
**Figura 5.13 - Modelo neural série-paralelo**

Em função dos resultados mostrados na Tabela 5.2, pode-se verificar que, como regra geral, o erro diminui com o aumento do número de ciclos. O menor erro (RMSE Tr = 0.0081 e 0.0060) acontece no caso IV (todo o conjunto de dados disponível). Em geral, melhores resultados são obtidos quanto maior o número de exemplos fornecidos à rede na fase de treinamento.

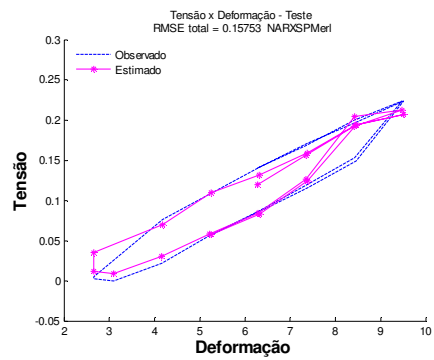
**Tabela 5.2 - Desempenho dos modelos série-paralelo**

Modelo	Casos	RMSE			
		Treinamento		Teste	
		Com Validacao	Sem Validação	Com Validacao	Sem Validacao
Série-Paralelo	I	0.0890	0.0158	0.1263	0.2456
	II	0.0667	0.0136	0.1575	0.2523
	III	0.0878	0.0419	0.1221	0.2261
	IV	0.0081	0.0060	0.0098	0.0087

Outro aspecto importante observado é o aumento do erro nas análises sem a parada antecipada (*early stopping*), caracterizando o *overtraining*. Como exemplo, no caso II, o erro de treinamento com parada antecipada foi de RMSE = 0.0667 enquanto que o erro do treinamento sem a parada antecipada foi de RMSE = 0.0136, muito menor. Já o erro de teste é bem menor quando existe o controle de *overtraining* (RMSE = 0,1221) comparado com o erro do teste sem parada antecipada (RMSE = 0,2261).

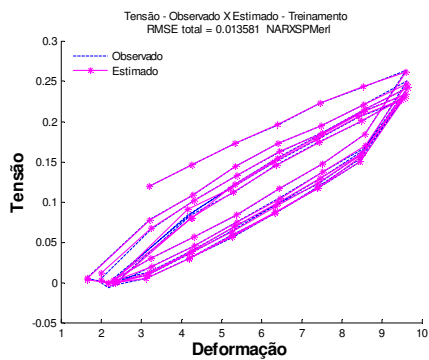


a

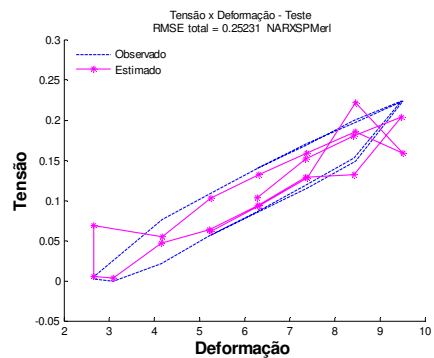


b

**Figura 5.14 – Resultados das análises com o caso II – treinamento com parada antecipada (a) Treinamento (b) Teste**



a



b

**Figura 5.15 – Resultados das análises com o caso II – treinamento sem parada antecipada – Overtraining. (a) Treinamento (b) Teste**

Como pode ser observado na Figura 5.14.b, que representa o teste, a rede generaliza melhor que a rede da Figura 5.15.b, onde ocorre o claramente o *overtraining*. As Figuras 5.14.a e 5.15.a representam o treinamento.



## 5.6 - Comparação dos modelos propostos

Neste problema, o passado só fica bem representado quando se usa o par tensão deformação. As redes TDNN, por definição, não incluem a tensão (saída) como entrada. Sendo assim, as redes TDNN não são um bom modelo para representar a memória nos sistemas histeréticos, como ficou demonstrado no item 5.5.2.

Considerando-se então as redes recorrentes, a diferença entre rede série-paralelo e paralelo pode ser analisada sob vários aspectos.

Um primeiro aspecto diz respeito à modelagem de redes. As redes recorrentes série-paralelo como já foi explicado no capítulo 3, apresentam melhor desempenho e estabilidade uma vez que elas usam os dados de saída observados como entrada, não introduzindo o erro dos dados de saída estimados pela rede neural. Diferentemente, as redes recorrentes paralelo, ao introduzir esses erros, possuem um desempenho inferior e são propensas à instabilidade. As Tabelas 5.1 e 5.2 ilustram a diferença de desempenho. Uma discussão mais detalhada sobre a convergência do modelo paralelo pode ser encontrada em (NARENDA&PARTHASARATHY, 1990) e um exemplo em (CALOBA, 2002). Por outro lado, o modelo paralelo traz a vantagem de ser capaz de captar a dinâmica do sistema em vez de simplesmente descobrir as regularidades ou periodicidades dominantes na série (VONZUBEN, 1996).

O segundo aspecto diz respeito aos custos de aquisição dos dados de saída ao empregar o modelo série-paralelo em produção. Para as redes em produção é fundamental a escolha entre as redes paralelo e série paralelo porque o uso de rede série-paralelo exige que os dados de saída sejam monitorados para servir de entrada. Custo esse que não existe quando é empregada a rede recorrente paralelo. Para efeito de treinamento, em ambos os casos será necessário a disponibilidade dos dados de saída, uma vez que ambos os modelos precisam desse dado para calcular o erro.

Uma outra forma de comparação é através do tempo de treinamento. As redes paralelo demandam muito mais tempo do que as redes série-paralelo. Por exemplo, para treinar o modelo série-paralelo com o caso IV levou-se em média 10 minutos, enquanto que

para o modelo paralelo foram gastos 3600 minutos (ambos os casos rodados sob as mesmas configurações de hardware e software).

## CAPÍTULO VI - ANÁLISES E RESULTADOS

### 6.1 - Introdução

Embora os dados de laboratório não tenham efeito dinâmico expressivo pelo fato de terem sido obtidos através de ensaios com carregamento cíclico quasi-estático, as estratégias empregadas para representar a histerese exibida pelo material podem ser as mesmas para representar a dinâmica do sistema devido à natureza do carregamento. Em ambos os casos as redes neurais mais adequadas são aqueles que representam memória (ou o tempo) através da inclusão de atrasos de tempo em sua estrutura. Logo, a metodologia empregada no capítulo anterior, cujo foco era a modelagem da histerese, serve para aplicação aos dados obtidos do sistema real em operação, em que a dinâmica e histerese estão presentes.

### 6.2 - O modelo de identificação

Após a modelagem da histerese, as características da resposta dinâmica do sistema podem ser avaliadas baseadas nos modelos identificados, conforme foi dito anteriormente. Os modelos matemáticos selecionados no capítulo anterior foram o NARX série-paralelo e paralelo para modelar essa resposta dinâmica do sistema histerético sujeito a cargas dinâmicas e aleatórias. O passo seguinte é definir quais são as entradas e saídas da rede para os dados reais.

Os dados de laboratório referem-se à tensão e deformação e os dados disponíveis do sistema real correspondem a esforços e deslocamentos (rotação). A maioria dos trabalhos envolvendo a modelagem da histerese com redes neurais (p.ex. (BASHEER *et al.*, 2002) (G. W. ELLIS. *et al.*, 1996) trabalha com a relação tensão-deformação ( $\sigma - \varepsilon$ ) de geomateriais sob carregamento cíclico, sendo a tensão a variável saída e a deformação a entrada, i.e.,  $\sigma = f(\varepsilon)$ . No trabalho (BASHEER *et al.*, *op. cit*) o autor diz que uma das

vantagens do mapeamento seqüencial é que a relação inversa pode ser facilmente derivada pela troca dos vetores de entrada pelos vetores de saída, ou seja, a deformação como saída e tensão como entrada ( $\varepsilon = f(\sigma)$ ).

Foram encontrados poucos trabalhos com a modelagem da relação força-deslocamento. No trabalho de (WAON-HO YI *et al.*, 2000), os autores adotam como entrada da rede neural os deslocamentos e como saída as forças atuantes em uma junta de teste viga-coluna sujeita a cargas cíclicas sob histerese.

Neste estudo foi adotado como entrada os esforços (momento e tração) e como saída o deslocamento (rotação) atuante na flexjoint, simplesmente por uma questão de lógica física mais comum, as forças como excitações e os deslocamentos como respostas (ver Figura 2.13). Ensaios comprovaram que a relação é inversível, ou seja, os deslocamentos tanto podem ser entrada quanto saída.

### 6.2.1 - As variáveis dos modelos

No capítulo 4 analisou-se o comportamento das variáveis importantes e disponíveis para a modelagem do sistema, analisando a influência das variações ocorridas nelas e suas conseqüências no comportamento das outras variáveis. Então, para os modelos propostos têm-se as seguintes variáveis de entrada e saída:

- Variáveis de entrada

Num0 – momento na direção zero

Num90 – momento na direção 90°

Tração – carga axial

- Variáveis de saída

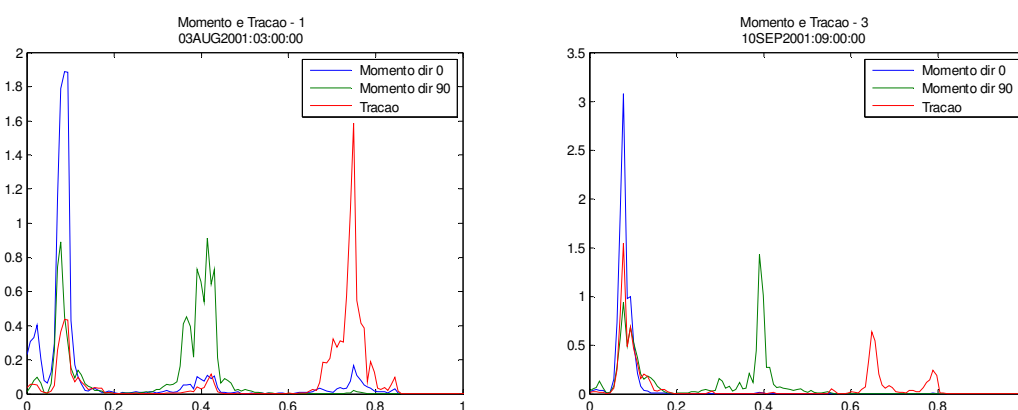
Apr0 – ângulo na direção 0°

Apr90 – ângulo na direção 90°

Apesar das medições do momento e do ângulo terem sido realizadas nas direções  $0^\circ$  e  $90^\circ$  por praticidade técnica e matemática, no início das análises, foram utilizadas as resultantes (tanto para o momento quanto para o ângulo) da composição dessas direções.

Conjeturou-se se a utilização das componentes no lugar das resultantes poderia fornecer um melhor desempenho do modelo. Comparando-se o modelo com as resultantes e com as componentes, verificou-se que a utilização das componentes realmente forneceu um ganho para o modelo.

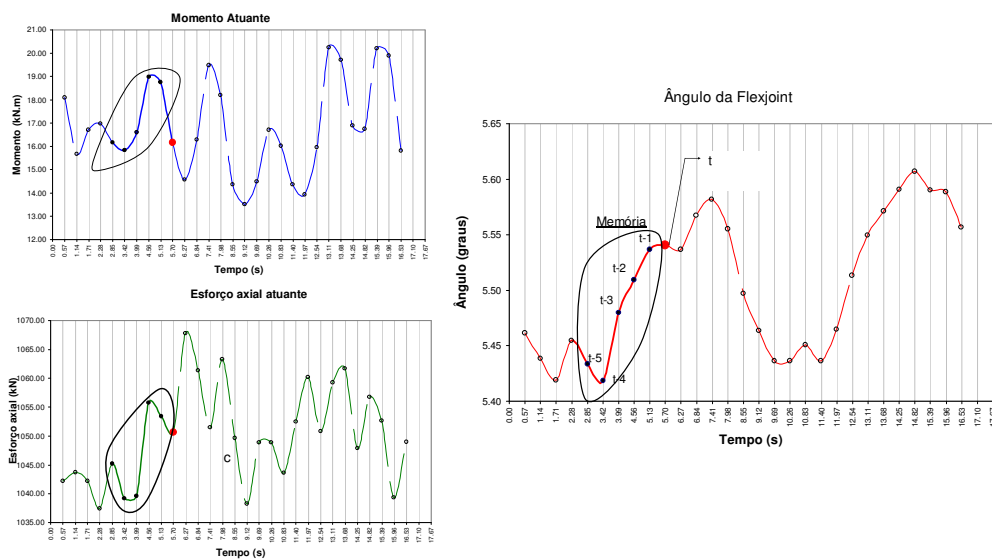
Uma possível explicação para este ganho pode estar associada à ação das ondas e da correnteza. Através de análises no domínio da frequência para o período em estudo, foi observado a predominância do fenômeno do VIV na direção  $90^\circ$  e a predominância das ondas na direção  $0^\circ$ . A adoção de uma resultante talvez mascare esses efeitos, homogeneizando variáveis que podem ter valores muito diferentes em função do fenômeno predominante na seqüência analisada. Na Figura 6.1 pode-se observar que o momento na direção  $0^\circ$  é predominante na região de frequência da onda e o momento na direção  $90^\circ$  mostra-se mais acentuado na faixa de frequência de VIV.



**Figura 6.1 – Espectro de energia da variável momento, nas direções  $0^\circ$  e  $90^\circ$  e da tração**

## 6.2.2 - Definição dos atrasos de tempo e parâmetros dos modelos

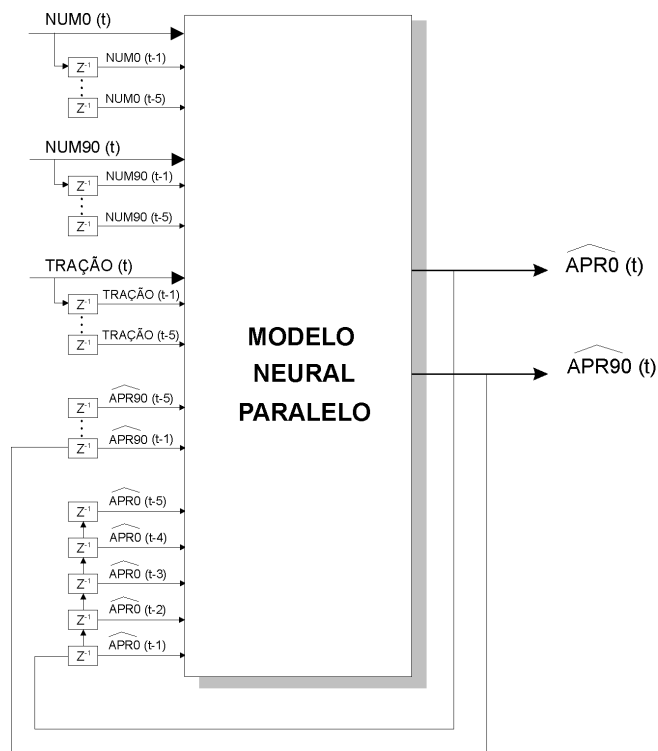
Os modelos propostos têm como premissa básica que as variáveis de saída tem influência dos valores atuais e passados de todas as variáveis de entrada externas e dos valores passados das próprias variáveis de saída. O processo de busca do atraso de tempo foi empírico, o mesmo adotado no capítulo 5, para os dados de laboratório. Durante a busca dos melhores resultados foi utilizada a rede série-paralelo com valores passados das variáveis entre 0 e 10, porém os melhores resultados foram obtidos com 5 valores passados. Cada variável de entrada pode influenciar as variáveis de saída com instantes de tempo diferentes, entretanto nesse estudo não foi investigada individualmente a contribuição de cada variável e sim de todo o conjunto de entradas. A Figura 6.2 é uma representação gráfica da memória (ou passado) das variáveis, necessária para obtenção do valor atual das variáveis de saídas.



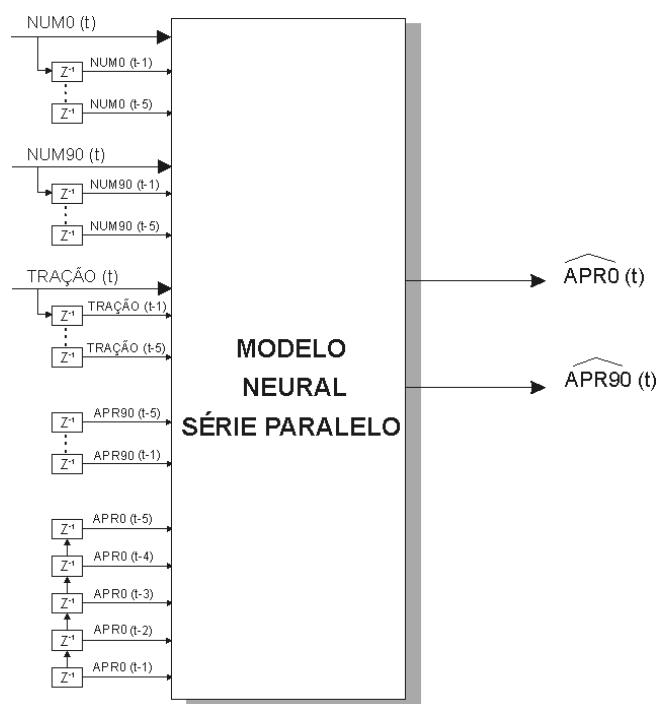
**Figura 6.2 – Representação da memória necessária para obtenção do valor atual da variável de saída.**

O algoritmo de treinamento utilizado foi o *backpropagation* com a utilização da aproximação do método de Newton Levenberg-Marquart do Matlab e o número de neurônios da camada escondida obtido foi 10. Outros parâmetros, a exemplo da taxa de aprendizado (0,01), foram adotados os padrões do software.

O conjunto de dados de entrada e da saída estão mostrados na Figura 6.3 para o modelo NARX paralelo e Figura 6.4 para o modelo NARX série-paralelo e a arquitetura utilizada é a apresentada na Figura 3.16.



**Figura 6.3 – Variáveis de entrada e saída para o modelo NARX paralelo**



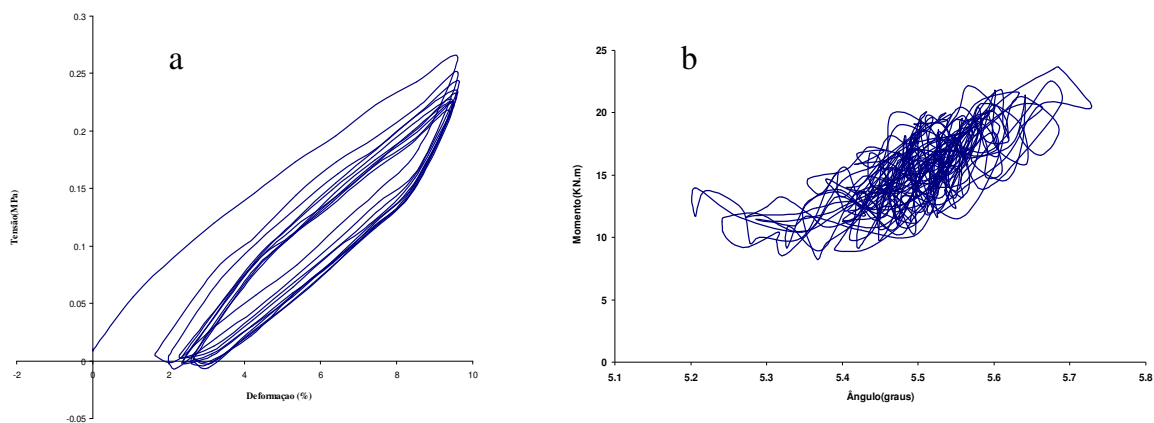
**Figura 6.4 – Variáveis de entrada e saída para o modelo NARX série-paralelo**

### 6.3 Conjuntos de treinamento, validação e teste

Os dados do comportamento rotacional da flexjoint vindos do processo de monitoração da plataforma P18 no período de 18/07/2001 a 20/11/2001 são usados como dados fonte para o treinamento, teste e validação da rede neural. Idealmente esses dados empregados para identificação deveriam ser capazes de refletir os diversos regimes que regem os comportamentos dinâmicos apresentados pelo sistema. Entretanto, o período estudado não contempla todas as estações do ano e a análise de cluster realizada no capítulo 4, mostra um conjunto de dados bastante homogêneo, o que pode estar refletindo uma faixa estreita de amplitudes e frequências do período em questão. Então, fica aqui registrado que o treinamento da rede não abrange todo o espectro da dinâmica ao qual o sistema pode estar sujeito. Surgindo novos dados, de outros períodos ou com características diferentes da apresentada, a rede deverá ser re-treinada.

A Figura 6.5 mostra os *loops* de histerese para os dados obtidos em laboratório e na Figura 6.2 pode ser visto o comportamento histeretico com a dinâmica do carregamento.





**Figura 6.5 - a) Ciclos de histerese para dados de laboratório e b) Ciclos de histerese para dados reais.**

Como foi descrito no capítulo 3, os dados são compostos de seqüências no tempo, cada seqüência representando um estado de mar, correnteza e vento, separadas umas das outras por um intervalo constante de tempo, correspondente a 2,5 horas (ver Figura 4.4). Algumas abordagens foram levantadas a fim de escolher como treinar, testar e validar a rede e foram separadas nos seguintes casos:

- Caso 1.** Considerar cada seqüência como um comportamento particular do mar (estado de mar, correnteza e vento) e conseqüentemente da flexjoint
- Caso 2.** Considerar todas as seqüências como sendo um único conjunto de dados.
- Caso 3.** Selecionar, aleatoriamente do período, uma coleção de séries composta de 2 grupos: o primeiro para treinamento e validação e o segundo para teste.

## Caso 1

Com relação à primeira abordagem, a fim de que o modelo pudesse representar o comportamento global do sistema flexjoint durante o período de tempo analisado, foram feitas algumas considerações:

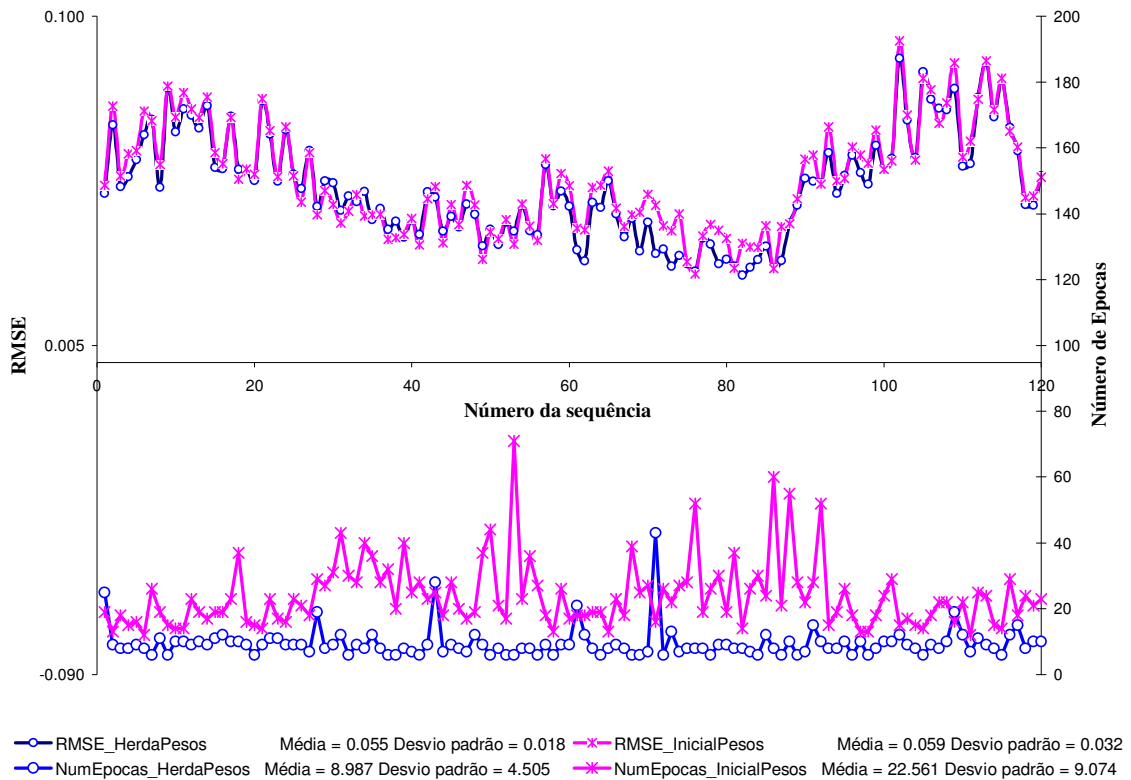
- As seqüências foram divididas em duas partes: 80% e 20%, através de amostragem aleatória. Cada uma das seqüências que compõem os 80% foi separada em treinamento e validação, observando-se agora a ordem. Os 70% iniciais da seqüência foram utilizados para treinamento e o restante final da seqüência para validação. Os 20% das seqüências foram utilizadas para teste.
- A rede é treinada e validada com todas as seqüências do conjunto treinamento-validação, sendo que a rede é treinada por uma seqüência de cada vez, aprendendo aquele comportamento apresentado. A próxima seqüência a ser apresentada para essa mesma rede, já treinada com a seqüência anterior, tem como parâmetros iniciais (pesos e *bias*) os valores do treinamento anterior que foram salvos e assim sucessivamente até finalizar todas as seqüências do treinamento.

Várias análises foram realizadas e a fim de comparar a influência de herdar ou não os parâmetros de uma seqüência para outra, as análises se dividiram em:

- Treinamento com herança dos pesos de uma seqüência para outra (1)
- Treinamento sem herança dos pesos, ou seja, os pesos são inicializados a cada seqüência apresentada para a rede (2)

No gráfico da Figura 6.6 são apresentados os RMSE's e número de épocas para 120 seqüências para o dois tipos de análises. É fácil observar que em termos de RMSE as duas abordagens têm o mesmo desempenho, com médias de 0,055 (1) e 0,059 (2) e desvio

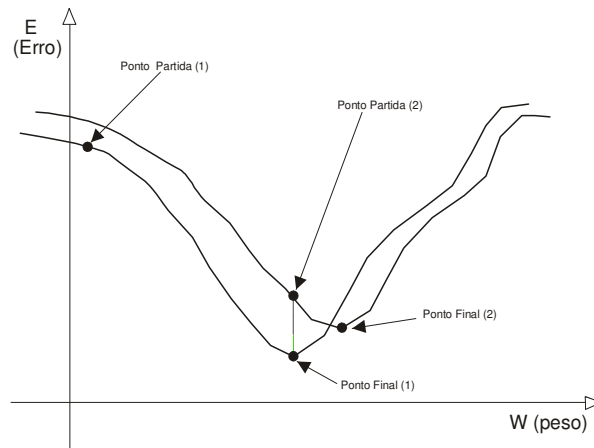
padrão de 0,018 (1) e 0,032 (2). Contudo, o número de épocas necessário para treinar uma rede com pesos inicializados a cada seqüência nova apresentada à rede é significativamente superior à abordagem que, a cada nova seqüência herda pesos do treinamento anterior. O número médio de épocas em (1) é 8,987 e em (2) é de 22,561 épocas.



**Figura 6.6 – RMSE e número de épocas de 120 séries**

Analisando os resultados, entende-se que cada série tem suas peculiaridades e o fato da rede ter seus parâmetros inicializados em função do conhecimento “aprendido” na série apresentada no passo anterior não minimiza o erro. Após 120 séries apresentadas seqüencialmente, esperava-se um RMSE decrescente, o que não foi verificado, como pode ser observado na parte superior da Figura 6.6. Contudo, o ponto de partida na curva de aprendizado pode ser otimizado, ou seja, em função da seqüência anterior a rede pode aprender mais rápido e esse fato explica o número de épocas menor no caso em que a rede herda os pesos de uma seqüência para outra. A Figura 6.7 ilustra esse comportamento em

um gráfico Erro x Peso. Em função do aprendizado anterior, o ponto de partida 2 está bem mais próximo de um mínimo (ponto final 2), alcançando-o com menor número de épocas.



**Figura 6.7 – Esquema da curva erro versus peso para duas seqüências**

Entendeu-se que rede “acumularia” conhecimento a cada nova seqüência apresentada. Entretanto, é conhecido que os modelos convencionais de redes neurais têm um problema chamado “esquecimento catastrófico” (*catastrophic forgetting*) ou também algumas vezes chamado de “interferência catastrófica” (*catastrophic interference*). O esquecimento catastrófico refere-se à inabilidade da rede neural de reter antigas informações na presença de novas. Ou seja, após treinar uma rede para uma seqüência de padrões, se ela for treinada para uma seqüência diferente, ela “desaprende” as antigas e substitui pelas novas (SEIPONE et al, 2005). Para modelos que representem o processo de aprendizado do cérebro (cognição humana) este fato é bastante crítico visto que o cérebro é capaz de aprender uma informação sem necessariamente esquecer outra. Os seres humanos lembram-se dos erros para compará-lo com os acertos.

## **Caso 2**

Existem alguns problemas na segunda abordagem. Primeiro, como os conjuntos de treinamento, teste e validação devem ser extraídos na ordem, implicaria que a rede seria

treinada para algumas estações do ano e testada e validada com outras às quais nunca foi apresentada.

Outro aspecto é referente à representação da memória no modelo. Cada seqüência representa 30 minutos de cada 3 horas. E, como as redes dinâmicas trabalham com atrasos de tempo, o início de uma nova seqüência não tem como passado o final da seqüência anterior, já que as condições de carga (onda, vento e correnteza) são teoricamente outras após 2,5 horas. A princípio este problema acarretaria alguma queda de desempenho da rede devido ao ruído presente no início de cada seqüência. Porém, essa perturbação, além de ser somente no início, as seqüências são longas o suficiente para dissipar esse efeito. Fato este comprovado em análises efetuadas. Mesmo sabendo que o efeito se dissipa, optou-se por eliminar os primeiros registros (o número de registros eliminados é função do *lag* adotado no modelo) de cada série sendo estes apresentados ao modelo apenas como o passado para os primeiros registros efetivamente apresentados para a rede neural.

Além dos problemas de modelagem, este caso esbarra num problema computacional. A versão do software utilizado para implementação dos modelos não foi capaz de gerenciar a quantidade de memória necessária para armazenar os dados representados segundo este caso.

### **Caso 3**

Este caso foi pensado a partir do problema de gerência de memória descrito no caso anterior. Como o MATLAB não foi capaz de tratar todo o conjunto de séries, foi necessário identificar até que volume o MATLAB seria capaz de tratar. Após alguns testes, verificou-se que este limite equivale a 11% do número total de séries. Além disso, foram feitas análises que mostraram que o uso de um subconjunto de 11% das séries tomadas aleatoriamente (distribuição uniforme) é representativo do comportamento da flexjoint, verificando-se que os dados são homogêneos o suficiente para tanto.

É importante observar que cada série é entendida como um comportamento único porque representa um estado de mar (referente a um período de 3 horas) que está sim associado ao período seguinte (pelo fato de ser o mar no mesmo local num momento posterior), mas que é independente estatisticamente deste. Por este motivo não é problema

para a solução do sistema a apresentação ao modelo de séries que não estejam na ordem em que foram observadas, i.e., em ordem aleatória.

De acordo com descrito no item 5.2, o modelo deve ser treinado, validado e testado com amostras nas seguintes proporções: 60%, 20% e 20% respectivamente. Sendo assim, o conjunto de dados usado para modelar o sistema escolhido do volume total de séries de forma que o conjunto para treinamento, validação e teste juntos some 11% do total e mantenha a devida proporção.

O problema de representação da “memória” ou “passado” no modelo, também mencionado no caso 2, foi resolvido da mesma forma.

Este é o caso implementado.

#### **6.4 - Comparação dos modelos série-paralelo e paralelo**

Nesta seção, são apresentados resultados de identificação do sistema dinâmico não linear flexjoint, através dos modelos escolhidos no item 6.2, modelos NARX série-paralelo e paralelo. Os resultados dos testes com relação à estrutura das arquiteturas descritas acima são apresentados a seguir.

Os parâmetros de treinamento das redes foram:

- Arquitetura da rede: camada de entrada: 24 unidades  
camada escondida: 10 unidades processadoras  
camada de saída: 1 unidade processadora
- Número de padrões de treinamento: 104 padrões (Séries)
- Número de épocas de treinamento: 600 épocas
- Função de ativação: sigmóide
- Número de atrasos de tempo (janela temporal): 5

- Apresentação seqüencial dos padrões para treinamento
- Algoritmo de treinamento: Levenberg-Marquard
- Critério de *overtraining*: *Early stopping*
- Software : MATLAB versão 2006Ra

Dado todos os parâmetros acima, foram utilizadas as funções do MATLAB *newnarxsp* e *newnarx* para criar as redes recorrentes série-paralelo e paralelo respectivamente.

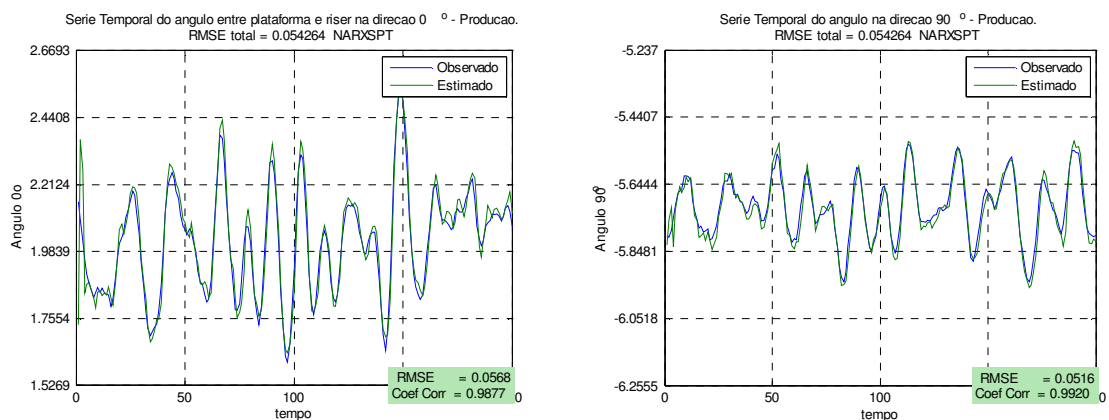
A Tabela 6.1, mostra os erros quadráticos médios (RMSE) de treinamento e teste de cada um dos modelos de redes neurais considerados.

**Tabela 6.1 – Erros médios quadráticos dos modelos**

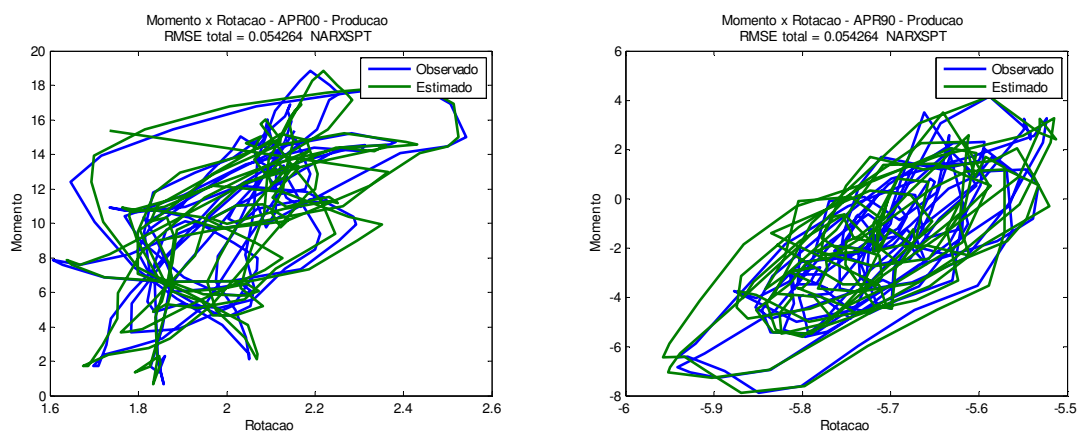
Modelo	Variavel Saída	RMSE Treinamento	RMSE Teste
Serie-Paralelo	Rotacao 0°	0.0238	0.0568
Paralelo	Rotacao 90°	0.0289	0.0516
Paralelo	Rotacao 0°	0.1621	0.3803
	Rotacao 90°	0.2652	0.3663

Os resultados obtidos para as estruturas série-paralelo e paralelo são ilustrados nas Figuras 6.8, 6.9, 6.10 e 6.11. Estes resultados representam os melhores resultados obtidos para cada estrutura após várias execuções, alterando os parâmetros associados a cada estrutura como taxas de aprendizado, inicialização dos pesos, atrasos de tempo, o conjunto de treinamento e considerando o melhor desempenho para cada modelo.

Na Figura 6.8, referente ao modelo NARX Série paralelo, os gráficos representam valores das séries estimadas (—) e observadas (—) das variáveis de saída, rotação nas direções 0° e 90°. E a figura 6.9 mostra a relação Momento x Rotação estimado e observado.



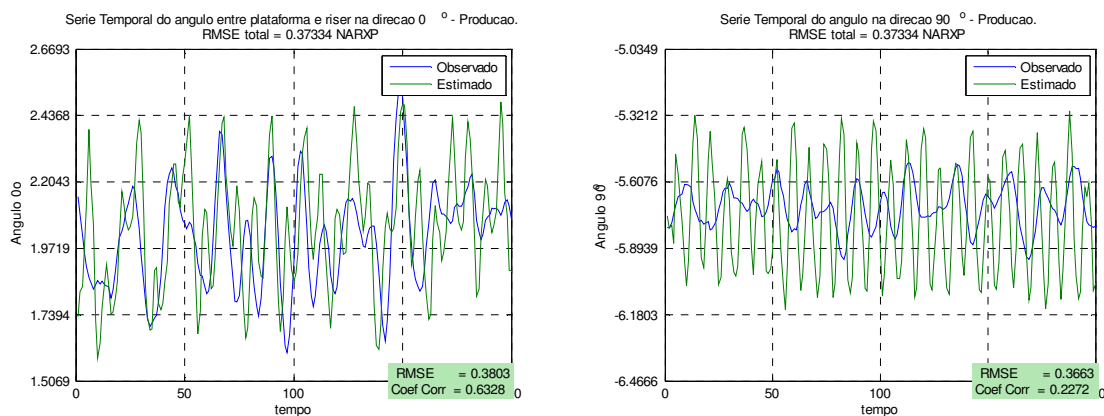
**Figura 6.8 – Amostra dos valores das séries observadas e estimadas da rotação (0° e 90°) – Modelo Série-paralelo**



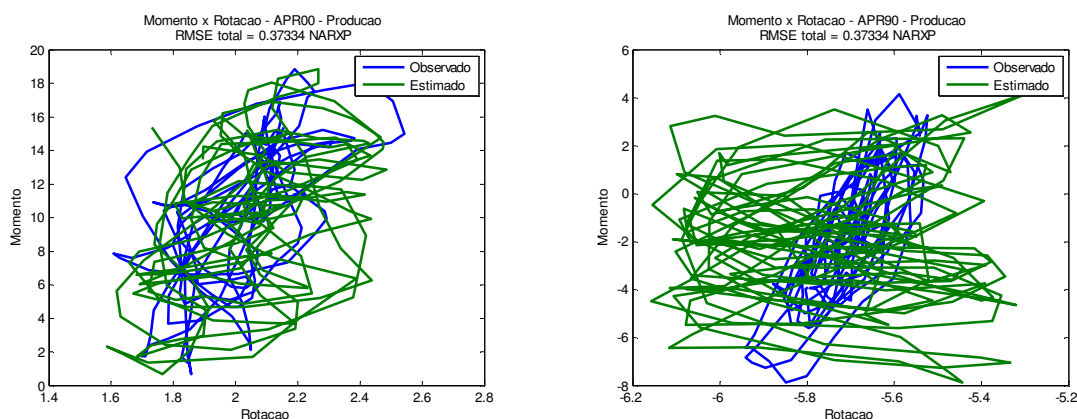
**Figura 6.9 – Relação momento X rotação observado e estimado (0° e 90°) – Modelo série-paralelo**

O mesmo conjunto de padrões gerou os resultados mostrados nas Figuras 6.10 e 6.11 para o modelo treinado paralelo e testado paralelo.





**Figura 6.10 – Amostra dos valores das séries observadas e estimadas da rotação (0° e 90°) – Modelo Paralelo**



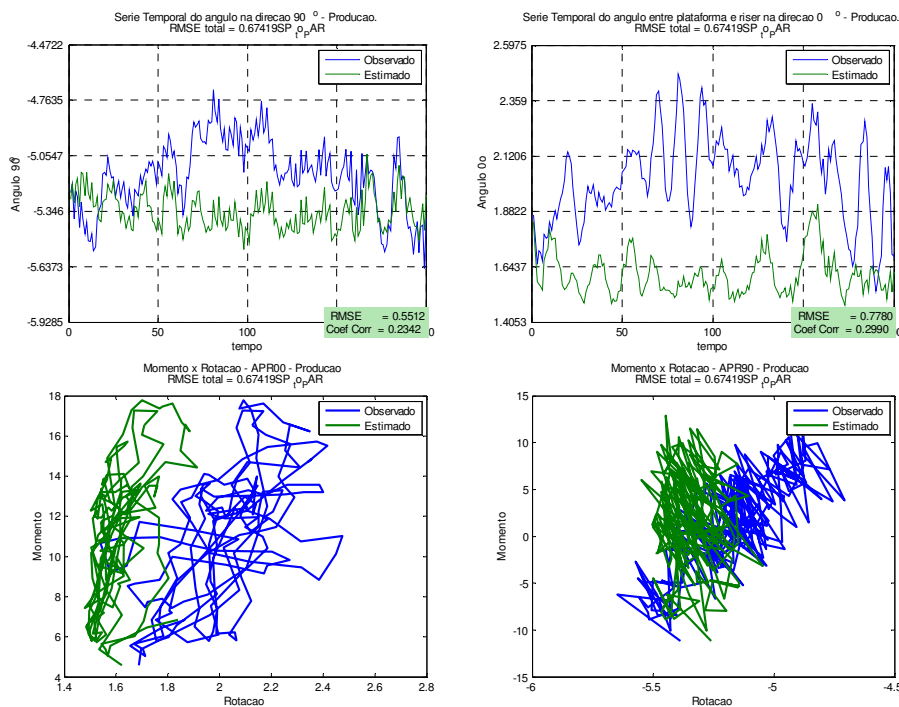
**Figura 6.11 – Relação momento X rotação observado e estimado (0° e 90°) – Modelo Paralelo**

Analisando os resultados, pode-se dizer que o modelo série-paralelo apresenta um RMSE de teste muito menor que o paralelo. O que era de se esperar, pois como foi visto no capítulo 5, os modelos paralelos não garantem a convergência e levam muito mais tempo para serem treinados.

Os resultados apresentados foram obtidos considerando o modelo paralelo treinado e trabalhando no modo paralelo. Contudo, quando o erro na saída da rede neural tende a valores pequenos, tal que  $y(t) \approx \hat{y}(t)$ , (NARENDA&PARTHASARATHY, *op. cit*) diz que

o modelo série-paralelo pode ser substituído pelo modelo paralelo sem consequências sérias.

Para o caso em estudo, os erros apresentados pelo treinamento com o modelo série-paralelo são considerados muito pequenos. A Figura 6.12 ilustra a saída da rede treinada em série-paralelo trabalhando em modo paralelo na identificação do sistema.



**Figura 6.12 – Modelo série-paralelo(treinamento) / paralelo (teste)**

Como pode ser observado, a rede neural treinada em série-paralelo e trabalhando em modo paralelo, tem um desempenho pior comparado com o paralelo-paralelo, obtendo neste caso, um RMSE de teste de 0.5734 para a rotação a  $0^\circ$  e de 0.4657 para rotação a  $90^\circ$ . É interessante notar que o modelo represserva a estabilidade nos primeiros passos e logo diverge, característica comum nos modelos paralelos (CALOBA,2002).

## 6.5 - Influência das variáveis externas nos modelos

Alguns testes foram realizados para avaliar a influência das variáveis externas (momento e tração) no comportamento do modelo através de re-análises utilizando a própria rede neural, retirando-se as variáveis externas do modelo e obtendo-se seu desempenho em termos de RMSE.

Os erros de treinamento e teste do modelo série-paralelo sem as variáveis externas (Modelo II) são apresentados na tabela 6.2 e comparados com o Modelo I, já apresentado anteriormente.

**Tabela 6.2 – Resultados do modelo série-paralelo com e sem variáveis externas**

	Variáveis de Entrada	Variável Saída	RMSE Treinamento	RMSE Teste
Modelo I	Momento	Rotação 0°	0.0238	0.0568
	Tracao Rotação	Rotação 90°	0.0289	0.0516
Modelo II	Rotação	Rotação 0°	0.0254	0.0453
		Rotação 90°	0.0266	0.0413

Como pode ser observado, há pouca influência das variáveis externas. Assim como todas as análises realizadas neste estudo, estes resultados representam os melhores resultados obtidos após várias execuções. Esse resultado é surpreendente devido ao conhecimento do fenômeno físico e da boa correlação entre as variáveis externas e as variáveis de saída.

Um das hipóteses levantadas para explicar a fraca contribuição das variáveis externas nos modelos série-paralelo foi a de que as seqüências poderiam ter sido amostradas com alta taxa de aquisição. A seção seguinte trata dessa investigação.

### 6.5.1 - Amostragem e Filtragem

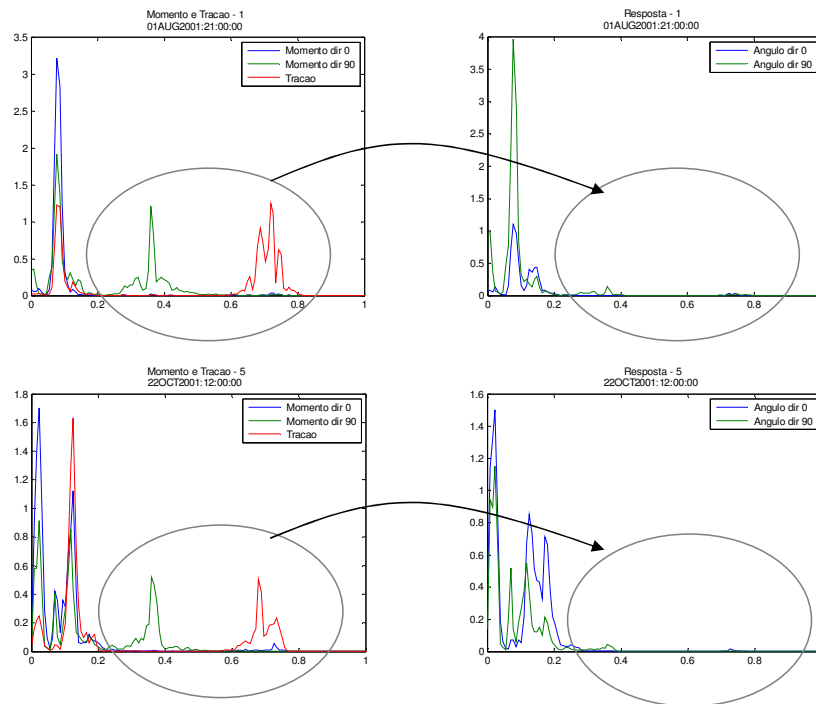
Decimação é uma operação que diminui a taxa de amostragem do sinal. Em outras palavras, significa eliminar dados. Decimar um sinal pelo fator 2, significa que cada 2 pontos, um é eliminado (sub-amostrado por um fator de 2), obtendo-se metade do sinal original. No processo de decimação, aumenta-se o tamanho do período de amostragem do sinal, e conseqüentemente, diminui-se a taxa de amostragem .

Para reduzir o número de pontos das seqüências sem deteriorar a qualidade do sinal, foi adotado o teorema desenvolvido por Shannon, que garante que se o critério de *Nyquist* (MATLAB *Help*, 2006) for satisfeito, o sinal pode ser reconstruído sem qualquer perda de informação. De acordo com o Teorema de *Nyquist*, a quantidade de amostras por unidade de tempo de um sinal, chamada taxa ou freqüência de amostragem, deve ser maior que o dobro da maior freqüência contida no sinal a ser amostrado, para que possa ser reproduzido integralmente sem erro de *aliasing*. A metade da freqüência de amostragem é chamada **freqüência de Nyquist** e corresponde ao limite máximo de freqüência do sinal que pode ser reproduzido.

Utilizando-se a função *Spectrum* do MATLAB, foram obtidos os sinais no domínio da freqüência. A freqüência de corte foi obtida através de análise visual dos espectros, e a nova taxa de amostragem foi definida com a freqüência o dobro dessa freqüência de corte. Segundo as análises, o tempo de amostragem dos sinais pode ser aumentado por um fator de 2. O intervalo de amostragem do processo de monitoração foi de 1s.

Entretanto, durante estas análises dos espectros de energia, foi observado em grande parte destes, que quando havia excitação em alta freqüência nos espectros dos esforços, possivelmente devido ao fenômeno de VIV, não havia resposta na rotação ou pelo menos não havia na intensidade esperada. Esse fato já foi relatado no capítulo 4 referente à descrição dos dados.

Na Figura 6.13 pode-se observar excitação nas faixas de freqüência de 0.4 Hz para o momento e entre 0.6 e 0.8 Hz para a tração e quase nenhuma excitação para a rotação correspondente, tanto nos dias 01/08/2001 às 21:00 e 22/10/2001 às 12:00.



**Figura 6.13 – Espectros de energia do momento (0o e 90°), tração e rotação em 01/08/2001 às 21:00 e 22/10/2001 às 12:00.**

A principal razão para esse fato é que o equipamento utilizado para medição das inclinações de topo não é capaz de capturar adequadamente as altas frequências.

Em função dessa observação, entendeu-se que o mapeamento dos esforços e rotação seria mais bem representado na faixa de frequência da onda, ou seja, em baixas frequências.

Portanto, antes de sub amostrar, foi passado um filtro passa-baixas, que permite a passagem de frequências até a frequência de corte, com a finalidade de eliminar as altas frequências. A Figura 6.14 mostra como o sinal será tratado. O sinal entra com uma frequência  $f$ , são eliminadas as altas frequências em função do limiar definido e então é feita a decimação, tendo como saída um sinal metade do original.



**Figura 6.14 – Esquema empregado para decimação do sinal**

Foram realizadas análises com os modelos serie-paralelo (SP) e paralelo (P) com as considerações relacionadas abaixo:

- Com Filtro e variáveis externas (SP\_CF)
- Com Filtro sem variáveis externas (SP\_CFAng)
- Com Filtro e decimação com variáveis externas (SP\_CFD, P\_CFD, SP\_P\_CFD)
- Com Filtro e decimação sem variáveis externas (SP\_CFDAng)
- Sem Filtro com variáveis externas (SP\_SF, P\_SF, SP\_P\_SF)
- Sem Filtro sem variáveis externas (SP\_SFAng)

Os resultados das análises encontram-se resumidas na Tabela 6.3.

Analisando os modelos série-paralelo com variáveis externas, os com filtro (SP\_CF) apresentam melhores resultados que aqueles sem o filtro (SP\_SF). Os RMSEs de teste são 0.0243 e 0.0568 respectivamente, para a variável de saída ângulo a 90°. Já os modelos com filtro (SF\_CFD) e decimação (RMSE=0.0560) não apresentam diferenças significativas com relação aos modelos sem filtro (SF\_SF) (RMSE=0.0568). Uma hipótese levantada é a da perda de qualidade do sinal com a decimação, apesar do ganho comprovado com a filtragem.

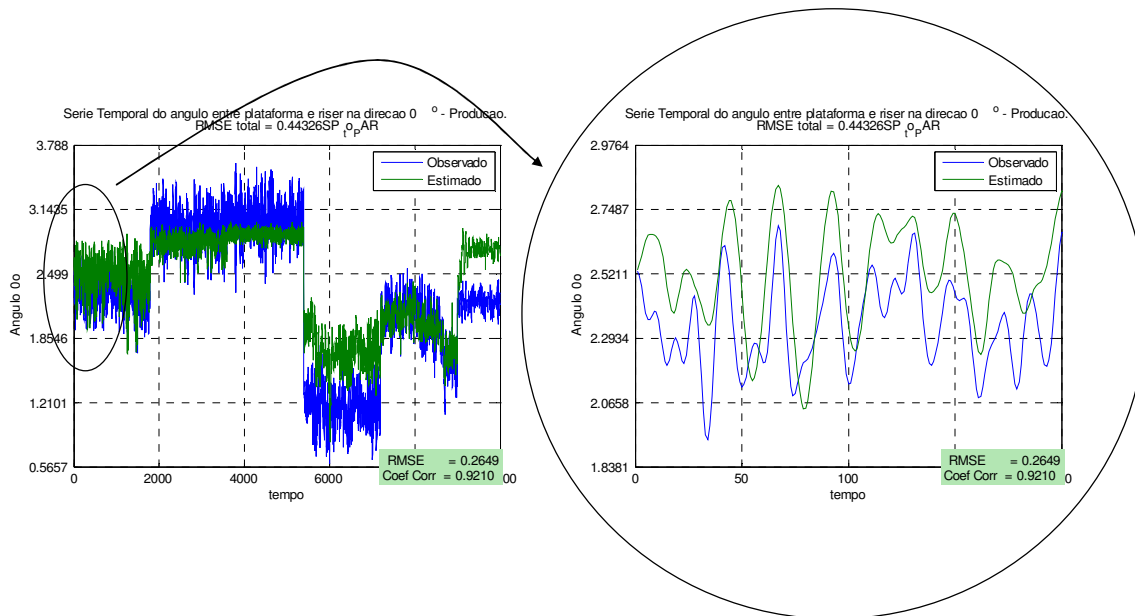
Todos os modelos sem as variáveis externas têm um desempenho levemente melhor que aqueles com as variáveis externas. Por exemplo, o modelo SP\_CFAng tem um RMSE de 0.0278 e sem filtro (SP\_SFAng) o RMSE é de 0.0453.

O desempenho dos modelos paralelo, treinado e testado paralelo, é significativamente inferior aos modelos série-paralelo. Aqui também a decimação parece ter piorado o desempenho do modelo. Para a variável externa ângulo a  $0^\circ$ , os RMSEs para os modelos P\_CFD e P\_SF são 0.6688 e 0.3803, respectivamente.

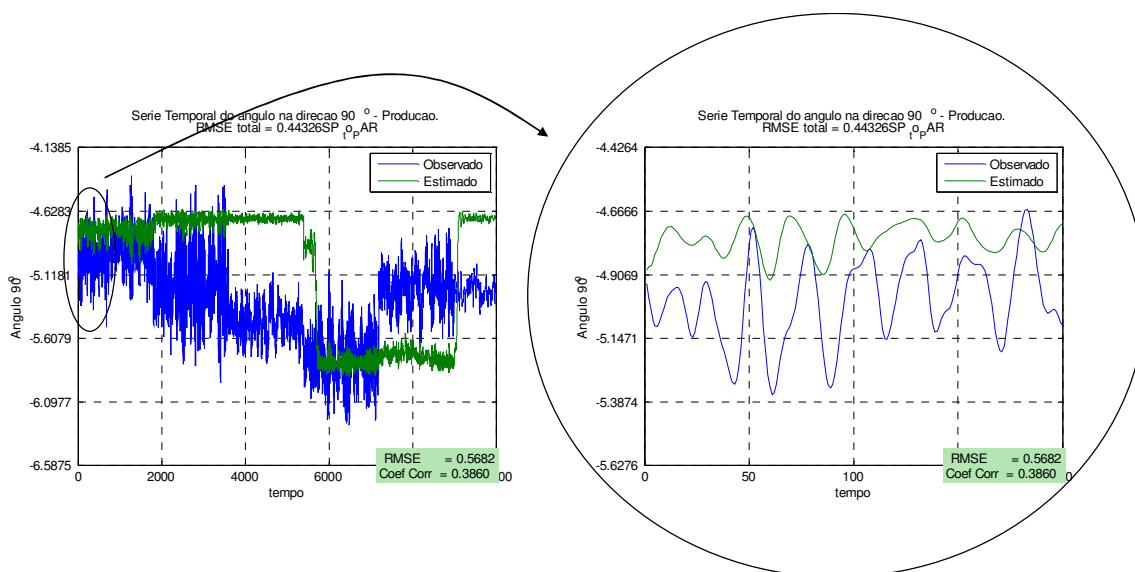
Tendo em vista que os erros do treinamento foram pequenos na arquitetura série-paralelo, esta pode ser convertida em paralelo e utilizada para predições a mais longo prazo. As arquiteturas série-paralelo empregadas neste estudo são para predições a “um-passo-a-frente”. Foram então gerados os modelos SP\_P\_CFD e SP\_P\_SF.

Os resultados são bastante estimuladores. Como mostra a Figura 6.15, apesar das discrepâncias, o modelo consegue acompanhar razoavelmente o comportamento do sistema até 2000 segundos (30 min) para a variável de saída ângulo a  $0^\circ$ . Entretanto, para a variável de saída ângulo a  $90^\circ$ , o desempenho não é tão bom, conforme pode ser visto através da Figura 6.16.

Uma possível explicação para essa diferença de desempenho pode ser devido ao poder de explicação das variáveis. A variável de saída APR0 (ângulo na direção  $0^\circ$ ) tem uma correlação com as variáveis externas (NUM0 e NUT) muito mais alta que a variável APR90 (ângulo na direção  $90^\circ$ ) com suas correlatas (NUM90 e NUT), conforme mostrado na Tabela 4.6.



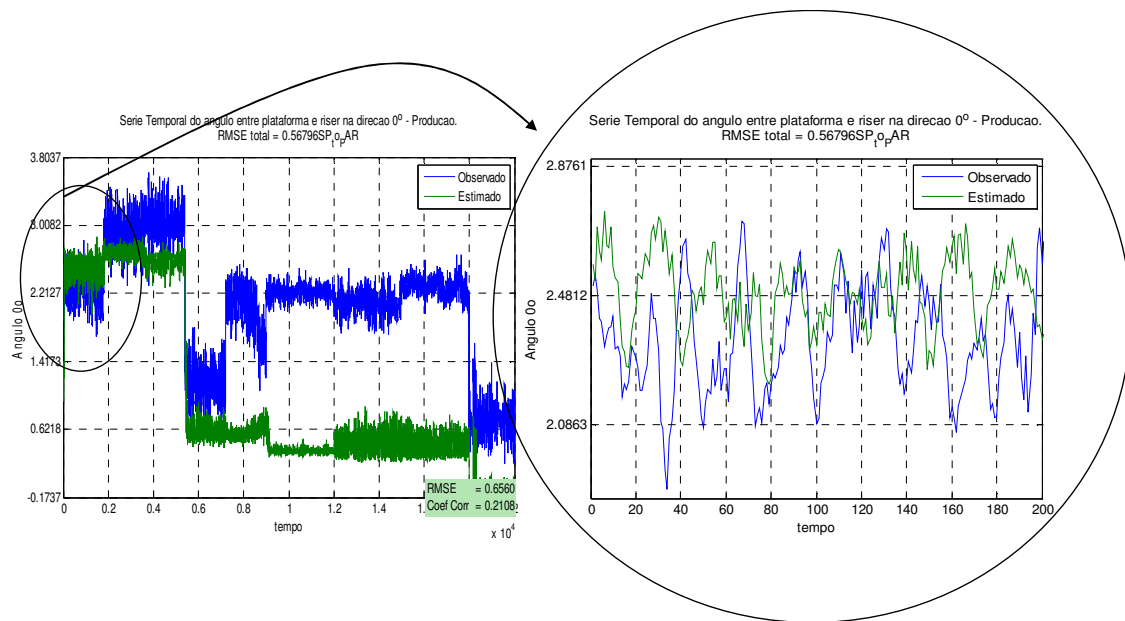
**Figura 6.15 - Observado versus Estimado das séries da variável saída APR0 para o modelo SP\_P\_CFD**



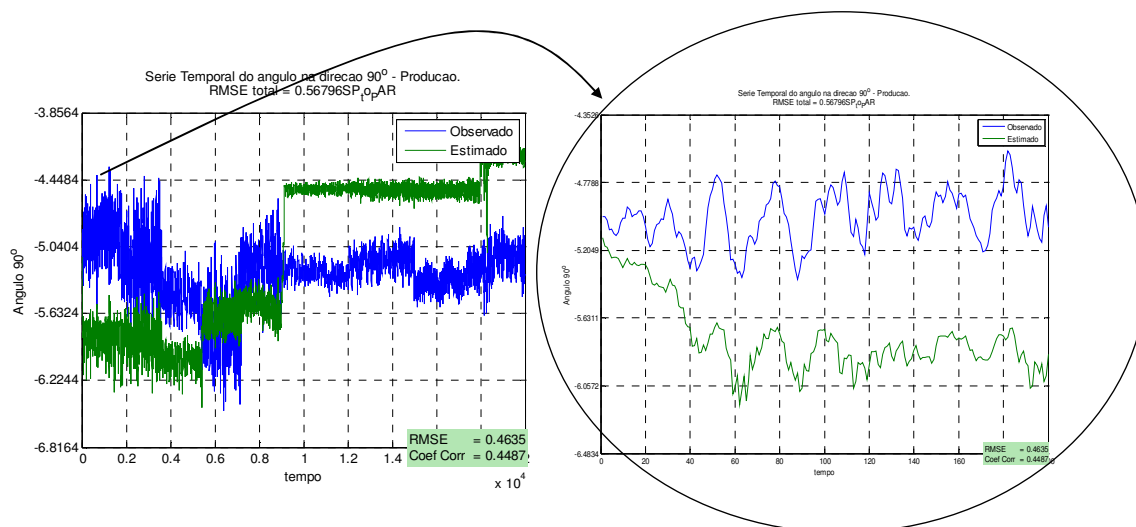
**Figura 6.16 - Observado versus Estimado das séries da variável saída APR90 para o modelo SP\_P\_CFD**



Os resultados para os modelos treinados em serie-paralelo e testado com o paralelo sem o filtro, são inferiores aos modelos com filtro e decimação. As Figuras 6.17 e 6.18 mostram o desempenho da rede para os primeiros 200 segundos. A variável de saída APR0 tem um desempenho melhor que a APR90, comportamento semelhante aos do modelo com filtro.



**Figura 6.17 - Observado versus Estimado das séries da variável saída APR0 para o modelo SP\_P\_SF**



**Figura 6.18 - Observado versus Estimado das séries da variável saída APR90 para o modelo SP\_P\_SF**

Houve um ganho com o processo de filtragem, entretanto quanto a questão da influencia das variáveis externas continuam as duvidas. O processo de filtragem e decimação não revelaram a influencia destas.

Resolveu-se então fazer as mesmas análises para os dados de laboratório para se investigar se esse resultado também acontece neles. A seção seguinte trata dessa investigação.

**Tabela 6.3: Resultados dos modelos testados**

Arquitetura	Modelo	Sinal	Numero de Atrasos	Variaveis de Entrada	Variaveis de Saida	RMSE	RMSE
						Treinamento	Teste
Serie-Paralelo	SP_CF	Com Filtro	5	Momento	Ângulo 0o	0.0109	0.0243
				Tracao	Ângulo 90o	0.0100	0.0168
				Ângulo			
	SP_CFD	Com Filtro e Decimacao (fator 2)	5	Momento	Ângulo 0o	0.0154	0.0560
				Tracao	Ângulo 90o	0.0143	0.0436
				Ângulo			
	SP_CFAng	Com Filtro	5	Ângulo	Ângulo 0o	0.0120	0.0278
					Ângulo 90o	0.0122	0.0216
	SP_CFDAng	Com Filtro e Decimacao	5	Ângulo	Ângulo 0o	0.0154	0.0350
					Ângulo 90o	0.0150	0.0253
SP_SF	Sem Filtro	5	Momento	Ângulo 0o	0.0238	0.0568	
			Tracao	Ângulo 90o	0.0289	0.0516	
			Ângulo				
SP_SFAng	Sem Filtro	5	Ângulo	Ângulo 0o	0.0254	0.0453	
				Ângulo 90o	0.0266	0.0413	
Uma serie		5	Momento	Ângulo 0o	0.0448	0.0708	
			Tracao	Ângulo 90o	0.0784	0.1109	
			Ângulo				
Paralelo	P_CFD	Com Filtro e Decimacao	5	Momento	Ângulo 0o	0.6241	0.6688
				Tracao	Ângulo 90o	0.8403	1.0124
				Ângulo			
	P_SF	Sem Filtro	5	Momento	Ângulo 0o	0.1621	0.3803
Tracao				Ângulo 90o	0.2652	0.3663	
Paralelo - Rede treinada com  Serie Paralelo	SP_P_CFD	Com Filtro e Decimacao	5	Momento	Ângulo 0o		0.2649
				Tracao	Ângulo 90o		0.5682
				Ângulo			
	SP_P_SF	Sem Filtro	5	Momento	Ângulo 0o		0.6560
				Tracao	Ângulo 90o		0.4635
			Ângulo				

### 6.5.2 - Análise da influência das variáveis externas com os dados de laboratorio

A idéia de desse modelo surgiu devido à necessidade de se avaliar a influência das variáveis externas no modelo. No caso estudado para os dados de laboratório, só há uma variável de entrada externa, a deformação, mas com os dados do problema da monitoração em escala real, tem-se mais de uma variável de entrada, o momento e a tração.

No caso das abordagens NARX série-paralelo e paralelo, a variável de saída é também variável de entrada, seja como variável observada ou a própria estimada, ambas defasadas no tempo. Em função dos resultados obtidos com os dados reais, resolveu-se fazer um teste, retirando a variável externa, deformação e avaliando o desempenho da rede somente com a tensão no tempo  $t$  como saída e a tensão em tempos anteriores defasados de 1, como entrada da rede.

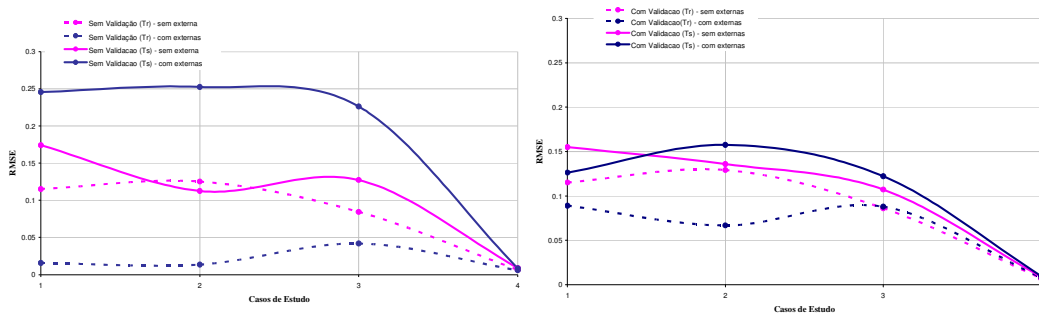
Foram mantidos os mesmos parâmetros para todos os modelos estudados para esses dados, tais como: o algoritmo de treinamento (Levenberg-Marquard), número de neurônios da camada escondida (10), número de épocas limite (600), critério de parada do treinamento (early stopping), função de ativação (sigmóide), janela temporal igual a 5.

Os mesmos conjuntos de dados definidos no capítulo 5, representados pelos casos I,II,III e IV, foram empregados para conduzir as análises. A Tabela 6.4 mostra o desempenho da rede em termos de RMSE.

**Tabela 6.4 – Desempenho do modelo sem as variáveis externas**

Modelo	Casos	RMSE			
		Treinamento		Teste	
		Com Validacao	Sem Validação	Com Validacao	Sem Validacao
Série-Paralelo (Dados-laboratorio)	I	0.1152	0.1152	0.1549	0.1741
	II	0.1291	0.1252	0.1359	0.1125
	III	0.0861	0.0845	0.1071	0.1273
	IV	0.0072	0.0065	0.0088	0.0092

Algumas questões foram levantadas em função dos resultados dessas análises.



**Figura 6.19: RMSEs das análises dos modelos série-paralelo com as variáveis externas (azul) e sem externas (rosa) (a) – Sem validação (b) – Com validação.**

Foram feitas comparações os modelos série-paralelo sem variáveis externas e com variáveis externas. As Figuras 6.19.a e 6.19.b são representações gráficas das Tabelas 6.4 e 5.x.

Observa-se que o desempenho do treinamento (linhas tracejadas) é geralmente superior ao do teste (linhas cheias), tanto nos modelos com e sem o controle de *overtraining*. Entretanto, um resultado inesperado é que os modelos sem as variáveis externas mostram um desempenho ligeiramente superior, no teste. Já no treinamento o desempenho é inferior ao do modelo com variáveis externas, porém não muito significativo, mas que parece mais coerente.

Um outro aspecto é a redução dos erros quanto ao número de padrões fornecidos para a rede. Os dados mostram um comportamento relativamente fácil de se aprender, pois não são grandes as discrepâncias de RMSE entre os casos estudados, entretanto, o caso IV que é composto de maior número de exemplos, apresenta o RMSE muito mais baixo, com um ajustamento dos dados quase perfeito e com excelente capacidade de generalização.

Mas, a questão mais relevante nessas comparações é o desempenho da rede sem as variáveis externas. Este modelo mostra um desempenho excelente sem a variável externa, a deformação. Ou seja, tanto os dados reais quanto o de laboratório apresentam uma mesma característica que não está relacionada à amostragem, por exemplo, como foi conjecturado. Uma hipótese levantada seria associada à natureza da relação tensão-deformação ou força-deslocamento.

## 6.6 - Resumo

Dos resultados apresentados, pode-se dizer que as redes série-paralelo propostas mostraram-se mais eficientes no processo de identificação do sistema que as redes modelo paralelo. Contudo, conforme abordado no capítulo anterior, a rede neural recorrente paralelo aproxima a dinâmica do sistema, ou seja, ela é capaz de descobrir as leis fundamentais que regem o comportamento do sistema, enquanto que, os modelos série-paralelo “descobrem” as regularidades ou periodicidades dominantes nas séries.

As redes treinadas através da arquitetura série-paralelo e convertidas em paralelo para a produção apresentaram resultados bastante promissores. A aumento da qualidade das variáveis pode melhorar sensivelmente este tipo de abordagem.

Outro aspecto importante como resultado das análises foi a investigação sobre a influência das variáveis externas. Tanto os dados de laboratório quanto os dados reais obtidos no processo de monitoração apresentaram o mesmo comportamento. Pode ser que esse problema tenha essa particularidade, ou seja, a resposta independe das variáveis externas. A resposta tem informação (memória) significativa para predizer seu próprio comportamento futuro.

## CAPÍTULO VII – CONCLUSÕES E RECOMENDAÇÕES

### 7.1 – Conclusões

Foi desenvolvido um modelo baseado em dados, empregando-se a técnica de redes neurais, objetivando a reprodução do comportamento histerético da junta de topo elastomérica, instalada no *riser* rígido de importação de gás da plataforma P18 da Petrobras.

A motivação do trabalho foi buscar um melhor conhecimento do comportamento da referida junta para verificar se a mesma estaria sofrendo alterações de suas características originais, com reflexos nos esforços do *riser* na região do topo.

O trabalho baseou-se em medições realizadas no *riser*, entre os anos de 1998 e 2002, com o objetivo de obter validações de metodologia e premissas de projeto do *riser*. Ou seja, na época do planejamento do sistema de monitoração, não existia a intenção de que os dados fossem utilizadas futuramente para geração de um sistema de modelagem empírica das propriedades físicas do equipamento.

Uma situação ideal utilizaria ensaios em laboratório, sob condições controladas, que propiciariam uma calibração inicial do modelo. E permitiriam filtrar possíveis respostas espúrias ou mesmo erros dos equipamentos de campo, quando do ajuste do modelo frente aos dados reais. Este procedimento serviria inclusive como uma forma de auxiliar no processo de crítica aos dados de campo obtidos.

Apesar das dificuldades, da elevada quantidade de dados e da complexidade envolvida no problema, foi empreendida uma iniciativa pioneira na área de projeto e desenvolvimento de *risers*, procurando caracterizar o comportamento da junta de topo, investigando-se as várias possíveis correlações com parâmetros que não são considerados no projeto tradicional do equipamento.

A caracterização da independência do fenômeno em relação à frequência de excitação foi sugerida pelos resultados das análises realizadas, o que permitiu inferir que o elastômero instalado trabalha em níveis de deformação relativamente contidos. Esta informação é relevante na medida em que altos níveis de deformação poderiam precipitar um processo de degradação do elastômero (BURTSCHER *et al.*, 1998)

Um outro resultado relevante sobre o comportamento da junta mostra que não se pode avaliar a dependência do comportamento da mesma em relação às flutuações de tração. No projeto do *riser* é assumida a hipótese de independência.

A investigação realizada abordou a questão das altas frequências, ou seja, a faixa de resposta dinâmica associada a vibrações induzidas por vórtices. A resposta do modelo, eliminando-se as altas frequências, acima de 0,25 Hz, apresentou um ajuste sensivelmente melhor.

Ao se tentar investigar as causas junto aos responsáveis pela monitoração, concluiu-se que o equipamento utilizado para medição das inclinações de topo, não era capaz de capturar adequadamente altas frequências. A partir desta conclusão, uma nova especificação para os futuros equipamentos de medição de inclinações já encontra-se em elaboração.

Em função dos referidos problemas, permanece a dúvida se o padrão de resposta da junta é alterado pelas vibrações de alta frequência do *riser*.

O modelo proposto, além de trazer informações relevantes para o projeto de risers e de sistemas de monitoração deverá ser utilizado em futuros trabalhos de monitoração do topo de risers rígidos.



## 7.2 - Recomendações para trabalhos futuros

Uma primeira sugestão é que sejam registrados os valores de temperatura interna do fluido e pressão de operação. No caso do *riser* analisado dispunha-se destas informações em termos de valores médios, porém não havia registros temporais disponíveis que permitissem um tratamento rigoroso para investigação da influência destas variáveis.

As possibilidades de utilização do modelo podem ser associadas à monitoração da região de topo de risers ou às simulações realizadas dentro do projeto.

Na monitoração, uma das possibilidades é que o modelo seja instalado e passe por um período de aprendizado, no início da operação de uma junta nova de topo. Após a fase de aprendizado, ele passaria a funcionar como um preditor cujos resultados seriam comparados com os medidos. Os desvios entre resultados simulados e medidos, seriam considerados para avaliar se estaria havendo alteração das características do equipamento do equipamento. A fase de aprendizado mencionada poderia até mesmo incluir a etapa de testes de aceitação na fábrica.

Uma outra possibilidade é o emprego de 2 modelos: um para o aprendizado na fase inicial de operação do equipamento, conforme mencionado acima, e um outro adaptativo mantido em constante aprendizado. A avaliação da alteração no equipamento seria então feita através da comparação entre os modelos. O modelo adaptativo poderia também ser substituído por um modelo que seria treinado num período diferente do original, supondo-se as mesmas condições operacionais.

A rede neural, uma vez treinada para um determinado modelo de junta de topo, poderia ser acoplada ao simulador do comportamento global do *riser*, possibilitando assim uma representação mais realística do comportamento da junta no modelo utilizado em projeto.

## Referências bibliográficas

AGUIRRE, L. A. (2004). **Introdução à Identificação de Sistemas: Técnicas Lineares e Não-Lineares Aplicadas a Sistemas Reais**, 2 edn, Editora UFMG, Belo Horizonte, BH.

ALMEIDA, LUIZ ALBERTO L. DE; GURDIP S. DEEP; ANTONIO MARCUS N. LIMA; HELMUT NEFF, **Um modelo diferencial para histerese magnética: representação algébrica recursiva**, Sba Controle & Automaçãov.14 n.1 Campinas mar. 2003

AMARI, S. N. MURATA, K.-R. MÜLLER, M. FINKE, AND H. YANG, **Statistical theory of overtraining—Is cross-validation effective?**, in NIPS'95: Advances in Neural Information Processing Systems 8, D. S. Touretzky, M. C. Mozer, and M. E. Hasselmo, Eds. Cambridge, MA: MIT Press, 1996.

BANKS, H.T., GABRIELLA A. PINTER AND LAURA K. POTTER, **Modeling of Nonlinear Hysteresis in Elastomers..** Center for Research in Scientific Computation, 1999.

BASHEER A., P.E., M.ASCE, **Stress-Strain Behavior of Geomaterials in Loading Reversal Simulated by Time-Delay Neural Networks**, ASCE - Journal of Materials in Civil Engineering, Vol.14 No.3 MAY-JUNE 2002

BERGSTRÖM, J. S. and Boyce, M. C. **Constitutive Modeling Of The Large Strain Time-Dependent Behavior Of Elastomers**, Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, USA. The final version of this draft working paper is published in: J. Mech. Phys. Solids., Vol. 46, pp. 931–954, 1998.

BRAGA, ANTÔNIO DE P., CARVALHO, ANDRÉ P. DE L, LUDERMIR, TERESA B. / **Fundamentos de Redes Neurais Artificiais**; 11º Escola de Computação - Rio de Janeiro. DCC/IM, COOPE/Sistemas, NCE/UFRJ, 1999

BURTSCHER, S., DORFMAN, A., BERGMEISTER, K., 1998. **Mechanical aspects of high damping rubber**. 2nd. Int. PhD Symposium in Civil Engineering, Budapest.

CALÔBA, L. P. Livro de Minicursos do XIV Congresso Brasileiro de Automática, Natal, 2002. **Introdução ao Uso de Redes Neurais na Modelagem de Sistemas Dinâmicos e Séries Temporais**. COPPE & EP – UFRJ (caloba@ufrj.br).

CAMPELLO, George Carneiro. **Avaliação da vida residual de risers de perfuração corroídos**. Rio de Janeiro: março de 2007

CHAKRABARTI, S. K., **Hydrodynamics of Offshore Structures**. Berlin, Springer-Verlang, 1987.

CHEN, S., BILLINGS, S.A.. **NEURAL NETWORKS FOR NONLINEAR DYNAMIC SYSTEM MODELING AND IDENTIFICATION**. International Journal of Control 56 319 - 346, 1992..

CHEW, BRUCE, “**DYNAMIC DATA PREDICT RESPONSE OF ELASTOMERIC ISOLATORS**”, E-A-R Specialty Composites Indianapolis, Indiana, www.earsc.com ,2003

COELHO, J.; DUTRA, R M; SPERANDIO, Maurício. “**O Método Ward de Agrupamento de Dados e sua aplicação em Associações com os Mapas Auto-Organizáveis de Kohonen**”. In: WorkComp-Sul, 2004, Florianópolis/SC. WorkComp-Sul, 2004.

COSTA NETO, Pedro Luiz de Oliveira Estatística. Edição: 2 Assunto: **Estatística - Probabilidade** - ISBN: 85-212-0097-8 Brochura, Editora: Edgard Blücher,1977

DALLAGNOL FILHO, Valdemar. **Identificação de modelos ARMAX e NARMAX para um poço de petróleo operando por injeção contínua de gás**. 2005. 121 f. Dissertação (Mestrado em Engenharia Elétrica) - Universidade Federal de Santa Catarina, Conselho Nacional de Desenvolvimento Científico e Tecnológico. Orientador: Daniel Juan Pagano.

DE JESÚS, O., AND M.T. HAGAN, **Backpropagation Through Time for a General Class of Recurrent Network**, Proceedings of the International Joint Conference on Neural Networks, Washington, DC, July 15-19, 2001, pp.2638-2642.

DE JESÚS, O., AND M.T. HAGAN , **Forward Perturbation Algorithm for a General Class of Recurrent Network**, Proceedings of the International Joint Conference on Neural Networks, Washington, DC, July 15-19, 2001, pp. 2626-2631.

ELMAN, J. L. (1990), **Finding structure in time**. Cognitive Science, 14(2):179-211.

FRANCISS, R. AND RIBEIRO, E. (2004). **Analyses of a large diameter steel lazy wave riser for ultra deepwater in campos basin**. Proc. OMAE 51176.

G. W. ELLIS, C. YAO, R. ZHAO, AND D. PENUMADU, "**Stress-strain modeling of sands using artificial neural networks**." International Journal of Rock Mechanics and Mining Sciences and Geomechanics Abstracts, Volume 33, Number 3, April 1996 , pp. 115A-115A(1)

GONZALEZ, E.C., MOURELLE, M.M., LIMA, J.M.T.G. AND MOREIRA, C.C.(2005). **Steel catenary riser design and analysis for Petrobras Roncador field development**, Proc. OTC 17670.

HARRIS, J. et al., **Relatório da MERL. Flexible Joint Integrity Assessment – Phase 1**. 20 de março de 2006.

HATTON, S., WILLIS, N. “**Steel Catenary Risers for Deepwater Environments**” OTC8607, 1998

HAYKIN, SIMON / **Redes Neurais - Princípios e pratica**; 2º edição; editora Bookman,2004

HIROSE, Y., YAMASHITA, K., HIJIYA, S. **Back-Propagation algorithm which varies the number of hidden units**. Neural Networks, v. 4, p. 61-66, 1991

JYH-DA WEI; CHUEN-TSAI SUN, **Large simulation of hysteresis systems using a piecewise polynomial function**, Signal Processing Letters, IEEE, Volume 9, Issue 7, Jul 2002 Page(s): 207 – 210

J. YI, M.C. BOYCE, G.F. LEEA AND E. BALIZER, **Large deformation rate-dependent stress–strain behavior of polyurea and polyurethanes**, *Polymer* Volume 47, Issue 1, 3 January 2006, Pages 319-329

K.P. SUDHEER AND S.K. JAIN, **Radial basis function neural network for modeling rating curves**, *ASCE, J. Hydrologic Eng.* **8** (2003) (3), pp. 161–164

LANG, K.J., AND G.A.E. HINTON, 1988. "**The development of the time-delay neural network architecture for speech recognition**," Technical Report CMU-CS-88-152. Carnegie Mellon University, Pittsburgh, PA.

LJUNG, L. (1987). **System Identification: Theory for the User**, Prentice-Hall, Englewood Cliffs, NJ.

LIMA, CLODOALDO, "**Emprego de Teoria de Agentes no Desenvolvimento de Dispositivos Neurocomputacionais Híbridos e Aplicação ao Controle e Identificação de Sistemas Dinâmicos**" Faculdade de Engenharia Elétrica e de Computação (FEEC) Universidade Estadual de Campinas (Unicamp), SP, Brasil, 2000.

MOURELLE, M. M.; GONZALEZ, E. C.; JACOB, B. T. Anflex: **Computational system for flexible and rigid risers analysis**. In: Proceedings of the 9th International Symposium on Offshore Engineering. Rio de Janeiro, RJ, Brasil, 1995.

MOURELLE, MÁRCIO MARTINS, SENRA, STAEL FERREIRA, TORRES, ANA LÚCIA FERNANDES LIMA. **Simulações de Situações Operacionais para o SCR da P-XVIII**. Rio de Janeiro, CENPES/DIPREX/SEDEM, 2000.

MACHADO, R.Z.; MOURELLE, M.M.; FRANCISS, R.; SILVA, R.M.; LIMA, C.S.; EISEMBERG, R.; OLIVEIRA, D. ,"**Monitoring Program for the First Steel Catenary Riser Installed in a Moored Floating Platform in Deep Water**", OCEANS apos;99 MTS/IEEE. Riding the Crest into the 21st Century Volume 2, Issue , 1999 Page(s):801 - 810 vol.2

MAKRIDAKIS, S., ANDERSON, A., CARBONE, R., FILDES, R., HIBDON, M., AND LEWANDOWSKI, R.(1982). **The accuracy of extrapolation (time series) methods: Results of a forecasting competition.** Journal of Forecasting 1(2), pgs. 111-153.

MANDIC, D.P. AND CHAMBERS, J.A. (2001): **Recurrent Neural Networks for Prediction: Architectures, Learning Algorithms and Stability.** John Wiley & Sons

MEKHA, B.B., **New Frontiers In The Design Of Steel Catenary Risers For Floating Production Systems,** International Conference on Offshore Mechanics and Arctic Engineering 20th Annual (2001) (OMAE001/OFT-1200)

NARENDRA, K. & PARTHASARATHY, K. **Identification and Control of Dynamical Systems using Neural Networks,** IEEE Transactions on Neural Networks 1(1): 4-27, (1990).

NERRAND, O., ROUSSEL-GAGOT, P., URBANI, D., PERSONNAZ, L., AND DREYFUS, G. (1994). **Training recurrent neural networks: Why and how? an illustration in dynamical process modeling.** *IEEE Transactions on Neural Networks,* 5(2):178-184.

NEWLAND, D. E., **Introduction to Random Vibrations and Spectral Analysis.** London, Longman Group, 1980.

OIL STATES INDUSTRIES, INC. **Factory Acceptance Test Specification for Petrobras 10” Catenary Riser Flexjoint.** Texas, 1996. Report.

PARE, T. E., & HOW, J. P. (1998). **Robust stability and performance analysis of systems with hysteresis nonlinearities.** In Proceedings of American control conference (pp. 1904-1908).

RELATÓRIO TÉCNICO RL-3534.00-6520-274-CUF-001. Projeto INVENTÁRIO DO SISTEMA DE MONITORAÇÃO. **Sistema RISER RÍGIDO EM CATENÁRIA (SCR)/FLEX JOINT DO GASODUTO P-26/P-18 NA UNIDADE MARÍTIMA P-18.**

ROVERI, F.E. AND PESSOA, P.R.F., (2005). **Free standing hybrid riser for 1800 m water depth**. Proc OMAE 67178.

ROWELL, DEREK, (2005), **State-Space Representation of LTI Systems**, Analysis and Design of Feedback Control Systems - Department of Mechanical Engineering

SARAIVA, J. M. F., **A utilização de redes neurais em conjunto com o Método de Monte Carlo na análise da confiabilidade de estruturas**, Rio de Janeiro: Programa de Pós-Graduação em Engenharia Civil, COPPE/UFRJ, 1997. (Tese de Doutorado).

SEIPONE, T., BULLINARIA, J. A., “Evolving Neural Networks That Suffer Minimal Catastrophic Forgetting”, PROGRESS IN NEURAL PROCESSING, 2005, VOL 16, pages 385-390

SERTA, M.M. MOURELLE ET AL, “Steel Catenary **Riser. for the Marlim Field FPS P-XVIII**”, Offshore Technology. Conf. Proc., 1996.

SUDHEER, K. P., JAIN, S. K., **Radial Basis Function Neural Network for Modeling Rating Curves**, J. Hydrologic Engrg., Volume 8, Issue 3, pp. 161-164 (May/June 2003)

SUN, M. , FUJIWARA ,H., **Data driven function approximation models for simulating magnetic hysteresis**, Journal of Applied Physics -- May 15, 2002 -- Volume 91, Issue 10, pp. 8325-8327

T. Z. TAN , C. QUEK, G. S. NG, “**Brain-inspired genetic complementary learning for stock market prediction**”, IEEE Congress of Evolutionary Computation, Vol. 3, Edinburgh, Scotland, 2-5 September 2005, pp. 2653-2660.

TANAKA, R.L., MARTINS, C.A., TAKAFUJI, F.C.M., SILVEIRA, L.M.Y. AND BALENA, R. (2005). **Dynamic analysis of lazy-wave steel risers**. Proc. COBEM 1848.

TAWFIK M, IBRAHIM A, FAHMY H. **Hysteresis sensitive neural network for modeling rating curve**. ASCE J Comput Civil Engng 1997;11(3): 206-11.

VOJTKO, J. ,**Neural Network, Component of Measuring Set for Error Reduction**, Measurement Science Review, Volume 4, Section 1, 2004

VON ZUBEN, F.J, “**Modelos Paramétricos e Não-Paramétricos de Redes Neurais Artificiais e Aplicações**” , Faculdade de Engenharia Elétrica e de Computação (FEEC) Universidade Estadual de Campinas (Unicamp), SP, Brasil, 1996.

YI, WAON-HO, HACK-SOO LEE, SEUNG-CHANG LEE, HO-SUNG KIM, **Development Of Artificial Neural Networks Based Hysteretic Model**, Fourteenth Engineering Mechanics Conference American Society of Civil Engineers May 21-24, 2000.

YI, J., BOYCE, M.C., BALIZER, E., LEE, G., "**Large Deformation Rate-Dependent Stress-Strain Behavior of Polyurea and Polyurethane**", Polymer, 47, 319-329, 2005.

WAIBEL, A., T. HANAZAWA, G. HINTON, K. SHIKANO, AND K.J. LANG, 1989. "**Phoneme recognition using time-delay neural networks.**" IEEE Transactions on Acoustics, Speech, and Signal Processing ASSP-37, 328-339.

WEI, J.D., SUN, C.T., “**Constructing Hysteretic Memory in Neural Networks**”, IEEE Transactions on Systems, Man and Cybernetics - Part B: Cybernetics, Vol.30, No.4, pp. 601-609, 2000.

WERBOS, P. J. (1990). **Backpropagation through time: what it does and how to do it.** *Proc. IEEE*, 78(10):1550-1560.



# Livros Grátis

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

Milhares de Livros para Download:

[Baixar livros de Administração](#)

[Baixar livros de Agronomia](#)

[Baixar livros de Arquitetura](#)

[Baixar livros de Artes](#)

[Baixar livros de Astronomia](#)

[Baixar livros de Biologia Geral](#)

[Baixar livros de Ciência da Computação](#)

[Baixar livros de Ciência da Informação](#)

[Baixar livros de Ciência Política](#)

[Baixar livros de Ciências da Saúde](#)

[Baixar livros de Comunicação](#)

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

[Baixar livros de Defesa civil](#)

[Baixar livros de Direito](#)

[Baixar livros de Direitos humanos](#)

[Baixar livros de Economia](#)

[Baixar livros de Economia Doméstica](#)

[Baixar livros de Educação](#)

[Baixar livros de Educação - Trânsito](#)

[Baixar livros de Educação Física](#)

[Baixar livros de Engenharia Aeroespacial](#)

[Baixar livros de Farmácia](#)

[Baixar livros de Filosofia](#)

[Baixar livros de Física](#)

[Baixar livros de Geociências](#)

[Baixar livros de Geografia](#)

[Baixar livros de História](#)

[Baixar livros de Línguas](#)

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