

Universidade de São Paulo
Instituto de Física

Previsão das Instabilidades de Disruptura
através de Redes Neurais Artificiais

Kenya Andréia de Oliveira
Orientador: Dr. Alvaro Vannucci

SBI-IFUSP



305M810T3314

Tese apresentada ao Instituto de Física da Universidade de São Paulo para a obtenção do título de Doutora em Ciências.

Comissão examinadora:

- Prof. Dr. Álvaro Vannucci - USP
- Prof. Dr. Iberê Luiz Caldas - USP
- Prof. Dr. Ivan Cunha Nascimento - USP
- Prof. Dr. Luiz Pereira Calôba - UFRJ
- Prof. Dr. Munemassa Machida - UNICAMP

Alvaro Vannucci



delino 2002

Prof. Armando Orlando Ferraz
Presidente da Comissão de Pós Graduação

São Paulo
2000

006, 3
048 p
D
ex. 1

FICHA CATALOGRÁFICA

Preparada pelo Serviço de Biblioteca e Informação
do Instituto de Física da Universidade de São Paulo

Oliveira, Kenya Andréia de

Previsão das Instabilidades de Disrupturas através de
Redes Neurais Artificiais. São Paulo, 2000.

Tese (Doutoramento) - Universidade de São Paulo.
Instituto de Física - Departamento de Física Aplicada

Orientador: Prof. Dr. Alvaro Vannucci
Área de Concentração: Física Nuclear

Unitermos: 1. Tokamak; 2. Plasmas;
3. Fusão Termonuclear Controlada;
4. Instabilidades de Disruptura;
5. Redes Neurais.

USP/IF/SBI-007/2000

*Ao meu pai (saudades ...),
À minha mãe, exemplo de vida,
Ao Elton, com todo o meu amor e gratidão.*

Agradecimentos

Impossível realizar qualquer trabalho sério, no meio científico, sem a colaboração e o companheirismo dos profissionais do dia-a-dia. À estes, e ao meu orientador, mestre e amigo de todos os momentos, os meus mais sinceros agradecimentos.

Não citarei nomes, em específico, para não cometer a injustiça de me esquecer eventualmente de alguém. Fica aqui, portanto, o meu grande carinho ao grupo de plasma, do qual tive o privilégio de participar durante tantos anos, não só pelo ótimo ambiente de trabalho, como também pela excelente oportunidade de convívio, o que contribuiu, sobremaneira, para a minha formação pessoal, inclusive.

À minha querida família, e amigos, agradeço o incentivo e, sobretudo, o respeito pela minha opção de vida, nem sempre muito compreendida.

À todos os demais, que de alguma forma contribuíram, direta ou indiretamente, para a realização deste trabalho.

À CAPES, pelo apoio financeiro.

Muita paz, saúde e justiça a todos.

Conteúdo

Lista de figuras	iv
Lista de tabelas	vi
1 Introdução	1
2 A Fusão Termonuclear Controlada	4
2.1 Princípios Básicos de Funcionamento do Tokamak	5
2.2 Instabilidades Importantes	9
2.2.1 Teoria MHD: modos de Dobra e de Troca	9
2.2.2 Modos de Ruptura	12
2.2.3 Instabilidades de Disruptura	12
2.3 Diagnósticos Relevantes	12
2.3.1 Bobinas magnéticas de Mirnov	13
2.3.2 Raios-X de baixas energias	14
3 Redes Neurais Artificiais	19
3.1 Princípios da Neurocomputação	21
3.2 O Funcionamento do Sistema Nervoso Biológico	22
3.2.1 Estrutura de um neurônio	22
3.2.2 Transmissão da informação	23
3.2.3 Aprendizado e plasticidade	24
3.3 Fundamentos das Redes Neurais Artificiais	25
3.3.1 O neurônio artificial	25
3.3.2 Rede artificial típica	28
3.3.3 Algoritmo “ <i>Error-Backpropagation</i> ”	30
3.4 Implementação Numérica	34
3.4.1 “Biasing”	36
3.4.2 Programas em uso	36
4 Previsão de Séries Caóticas com Redes Neurais	38
4.1 Reconstruindo Estados com Coordenadas de Atraso	39
4.2 O Conceito de Imersão (“ <i>Embedology</i> ”)	40
4.3 Sistemas Testados	41
4.3.1 Atrator de Lorenz	41
4.3.2 Mapa de Henon	44
4.3.3 Mapa Logístico	49

5	Redes Neurais na Previsão das Disrupturas	52
5.1	Processo de Previsão Contínua	53
5.1.1	Arquitetura 15:9:3:1	56
5.1.2	Prevendo a descarga inteira	62
5.1.3	Buscando novas arquiteturas	64
5.2	Arquitetura 15:9:3:1, com Previsão por <i>STEP</i>	70
5.3	Arquitetura final 15:30:15:1	72
6	Sistema de Detecção de Raios-X (Situação Atual)	78
6.1	Câmaras de Raios-X	78
6.2	Detectores Utilizados	79
7	Tomografia de Plasmas por Raios-X	81
7.1	Reconstrução Tomográfica	81
7.2	Técnicas de Reconstrução	83
7.2.1	Polinômios de Zernicke	86
7.2.2	As Funções de Bessel	87
7.3	Implementação Numérica	90
7.3.1	Coefficientes de Zernicke e de Bessel	91
7.3.2	Programa TOMOG	92
8	Conclusões	98
	Bibliografia	101

Lista de Figuras

2-1	Figura esquemática de um tokamak	6
2-2	Desenho ilustrativo do tokamak TCABR, do IFUSP	7
2-3	Método dos Absorvedores	17
3-1	Estrutura típica de um neurônio biológico	23
3-2	Mapa cortical humano	25
3-3	Analogia entre neurônios naturais e neurônios artificiais	26
3-4	Transmissão da informação num neurônio artificial	27
3-5	Funções de ativação típicas	28
3-6	Símbolo universal de um neurônio artificial	29
3-7	Desenho ilustrativo de uma rede neural artificial	30
4-1	Séries de Lorenz utilizadas na rede neural artificial.	43
4-2	Previsão da série de Lorenz	45
4-3	Parâmetros VRM e ρ para o sistema de Lorenz	46
4-4	Diferentes arquiteturas para a previsão da série de Lorenz	46
4-5	Série, em y , do mapa de Henon	47
4-6	Previsão da rede artificial para o mapa de Henon.	48
4-7	Parâmetro VRM em função das épocas utilizadas.	50
4-8	Previsão da rede artificial para mapas Logísticos.	51
5-1	Pulso do tokamak TEXT: 133657	53
5-2	Pulsos do tokamak TEXT: 133660 e 133665	54
5-3	Pulsos do tokamak TEXT: 133673 e 133677	55
5-4	Previsão da rede para o pulso 133665 (TEXT)	58
5-5	Previsão da rede para o pulso 133660 (TEXT)	59
5-6	Sinais de raios-X e de bobinas magnéticas do pulso 133665 (TEXT)	61
5-7	Sinais de raios-X e de bobinas magnéticas do pulso 133660 (TEXT)	61
5-8	Previsão do pulso rx133677, inteiro	63
5-9	Detalhe da previsão do pulso rx133677	64
5-10	Previsão da rede $15:9:3:1$ para uma função teste.	66
5-11	Comparando arquiteturas diferentes	67
5-12	Previsão da rede $33:17:3:1$ para uma função teste.	68
5-13	Arquiteturas $33:17:3:1$ e $150:50:10:1$	69
5-14	Arquitetura $15:9:3:1$ com previsão por $STEP$	70
5-15	Previsão do pulso rx133677 inteiro	71
5-16	Detalhe da previsão do pulso rx133677	72
5-17	Previsão do pulso rx133660, inteiro, com a rede $15:30:15:1$	74

5-18	Previsão do pulso rx133677, inteiro, com a rede 15:30:15:1	75
5-19	Detalhe da previsão da rede 15:30:15:1 para o pulso rx133660, inteiro	76
5-20	Detalhe da previsão da rede 15:30:15:1 para o pulso rx133677, inteiro	77
6-1	Sistema de detecção de raios-X do TCABR	80
7-1	Esquema do sistema de detecção de raios-X para reconstrução tomográfica	82
7-2	Ângulo sólido de um detector	83
7-3	Linha de visualização de cada detector	84
7-4	Coordenadas p, ϕ dos detectores utilizados	93
7-5	Gráfico das brilhanças calculadas pelo TOMOG	94
7-6	Emissividades verticais e horizontais	95
7-7	Harmônicos de Fourier da emissividade obtida.	96
7-8	Imagem 3D da emissividade obtida.	96
7-9	Curvas de nível da emissividade obtida.	97

Lista de Tabelas

2-1	Parâmetros de máquina de diversos tokamaks	8
4-1	Diferentes arquiteturas para previsões do mapa de Henon	49

RESUMO

Redes neurais artificiais, tipo “*feedforward*”, de duas camadas, foram utilizadas neste trabalho para fazer previsões das instabilidades de disruptura que surgem nas descargas de plasma do tokamak TEXT (E.U.A.), obtendo-se resultados bastante encorajadores.

Verificou-se que uma arquitetura de rede, do tipo $m:2m:m:1$, onde m é dimensão de imersão do atrator do sistema dinâmico em estudo, costuma ser um bom *chute* inicial para a escolha da arquitetura ideal de trabalho, que costuma ser livre e, não raro, trabalhosa.

Utilizando-se, em sinais de raios-X, uma rede neural artificial com arquitetura $15:30:15:1$, por exemplo, conseguiu-se fazer previsões com uma antecipação de até 4 *ms* das instabilidades de disruptura, tempo quatro vezes maior do que o obtido utilizando-se sinais magnéticos das bobinas de Mirnov.

Tal antecipação é bastante significativa e abre a possibilidade de, no futuro, utilizarem-se mecanismos de defesa da máquina, tais como injeção de partículas neutras (ou “*pellets*”), aplicação de campos magnéticos externos, etc, no sentido de se tentar evitar a ocorrência destas instabilidades, ou, pelo menos, minimizar os seus efeitos nocivos. Isto certamente contribuirá significativamente para a viabilização dos futuros reatores de fusão à plasma.

Finalmente, o sistema de diagnóstico de raios-X de baixas energias do tokamak TCABR, que foi projetado, e já se encontra em fase de instalação, para fornecer sinais que servirão para alimentar a rede neural, também possibilitará a reconstrução tomográfica das regiões de mesma emissividade da coluna de plasma. A análise tomográfica, utilizando-se os sinais de dois conjuntos de detectores de raios-X moles, também será muito útil na investigação dos mecanismos físicos que dão surgimento às instabilidades de disruptura, além de permitir, ainda, a medida da temperatura eletrônica do plasma, através do método dos absorvedores.

ABSTRACT

Two-layer feedforward neural network has been used in this work to forecast the disruptive instabilities that occur in the TEXT tokamak plasma discharges. For this task, soft X-ray experimental signals were used with very promising results.

It was verified that a neural net with an architecture of the type $m:2m:m:1$, where m is the embedding dimension of the attractor of dynamical system in focus, is usually a good initial guess in the searching process of finding the ideal architecture.

A neural network with architecture $15:30:15:1$ was capable, for example, to forecast the disruptive instabilities up to 4 ms in advance. This period of time is four times larger than the one obtained when magnetic signals from Mirnov coils were used.

This forecasting time is quite significant and opens up the possibility of using defensive mechanisms, such as the injection of neutral particles (or pellets), the application of external magnetic fields, etc, with the objective of avoiding the occurrence of the disruptions or, at least, to minimize their harmful effects. This achievement certainly would be an important contribution to the development of the next generation fusion devices.

Finally, the soft X-ray diagnostic system for the TCABR was projected and it is already being installed. This system will provide experimental signals that will be analysed by neural networks and will be also used to identify, through tomographic image reconstructions, the regions of the plasma that have the same soft X-ray emissivity. The tomography analysis of the plasma, that will be carried out by using the signals of two soft X-ray detectors arrays, will be also very useful for investigating the triggering mechanism of disruptions and will also allow the determination of the plasma electron temperature through the two foil absorbing method.

Capítulo 1

Introdução

A fusão termonuclear controlada pode significar a esperança de um suprimento ilimitado de energia com o qual, num futuro não muito distante, qualquer nação possa contar. Nesse sentido, vários dispositivos de confinamento de plasmas a altas temperaturas, onde ocorreriam as reações nucleares de fusão, têm sido estudados e o mais promissor de todos ainda parece ser o tokamak [1, 2, 3]. Mas, mesmo nesse tipo de máquina, surgem muitos processos físicos, durante o processo de confinamento do plasma, que interferem no seu equilíbrio e que precisam de ser controlados.

Muitos desses fenômenos, como as instabilidades de disruptura, por exemplo, são decorrentes de perturbações nos parâmetros macroscópicos do plasma, propagam-se através de ondas e são usualmente estudados através das equações da magnetohidrodinâmica (modelo MHD), que descreve o plasma como um sistema de grande escala, não se importando com os aspectos microscópicos do meio [4, 5, 6].

De uma forma geral, as instabilidades de disruptura são consideradas, ainda, verdadeiros entraves para a operação dos futuros reatores de fusão, uma vez que elas podem acarretar a perda do confinamento do plasma e provocar, conseqüentemente, a indução de altas tensões elétricas nos diversos componentes eletro-mecânicos do tokamak, podendo causar-lhe sérios danos [6, 7]. Em máquinas de grande porte, como o JET, o TFTR e o JT-60U, forças de até 2 MN e cargas térmicas da ordem de $2\text{ MJ}/\text{m}^2$ têm causado deformações e danos nas estruturas que entram em contato direto com o plasma, quando uma disruptura ocorre [8]. E espera-se que este panorama seja ainda pior para os futuros reatores de fusão, de gigantesco porte, como o ITER, no qual a magnitude das forças e cargas térmicas deverá ser duas ordens de grandeza maior [9, 10]!

Apesar dos vários esforços na tentativa de se compreender os mecanismos que desen-

cadeiam os fenômenos disruptivos, ainda não se chegou a um senso comum sobre como são formados e como controlá-los, ou evitá-los. Diante do exposto, e dada a relevância do problema, surgiu a idéia de se tentar prever, ainda que não exatamente, o instante em que as disrupturas ocorreriam para que se conseguisse criar possíveis mecanismos de controle que talvez permitissem alterar as condições de confinamento do plasma em tempo hábil para evitar ou, pelo menos, atenuar os efeitos nocivos dessas perigosas instabilidades.

Com esse intuito, escolheu-se uma técnica relativamente recente, para aplicações na área da física de plasmas, que se constitui na utilização de redes neurais artificiais com o objetivo de se tentar fazer previsões dessas instabilidades de disrupturas através da utilização dos sinais de raios-X de baixas energias emitidos pelo plasma [11].

Inicialmente, o projeto de pesquisa estava voltado para se fazer esse estudo utilizando-se dados do tokamak TCABR, do Instituto de Física, da USP. Mas, por diversos problemas, inclusive financeiros, atrasou-se a construção e operação do mesmo, o que forçou a utilização de dados experimentais de um outro tokamak, TEXT, gentilmente cedidos pelo grupo de plasma da Universidade do Texas, em Austin (E.U.A.).

As bases de fundamentação das redes neurais artificiais estão na neurotransmissão que ocorre no cérebro, onde se procura fazer uma analogia entre células nervosas vivas e o algoritmo computacional utilizado, enfatizando-se o aprendizado como forma de captação de conhecimento. Suas aplicações têm sido as mais diversas possíveis, nos vários segmentos da ciência, e com resultados bastante encorajadores [12]. Não obstante, um árduo trabalho é necessário para se encontrar as condições ideais do sistema neural artificial adotado para que ele realize, com sucesso, a tarefa desejada.

Normalmente, a forma de se estruturar a arquitetura de uma rede costuma ser livre, o que aumenta as variáveis a serem ajustadas para a otimização do sistema, tornando a sua escolha ainda mais difícil [13]. No presente trabalho, após um considerável esforço, verificou-se que uma arquitetura inicial, do tipo $m:2m:m:1$, onde m é a dimensão de imersão do atrator do sistema dinâmico em estudo, costuma ser um bom *chute* inicial, pois a rede, depois de otimizada, costuma apresentar uma estrutura final não muito diferente daquela escolha primeira [14].

Essa abordagem permitiu obter resultados de previsão bastante satisfatórios, se comparados com outros trabalhos realizados com previsões de séries temporais. Utilizando-se uma série temporal caótica do sistema de equações de Lorenz, por exemplo, conseguiu-se realizar previsões de até 25 *steps* (iterações da rede artificial) de antecedência, enquanto

outros autores não passaram de 3 ou 10 *steps* [15, 16], resultados estes que serão discutidos nos capítulos seguintes.

Para a utilização de redes neurais artificiais na previsão de disrupturas no tokamak TCABR, projetou-se e construiu-se um diagnóstico de raios-X de baixas energias, emitidos pelo plasma, que ainda está em fase de implementação.

Dos vários sistemas de diagnósticos utilizados em tokamaks, o que mede a emissão dos raios-X de baixas energias revela-se uma ferramenta valiosa no estudo de plasmas magneticamente confinados. Isso porque a magnitude e o espectro do fluxo irradiado estão diretamente relacionados com as características básicas (composição química, presença de impurezas, densidade e temperatura eletrônica) do plasma. Além disso, o plasma é opticamente fino para suas próprias emissões de raios-X, o que faz com que a radiação de qualquer parte do mesmo alcance um detector externo essencialmente inalterada. Consequentemente, a análise das emissões de raios-X de baixas energias proporcionam um método natural de estudo da evolução temporal e espacial dos muitos fenômenos macroscópicos e microscópicos observados durante uma descarga do plasma, auxiliando no estudo, tanto das características de equilíbrio, quanto dos fenômenos transientes [17].

Além disso, o projeto do diagnóstico de detecção de raios-X, no TCABR, foi elaborado de tal forma a possibilitar, ainda, a utilização da técnica da tomografia do plasma como uma outra forma de se estudar e analisar as instabilidades de disrupturas. O objetivo principal, neste caso, será investigar e monitorar as flutuações MHD relativas ao modo $m = 1$, uma vez que parece ser essa componente a principal responsável pela formação dos *dentes-de-serra* na região central da coluna de plasma [18]. Com essa técnica, abre-se também a possibilidade de se localizar as ilhas em $q = 1$ e $q = 2$, medir as suas respectivas larguras, além de, possivelmente, obter uma medida espacial da temperatura eletrônica do plasma, através do método dos absorvedores [19, 20].

Capítulo 2

A Fusão Termonuclear Controlada

A busca de fontes alternativas de energia se faz cada vez mais necessária, dado o crescente, e talvez alarmante, aumento do consumo mundial que se observa. Combustíveis fósseis vêm sendo historicamente utilizados para suprir parte dessas necessidades, mas suas fontes de extração, ainda que eventualmente grandes, são certamente finitas, o que apenas transfere o problema do suprimento de energia para um futuro, talvez, não muito distante. As usinas hidrelétricas também se constituem em fontes alternativas de energia, contudo, em ambos os casos, os impactos ambientais, diretos ou indiretos, advindos desses sistemas, são reconhecidamente desastrosos.

Dentro deste contexto, a pesquisa na área da física de plasmas para a obtenção de energia a partir da fusão termonuclear controlada se justifica uma vez que, além de ser bastante segura, já que não apresenta o perigo de escapar ao controle, é também relativamente *limpa*, se comparada aos processos tradicionais de obtenção de energia a partir da fissão nuclear, uma vez que a quantidade de resíduos radioativos liberados é significativamente menor. Some-se a isso o fato de as fontes de matéria prima para as reações de fusão (deutério e trítio) serem praticamente inesgotáveis: o deutério pode ser extraído abundantemente da água do mar, e o trítio, embora raro na natureza, pode ser obtido a partir de reações nucleares que envolvem o bombardeamento de átomos de Li^6 com neutrons oriundos do próprio processo de fusão deutério-trítio [1, 2].

Dos vários dispositivos de confinamento de plasmas a altas temperaturas já pesquisados, o mais promissor ainda parece ser o tokamak, cujos estudos vêm se desenvolvendo desde a década de 60. A obtenção de plasma em tokamaks é resultado da aplicação de campos elétricos e magnéticos específicos num meio gasoso, a uma determinada pressão, com o objetivo de mantê-lo a altas temperaturas por um tempo suficientemente longo de

forma a permitir o seu aquecimento, da ordem de 100 milhões de graus Celsius, para que os íons do plasma adquiram energia cinética suficientemente grande que lhes possibilite vencer a barreira de repulsão coulombiana e sofram, então, o processo de fusão.

Até o momento, somente o TFTR (*Tokamak Fusion Test Reactor* - E.U.A.), hoje desativado, e, mais recentemente, o JET (*Joint European Torus* - Comunidade Européia), foram operados com descargas onde se introduziu uma mistura mais ou menos equitativa de deutério e trítio, com resultados bastante encorajadores. No TFTR, conseguiu-se produzir potências de $\sim 10,7$ MW trabalhando-se com um campo magnético toroidal de $B_T = 5,6$ T e corrente de plasma $I_P = 2,7$ MA. O JET apresentou resultados ainda melhores, já tendo fornecido uma potência de fusão de 16 MW, correspondendo a uma razão de 65% entre a energia de fusão produzida e a energia total consumida para manter o tokamak em operação [21].

Um outro projeto, chamado ITER (*International Thermonuclear Experimental Reactor*) constitui-se em um empreendimento internacional de grandes proporções, envolvendo várias comunidades científicas, como a do Japão, Rússia e Comunidade Européia, com o objetivo de comprovar, de forma definitiva, a viabilidade da fusão termonuclear controlada para fins pacíficos de produção de energia. Atualmente este projeto se encontra em fase de detalhamento técnico, estando o local e o início de sua construção em fase de definição [9, 22].

2.1 Princípios Básicos de Funcionamento do Tokamak

A figura 2-1 mostra, esquematicamente, um dispositivo de confinamento de plasma do tipo tokamak. A corrente de plasma (I_p) é formada por intermédio da aplicação de um campo elétrico (E), na direção toroidal (ver figura), no interior da câmara de vácuo do vaso do tokamak. Esse campo elétrico é obtido a partir da variação temporal do fluxo magnético (Φ_M) induzido por uma corrente elétrica (I_{TAO}) que circula em espiras colocadas na parte central do toróide, constituindo o transformador de aquecimento ôhmico. Dessa forma, a coluna de plasma, em forma de anel, atua como se fosse o secundário de um transformador.

O campo magnético toroidal (B_T) é criado pela corrente (I_T) que circula pelas bobinas colocadas ao redor do vaso do tokamak. A própria corrente de plasma (I_P) gera, por sua vez, um outro campo magnético, chamado campo magnético poloidal (B_p) que, somando-se ao campo toroidal, cria linhas de campo magnético com simetria helicoidal. Essa

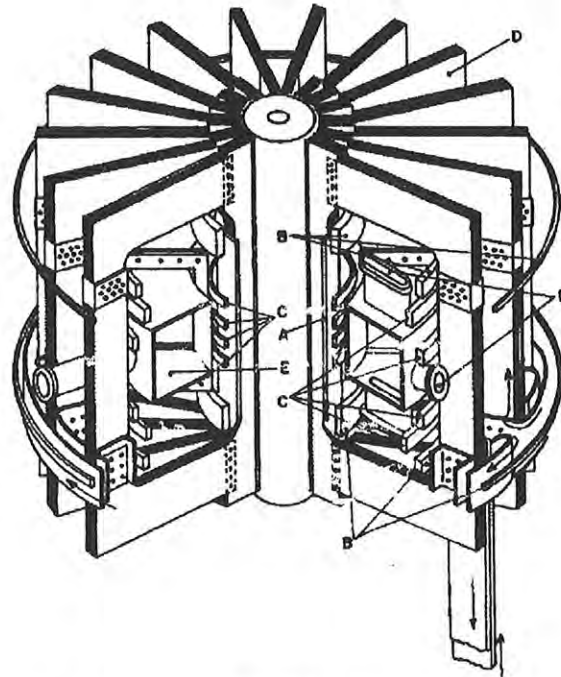


Figura 2-2: Desenho ilustrativo do tokamak TCABR, do Instituto de Física, da Universidade de São Paulo, onde A: bobinas para aquecimento ôhmico, B: bobinas auxiliares (6), para minimizar o campo de fuga da bobina A, C: bobinas do campo vertical (8), D: bobinas do campo toroidal, E: câmara de vácuo e F: janelas para diagnóstico. O TCABR se encontra em operação desde o mês de outubro de 1999.

estudo do aquecimento do plasma por ondas de Alfvén; o TEXT, tokamak da Universidade do Texas (E.U.A.), cujos dados experimentais foram utilizados, no presente trabalho, em estudos envolvendo a utilização de redes neurais artificiais; o TFTR, máquina de grande porte (E.U.A.) e a primeira a trabalhar com quantidades equitativas de deutério e trítio para as reações nucleares de fusão; o JET (Comunidade Européia), maior tokamak construído, e em operação, até o momento; e o ITER, tokamak de proporções gigantescas, a ser construído por comunidades científicas de vários países, nos próximos anos.

Os valores do raio maior (R_0 - raio do toróide) e do raio menor (a - raio da coluna de plasma), em tokamaks, estão geralmente relacionados com uma grande *razão de aspecto*: $R_0/a \gg 1$. Com esta condição, simplificações teóricas são possíveis, permitindo o estudo de importantes características do plasma magneticamente confinado.

Considerando-se, por exemplo, a seguinte correspondência de variáveis [4, 26, 27]:

$$(r, \theta, \phi) \longrightarrow (r, \theta, z = R_0\phi) \quad (2-1)$$

onde θ refere-se à direção poloidal e ϕ à direção toroidal do tokamak, podemos escrever a

Parâmetros de Máquina	TBR-1	TCABR	TEXT-U	TFTR	JET	ITER
Raio Maior: $R(m)$	0,30	0,61	1,05	2,60	2,96	8,14
Raio Menor: $a(m)$	0,08	0,18	0,27	1,96	2,10	2,80
Campo Magn. Toroidal: $B_\phi (T)$	0,4	1,2	3,0	6,0	3,8	5,7
Corrente de Plasma: $I_p(MA)$	0,012	0,12	0,4	3,0	7,0	21,0
Duração das Descargas: $\tau_D(s)$	0,012	0,10	0,50	5,0	60	1300

Tabela 2-1: Dados comparativos dos parâmetros de máquina dos tokamaks TBR-1 (Brasil), TCABR (Brasil), TEXT-U (E.U.A.), TFTR (E.U.A.), JET (Comunidade Européia) e ITER (Colaboração Internacional).

equação das linhas de campo da seguinte maneira:

$$\frac{r d\theta}{B_\theta} = \frac{R_0 d\phi}{B_\phi} = \frac{dl}{B}, \quad (2-2)$$

onde dl corresponde a um deslocamento infinitesimal ao longo da linha do campo magnético total, $\vec{B} = \vec{B}_\theta + \vec{B}_\phi$.

Seguindo-se uma linha de campo em seu percurso, ao redor do toróide, obtém-se:

$$d\phi = \frac{r B_\phi}{R_0 B_\theta} d\theta. \quad (2-3)$$

Definindo-se o fator de segurança, q , como sendo:

$$q = \frac{r B_\phi}{R_0 B_\theta}, \quad (2-4)$$

obtem-se:

$$d\phi = q(r) d\theta. \quad (2-5)$$

Uma linha do campo magnético \vec{B} pode ser percorrida infinitas vezes, ao redor do toróide, sem que se retorne ao ponto inicial de partida. Dessa forma, a essa linha de campo pode-se associar uma superfície magnética, caracterizada por um determinado valor de q . Outras linhas de campo magnético estarão associadas a outras superfícies magnéticas, apresentando o mesmo comportamento.

Duas intersecções sucessivas de uma linha de campo com um plano poloidal, referente a um determinado ϕ , após uma volta toroidal completa ($\Delta\phi = 2\pi$), determinam uma variação $2\pi/q$ na coordenada angular θ . Para o caso especial $q = m/n$, onde m e n são números inteiros e positivos, a linha de campo fecha-se sobre si mesma após n voltas na direção poloidal e m voltas na direção toroidal terem sido completadas.

No caso particular das *superfícies racionais*, para as quais $q = m/n$, as linhas de campo dão um número finito de voltas até atingirem o ponto inicial. Essa situação é

particularmente interessante porque é justamente nesta região onde se formam as *ilhas magnéticas*, que por sua vez desempenham um papel importante nos processos de formação das instabilidades de disruptura. Por essa razão, estas superfícies magnéticas também são conhecidas como *superfícies ressonantes* [27].

2.2 Instabilidades Importantes

De uma forma geral, podem-se dividir as várias instabilidades macroscópicas que surgem durante o confinamento magnético do plasma em tokamaks, usualmente tratadas pela teoria MHD (magnetohidrodinâmica), em duas classes distintas: as *ideais*, que consideram o plasma como um fluido condutor perfeito, e as *resistivas*, que levam em consideração a resistividade do plasma.

2.2.1 Teoria MHD: modos de Dobra e de Troca

Os fundamentos básicos da teoria MHD são descritos pelas equações de fluido e pelas equações de Maxwell, apresentadas a seguir [4, 27, 28]:

- conservação do momento:

$$\rho \frac{d\vec{v}}{dt} = -\vec{\nabla}P + \vec{j} \times \vec{B} \quad (2-6)$$

- equação da continuidade:

$$\frac{\partial \rho}{\partial t} + \vec{\nabla} \cdot (\rho \vec{v}) = 0 \quad (2-7)$$

- conservação da energia (processos adiabáticos):

$$\frac{d}{dt}(P\rho^{-\Gamma}) = 0 \quad (2-8)$$

- lei de Ohm ideal:

$$\vec{E} + \vec{v} \times \vec{B} = 0 \quad (2-9)$$

- lei de Faraday:

$$\vec{\nabla} \times \vec{E} = -\frac{\partial \vec{B}}{\partial t} \quad (2-10)$$

- lei de Ampère:

$$\vec{\nabla} \times \vec{B} = \mu_0 \vec{j} \quad (2-11)$$

- inexistência de monopolos magnéticos:

$$\vec{\nabla} \cdot \vec{B} = 0, \quad (2-12)$$

nas quais $\frac{d}{dt} = \frac{\partial}{\partial t} + \vec{v} \cdot \vec{\nabla}$ é a derivada convectiva e Γ a razão dos calores específicos.

Estuda-se a estabilidade linear do sistema supondo o plasma inicialmente em equilíbrio estático (velocidade de fluido nula) e investiga-se a evolução temporal de pequenas perturbações que surgem em torno dessa situação de equilíbrio, através da linearização das equações acima descritas (2-6 a 2-12). Entretanto, essa avaliação pode tornar-se bastante complexa quando se trata com geometrias não muito simples.

Uma outra forma de se abordar o problema é fazendo-se a aproximação da coluna de plasma para um cilindro periódico e utilizando-se o inverso da razão de aspecto ($\varepsilon = a/R_0 \ll 1$) como um parâmetro de expansão para a variação da energia potencial do plasma, num critério conhecido como *Princípio de Energia*. Este princípio está relacionado com a variação da energia potencial do sistema, mediante uma perturbação inicial, $\vec{\xi}$, de tal forma que o equilíbrio será considerado estável se, para qualquer deslocamento $\vec{\xi}$, a variação da energia potencial for positiva ($\delta W > 0$). Por outro lado, se a variação provocada for negativa ($\delta W < 0$), então o equilíbrio será considerado instável.

Com essas considerações, e tomando-se uma perturbação radial $\xi(\vec{r})$, obtém-se a seguinte expressão para a variação da energia potencial [29, 30]:

$$\begin{aligned} \delta W = & \frac{\pi^2 B_\phi^2}{R_0} \left\{ \int_0^a \left[\left(r \frac{d\xi}{dr} \right)^2 + (m^2 - 1) \xi^2 \right] \left(\frac{n}{m} - \frac{1}{q} \right)^2 r dr + \right. \\ & \left. + \left[\frac{2}{q_a} \left(\frac{n}{m} - \frac{1}{q_a} \right) + \left(1 + m \frac{1 + (a/b)^{2m}}{1 - (a/b)^{2m}} \right) \left(\frac{n}{m} - \frac{1}{q_a} \right)^2 \right] a^2 \xi_a^2 \right\}, \quad (2-13) \end{aligned}$$

onde δW é a variação da energia potencial, sofrida pelo sistema, e a , b e R_0 correspondem, respectivamente, ao raio da coluna de plasma, raio menor e raio maior do toróide.

Instabilidades de Dobra

A condição de estabilidade do plasma é determinada pelo segundo termo da equação 2-13 e as instabilidades, quando surgem, provocam deformações na superfície externa da coluna. Por essa razão, essas instabilidades responsáveis pela referida deformação são denominadas *instabilidades de dobra* (“*kink instabilities*”). Analisando-se melhor essa expressão, observa-se que o equilíbrio é sempre instável para $q_a < 1$ (limite de *Krushal-Shafranov*) e que a condição suficiente para que haja estabilidade é $q_a > m/n$.

Escrevendo-se a expressão 2-13 de outra forma, obtém-se [29]:

$$\delta W = \pi^2 R_0 \int_0^a b \left[B_1^2 + B_{0\theta} \left(1 - \frac{nq}{m} \right) \frac{dj_\phi}{dr} \xi^2 \right] r dr, \quad (2-14)$$

onde B_0 e B_1 são, respectivamente, os termos de ordem zero e de primeira ordem, do campo magnético perturbado, e j_ϕ é a densidade de corrente do plasma, na direção toroidal.

Analisando-se essa expressão, pode-se observar que os modos de dobra também dependem do perfil de corrente, ou, equivalentemente, do perfil do fator de segurança $q(r)$, principalmente nas imediações da superfície ressonante, para a qual $q(r_s) = m/n$.

No caso mais simples, de uma corrente uniforme ($q(r) = q_a = \text{constante}$, para $r < a$), as taxas de crescimento da amplitude das instabilidades correspondentes são dadas por [30]:

$$\gamma^2 = \frac{2m B_{\theta a}^2}{\rho a^2} \left(1 - \frac{nq_a}{m} \right) \left[1 - \frac{m - nq_a}{1 - (a/b)^{2m}} \right], \quad (2-15)$$

onde ρ é a densidade de partículas, considerada constante, e $B_{\theta a}$ é o campo magnético poloidal na borda do plasma.

Instabilidades de Troca

O modelo MHD prevê, ainda, um outro tipo de instabilidade, mais inofensiva, denominada *instabilidade de troca*, associada a gradientes de pressão no plasma. Para a geometria toroidal, as condições de estabilidade de perturbações deste tipo são dadas pelo critério de Mercier [29]:

$$\left(-\frac{dp}{dr} \right) (q^2 - 1) + \frac{r b_\theta^2}{8} \left(\frac{1}{q} \frac{dq}{dr} \right)^2 > 0 \quad (2-16)$$

o que mostra o tokamak como sendo bastante estável para essas perturbações, a não ser para a situação em que $q(0) < 1$, que ocasionaria a existência de uma pequena região do plasma, próxima a $r = 0$, onde o critério de Mercier seria violado.

Pelo exposto, poder-se-ia pensar que plasmas confinados em tokamaks de alta razão de aspecto atingiriam uma situação de equilíbrio satisfatória, sob as condições $q(0) > 1$ e $q(a) > m/n$. Ocorre que, na verdade, o plasma não é um condutor ideal, como suposto pela teoria MHD, e a inclusão de características resistivas nas equações da magnetohidrodinâmica implicam em condições de estabilidade mais severas [31].

2.2.2 Modos de Ruptura

Ao se levar em consideração os efeitos resistivos do plasma, surgem outras instabilidades, denominadas *instabilidades de ruptura* (“tearing instabilities”), cuja principal característica reside na quebra e posterior reconexão das linhas de força do campo magnético, formando as chamadas *ilhas magnéticas*. Isso ocorre numa estreita região em torno das superfícies racionais, onde se tem $\vec{k} \cdot \vec{B} = 0$, sendo \vec{k} o vetor de onda da perturbação. O efeito da resistividade é particularmente importante nessa região do plasma, em torno das superfícies racionais. Fora dela, as equações MHD, na forma ideal, são usualmente utilizadas [30].

2.2.3 Instabilidades de Disruptura

Instabilidade de disruptura é o nome que se dá para a instabilidade macroscópica que, quando surge, durante uma descarga do plasma, provoca a súbita perda do confinamento do mesmo. De um modo geral, o evento se dá através de uma rápida perda da energia interna do plasma, uma abrupta expansão do raio menor e um deslocamento da coluna, como um todo, em direção às paredes internas da câmara de vácuo, além do surgimento de um transiente característico (pico negativo) na tensão de enlace. Existem evidências que mostram que o mecanismo físico, responsável pelo surgimento dessas instabilidades, está relacionado com o acoplamento entre os modos $m/n = 2$ e $m/n = 1$ [32, 33, 34].

O surgimento dessas disrupturas pode ser evitado operando-se os atuais tokamaks em limites experimentalmente conhecidos, de modo a evitar o *limite de densidade* (“density limit disruption”), *limite de baixo q* (“low-q disruption”), *limite de alto β* (“high- β disruption”), etc [35]. Não obstante, como o tokamak ITER terá que operar próximo do limite de alto β , ao mesmo tempo em que excederá o limite de densidade, espera-se que os fenômenos disruptivos venham a se tornar eventos regulares, ao menos até que medidas preventivas possam ser encontradas [36].

2.3 Diagnósticos Relevantes

Dos vários diagnósticos existentes, os sistemas que medem a flutuação do campo magnético poloidal e os sinais de raios-X de baixas energias, emitidos pelo plasma, são os mais apropriados para a investigação das instabilidades de disruptura em tokamaks, uma vez que possibilitam a monitoração dos modos $\frac{m}{n} = \frac{2}{1}$ e $\frac{m}{n} = \frac{1}{1}$, cujo acoplamento parece estar

associado ao desencadeamento do evento disruptivo [32, 33].

2.3.1 Bobinas magnéticas de Mirnov

A intensidade das flutuações do campo magnético poloidal pode ser medida através do posicionamento de pequenas bobinas magnéticas no interior da câmara de vácuo do tokamak. Essas flutuações magnéticas são também conhecidas como *oscilações de Mirnov*, em homenagem ao cientista russo, S.V.Mirnov, por ter sido ele o primeiro a realizar medidas dessa natureza, cuja faixa de frequência varia para cada tokamak [37, 38]. Geralmente, para tokamaks de grande e médio porte, a frequência típica é de 1 a 20 kHz , enquanto que, para tokamaks menores, a faixa medida situa-se entre 20 e 60 kHz .

Acredita-se que as oscilações de Mirnov estejam relacionadas com as estruturas espaciais helicoidais da forma $e^{i(m\theta-n\phi)}$, consistentes com a rotação dos modos de ruptura (“*tearing instabilities*”), onde m e n , já definidos anteriormente, correspondem à helicidade dessas estruturas. Desse modo, informações relevantes sobre os mecanismos físicos responsáveis pelo surgimento dos fenômenos disruptivos podem ser obtidos através da análise das oscilações de Mirnov. Tem-se observado, por exemplo, que a amplitude do campo magnético poloidal de perturbação cresce exponencialmente nos instantes que antecedem a ocorrência de uma disruptura maior [5, 32, 33].

Determinação dos modos m e n

Experimentalmente, determina-se o modo m colocando-se várias bobinas magnéticas internamente à câmara do tokamak, separadas angularmente de θ , ao longo da direção poloidal, para um dado valor de ϕ . Para a determinação do modo n , outras tantas bobinas são angularmente distribuídas ao longo da direção toroidal. Assim, quando se deseja investigar as instabilidades que ocorrem num determinado momento do pulso de plasma, analisam-se os sinais de todas as bobinas conjuntamente, em termos de suas amplitudes relativas e da variação de fase que apresentam entre si.

Esta análise dos sinais experimentalmente medidos é usualmente feita através da decomposição dos modos em série de Fourier, obtida através de algoritmos numéricos que fornecem os coeficientes a_0 , a_m e b_m da equação [39]:

$$\dot{B}_\theta = \frac{a_0}{2} + \sum_{m=1}^{N-1} [a_m \cos(m\theta) + b_m \sin(m\theta)] + \frac{a_N}{2} \cos(N\theta), \quad (2-17)$$

onde $a_m = \frac{2}{N} [\sum_{k=0}^{N-1} \dot{B}(k) \cos(ksm)]$ e $b_m = \frac{2}{N} [\sum_{k=0}^{N-1} \dot{B}(k) \sin(ksm)]$, com N correspondendo ao número de bobinas magnéticas utilizadas e s o espaçamento existente entre elas.

Os coeficientes $c_m = 1/2(a_m^2 + b_m^2)^{1/2}$, resultantes dessa análise, quantificam a amplitude das perturbações correspondentes aos números de onda m presentes num determinado instante da descarga do plasma.

2.3.2 Raios-X de baixas energias

O espectro de emissão de raios-X de baixas energias (0,1 – 10,0 keV) em plasmas confinados por tokamaks consiste-se, basicamente, de um contínuo com picos característicos superpostos. A parte contínua é composta fundamentalmente por dois componentes:

- *Radiação de “bremsstrahlung”*, originada na interação dos elétrons energéticos com os íons de impurezas e isótopos do hidrogênio.
- *Recombinação elétron-íon*, que se consiste na captura de um elétron livre pelos íons do plasma.

No primeiro caso, supondo-se que os elétrons possuam velocidades muito superiores às dos íons, a taxa de interação dependerá basicamente da temperatura dos elétrons (T_e). A produção da radiação por bremsstrahlung, na faixa de energia relativa aos raios-X, pode então ser escrita como [17]:

$$\left(\frac{\Delta W}{\Delta E}\right)_{br} \propto n_e \sum_i n_i Z_i^2 T_e^{-\frac{1}{2}} e^{-\frac{E}{T_e}}, \quad (2-18)$$

onde:

- $\Delta W \equiv$ potência irradiada, por cm^3 , correspondente aos fótons de energias pertencentes ao intervalo ΔE .
- n_e e $n_i \equiv$ densidades dos elétrons e íons, respectivamente.
- $Z_i \equiv$ número atômico dos íons.
- $E \equiv$ energia do fóton emitido.

A expressão 2-18 indica que o formato do espectro da emissão depende exponencialmente da temperatura dos elétrons, fornecendo, portanto, um meio de se determinar essa

temperatura. Devida à dependência direta com a densidade, e uma vez que a parte mais densa do plasma é justamente a da região central, conclui-se que a emissão dos ftons se dará principalmente pela região central da coluna de plasma. O diagnóstico por raios-X torna-se, portanto, uma ótima ferramenta na análise dos fenômenos que surgem nesta região do plasma magneticamente confinado.

A radiação devida à recombinação elétron-íon, na faixa de energia acima do potencial de ionização do átomo que participa do processo, pode ser escrita como [17]:

$$\left(\frac{\Delta W}{\Delta E}\right)_{rec} = \left(\frac{\Delta W}{\Delta E}\right)_{br} (\gamma - 1), \quad (2-19)$$

uma vez que os dois processos de emissão demonstram possuir o mesmo formato espectral. O parâmetro γ (equação 2-19) corresponde ao acréscimo no espectro contínuo emitido por bremsstrahlung, devida à radiação de recombinação. A energia do fóton emitido, nesse processo de recombinação elétron-íon, corresponde à energia cinética do elétron mais a energia devida ao decaimento desse elétron para o estado ligado final.

Os picos superpostos ao contínuo correspondem ao decaimento espontâneo dos íons, excitados, para um estado menos energético, e tendem a possuir comprimentos de onda na região do ultra-violeta (VUV) do espectro.

Devida à forte dependência entre a emissão dos raios-X de baixas energias com a temperatura do plasma, a utilização do sistema de detecção destes sinais está mais voltada para a observação do centro da coluna de plasma, o que é perfeitamente possível, uma vez que o plasma é opticamente fino para as suas próprias emissões de raios-X.

Dentro deste contexto, a observação de flutuações na emissão de raios-X, denominadas *dentes-de-serra*, assim chamadas por apresentarem oscilações periódicas semelhantes a dentes-de-serra, são de extrema importância, uma vez que elas estão intrinsecamente associadas a um processo de expansão e contração cíclica da região central da coluna de plasma. Estas oscilações, por afetarem apenas a região interna do plasma, também recebem o nome de *disrupturas internas* [17, 18].

O vetor de onda \vec{k} , dessas perturbações, é usualmente escrito como $\vec{k} = \frac{m}{R} \vec{e}_\theta - \frac{n}{R} \vec{e}_z$, e os parâmetros m e n podem ser obtidos através da análise das diferenças de fase dos sinais experimentais medidos por sensores de raios-X convenientemente instalados em diferentes posições poloidais e toroidais da câmara do tokamak.

É também comumente observada uma oscilação MHD superposta aos *dentes-de-serra*, que parece ser a responsável pelas quedas abruptas das disrupturas internas. Medidas simultâneas da radiação emitida por diferentes regiões poloidais e toroidais do plasma indicam que essas oscilações correspondem ao modo $m = n = 1$. Portanto, as disrupturas internas estariam associadas a uma perturbação existente nas imediações da superfície magnética $q = 1$ [17, 18].

O diagnóstico de raios-X de baixas energias também pode ser utilizado para medir a temperatura dos elétrons através do método dos absorvedores, que já foi utilizado, com sucesso, no tokamak TBR-1, do Instituto de Física, da USP, devendo também ser implantado no tokamak TCABR, deste mesmo instituto [19, 20].

Método dos absorvedores

O método dos absorvedores consiste em calcular a razão dos fluxos de raios-X moles medidos por um par de detectores barreira-de-superfície, por exemplo, que visualizam, através de ângulos sólidos semelhantes, a mesma região térmica da coluna central do plasma. No caminho óptico de cada detector é inserida uma folha de material absorvedor, de determinada espessura (fig.2-3), de modo que somente uma fração I , da radiação incidente I_0 , consiga atravessar cada lâmina absorvedora. A partir da razão obtida, e de uma curva de calibração apropriada, obtém-se a temperatura eletrônica desejada.

Os parâmetros de intensidade I e I_0 estão relacionados entre si pela equação [40]:

$$I = I_0 e^{-\mu x}, \quad (2-20)$$

onde:

- $I_0 \equiv$ intensidade da radiação incidente.
- $I \equiv$ intensidade da radiação transmitida.
- $\mu \equiv$ coeficiente de absorção de massa do material.
- $x \equiv$ espessura do anteparo.

Por outro lado, o fluxo inicial I_0 , originado nos processos atômicos de bremsstrahlung e de captura eletrônica, supondo-se um plasma maxwelliano, é de natureza contínua, e sua produção é descrita através das equações 2-18 e 2-19, podendo-se escrever, portanto:

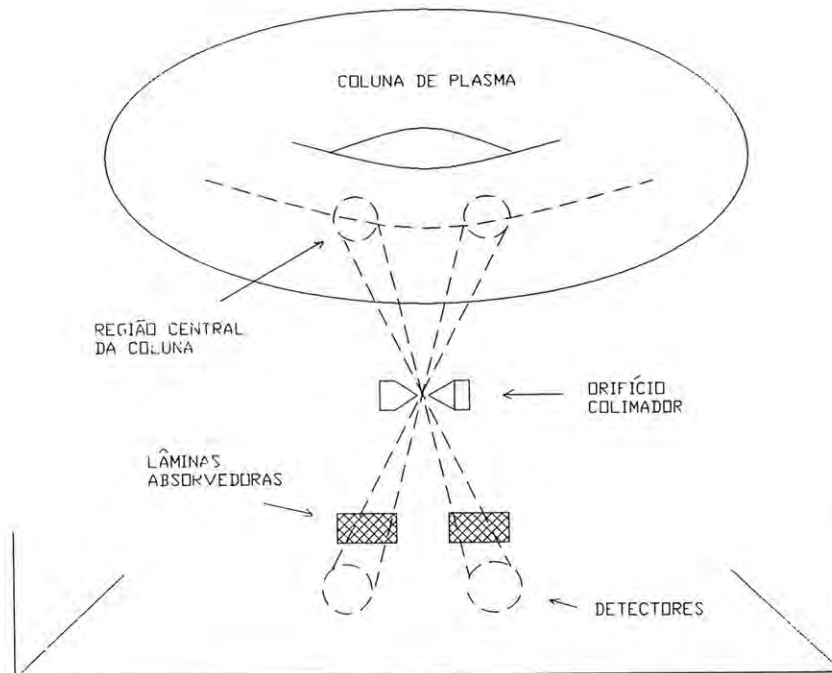


Figura 2-3: Esquema do diagnóstico de raios-X moles utilizando-se dois detectores barreira-de-superfície que visualizam a região central da coluna de plasma do tokamak.

$$I(E) \propto e^{\frac{-E}{T_e} - \mu(E)x} \quad (2-21)$$

A constante de proporcionalidade desaparece quando se toma a razão entre os sinais dos dois detectores, não sendo necessária nenhuma medida absoluta de intensidade.

Levando-se em conta a existência de outros materiais absorvedores, como o alumínio, que compõe o eletrodo frontal do detector, e a camada da zona morta de silício, obtém-se a seguinte expressão para a razão dos fluxos de raios-X moles, a partir da qual se pode calcular a temperatura eletrônica do plasma para cada par de pulsos de raios-X medidos [20, 41, 42]:

$$R = \frac{I_1(E)}{I_2(E)} = \frac{\int \exp\left[\frac{-hc}{\lambda T_e} - \mu_F x_1 - \mu_{Al} x_{Al} - \mu_{Si} x_{Si}\right] E_f(\lambda) \lambda^{-2} d\lambda}{\int \exp\left[\frac{-hc}{\lambda T_e} - \mu_F x_2 - \mu_{Al} x_{Al} - \mu_{Si} x_{Si}\right] E_f(\lambda) \lambda^{-2} d\lambda}, \quad (2-22)$$

onde:

- $\lambda \equiv$ comprimento de onda da radiação incidente
- $E_f(\lambda) \equiv [1 - \exp(-0,01\mu_{Si})]$, eficiência do detector.
- $hc \equiv 12398$ (eVÅ).
- x_1 e $x_2 \equiv$ espessura dos filtros utilizados.
- $x_{Si} \equiv$ espessura da zona morta de silício.
- $x_{Al} \equiv$ espessura do depósito de alumínio.
- μ_{Al} , μ_{Si} e $\mu_F \equiv$ coeficientes de absorção de massa dos elementos Al, Si e do filtro empregado no método, respectivamente.

Capítulo 3

Redes Neurais Artificiais

As redes neurais artificiais foram desenvolvidas, originalmente, em 1943, por Warren McCulloch, médico, filósofo, poeta e matemático, juntamente com o estatístico Walter Pitts, num trabalho que se consistia no modelo de resistores variáveis e amplificadores representando conexões sinápticas de um neurônio. Como resultado deste trabalho, os autores publicaram o artigo “*A Logical Calculus of the Ideas Immanent in Nervus Activity*” no *Bulletin of Mathematical Biophysics*, sendo considerada a primeira referência nessa área [43].

A essência da proposta de McCulloch e Pitts era a seguinte: “*A inteligência é equivalente ao cálculo de predicados, que por sua vez pode ser implementado por funções booleanas. Por outro lado, o sistema nervoso é composto de rede de neurônios que, com as devidas simplificações, têm a capacidade básica de implementar estas funções booleanas. Conclusão: a ligação entre inteligência e atividade nervosa fica estabelecida de forma científica*” [43].

Anos depois, a publicação do livro intitulado ‘*Perceptron*’, em 1969, de autoria de dois renomados cientistas da época: Marvin Minsky e Seymour Papert, causou um grande impacto, negativo, no estudo das redes neurais artificiais, pois os autores criticavam seriamente os trabalhos envolvendo esses conceitos, alegando que os processos de aproximação adotados eram empíricos e algumas das operações utilizadas não eram booleanas. Desnecessário dizer que, dada a popularidade dos autores, naquela ocasião, houve, com suas críticas, uma grande retração nos programas de pesquisa e investimentos nessa área, até então, emergentes. Somente em 1982, o físico e biólogo John Hopfield, do Instituto de Tecnologia da Califórnia, através de trabalhos baseados num modelo de funcionamento do neurônio, a partir do estudo do sistema neurológico de uma lesma de jardim, contestou,

com sucesso, as teses matemáticas de Minsky e Papert, dando um novo impulso à pesquisa nesta área [43, 44]. Entretanto, somente a partir de 1986, com a publicação do hoje famoso trabalho “*Parallel Distributed Processing*”, de Rummelhart e MacClelland, do PDP Research Group da Universidade da Califórnia, em San Diego [43], é que a área de redes neurais teve seu desenvolvimento decididamente impulsionado. Desde então, suas aplicações têm sido as mais diversas possíveis, podendo-se citar, dentre tantas, aplicações no setor financeiro, engenharia civil, sistemas de energia, biomédicas, processamento de imagem (e de som), processos industriais, robótica, classificadores e, naturalmente, aplicações na área da física, revelando, assim, o seu carácter multidisciplinar.

O Cérebro e o Processamento Paralelo

Em termos de ciência computacional, podemos descrever o cérebro como um sistema paralelo de $\sim 10^{11}$ processadores. De acordo com um modelo matemático simplificado de um neurônio, cada processador executa uma tarefa muito simples: ele calcula a soma ponderada dos dados de entrada, proveniente de outros processadores, e fornece um número como saída, que corresponde ao resultado de uma função não linear aplicada àquela soma ponderada. Esta saída é então enviada para outros processadores que estão continuamente fazendo o mesmo tipo de cálculo.

O fato de haver muitos termos nas somas envolvidas, o que significa, falando-se em termos de redes neurais, haver uma alta conectividade, faz com que alguns erros em poucos termos sejam provavelmente irrelevantes, o que lhe dá a característica de robustez, permitindo que a performance da rede sofra uma *degradação graciosa* na presença de ruídos, ou erros [43, 44]. As próprias células do cérebro morrem diariamente, sem que isto afete significativamente o seu desempenho. Esse tipo de robustez em redes neurais biológicas tem sido, provavelmente, essencial para a evolução da inteligência.

O procedimento descrito contrasta com o tipo convencional (von Neumann), no qual um ou mais processadores executam programas muito complicados, e, em contraste à robustez do sistema paralelo, uma sequência computacional usual pode facilmente tornar-se ruidosa por um erro de um simples bit [43, 44] !

Nunca é demais lembrar que, embora o período típico dos neurônios para processarem informações seja milhões de vezes mais lento do que o de um computador, o cérebro pode processar várias tarefas muito mais rápido do que qualquer supercomputador existente nos dias de hoje, tais como o controle motor, a visão, e decisões baseadas em dados

incompletos ou ruidosos, só para citar alguns exemplos. E isso, obviamente, só é possível porque bilhões de neurônios estão simultaneamente em operação.

3.1 Princípios da Neurocomputação

Não há a menor dúvida de que o cérebro humano é superior a um computador digital em diversas tarefas como, por exemplo, o controle de sistema complexos, pois até mesmo uma pequena criança pode andar de bicicleta sem precisar de conhecer a Mecânica de Newton!

Outras qualidades que o cérebro humano possui são igualmente notáveis, e, portanto, desejáveis em sistemas artificiais, tais como o fato de ele ser fortemente *paralelo* (1), *pequeno* (2), *compacto* (3) e de dissipar *pouca potência* (4). Como já mencionamos antes, células nervosas morrem no cérebro todos os dias sem afetar a sua performance significativamente, o que lhe dá, ainda, a característica de *robustez e tolerância à falhas* (5). Além do mais, o cérebro também se adapta facilmente às novas situações através do *aprendizado* (6), sem precisar de ser previamente programado para isso, o que resulta em *flexibilidade* (7). E, finalmente, o cérebro tem, ainda, a grande habilidade de lidar com informações vagas, probabilísticas, ruidosas ou inconsistentes, o que lhe confere, também, uma grande *capacidade de abstração* (8). Nosso “*amigo de silício*” só consegue mesmo superar o cérebro humano em tarefas baseadas simplesmente em aritmética, real motivação para o estudo e desenvolvimento da neurocomputação.

Estamos falando, portanto, de um *novo paradigma computacional* que, ao contrário do paradigma atual, baseado na instrução programada sequencial, introduzido por Neumann, inspira-se num modelo biológico para a inteligência, ou seja, na maneira como o cérebro é organizado, em sua estrutura elementar, e em como ele é capaz de executar tarefas computacionais, sem, contudo, pretender ser biologicamente realístico em seus detalhes. Os modelos utilizados são extremamente simples, quando vistos do ponto de vista neurofisiológico. Entretanto, os detalhes das células nervosas individuais podem ser irrelevantes quando se pensa em compreender uma rede neural do ponto de vista do *comportamento coletivo* [14, 44].

Dentro deste contexto, podemos definir, então, uma rede neural artificial como “*uma rede massivamente paralela, interconectada de elementos, e suas organizações hierárquicas, que estão intencionados para interagir com objetos do mundo real do mesmo modo que um sistema nervoso biológico faz*” [44].

3.2 O Funcionamento do Sistema Nervoso Biológico

O sistema nervoso é responsável pela integração do organismo com o seu meio ambiente. Sua função é detectar estímulos internos e externos, tanto físicos, quanto químicos, desencadeando as respostas glandulares e musculares correspondentes a estes estímulos. De um modo simples, pode-se dizer que ele é formado por células nervosas, interconectadas de forma específica e precisa, constituindo-se o que se chama de circuitos neurais. É através desses circuitos que o organismo consegue produzir respostas estereotipadas, que constituem os comportamentos fixos e invariantes, como por exemplo os reflexos, ou, então, produzir outros comportamentos variáveis em maior ou menor grau [45].

3.2.1 Estrutura de um neurônio

O neurônio, ou célula nervosa, é o principal componente do sistema nervoso cerebral. Trata-se de uma célula especializada que transmite impulsos nervosos, que correspondem à transmissão de um sinal codificado de um dado estímulo, ao longo da membrana do neurônio, a partir do ponto em que ele foi estimulado [45].

Assim como as demais células, os neurônios se alimentam, respiram, têm os mesmos genes, os mesmos mecanismos bioquímicos e as mesmas organelas. A única grande diferença que se observa entre um neurônio e as demais células reside num fator muito importante: os neurônios são os **únicos** capazes de **processar informação**.

Um neurônio típico tem quatro regiões morfológicamente definidas (figura 3-1). São elas:

- **corpo do neurônio** (também chamado de **soma**): contém o núcleo, no qual residem as informações genéticas, sendo também responsável pelo suprimento metabólico de toda a célula.
- o **axônio** (do grego “*axoon*”, ou eixo): fibra nervosa constituída de um único prolongamento do corpo celular, na forma de um eixo cilíndrico, que pode ser curto ou longo. É o principal responsável pela condução do impulso nervoso para outros neurônios, sendo capaz de conduzir sinais elétricos à distâncias que podem variar de 0,1 mm até 2 m !
- **dendritos** (do grego “*dendron*”, ou árvore): milhares de ramificações (arborizações terminais) de finas estruturas de fibras nervosas conectadas ao corpo da célula. Elas

funcionam como ‘antenas’ dos neurônios, sendo o principal aparato para receber os sinais dos neurônios vizinhos.

- **terminais pré-sinápticos** (ou terminal nervoso): ramificações terminais de um axônio que podem formar sinapses (região onde ocorre a transmissão e o armazenamento da informação) com até 1000 outros neurônios.

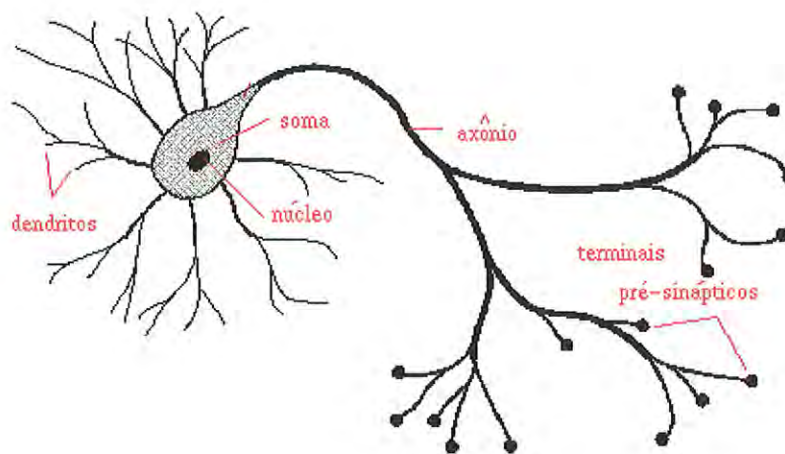


Figura 3-1: Estrutura típica de um neurônio biológico.

3.2.2 Transmissão da informação

A transmissão da informação no cérebro biológico se dá através do impulso nervoso, cujo processamento envolve fenômenos elétricos e químicos. Os eventos elétricos propagam o sinal *dentro* do neurônio, enquanto os processos químicos são responsáveis pela *transmissão* do sinal de um neurônio para outro, na sinapse.

A sinapse é o nome dado à região onde os ramos do axônio de um neurônio (o neurônio pré-sináptico) transmitem sinais a um outro neurônio (o neurônio pós-sináptico) através da liberação de substâncias químicas chamadas *neurotransmissoras*. Embora os neurônios estejam intimamente próximos entre si, não existe continuidade de material entre as células, havendo, entre ambos, um pequeno espaçamento denominado *fenda sináptica*, na qual as substâncias neurotransmissoras do neurônio pré-sináptico ligam-se quimicamente

a receptores específicos do neurônio pós-sináptico, dando continuidade à propagação do sinal.

A transmissão do sinal de uma célula para a outra, na sinapse, é, de fato, um processo químico complexo, cujo efeito é aumentar ou diminuir o potencial elétrico dentro do corpo da célula receptora. Cada neurônio envia picos de atividade elétrica para um outro neurônio através do axônio. Quando um neurônio recebe um sinal excitatório, que deve ser suficientemente maior do que um sinal inibitório, fazendo com que o potencial da célula receptora alcance um determinado limiar, então, um pulso de intensidade e duração fixos é enviado para o axônio. Diz-se, então, que a célula foi ativada. Este pulso se distribui, através das ramificações do axônio, para as junções sinápticas de outras células. Depois de ativada, a célula tem que esperar um determinado tempo, denominado período refratário, antes de ser excitada novamente. O aprendizado se dá através das alterações das atividades das sinapses, assim como através da influência de um neurônio nos demais.

O número e a qualidade das sinapses em um neurônio pode variar, entre outros fatores, pela experiência e aprendizagem, demonstrando a capacidade plástica do sistema neural [14, 44].

3.2.3 Aprendizado e plasticidade

Todo o córtex cerebral está organizado em áreas funcionais que desempenham tarefas receptivas, integrativas ou motoras no comportamento. Essas regiões funcionais respondem por todos os nossos atos conscientes, nossos pensamentos e pela capacidade que possuímos de responder a qualquer estímulo ambiental de forma voluntária. Existe um verdadeiro *mapa cortical* (figura 3-2) com divisões precisas a nível anatomo-funcional que tornam-se mais ou menos ativadas, dependendo da atividade que o cérebro esteja desempenhando.

Através do estudo e conhecimento morfológico-estrutural do neurônio, da natureza de suas conexões sinápticas e da organização das áreas associativas cerebrais, observou-se que a aprendizagem pode levar a alterações estruturais do cérebro. À essa habilidade do sistema nervoso, que é única em relação a todos os outros sistemas orgânicos, dá-se o nome de *plasticidade*. A cada nova experiência do indivíduo, portanto, redes de neurônios são rearranjadas, outras tantas sinapses são reforçadas e múltiplas possibilidades de respostas ao ambiente tornam-se possíveis. Portanto, o *mapa cortical* está sujeito a constantes

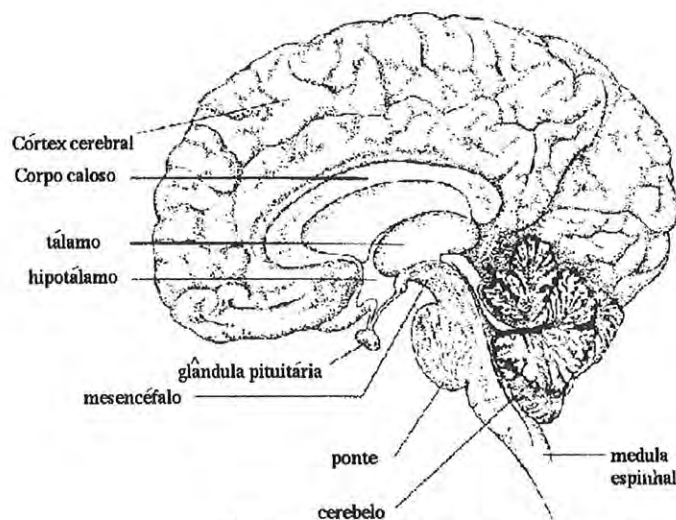


Figura 3-2: Mapa cortical humano

alterações com base no uso ou atividade de seus caminhos sensoriais periféricos [45].

3.3 Fundamentos das Redes Neurais Artificiais

As redes neurais artificiais são uma tentativa de modelar a capacidade que o cérebro tem de processar informações. Assim, para tentarmos construir modelos abstratos dessas redes, a primeira coisa a se fazer é considerar as propriedades essenciais das redes neurais biológicas do ponto de vista do *processamento das informações*.

Como vimos, em redes neurais biológicas, a informação é armazenada nas sinapses, pontos de contacto entre os diferentes dendritos, que são os canais de transmissão da informação. Organelas presentes no corpo da célula produzem toda a química necessária para o trabalho contínuo do neurônio. Os sinais de saída são transmitidos pelo axônio. Estes quatro elementos: **dendritos, sinapses, corpo da célula e axônio** formam a estrutura mínima que nós devemos adotar para o modelo biológico [13].

3.3.1 O neurônio artificial

Redes neurais artificiais são tipicamente compostas por unidades interconectadas, com uma estrutura lógico-matemática interna, que procuram simular a forma, o comportamento e as funções dos neurônios biológicos. Já a função da sinapse é modelada por um peso variável que está associado a cada uma das conexões.

A maioria das redes artificiais não refletem a geometria detalhada dos dendritos e axônios (figura 3-3). Elas expressam o sinal elétrico, emitido pelo neurônio, através de um único número, que representa a sua atividade [46]. De modo simples, podemos dizer que a **unidade**, o **peso** e o **número**, na rede artificial, estão associados, respectivamente, ao **neurônio**, à **sinapse** e à **atividade da célula**, da rede biológica.

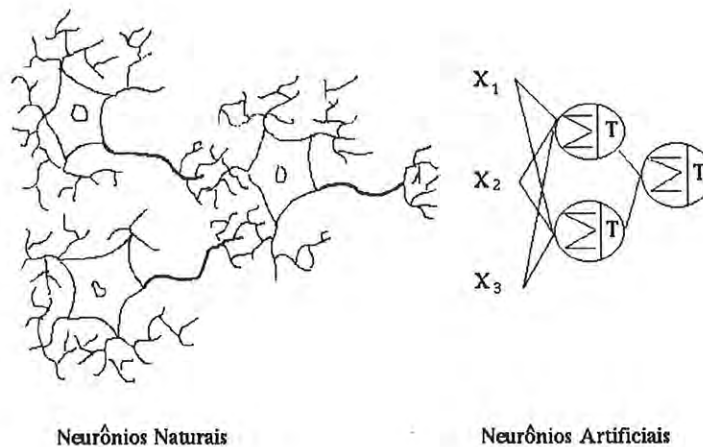


Figura 3-3: Analogia entre neurônios naturais e neurônios artificiais.

Cada unidade da rede artificial converte o padrão de atividades de entrada num único sinal de atividade de saída, que se difunde para as outras demais unidades da camada posterior, e assim por diante. Esta conversão se dá em dois estágios (figura 3-4):

1º estágio: a unidade multiplica cada atividade de entrada por um *peso*, na conexão, e adiciona todas elas para obter uma quantidade chamada *atividade total de entrada*.

2º estágio: a unidade utiliza uma função *entrada-saída* (ou, função de ativação) que transforma a atividade total de entrada numa atividade chamada *atividade total de saída*.

Dessa forma, o comportamento da rede neural depende tanto dos pesos, quanto do tipo da função de ativação utilizada, que, por sua vez, normalmente recai numa das três seguintes categorias: função linear, função tangente hiperbólica e função sigmóide (figura 3-5) [46]. Embora o terceiro tipo de função pareça representar melhor o comportamento dos neurônios, todas elas são, na verdade, aproximações grosseiras dos mesmos.

Convém lembrar que neurônios reais apresentam comportamentos mais complicados, que foram omitidos neste modelo simples. Os mais importantes são [14]:

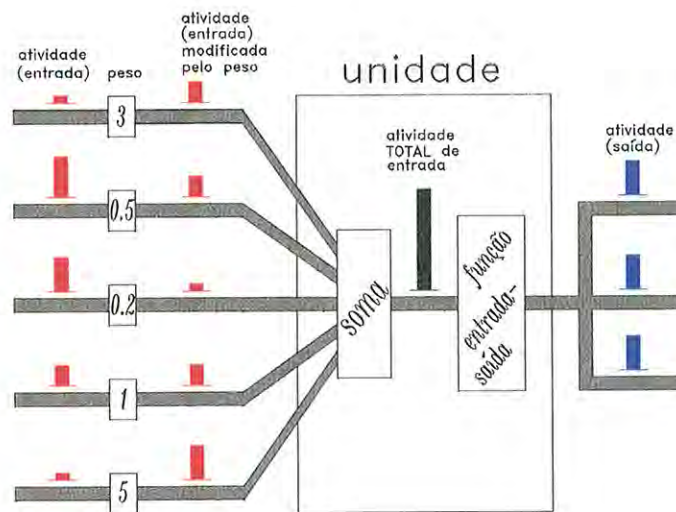


Figura 3-4: Esquema da transmissão da informação num neurônio artificial.

- neurônios reais não trabalham em sistemas tipo “*threshold*” (limiar). Ao invés disso, eles respondem aos estímulos de entrada de forma contínua, o que parece ser um comportamento universal. E é esse comportamento universal que nos interessa, e não a sua forma específica.
- muitas células reais efetuam a soma de suas entradas de uma maneira não linear. Parece haver processos lógicos significativos dentro das próprias ramificações dos dendritos.
- um neurônio real produz uma sequência de pulsos, e não um único nível de saída.
- os neurônios não trabalham de forma sincronizada, ou seja, não são ativados todos ao mesmo tempo.
- a quantidade de substância química transmissora, liberada na sinapse, pode variar de forma imprevisível.

O neurônio artificial passa a assumir, a partir desse momento, seu símbolo universal, apresentado na figura (3-6).

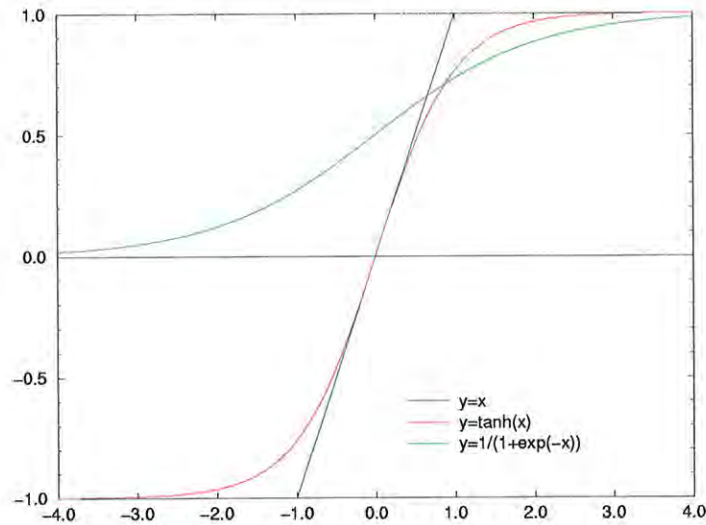


Figura 3-5: Funções de ativação tipicamente utilizadas em redes neurais artificiais.

3.3.2 Rede artificial típica

Para fazer com que uma rede neural artificial realize uma tarefa específica, deve-se definir **como** as unidades deverão estar conectadas entre si e devemos estabelecer, aproximadamente, os **pesos** nas conexões entre elas. Estas conexões determinam se uma dada unidade vai conseguir influenciar uma outra e os pesos determinam a força desta influência.

Uma rede neural típica consiste-se de três grupos, ou camadas, de unidades (figura 3-7) [46]:

1-) **camada de entrada:** as unidades pertencentes a essa camada são conhecidas como *unidades de entrada* e suas atividades são representadas pelos *dados iniciais*, utilizados para alimentar a rede.

2-) **camadas escondidas:** a rede neural artificial pode ter uma ou mais camadas escondidas, e suas atividades são determinadas pelas atividades das unidades de entrada e os pesos associados às conexões existentes entre elas.

3-) **camada de saída:** similarmente ao caso anterior, o comportamento das *unidades de saída* dependem da atividade das unidades escondidas e dos pesos relativos a essas unidades escondidas e às unidades de saída.

No resente trabalho, utilizou-se uma rede de 4 camadas, sendo 1 camada de entrada (esta camada contém os dados experimentais), 2 camadas escondidas e 1 camada de saída,

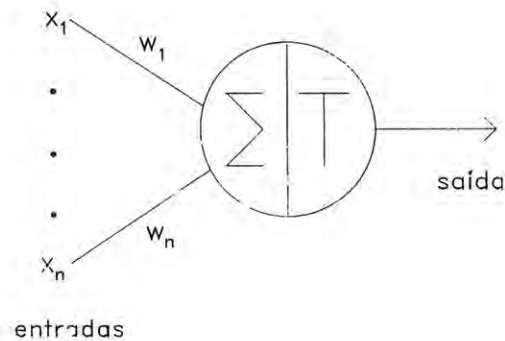


Figura 3-6: Símbolo universal de um neurônio artificial.

que contém o resultado da rede, (figura 3-7).

Para ensinar uma rede artificial a realizar uma tarefa particular, deve-se, primeiro, treiná-la adequadamente com exemplos conhecidos. Calcula-se, então, o erro entre o resultado obtido (fornecido pela rede) e o resultado real (valor experimental). No presente trabalho, o erro foi definido pelo quadrado da diferença entre a atividade fornecida pela rede e a atividade real, como será discutido, mais adiante. Em seguida, mudam-se os pesos de cada uma das conexões, durante uma determinada quantidade de iterações, chamada *época de treinamento*, até que o erro seja reduzido. O processo de treinamento se dá repetindo-se várias vezes este procedimento, com a contínua atualização dos pesos, a cada iteração.

Assim como em neurônios reais, nos quais o aprendizado se dá nas mudanças de intensidade das conexões sinápticas, provocando uma mudança no comportamento das ativações de uma célula para outra, em redes neurais artificiais o aprendizado ocorre quando os **pesos** de cada uma das conexões entre as unidades são **modificados**.

Para implementar este procedimento, é preciso mudar cada um dos pesos por uma quantidade que seja proporcional à taxa com que cada erro muda quando o peso é alterado. Esta quantidade, chamada *erro derivado do peso* (E_w), pode ser obtida através de um algoritmo apropriado que, no presente trabalho, tratou-se do algoritmo de *retropropagação* (“*backpropagation*”) [14, 46].

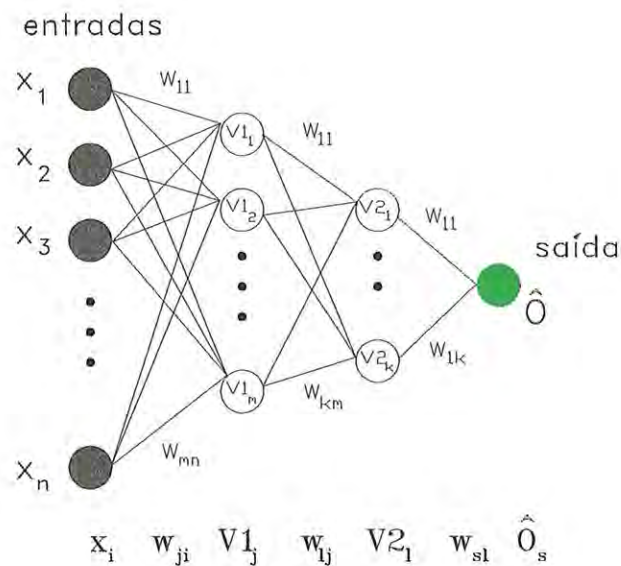


Figura 3-7: Desenho ilustrativo de uma rede neural artificial, tipo "feedforward", de 4 camadas: 1 camada de entrada, que contém os dados experimentais, 2 camadas escondidas e 1 camada de saída, que contém o resultado da rede.

Num cérebro humano, os neurônios apresentam uma alta conectividade. Isso também é possível de se fazer em redes neurais artificiais, mas a interconexão de todas as unidades levaria a um sistema muito complexo, chamado de "Hopfield Network", onde frequentemente ocorrem "loops" e oscilações que tornam o sistema lento [14]. Optou-se, então, por utilizar uma rede na qual o fluxo de informações fosse linear, ou seja, caminhasse num único sentido, o que é obtido arranjando-se as unidades em camadas (figura 3-7) e calculando-se suas atividades através das unidades de saída da camada imediatamente anterior. Desta forma, durante uma iteração, cada unidade é requisitada uma única vez, o que torna o funcionamento do programa muito mais rápido. Redes deste tipo, onde as informações caminham numa única direção e sentido, de uma camada para outra, são chamadas de "Feedforward Network" [13].

3.3.3 Algoritmo "Error-Backpropagation"

Para facilitar a compreensão da rede neural artificial em termos matemáticos, vamos primeiro identificar os termos utilizados nas equações que serão desenvolvidas daqui em diante (figura 3-7):

- $x \equiv$ dado de entrada da primeira camada.
- $\Upsilon \equiv$ atividade total de entrada em uma dada unidade.
- $w \equiv$ peso da conexão entre as unidades.
- $g \equiv$ função de ativação utilizada.
- $V \equiv$ saída de uma dada unidade.
- $\mu \equiv$ padrão de dados apresentado à rede.
- $\zeta \equiv$ dado experimental (alvo).
- $\hat{O} \equiv$ saída da rede.

Para cada padrão (μ) de dados apresentados à rede, o processamento da informação ocorre da seguinte forma: primeiro, a rede calcula a atividade total de entrada ($\Upsilon 1$) de cada uma das unidades (j), da primeira camada escondida, a partir da soma dos dados experimentais (x_i^μ) da camada de entrada, ponderada pelos respectivos pesos (w_{ji}), nas conexões, através da seguinte expressão:

$$\Upsilon 1_j^\mu = \sum_i w_{ji} x_i^\mu \quad (3-1)$$

Depois, a rede calcula a atividade de saída ($V 1_j$), correspondente à cada unidade da referida camada, aplicando a função de ativação ($g 1$) ao resultado de cada entrada:

$$V 1_j^\mu = g 1(\Upsilon 1_j^\mu) = g 1 \left(\sum_i w_{ji} x_i^\mu \right) \quad (3-2)$$

Para a segunda camada escondida, o procedimento é o mesmo: calcula-se a atividade total de entrada ($\Upsilon 2$) para cada uma das unidades (l), da segunda camada escondida, através da equação:

$$\Upsilon 2_l^\mu = \sum_j w_{lj} V 1_j^\mu = \sum_j w_{lj} g 1 \left(\sum_i w_{ji} x_i^\mu \right) \quad (3-3)$$

Em seguida, calcula-se a atividade de saída ($V 2_l$), das respectivas unidades da segunda camada, através da função de ativação ($g 2$), pela expressão:

$$V 2_l^\mu = g 2(\Upsilon 2_l^\mu) = g 2 \left[\sum_j w_{lj} g 1 \left(\sum_i w_{ji} x_i^\mu \right) \right] \quad (3-4)$$

A atividade total de entrada ($\Upsilon 3$), da camada de saída, é calculada de modo análogo:

$$\Upsilon 3_s^\mu = \sum_l w_{sl} V 2_l^\mu = \sum_l w_{sl} \left\{ g 2 \left[\sum_j w_{lj} g 1 \left(\sum_i w_{ji} x_i^\mu \right) \right] \right\} \quad (3-5)$$

Finalmente, a saída fornecida pela rede, \hat{O} (resultado final), é dada pela expressão:

$$\hat{O}_s^\mu = g 3(\Upsilon 3_s^\mu) = g 3 \left\{ \sum_l w_{sl} \left\{ g 2 \left[\sum_j w_{lj} g 1 \left(\sum_i w_{ji} x_i^\mu \right) \right] \right\} \right\} \quad (3-6)$$

No presente trabalho, optou-se por utilizar as funções de ativação da seguinte maneira: $g 1 = g 2$ correspondendo à função tangente hiperbólica e $g 3$ à função identidade. Desse modo, a expressão 3-6 passa a ter a seguinte forma:

$$\hat{O}_s^\mu = \sum_l w_{sl} \left\{ \tanh \left[\sum_j w_{lj} \tanh \left(\sum_i w_{ji} x_i^\mu \right) \right] \right\} \quad (3-7)$$

Uma vez que as atividades de todas as unidades de saída tenham sido determinadas, o programa computacional, desenvolvido para este trabalho, calcula o erro E_w , definido pela expressão:

$$E_w = \frac{1}{2} \sum_{\mu s} (\zeta_s^\mu - \hat{O}_s^\mu)^2, \quad (3-8)$$

onde \hat{O}_s^μ é a atividade de saída da última camada da rede artificial e ζ_s^μ é o dado experimental (resultado conhecido), pertencente ao padrão (μ) apresentado.

Como a saída da rede é uma função explícita dos valores de entrada, bem como dos pesos, nas conexões (isso não ocorre, em geral, nas redes Hopfield), podemos utilizar o método do gradiente descendente, associado ao algoritmo “*backpropagation*”, para minimizar o erro E_w [14, 46].

Para as conexões existentes entre a camada de saída e a segunda camada escondida, da rede, a regra do gradiente descendente nos dá:

$$\begin{aligned} \Delta w_{sl} &= - \frac{\partial E_w}{\partial w_{sl}} \\ &= - \frac{\partial E_w}{\partial \hat{O}_s^\mu} \cdot \frac{\partial \hat{O}_s^\mu}{\partial \Upsilon 3_s^\mu} \cdot \frac{\partial \Upsilon 3_s^\mu}{\partial w_{sl}} \\ &= \sum_\mu \delta_s^\mu V 2_l^\mu, \end{aligned} \quad (3-9)$$

onde, $\delta_s^\mu \equiv \zeta_s^\mu - \hat{O}_s^\mu$ corresponde à diferença entre o dado real (alvo) e o resultado apresentado pela rede.

Aplicando-se a mesma regra para as conexões existentes entre as unidades correspondentes às duas camadas escondidas da rede, obtém-se:

$$\begin{aligned}\Delta w_{lj} &= -\frac{\partial E_w}{\partial w_{lj}} \\ &= -\frac{\partial E_w}{\partial \hat{O}_s^\mu} \cdot \frac{\partial \hat{O}_s^\mu}{\partial \Upsilon 3_s^\mu} \cdot \frac{\partial \Upsilon 3_s^\mu}{\partial V 2_l^\mu} \cdot \frac{\partial V 2_l^\mu}{\partial \Upsilon 2_l^\mu} \cdot \frac{\partial \Upsilon 2_l^\mu}{\partial w_{lj}} \\ &= \sum_\mu \delta_l^\mu V 1_j^\mu,\end{aligned}\quad (3-10)$$

onde, $\delta_l^\mu \equiv \delta_s^\mu g'(\Upsilon 2_l^\mu) w_{sl}$.

Finalmente, trabalhando-se com as conexões existentes entre as unidades da primeira camada escondida e as da camada de entrada, da rede, obtém-se :

$$\begin{aligned}\Delta w_{ji} &= -\frac{\partial E_w}{\partial w_{ji}} \\ &= -\frac{\partial E_w}{\partial \hat{O}_s^\mu} \cdot \frac{\partial \hat{O}_s^\mu}{\partial \Upsilon 3_s^\mu} \cdot \frac{\partial \Upsilon 3_s^\mu}{\partial V 2_l^\mu} \cdot \frac{\partial V 2_l^\mu}{\partial \Upsilon 2_l^\mu} \cdot \frac{\partial \Upsilon 2_l^\mu}{\partial V 1_j^\mu} \cdot \frac{\partial V 1_j^\mu}{\partial \Upsilon 1_j^\mu} \cdot \frac{\partial \Upsilon 1_j^\mu}{\partial w_{ji}} \\ &= \sum_\mu \delta_j^\mu V 0_i^\mu,\end{aligned}\quad (3-11)$$

onde, $\delta_j^\mu \equiv \delta_l^\mu g'(\Upsilon 1_j^\mu) w_{lj}$ e $V 0_i^\mu \equiv x_i^\mu$.

Observe que as equações 3-9, 3-10 e 3-11 têm a mesma forma, embora com diferentes definições para o δ , conforme a situação. De um modo geral, a regra para a atualização dos pesos, pelo "backpropagation", possui a seguinte expressão :

$$\Delta w_{qt} = \sum_{padroes} \delta_{saida} \times V_{entrada} \quad (3-12)$$

onde a *saída* e a *entrada*, nesta última equação, referem-se às extremidades q e t de cada conexão, respectivamente.

A equação 3-11 permite-nos determinar o valor do δ , para cada unidade l , da segunda camada escondida, a partir dos dados experimentais utilizados para alimentar a rede, e dos valores de δ obtidos a partir da camada de saída. De modo análogo, obtém-se os valores de δ para as unidades da primeira camada escondida. Os coeficientes w_{qt} , ou pesos, são os mesmos utilizados na propagação da informação, mas, aqui, eles propagam os erros no sentido contrário ao da informação, o que justifica o nome do método empregado para fazer as atualizações dos pesos: "error-backpropagation".

Neste método, os cálculos das atualizações dos pesos ocorrem através de incrementos: um padrão de dados (μ) é fornecido à entrada da rede e os pesos são então atualizados antes que um próximo padrão seja apresentado. Dessa forma, a cada iteração, a função erro (ou função custo) vai diminuindo, de forma a permitir uma gradativa adaptação, nos próximos passos, ao gradiente local. Escolhendo-se os padrões, dentre os dados que servirão como

exemplo, de forma aleatória, permite-se que se percorra um caminho estocástico no espaço dos pesos, possibilitando uma maior exploração da superfície de custo.

O fato de se poder calcular as derivadas de uma função de custo apropriada, através da retropropagação dos erros, é claramente atrativa, mas existem outras duas vantagens a respeito do citado método: 1- A atualização é local, o que torna a retropropagação apropriada para computação paralela; 2- o trabalho computacional também é menor, diminuindo a quantidade de operações efetuadas.

Para o processo de atualização dos pesos, incluíram-se, ainda, dois outros parâmetros: um *termo inercial* (α) e um *parâmetro de aprendizagem* (η), que são utilizados para acelerar a convergência do programa de aprendizagem, obtendo-se a seguinte expressão [14, 47]:

$$w_{qt}^{\text{nov}} = \alpha w_{qt}^{\text{antigo}} - \eta \Delta w_{qt}, \quad (3-13)$$

onde:

$$\Delta w_{qt} = (\partial E_w) / (\partial w_{qt}),$$

$$0 < \eta \leq 1 \equiv \text{taxa de aprendizado},$$

$$0 < \alpha < 1 \equiv \text{termo inercial}$$

Para se poder avaliar a eficiência do desempenho da rede neural, dois outros parâmetros de averiguação de resultados foram também introduzidos: VRM (*variação relativa da média*) e ρ (*coeficiente de correlação*), dados pelas seguintes expressões, respectivamente [47]:

$$VRM = \frac{\sum_{\mu s} (\zeta_s^\mu - \hat{O}_s^\mu)^2}{\sum_{\mu s} (\zeta_s^\mu - \langle \zeta_s^\mu \rangle)^2} \quad (3-14)$$

$$\rho = \frac{\sum_{\mu s} (\zeta_s^\mu - \langle \zeta_s^\mu \rangle)(\hat{O}_s - \langle \hat{O}_s \rangle)}{\sigma_{\zeta_s^\mu} \sigma_{\hat{O}_s}} \quad (3-15)$$

3.4 Implementação Numérica

Dada uma rede artificial de M camadas, tomemos V_q^m como sendo a saída de uma unidade q , arbitrária, de uma camada m ($m = 1, 2, 3, \dots, M$), qualquer, sendo $V_q^0 \equiv x_i$ e w_{qt}^m a conexão entre as unidades V_i^{m-1} e V_q^m .

Nesta situação, o procedimento adotado para a implementação numérica é simples e obedece à seguinte sequência:

- inicializa-se a matriz dos pesos com números aleatórios e pequenos (menores do que 1), para fugir-se dos mínimos locais.
- apresenta-se um padrão de dados (μ) à rede inserindo-os na primeira camada (camada de entrada).
- propaga-se a informação, *para frente*, através da expressão:

$$V_q^m = g\left(\sum_t w_{qt} V_t^{m-1}\right) = g(\Upsilon_q^m), \quad (3-16)$$

que é aplicada para cada unidade q , da camada m , até que todas as saídas finais (V_q^m) tenham sido obtidas.

- calculam-se os valores de δ , se houver mais do que uma saída, da camada de saída, através da equação:

$$\delta_q^M = \zeta_q^M - V_q^M \quad (3-17)$$

comparando-se a saída da rede (V_q^M) com o dado experimental (ζ_q^M), do padrão (μ), apresentado.

- calculam-se os valores de δ , para as unidades das camadas interiores, propagando-se o erro *para trás* pela expressão:

$$\delta_t^{m-1} = g'(\Upsilon_t^{m-1}) \sum_q w_{qt}^m \delta_q^m, \quad (3-18)$$

para $m = M, M - 1, M - 2$, até que todos os valores de δ tenham sido obtidos para todas as unidades.

- utiliza-se a expressão : $\Delta w_{qt}^m = \delta_q^m V_t^{m-1}$ para atualizar todas as conexões através da expressão:

$$w_{qt}^{m(\text{nov})} = \alpha w_{qt}^{m(\text{antigo})} + \eta \Delta w_{qt}^m \quad (3-19)$$

- repete-se o procedimento, a partir do segundo item, para o próximo padrão (μ), e assim, sucessivamente.

3.4.1 “Biasing”

Como visto, anteriormente, um neurônio biológico consegue ativar um seu vizinho enviando-lhe um sinal excitatório, que deve ser suficientemente maior do que um sinal inibitório, fazendo com que o potencial da célula atinja um determinado limiar (“*threshold*”). Embora esta situação não tenha sido levada em conta, nos cálculos até aqui desenvolvidos, pode-se perfeitamente implementá-la numericamente através do acréscimo de uma unidade extra, com um valor fixo e igual a 1 (um), e conectá-la a todas as demais unidades.

Qualquer unidade, com um “*threshold*” positivo (T) e um certo número de entradas (n), pode ser substituída por uma unidade com um “*threshold*” zero e $n+1$ entradas, onde a *entrada extra* recebe um peso $-T$, que provém de uma outra unidade que está sempre ativada e que produz uma saída sempre igual a -1 . Dessa forma, o “*threshold*” pode ser modificado, a cada iteração, através dos pesos atualizados. Esse truque matemático é conhecido como “*biasing*” e a unidade, que permanece sempre ativada, com o valor -1 , recebe o nome de “*bias*”.

3.4.2 Programas em uso

O programa original, para treinamento e previsão dos sinais contendo disrupturas, foi nos cedido pelo grupo de plasma da Universidade do Texas (Austin - E.U.A.), mas foi por nós profundamente modificado para torná-lo mais compatível com as teorias correntes.

No momento, existem dois programas computacionais, feitos em FORTRAN, em uso no grupo de Plasma do Instituto de Física, da USP: um de treinamento da rede e um outro desenvolvido somente para previsões de séries temporais.

Embora também existam vários programas profissionais envolvendo redes neurais artificiais, como o excelente programa SNNsv4.2 (“*Stuttgart Neural Network Simulator, versão 4.2*”), que pode ser obtido, gratuitamente, pela Internet [48], optou-se por trabalhar com um programa mais *caseiro*, por possibilitar maior intervenção em sua estrutura lógica e agilizar o manuseio do mesmo, uma vez que o tempo de treinamento de uma rede é relativamente longo, principalmente quando ainda se está procurando a configuração ideal de trabalho.

No presente momento, a rede por nós utilizada trabalha com 4 camadas de neurônios, sendo 1 camada de entrada, 2 camadas escondidas e 1 camada de saída. O número de unidades, por camada, varia de acordo com a série temporal em estudo. As funções de ativação utilizadas são do tipo $g(x) = x$, para a unidade neural de saída, e $g(x) = \tanh(x)$

para as demais unidades neurais.

Os parâmetros livres, que o programa permite ao usuário escolher, são: a própria arquitetura da rede, o número de épocas de treinamento, o termo de inércia (α), a taxa de aprendizagem (η), o intervalo de tempo da sequência temporal dos dados experimentais, com os quais se deseja trabalhar, o intervalo de pontos que será tomado para a geração de padrões de treinamento, validação e previsão, quantos (e quais) dados experimentais serão utilizados em todo o processo e, finalmente, o tempo de antecipação (“*time-step*”) com o qual se deseja fazer a previsão da série temporal em estudo.

Efetuada o treinamento, dentre do número de épocas escolhido, o programa seleciona a melhor matriz de pesos, ou seja, escolhe aquele conjunto que apresentou os melhores resultados de previsão nos padrões utilizados para a **validação** da rede. O critério de escolha da matriz ideal baseia-se no *menor erro* (equação 3-8), *menor variação relativa da média* (equação 3-14) e *maior coeficiente de correlação* (equação 3-15) obtidos.

O programa desenvolvido, *teoricamente*, permite usar quantas séries temporais forem desejadas, tanto para o treinamento, quanto para a validação. Entretanto, existe uma limitação de memória para o compilador por nós adotado, que impõe a utilização de, no máximo, 5 pulsos para o treinamento e validação. Já existem, hoje em dia, compiladores FORTRAN com alocação dinâmica de memória, que permitiriam trabalhar com um número maior de pulsos. Mas essa adaptação não foi feita, ainda, porque o estudo da previsão das instabilidades de disrupturas, por nós realizado, ainda não exigiu tal alteração.

Capítulo 4

Previsão de Séries Caóticas com Redes Neurais

Uma necessidade bastante comum, encontrada nas várias áreas da física, é a de se prever a evolução futura de um sistema dinâmico a partir de suas medidas realizadas em tempos passados. O procedimento tradicionalmente seguido é aquele no qual se constrói um modelo matemático do determinado sistema e, a partir deste, obtêm-se as equações de movimento tentando-se, então, integrá-las no tempo para prever a situação futura daquele estado. Matematicamente, pode-se dizer que o estado de um sistema é descrito por um ponto p , num espaço multidimensional Γ . As evoluções dinâmicas do sistema são, então, descritas pelo movimento de p nesse espaço Γ .

A aplicação deste procedimento complica-se muito, entretanto, quando se trabalha com sistemas não-lineares com muitos graus de liberdade, como o plasma, por exemplo, pois torna-se impraticável resolver explicitamente todas as equações da dinâmica, a não ser que várias simplificações sejam feitas. Não se consegue, simplesmente, traçar o movimento num espaço de tantas dimensões !

Estudos da dinâmica de sistemas aparentemente caóticos, com muitos graus de liberdade, entretanto, revelam que, em havendo dissipação, o número desses graus de liberdade pode ser reduzido para um valor efetivamente relevante, o que significa dizer que, depois de um certo tempo, o espaço Γ dos possíveis estados do sistema reduz-se para um sub-espaço Γ_A , com uma dimensionalidade menor, denominada *atrator*. Caso se consiga, de algum modo, identificar as variáveis que caracterizam esse atrator Γ_A , a questão fica então reduzida a um problema físico simples, com alguns poucos graus de liberdade e, portanto,

passível de resolução [14, 49].

Existe um teorema, rigorosamente provado por Takens, que afirma o seguinte: dada uma medida temporal de um determinado evento, existe uma função suave, de no máximo $2d+1$ valores passados (d é a dimensão do atrator) da série medida, que corretamente prevê o valor futuro da variável em questão; e esta previsão é tão boa quanto à que obteríamos, se fôssemos capazes de resolver o sistema completo, com seus muitos graus de liberdade. Mais ainda, se os valores medidos puderem ser obtidos com uma precisão infinita, então o resultado será insensível, tanto ao intervalo de tempo, tomado entre as medidas passadas, quanto ao tempo de antecipação, com o qual se pretende prever o estado futuro daquele evento. Na prática, porém, ruídos e imprecisões nas medidas limitam o quão livremente essas quantidades possam ser escolhidas [50].

Sabemos, portanto, que, em princípio, é perfeitamente possível reduzir o problema de previsão de um sistema dinâmico complexo, cujo movimento resida num atrator de menor dimensão, para um problema físico mais simples. O que o teorema de Takens não nos fornece é a *forma explícita da função que executa a extrapolação desejada*. E é neste momento que a rede neural artificial entra em cena.

A idéia básica é utilizar uma rede neural, tipo “*feedforward*”, usando-se como padrão de entrada ($\{x_i^\mu\}$, da equação 3-1) os valores passados (*coordenadas de atraso*) da série temporal $\{x(t), x(t - \tau), \dots\}$; e o valor conhecido, $x(t + \Lambda)$, como o alvo da rede (o valor ζ_s^μ , da equação 3-8). Matematicamente:

$$x(t), x(t - \tau), x(t - 2\tau), \dots, x(t - (m - 1)\tau) \longrightarrow x(t + \Lambda), \quad (4-1)$$

onde m^1 é o número de coordenadas de atraso, utilizadas para alimentar a rede, e Λ é o tempo de antecedência, com o qual se deseja prever o estado futuro do sistema dinâmico [14].

4.1 Reconstruindo Estados com Coordenadas de Atraso

A reconstrução de estados, a partir de suas coordenadas de atraso, é um artifício gráfico interessante, conhecido como “*delay plot*”, que permite inferir a dimensão do atrator de

¹Importante notar que a variável m , aqui apresentada, não guarda nenhuma relação com o *modo* m , descrito no capítulo 2 !

um dado sistema [49]. Para poder descrever esse procedimento, comecemos por definir um estado, que é *toda informação necessária para descrever o sistema num instante futuro*.

Para se obter o “*delay plot*” de uma série temporal, basta fazer o gráfico da série em função de sua própria versão, atrasada no tempo, onde, para cada par ordenado (x, y) do gráfico, x corresponde à S_t , valor da série S , no instante t , e y corresponde à $S_{t-\tau}$, valor da série S , no instante de tempo t , atrasado de τ .

Processos periódicos passam por muitos estados diferentes, mas todos caem numa curva unidimensional fechada, no “*delay plot*”. Neste tipo de movimento, o sistema está restrito a um subconjunto unidimensional do seu espaço de estados possíveis. Entretanto, o uso das coordenadas de atraso não se restringe a gráficos bidimensionais. Em geral, pode-se obter um “*delay plot*”, de dimensão m , registrando o vetor das coordenadas de atraso $\{S(t), S(t - \tau), S(t - 2\tau), \dots, S(t - (m - 1)\tau)\}$ para cada instante (t) da série temporal. Atratores, que são substancialmente mais complicados do que simples curvas fechadas, precisarão de mais dimensões para serem descobertos [49].

Muitas vezes, a curva obtida pode apresentar auto-intersecções, provocando uma ambiguidade na interpretação da trajetória, no ponto de interseção. Outra medida, portanto, é necessária para se descobrir completamente o atrator.

4.2 O Conceito de Imersão (“*Embedology*”)

Suponha que R^k seja o espaço dos estados de um sistema dinâmico, e que as trajetórias desse espaço sejam atraídas para uma variedade A , de dimensão d . Seja R^m o espaço de reconstrução. Se nós pudermos fazer várias medidas m , simultâneas e independentes, a qualquer instante de tempo t , teremos então um vetor, em R^m , para cada um dos estados do sistema. Sendo F a função que representa o processo de medidas realizadas, então, para qualquer instante de tempo t , onde o estado é um ponto de A , em R^k , nós poderemos calcular F , naquele ponto, fazendo m medidas e construindo o vetor que a represente.

Existe um teorema que diz o seguinte [49]: *Se A for uma variedade, de dimensão d , em R^k , e se $m > 2d$ e $F : R^k \rightarrow R^m$ for genérica, então F é biunívoca em A .* Em outras palavras, dois estados diferentes em A permanecem diferentes quando mapeados em R^m , o que significa dizer que $F(A)$ não pode apresentar auto-intersecções! Se isto acontecer, é porque estamos trabalhando numa dimensão inferior à necessária para visualizar a função F . A função contínua e biunívoca, que leva um ponto do conjunto R^k para R^m , é chamada

função de imersão desse conjunto, ou, às vezes, *topologia de imersão*, e o número m é chamado *dimensão de imersão*. É essa a função que se espera que a rede neural artificial seja capaz de descobrir a partir dos dados experimentais à ela apresentados.

Convém ressaltar que o teorema descrito não proíbe uma dimensão de imersão menor do que $2d + 1$, mas, simplesmente diz que $2d + 1$ é o suficiente para casos genéricos. Mais tarde, foi demonstrado que é suficiente escolher a função F a partir de uma classe especial de funções obtidas estritamente a partir das coordenadas de atraso, de reconstrução, usando uma série temporal de uma única medida [50].

Do exposto, podemos destacar três informações importantes:

- a propriedade biunívoca implica que todo estado, no espaço de estados do sistema, pode ser univocamente representado através de medidas realizadas sobre ele.
- um atrator de dimensão finita pode sempre ser imerso em algum R^m .
- um atrator está imerso num espaço de dimensão m quando a curva, obtida com as coordenadas de atraso, não apresenta auto-intersecções.

É importante notar que, se os estados vizinhos, do sistema em estudo, divergirem exponencialmente no tempo, um erro experimental poderá aumentar essa diferença, tornando-a comparável ao tamanho do próprio atrator. Se isto acontecer durante o intervalo de tempo abrangido por um único vetor de coordenadas de atraso, então o estado que este vetor especifica terá sua valia limitada para aplicações do tipo previsão. Ou seja, a previsão é limitada para sistemas caóticos.

4.3 Sistemas Testados

Para investigar a viabilidade de se fazer previsões de séries temporais com redes neurais, utilizando-se os conceitos até aqui desenvolvidos, resolvemos, primeiramente, testar essas idéias com exemplos amplamente conhecidos: o atrator de Lorenz, o mapa de Henon e o mapa logístico.

4.3.1 Atrator de Lorenz

Um meteorologista, chamado Lorenz, elaborou um modelo atmosférico, bastante simplificado, para descrever o movimento de convecção de um fluido aquecido em baixo e

esfriado em cima, como ocorre na atmosfera, chegando às seguintes equações diferenciais ordinárias, não-lineares, de primeira ordem [51]:

$$\begin{aligned}\dot{x} &= -\sigma x + \sigma y \\ \dot{y} &= -xz + rx - y \\ \dot{z} &= xy - bz,\end{aligned}\tag{4-2}$$

onde σ (número de Prandtl), r (número de Rayleigh) e b são parâmetros do sistema.

No presente trabalho, as séries temporais escolhidas para alimentar a rede foram aquelas relacionadas às soluções em x , do sistema descrito em 4-2, com os seguintes parâmetros: $\sigma = 3,0$, $r = 26,5$ e $b = 1,0$, por exemplo.

Inicialmente, um exaustivo conjunto de testes foram tentados com inúmeras séries, obtidas a partir de diferentes condições iniciais (x_0, y_0, z_0) , trabalhando-se com diferentes tipos de arquiteturas em diferentes condições de treinamento e previsão.

Os resultados aqui apresentados referem-se a um particular grupo, arbitrariamente escolhido, no qual utilizou-se a série $(1,0,0)$ para o treinamento da rede, a série $(0,2,0)$ para a validação e a série $(1,1,1)$, sobre a qual se fez a previsão (figura 4-1).

Das várias arquiteturas tentadas, as que ofereceram os melhores resultados foram aquelas para as quais se utilizou uma configuração próxima à relação $m:2m:m:1$, onde m é a dimensão de imersão do sistema dinâmico em estudo. Desta forma, um chute inicial para se trabalhar com as soluções em x , do sistema de Lorenz, descrito em 4-2, foi $3:6:3:1$, uma vez que é sabido que, para este sistema, a dimensão de imersão é $m = 3$ [49]. As condições de treinamento da rede artificial foram as seguintes: número de épocas de treinamento = 100; $\alpha = 0$ (zero); $\eta = 0,001$ e $\tau = 5$.

Com este tipo de arquitetura, conseguiu-se prever, com bastante sucesso, os pontos futuros da série temporal $(1,1,1)$ para diversos *steps* escolhidos. Até o *step* 15 ($\Lambda = 0,3$), por exemplo, o resultado do processo de previsão é praticamente perfeito, quando comparado aos dados reais, conforme observado na figura 4-2a. Para o *step* 25 ($\Lambda = 0,5$), embora o ajuste entre a previsão da rede neural e os dados reais não sejam tão perfeitos (figura 4-2b), o padrão da série de Lorenz continua sendo muito bem reproduzido, podendo-se considerar esse resultado também muito bom. *Koga*, por exemplo, conseguiu fazer previsões desse tipo de série **somente até 3 steps** ($\Lambda = 0,03$), para uma taxa de amostragem de 0,01 [15]. Embora ele também tenha utilizado redes neurais artificiais para fazer a previsão, a estrutura, por ele adotada, é diferente desta que aqui apresentamos. *Diambra*, por sua vez, conseguiu fazer previsões, com uma taxa de amostragem das séries

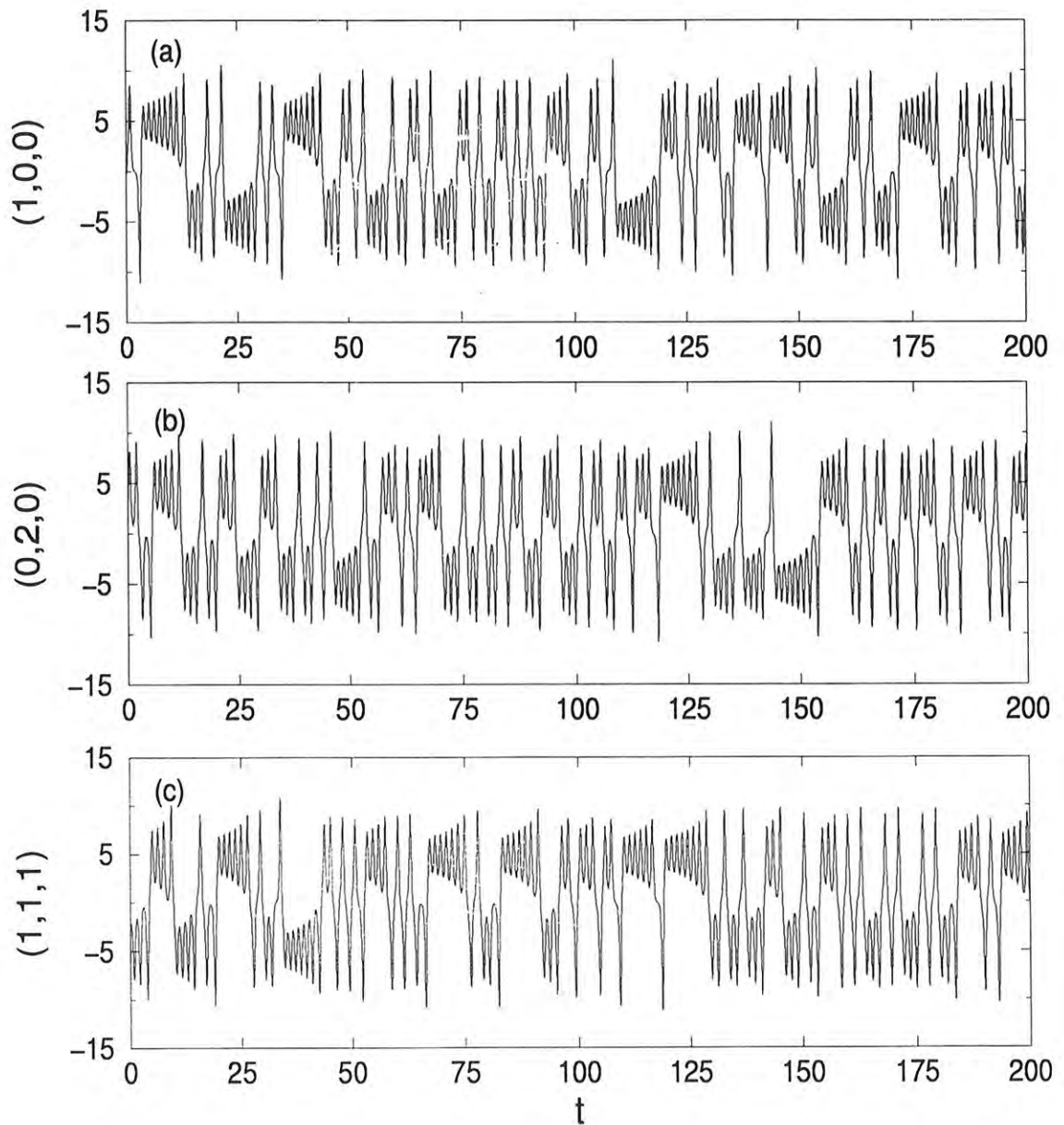


Figura 4-1: Soluções do sistema de Lorenz, em x , utilizadas para (a) treinamento, (b) validação da rede neural e (c) previsão.

utilizadas de 0,02, **somente até 10 steps** ($\Lambda = 0,2$), utilizando um outro modelo, baseado em *teoria de informações* [16].

Para *steps* maiores do que 25, os resultados começam a ficar distorcidos, mas, ainda, o padrão da série continua sendo bem representado, como pode ser observado pela figura 4-2c, na qual é apresentada a previsão da rede neural para 45 *steps* ($\Lambda = 0,9$).

Uma boa maneira de se avaliar a performance da rede neural, quanto às suas previsões, é registrar os valores de *VRM* (equação 3-14) e ρ (equação 3-15) em função dos *steps* utilizados para fazer as antecipações no tempo. Os resultados obtidos para a arquitetura 3:6:3:1 são mostrados na figura 4-3, na qual se pode observar que a degradação nas previsões tende a aumentar exponencialmente à medida em que se aumentam os *steps* de previsão.

Como dito anteriormente, outras arquiteturas, com outras proporções entre as unidades utilizadas pela rede neural, foram também testadas, porém, não com melhores resultados daqueles aqui apresentados. A figura 4-4 mostra os valores do parâmetro *VRM* obtidos para as previsões da série (1,1,1) em função de vários *steps* para diferentes arquiteturas. Como se pode notar, aquela que forneceu os menores erros foi a arquitetura 3:6:3:1.

4.3.2 Mapa de Henon

Um outro teste da performance da rede, quanto à previsão de séries caóticas, foi efetuado com o mapa de Henon que, como se sabe, é dado por [49]:

$$\begin{aligned}x_{n+1} &= a - x_n^2 + by_n \\ y_{n+1} &= x_n\end{aligned}\tag{4-3}$$

Várias séries temporais da variável y , por exemplo, foram geradas a partir das equações acima, com diversas condições iniciais (x_0, y_0) arbitrariamente escolhidas, onde somente os parâmetros $a = 1,4$ e $b = 0,3$ foram mantidos fixos. A figura 4-5, por exemplo, mostra uma das séries obtidas em y .

Através do “*delay plot*”, conseguiu-se construir o atrator desse sistema com $\tau = 1$, para o qual a dimensão de imersão encontrada foi $m = 2$. A primeira arquitetura testada foi, portanto, 2:4:2:1, com os seguintes parâmetros de treinamento: número de épocas de treinamento = 100; $\alpha = 0$ (zero); $\eta = 0,001$; $\tau = 1$. Nestas condições, o melhor resultado

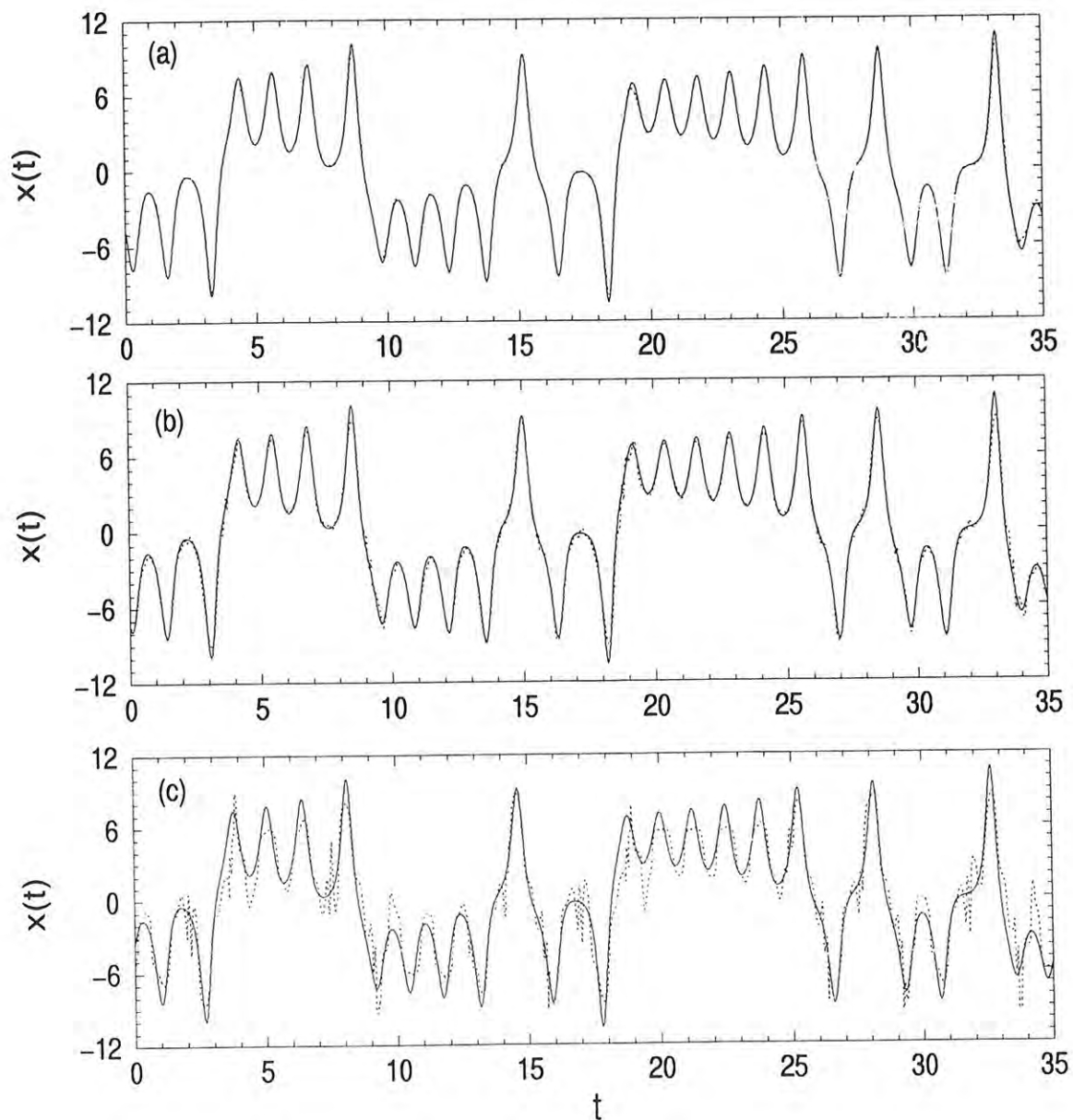


Figura 4-2: Previsão da série de Lorenz (1,1,1) usando-se a arquitetura 3:6:3:1 (linha pontilhada) comparada com os dados reais (linha sólida) para (a) 15 steps ($\Lambda = 0, 3$), (b) 25 steps ($\Lambda = 0, 5$) e (c) 45 steps ($\Lambda = 0, 9$). Como pode ser observado, a previsão é praticamente perfeita para 15 steps, o que indica que a rede neural aprendeu o sistema dinâmico, em questão.

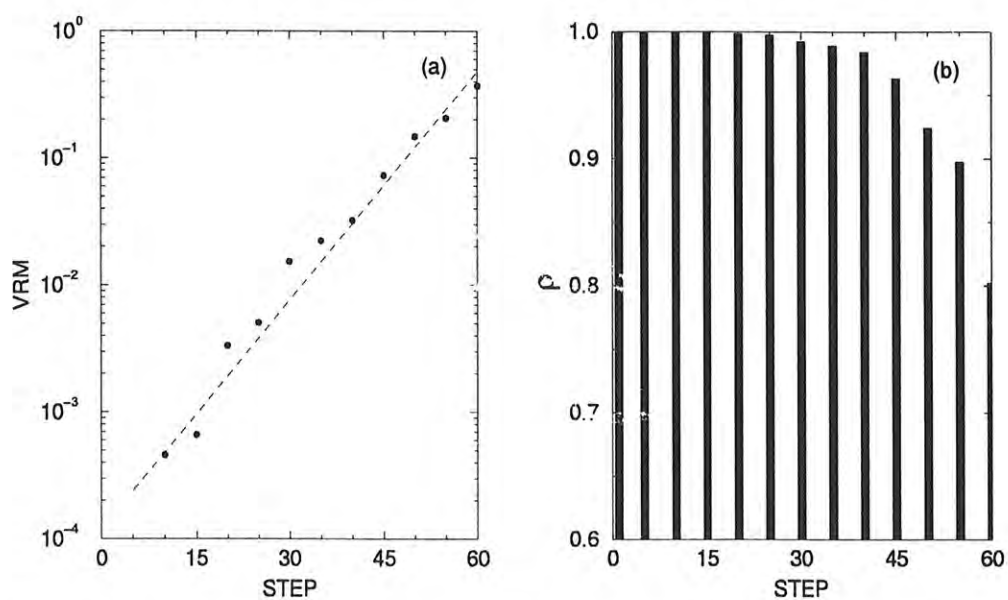


Figura 4-3: Gráficos mostrando (a) o parâmetro VRM e (b) o coeficiente de correlação, ρ , em função do *steps* utilizados para se fazer as antecipações na série temporal, em x , do sistema de Lorenz. Como se pode notar, o resultado da previsão, fornecido pela rede, degrada-se exponencialmente à medida em que se aumentam os *steps* de previsão.

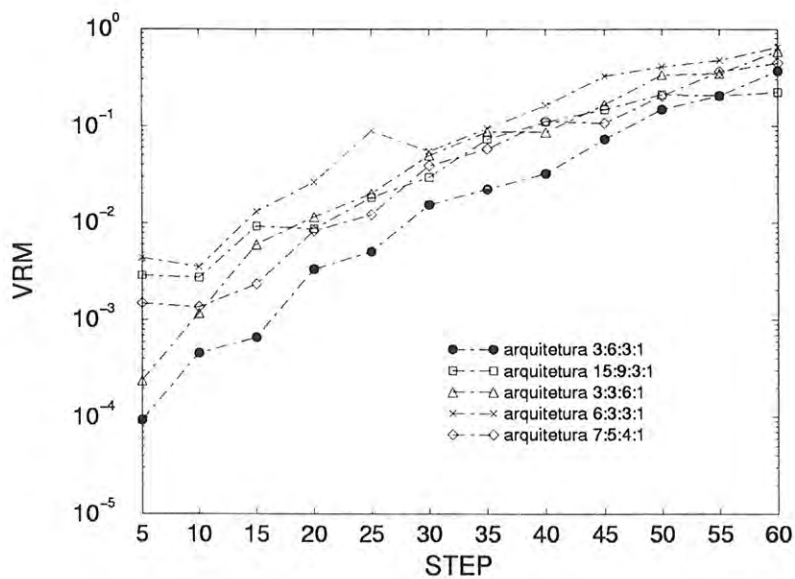


Figura 4-4: Diferentes arquiteturas utilizadas para a previsão da série (1,1,1), do sistema de Lorenz. Como se pode observar, a configuração que apresenta os melhores resultados é 3:6:3:1.

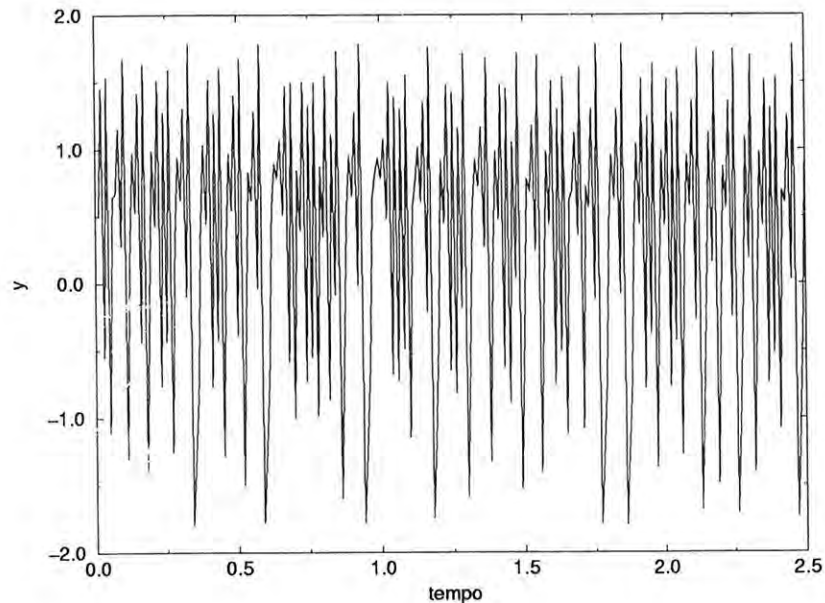


Figura 4-5: Série, em y , do mapa de Henon, obtida a partir da condição inicial $(0,5; 1,0)$, com os parâmetros $a = 1,4$ e $b = 0,3$.

obtido foi somente para o *step* até 3 ($\Lambda = 0,03$).

A figura 4-6 mostra o resultado da previsão da rede para série $(-0,5; 1,0)$, cujos treinamento e validação foram efetuados, respectivamente, com a série temporal $(0,5; 1,0)$ e a série $(0,6; 1,2)$.

Como pode ser observado na figura 4-6a, os resultados obtidos não foram plenamente satisfatórios, uma vez que a rede, de arquitetura $2:4:2:1$, não conseguiu ajustar perfeitamente os dados da série $(-0,5; 1,0)$. Diante disso, tentou-se, novamente, utilizar outras arquiteturas buscando melhorar a performance da previsão. A tabela 4-1 mostra alguns dos resultados obtidos, através da qual pode-se observar que o melhor resultado foi para a arquitetura $6:12:6:1$, em cujo treinamento se utilizou $\alpha = 0,9$.

É interessante notar que o simples fato de se aumentar a quantidade de unidades nas camadas da rede neural não implica diretamente numa melhora de sua performance. Os parâmetros de previsão apresentados para a rede neural com arquitetura $7:14:7:1$, por exemplo, não foram melhores daqueles obtidos com a arquitetura $6:12:6:1$.

A (figura 4-6b) mostra os resultados da previsão fornecidos pela rede neural de arquitetura $6:12:6:1$ realizados sobre a série $(-0,5; 1,0)$, onde se observa uma melhor performance

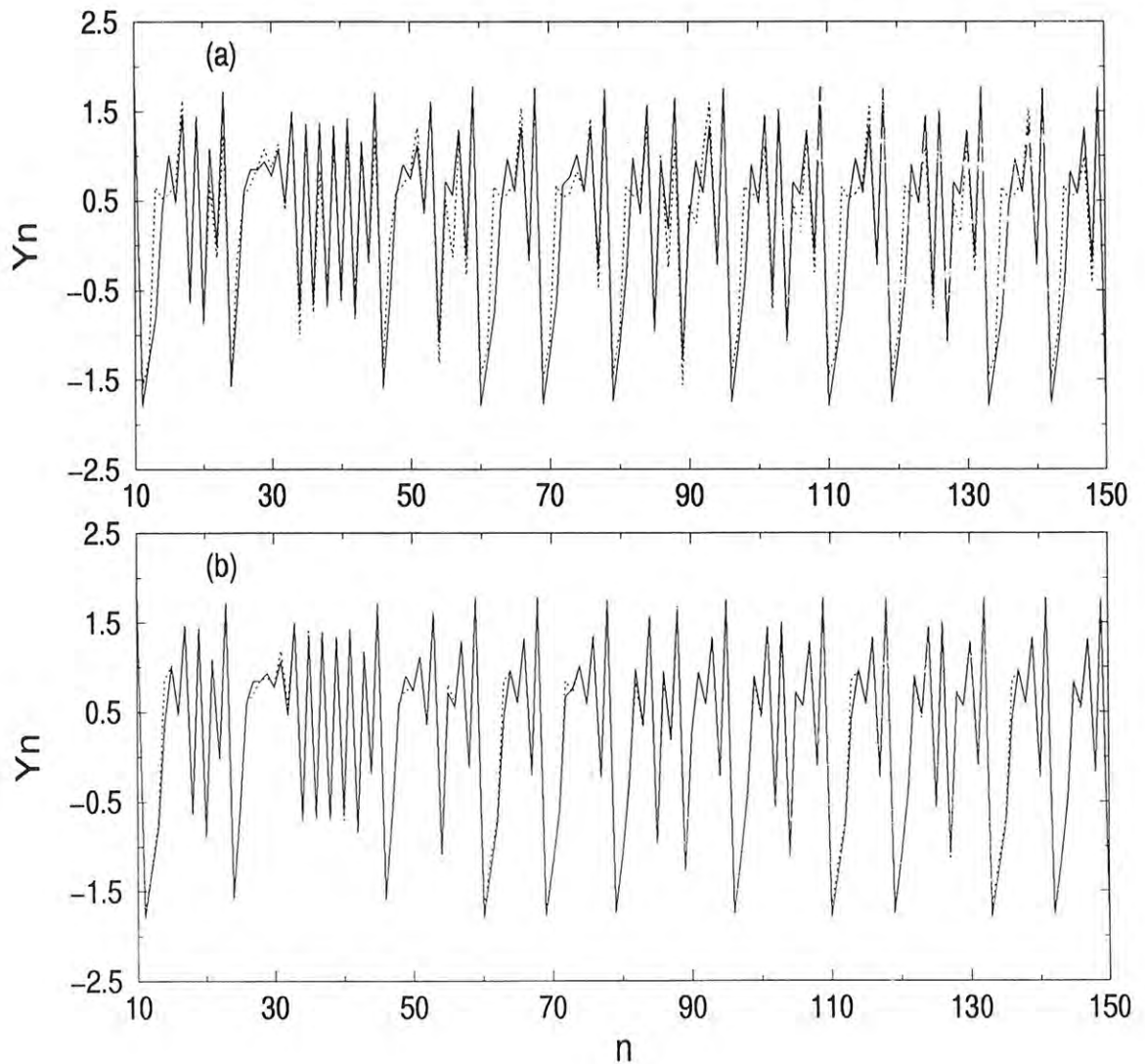


Figura 4-6: Comparação entre os resultados das previsões da rede neural artificial (linha pontilhada), em relação aos dados reais (linha sólida), realizadas na série $(-0,5; 1,0)$ para as arquiteturas (a) $2:4:2:1$ e (b) $6:12:6:1$, com *step* de previsão igual a 3.

Arquitetura	Melhor época	Erro	VRM	ρ	Treinamento	Validação	Teste
2-4-2-1	98	0,1176	0,1157	0,9404	(0,5; 1,0)	(0,6; 1,2)	(-0,5; 1,0)
5-10-5-1	99	0,0450	0,0440	0,9778	(0,5; 1,0)	(0,6; 1,2)	(-0,5; 1,0)
6-12-6-1	99	0,0191	0,0188	0,9905	(0,5; 1,0)	(0,6; 1,2)	(-0,5; 1,0)
7-14-7-1	99	0,0402	0,0395	0,9800	(0,5; 1,0)	(0,6; 1,2)	(-0,5; 1,0)
6-12-6-1	100	0,0184	0,0181	0,9900	(0,7; 0,2)	(1,0; -0,5)	(-1,0; 1,5)
3-6-3-1	100	0,0550	0,0540	0,9723	(0,7; 0,2)	(1,0; -0,5)	(-1,0; 1,5)
3-6-3-1 (*)	92	0,0117	0,0117	0,9941	(0,7; 0,2)	(1,0; -0,5)	(-1,0; 1,5)
6-12-6-1(*)	93	0,0056	0,0056	0,9972	(0,5; 1,0)	(0,6; 1,2)	(-0,5; 1,0)

Tabela 4-1: Diferentes arquiteturas testadas, nas mesmas condições de treinamento, com exceção dos resultados em (*), para os quais se utilizou $\alpha = 0,9$. Todos os testes foram realizados com STEP=3 de previsão.

da rede, muito embora o melhor *step* de previsão tenha sido novamente 3.

4.3.3 Mapa Logístico

Finalmente, como um último teste do desempenho da rede neural na previsão de sistemas caóticos, utilizou-se o mapa logístico, cujas séries temporais foram obtidas a partir da iteração da seguinte equação [49]:

$$x_{n+1} = ax_n(1 - x_n). \quad (4-4)$$

Neste teste, em particular, trabalhou-se com dois tipos de mapas: um obtido a partir do parâmetro $a = 3,7$, para o qual o coeficiente de Lyapunov é dado por $k = 0,35$, e um outro bem mais caótico ($k = 0,69$), no qual se utilizou o parâmetro $a = 4,0$ [49].

Com estes sistemas, obteve-se bons resultados de previsão da rede para até 2 *steps* ($\Lambda = 0,02$), com o mapa $a = 4,0$; e 3 *steps* ($\Lambda = 0,03$), com o mapa $a = 3,7$; utilizando-se, em ambos os mapas, uma rede com arquitetura 4:8:4:1. Embora a dimensão de imersão para esses sistemas seja $m = 2$, o que sugeriria uma arquitetura inicial 2:4:2:1 para se trabalhar, os resultados de previsão obtidos por esse tipo de rede não foram muito satisfatórios, forçando a busca de uma outra combinação mais apropriada. Neste estudo do mapa logístico, teve-se que adotar, ainda, um número de épocas bem maior, em torno de ~ 500 , para se poder *fugir* dos mínimos locais [13, 14]. A figura 4-7, por exemplo, mostra os valores do parâmetro *VRM* em função do número de épocas adotado para o treinamento da rede neural de arquitetura 4:8:4:1. Para o treinamento, utilizou-se a série

temporal (0,2) obtida a partir da iteração do mapa logístico $a = 3,7$ (para a validação deste sistema utilizou-se a série (0,3)). Como se pode observar, pela figura, os melhores resultados de previsão só puderam ser obtidos a partir de ~ 200 épocas de treinamento. Interessante observar que, para ambos os mapas, obteve-se a mesma qualidade de previsão para séries obtidas a partir de diferentes condições iniciais (x_0), não havendo a necessidade de novos treinamentos. Esses resultados revelam que, uma vez que a rede neural tenha aprendido o sistema dinâmico em questão, ela é capaz de efetuar o mesmo tipo de previsão sobre quaisquer outras séries temporais pertencentes àquela mesma dinâmica, o que já não ocorre com outros métodos de previsão usualmente adotados em sistemas caóticos, como o método da *aproximação local*, por exemplo [52].

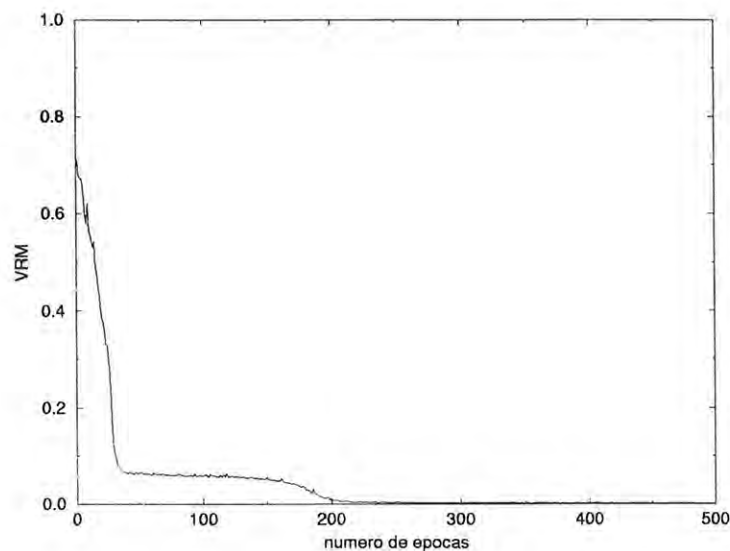


Figura 4-7: Valores do parâmetro VRM em função do número de épocas adotado no treinamento da rede neural artificial para o mapa logístico $a = 3,7$. Observe que a rede converge para os menores valores de VRM somente a partir de 200 épocas

A figura 4-8 mostra o resultado da previsão da rede neural, de arquitetura $4:8:4:1$, para os dois mapas aqui utilizados. Observa-se que, embora o ajuste entre os dados reais e aqueles fornecidos pela rede seja praticamente perfeito, sua capacidade de antecipação no tempo torna-se mais limitada à medida em que se trabalha com sistemas cada vez mais caóticos. Para o mapa $a = 4,0$, por exemplo, a rede foi capaz de antecipar apenas 2 steps, enquanto que, para o mapa $a = 3,7$, a rede foi capaz de antecipar 3 steps.

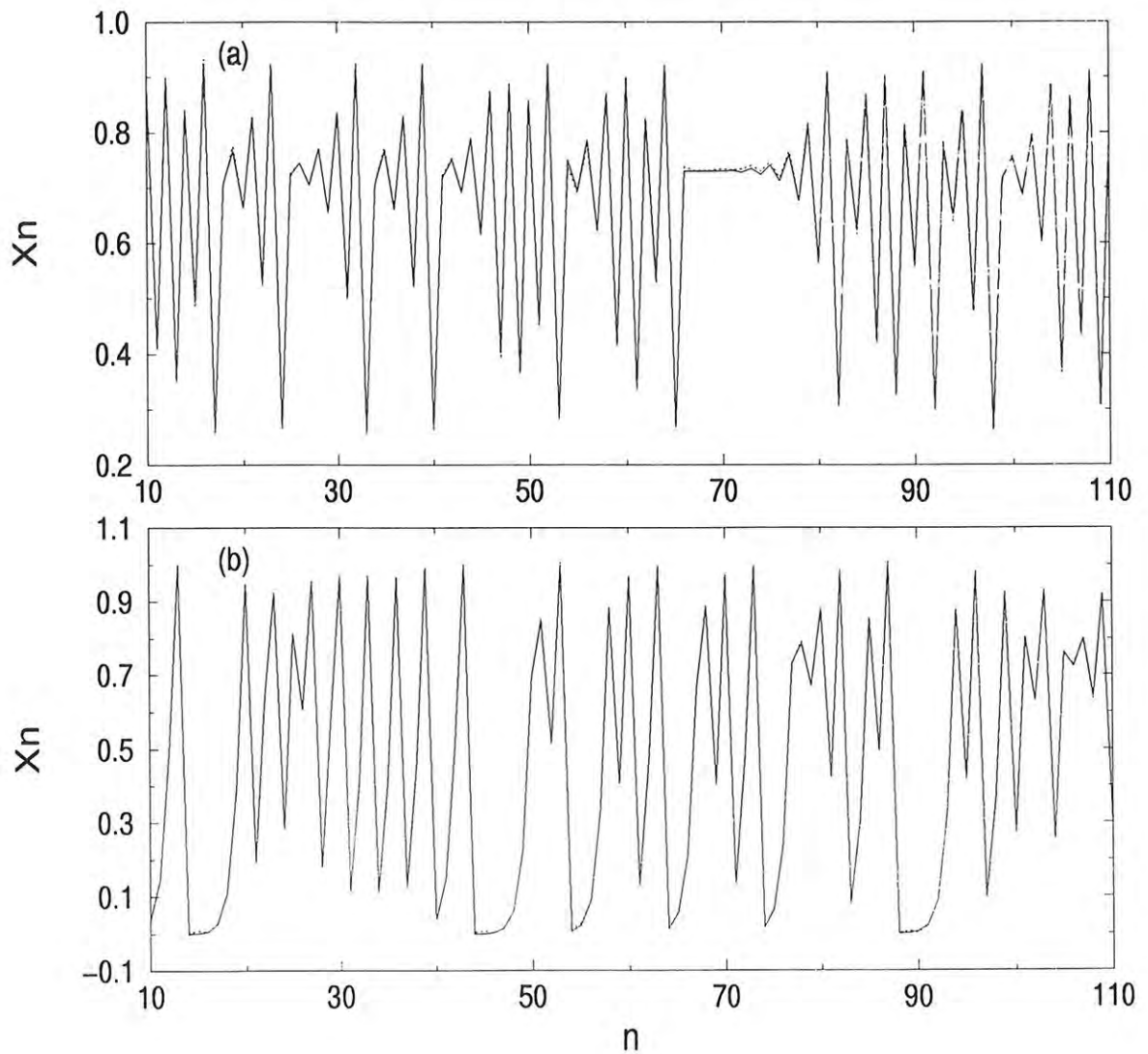


Figura 4-8: Comparação entre os resultados das previsões da rede neural artificial (linha pontilhada) em relação aos dados reais (linha sólida) realizadas para os mapas logísticos (a) $a = 3,7$ ($step = 3$) e (b) $a = 4,0$ ($step = 2$). Como se observa, o resultado de previsão é praticamente perfeito.

Capítulo 5

Redes Neurais na Previsão das Disrupturas

Em função dos bons resultados obtidos na previsão das instabilidades de disruptura do tokamak TEXT, usando-se, como sinais de entrada da rede neural, as oscilações magnéticas das bobinas de Mirnov [47], ocorreu-se a idéia de se trabalhar com os sinais de raios-X na tentativa de melhorar o tempo de previsão o máximo possível. Estima-se que um tempo de previsão da ordem de 3 – 4 *ms* talvez seja o suficiente para permitir que medidas preventivas, tais como injeção de partículas neutras (ou “*pellets*”), aplicação de campos helicoidais ressonantes, etc, possam ser tomadas no sentido de se evitar a ocorrência das instabilidades de disrupturas ou, pelo menos, atenuar os seus efeitos nocivos.

Com os sinais das bobinas magnéticas, conseguiu-se antecipar o instante da disruptura das descargas do TEXT em até 1 *ms*, e o motivo de se ter escolhido, desta vez, o diagnóstico de raios-X, para tentar aumentar esse tempo de previsão, dá-se pelo fato de eles terem sido medidos na região central da coluna de plasma, onde se supõe que o fenômeno da disruptura poderia começar a se desencadear.

Para o tratamento dos sinais de raios-X, utilizando-se redes neurais artificiais, adotaram-se, inicialmente, as condições de treinamento semelhantes às utilizadas por *Hernandez e Vannucci* [47], quando do tratamento dos sinais das bobinas magnéticas, que diferem bastante da abordagem descrita no capítulo anterior. Isso porque somente algum tempo depois, após um árduo trabalho na tentativa de melhorar o tempo de previsão das disrupturas, é que se descobriu que a melhor maneira de se utilizar a rede neural artificial seria aquela que levasse em consideração a dinâmica do sistema, em estudo, no momento de se

confeccionar a sua arquitetura, bem como gerar seus padrões de treinamento e previsão.

Todo o processo adotado na busca das melhores condições de previsão destas instabilidades será descrito, neste capítulo, na sequência em que foi efetuado.

As figuras 5-1, 5-2 e 5-3, mostram os sinais básicos dos pulsos de plasma do tokamak TEXT e, dentre eles, os sinais de raios-X, provenientes da região central da coluna de plasma, utilizados no presente estudo da previsão das instabilidades de disrupturas.

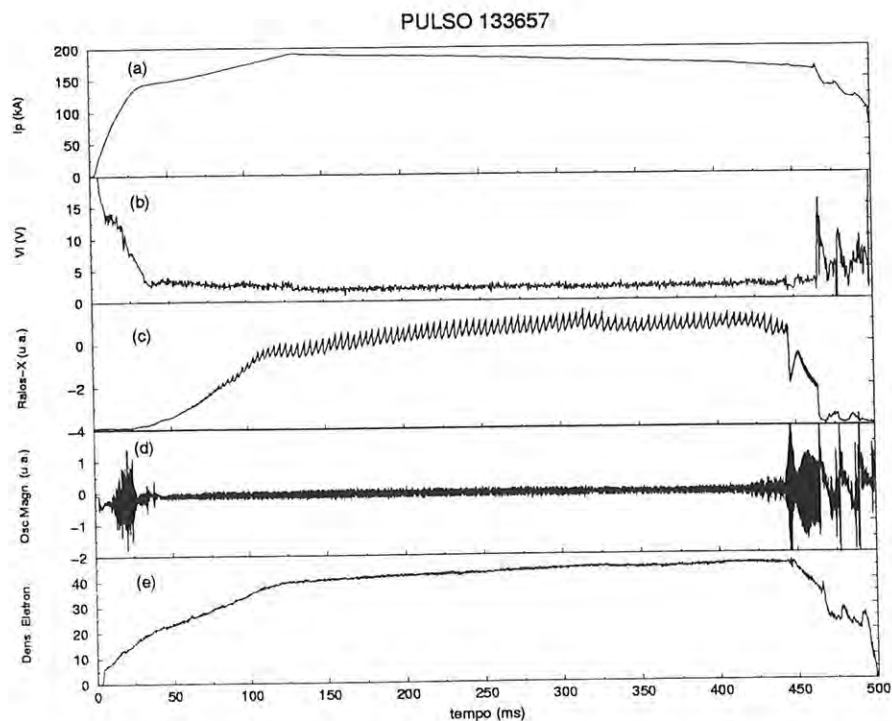


Figura 5-1: Sinais básicos correspondentes a uma descarga do tokamak TEXT: (a) corrente de plasma, (b) tensão de enlace, (c) sinais de raios-X, (d) oscilações magnéticas de Mirnov e (e) densidade eletrônica, em 10^{12} cm^{-3} .

5.1 Processo de Previsão Contínua

O processo de “*previsão contínua*” foi aqui denominado para diferenciá-lo de um outro tipo de previsão, “*por step*”, também utilizado no presente trabalho. Na previsão contínua, a

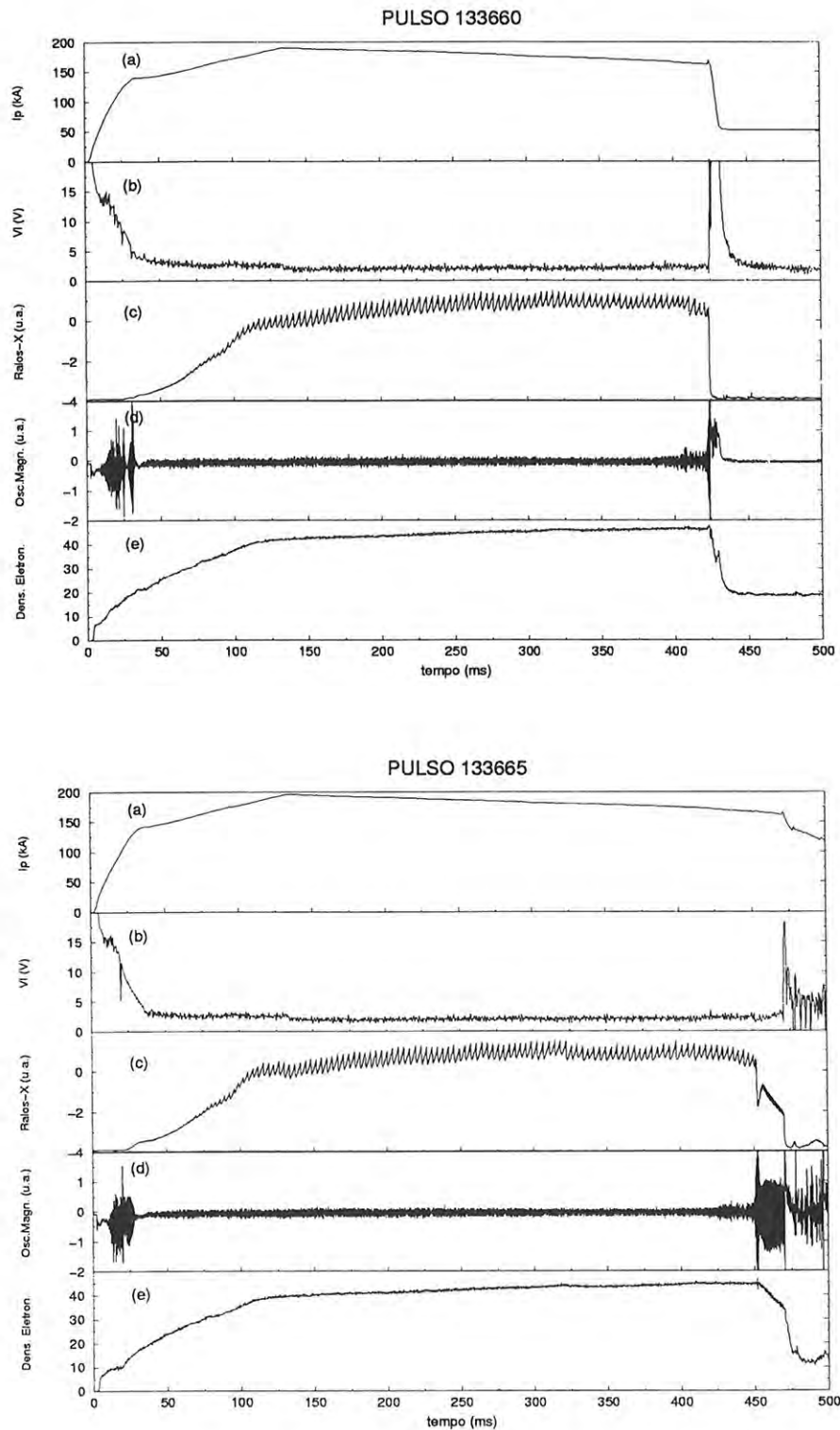


Figura 5-2: Sinais básicos correspondentes a duas descargas do tokamak TEXT: (a) corrente de plasma, (b) tensão de enlace, (c) sinais de raios-X, (d) oscilações magnéticas de Mirnov e (e) densidade eletrônica, em 10^{12}cm^{-3} .

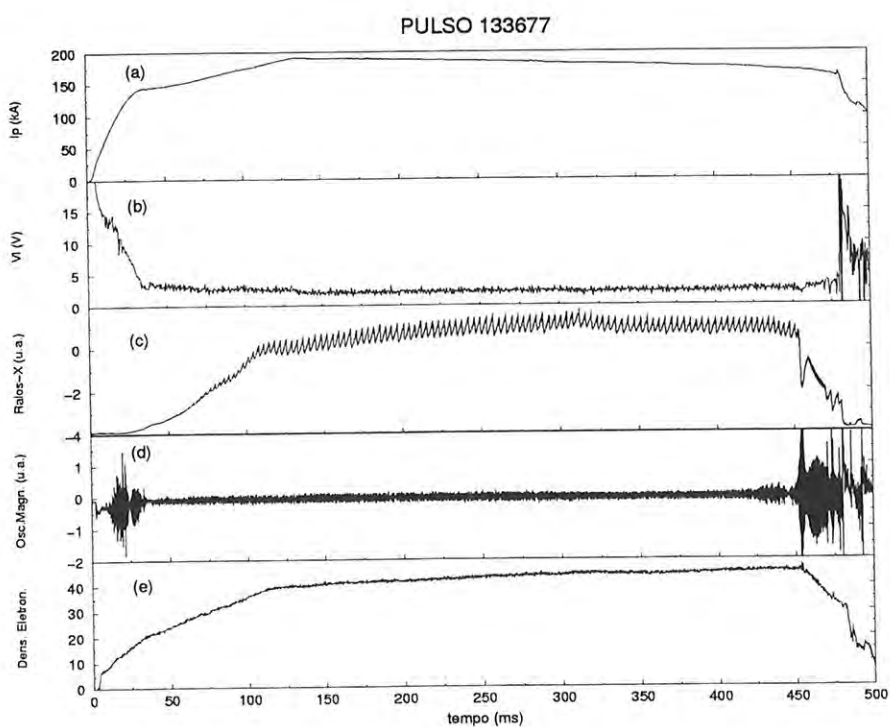
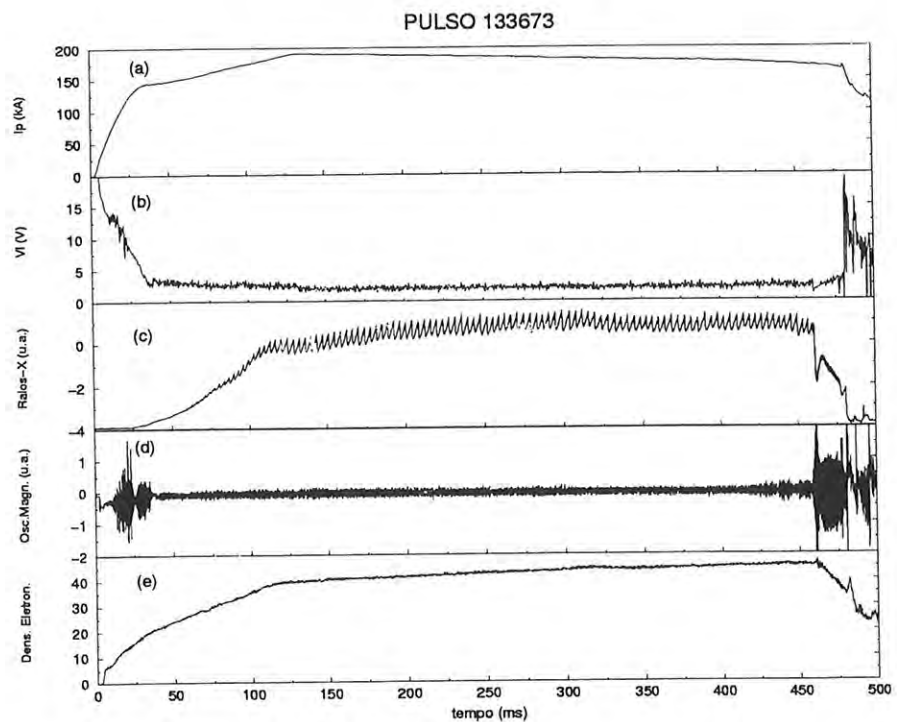


Figura 5-3: Sinais básicos correspondentes a duas descargas do tokamak TEXT: (a) corrente de plasma, (b) tensão de enlace, (c) sinais de raios-X, (d) oscilações magnéticas de Mirnov e (e) densidade eletrônica, em $10^{12}cm^{-3}$.

rede sempre trabalha com pontos consecutivos, passados, no sentido de prever o instante posterior da série a partir da realimentação da rede com pontos oriundos de sua própria previsão, para os casos nos quais o *step* seja maior do que 1. Para ficar mais claro, utilizemos um exemplo: seja $R(t)$ o conjunto de pontos passados, de uma determinada série temporal, digamos, um sinal de raio-X, que é utilizado para alimentar uma rede neural que convenha 6 unidades em sua camada de entrada, e $\hat{R}(t)$ o resultado da previsão dessa rede. Para uma previsão de um único *step*, teríamos a seguinte configuração:

$$R(1), R(2), R(3), R(4), R(5), R(6) \longrightarrow \hat{R}(7).$$

Para uma previsão de 5 *steps*, por exemplo, a rede seria realimentada, para cada iteração de treinamento, da seguinte forma:

$$R(1), R(2), R(3), R(4), R(5), R(6) \longrightarrow \hat{R}(7)$$

$$R(2), R(3), R(4), R(5), R(6), \hat{R}(7) \longrightarrow \hat{R}(8)$$

$$R(3), R(4), R(5), R(6), \hat{R}(7), \hat{R}(8) \longrightarrow \hat{R}(9)$$

$$R(4), R(5), R(6), \hat{R}(7), \hat{R}(8), \hat{R}(9) \longrightarrow \hat{R}(10)$$

$$R(5), R(6), \hat{R}(7), \hat{R}(8), \hat{R}(9), \hat{R}(10) \longrightarrow \hat{R}(11),$$

ou seja, ao final do processo, dos seis dados utilizados para alimentar a rede, quatro seriam originados sempre da sua própria previsão.

5.1.1 Arquitetura 15:9:3:1

Como já mencionado anteriormente, a tentativa de se estender o tempo de previsão das instabilidades de disruptura, utilizando-se sinais de raios-X moles, foi iniciada adotando-se uma rede tipo “*feedforward*”, de duas camadas escondidas, com uma arquitetura 15:9:3:1, e previsão contínua, idêntica à utilizada no tratamento dos sinais das bobinas de Mirnov por Hernandez e Vannucci [47].

Para o treinamento da rede, utilizou-se o pulso rx133657 (fig.5-1) e, para a validação, rx133673 (fig.5-3), ambos disruptivos. Os padrões de treinamento foram escolhidos aleatoriamente, pelo próprio programa, dentro de um intervalo de tempo referente aos últimos 200 *ms* dos sinais experimentais. O treinamento foi efetuado apenas com 35 épocas e com os seguintes parâmetros de treinamento: $\alpha = 0,9$ e $\eta = 0,001$.

A figura 5-2 mostra os sinais básicos do pulso disruptivo 133665, no qual se utilizou a rede neural para se fazer a previsão contínua da disruptura. Este pulso corresponde a uma descarga com corrente de plasma $I_p \sim 170 \text{ kA}$, cuja disruptura ocorreu em $t \sim 470 \text{ ms}$. Cerca de 18 *ms* antes desta disruptura maior, observa-se a ocorrência de uma disruptura

menor (sinal d), responsável pela sensível queda na densidade média dos elétrons (sinal e) e na temperatura eletrônica, verificada pela emissão dos raios-X (sinal c). Esse mesmo comportamento, ou seja, uma disruptura menor precedendo uma disruptura maior, instantes antes do colapso da corrente de plasma, também é observado nas descargas de plasma utilizadas no treinamento e validação da rede neural.

O resultado do processo de previsão pela rede neural, dessa particular descarga de plasma, é mostrada na figura 5-4 para vários *steps* diferentes. Como se pode observar, para um *step* de previsão (fig.5-4a), o resultado obtido da rede neural ajusta-se perfeitamente aos dados experimentais de raios-X. Aumentando-se o número de *steps* de previsão para 25 (fig.5-4c), por exemplo, observa-se que a forma do sinal dente-de-serra deforma-se um pouco, mas a rede continua sendo capaz de detectar a disruptura menor, que ocorre em $t \sim 452 \text{ ms}$. O tempo mais longo de previsão da disruptura que se conseguiu obter foi com um *step* igual a 63 (fig.5-4d) para o qual a rede neural foi capaz de prever, com precisão, o instante de ocorrência da disruptura menor. Isto corresponde a uma previsão de 2,52 *ms* antes que ela de fato ocorra no tokamak TEXT. Já para *steps* maiores, a rede apresenta um atraso na previsão em relação ao instante em que de fato a disruptura ocorre (fig.5-4e). É interessante notar que, mesmo alimentando a rede apenas com trechos não disruptivos do sinal, ela não apresenta *alarmes falsos*, indicando que o conjunto de pesos encontrado pela rede, durante o processo de treinamento, foi adequado.

Examinando em detalhe o resultado obtido para a previsão com 63 *steps* (fig.5-4d), observa-se que, para fazer a previsão, a rede utiliza 15 pontos experimentais que estão posicionados anteriormente ao instante $t \sim 449,5 \text{ ms}$. Este particular conjunto de pontos está localizado exatamente na região oscilatória da última oscilação típica de dente-de-serra, quando ainda não há sinais visíveis de que alguma instabilidade viria a ocorrer [11]. Somente após a queda deste dente-de-serra, em $t \sim 450,0 \text{ ms}$, quando uma forte oscilação começa a ocorrer, é que se pode dizer que se deu início à uma instabilidade disruptiva.

Como um outro teste da rede neural, utilizou-se uma segunda descarga de plasma (pulso 133660, da figura 5-2) para se fazer a previsão, adotando-se o mesmo conjunto de pesos (mesmo treinamento) usados na previsão do pulso 133665 (fig.5-2). Essa descarga também corresponde à uma corrente de plasma de $t \sim 170 \text{ kA}$, cuja disruptura ocorre em torno de $t \sim 424 \text{ ms}$. Interessante observar que, diferentemente da primeira descarga analisada, este particular pulso apresenta apenas a ocorrência de uma disruptura maior, sem qualquer outra instabilidade precursora.

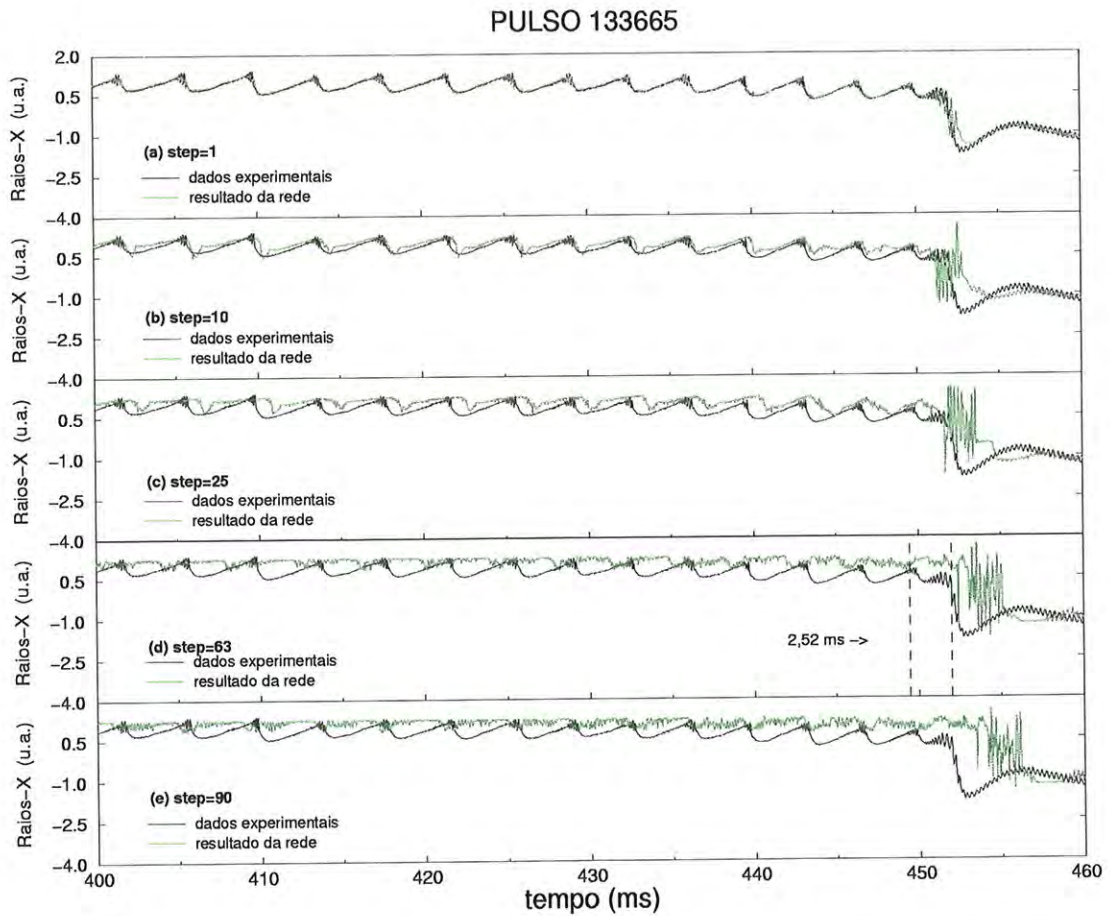


Figura 5-4: Resultado da previsão da rede, para o pulso de plasma 133665 (fig.5-2), para diferentes *steps*. Depois de devidamente treinada, a rede neural é capaz de prever a ocorrência da disruptura com até 2,52 *ms* de antecedência, o que corresponde a um *step*=63 (d). Para *steps* maiores, um atraso entre o resultado fornecido pela rede e os dados experimentais é observado (e).

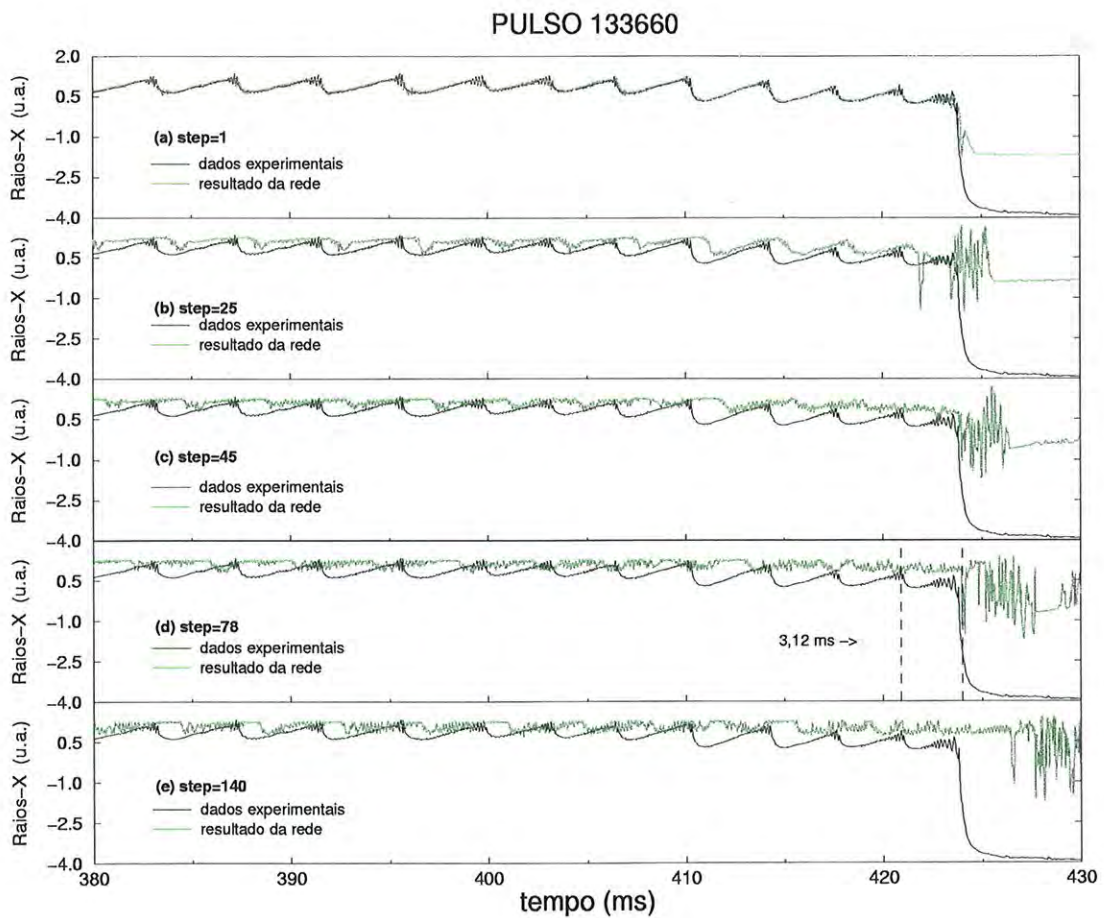


Figura 5-5: Resultado da previsão da rede para o pulso 133660 (fig.5-2). Neste caso, a rede foi capaz de prever a disruptura do sinal com uma antecedência de 3,12 ms, correspondendo à 78 steps (d). Novamente, para steps maiores, um atraso entre o resultado da rede e os dados experimentais é observado (e).

Os resultados de previsão para esse segundo caso é mostrado na figura 5-5. Novamente, para $step = 1$, observa-se um bom ajuste entre os dados fornecidos pela rede e os pontos experimentais (fig.5-5a). Para $steps$ maiores, a rede neural continua fazendo uma boa previsão da disruptura maior até o $step = 78$ (fig.5-5d), que corresponde à uma antecipação do instante da disruptura em $3,12 ms$. Este resultado, por nós obtido, é 3 vezes maior do que o conseguido anteriormente com sinais de bobinas magnéticas [47], e foi objeto de publicação em revista científica [11]. Para $steps$ maiores de previsão, no entanto, uma defasagem entre o sinal medido e o resultado fornecido pela rede novamente começa a aparecer (fig.5-5e).

Analisando-se com maior cuidado este último resultado, pode-se verificar que os pontos experimentais, utilizados pela rede neural para fazer a previsão, correspondem à um intervalo de tempo anterior a $t \sim 420,6 ms$, ou seja, $0,5 ms$ antes da queda do último dente-de-serra típico, após o qual uma forte oscilação MHD, precursora, é também observada nos sinais de Mirnov correspondentes.

É interessante notar que, em ambas as previsões, os pontos experimentais utilizados pela rede neural, para antecipar as disrupturas menor (fig5-4) e maior (fig5-5), estavam localizados num tempo anterior ao aumento das amplitudes das flutuações magnéticas dos correspondentes sinais das bobinas de Mirnov, como pode ser observado nas figuras 5-6 e 5-7, respectivamente.

Como sabemos, a grande desvantagem em se trabalhar com redes neurais é que não nos fica transparente a maneira pela qual os resultados são por ela obtidos. Consequentemente, no nosso caso, nenhuma hipótese pode ser diretamente levantada no sentido de procurar saber qual mecanismo pode ser o maior responsável pelo desencadeamento das instabilidades de disruptura. Não obstante, uma vez que a rede neural foi capaz de prever estas instabilidades utilizando os dados experimentais dos sinais de raios-X, em instantes de tempo anteriores ao do crescimento das amplitudas da atividade MHD, esta observação permite-nos especular que, talvez, não seja o crescimento das ilhas magnéticas, relacionadas à superfície magnética $q = 2$, que poderia, conseqüentemente, perturbar a ilha interna na superfície $q = 1$, como comumente aceito [33, 34].

Pelos resultados que se obteve, poder-se-ia talvez supor que algum tipo de distúrbio ocorreria **primeiro** no centro da coluna de plasma (monitorado pelos detectores de raios-X moles) e, **depois**, a região do plasma, dentro da superfície magnética $q = 2$ (monitorada pelas bobinas magnéticas), é que seria então afetada, desestabilizando o modo (MHD)

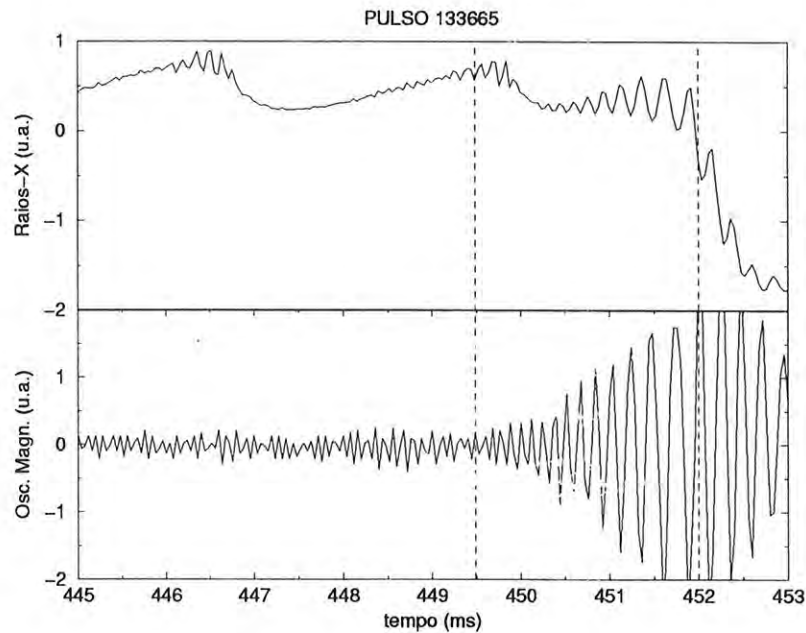


Figura 5-6: Ampliação dos sinais de raios-X moles (a) e das bobinas magnéticas de Mirnov (b), para o pulso 133665 (fig.5-2), instantes antes da ocorrência da disruptura menor ($t \sim 452$ ms). Como pode ser observado, os dados experimentais utilizados pela rede neural, em sua melhor performance, estão localizados numa região justamente anterior ao início do crescimento da amplitude da atividade MHD.

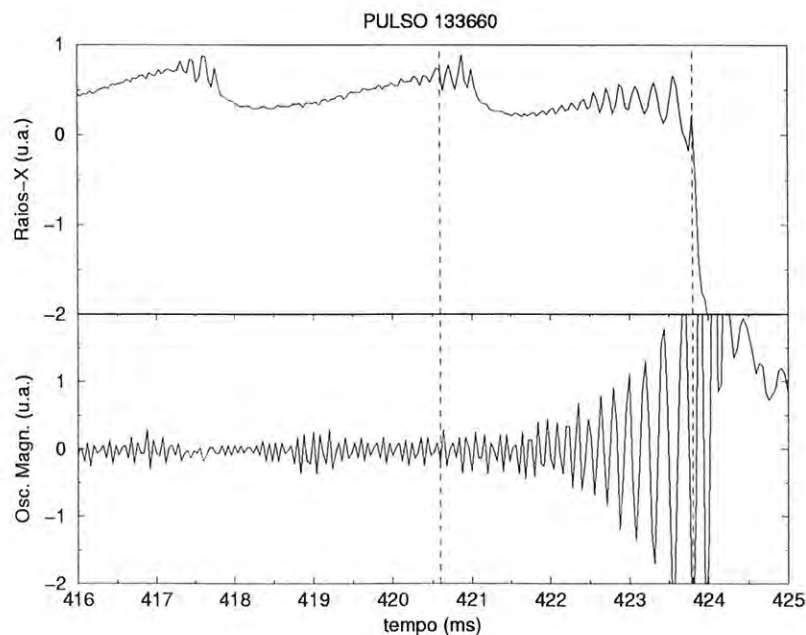


Figura 5-7: Ampliação dos sinais de raios-X moles (a) e das bobinas magnéticas de Mirnov (b), para pulso 133660 (fig.5-2), instantes antes da ocorrência da disruptura maior ($t \sim 424$ ms). Mais uma vez, observa-se que os dados experimentais utilizados pela rede neural, em sua melhor performance, também estão localizados numa região justamente anterior ao início do crescimento da amplitude da atividade MHD.

$m = 2$, desencadeando, conseqüentemente, a disruptura. Esta interpretação, inclusive, reporta-nos a alguns trabalhos publicados que sugerem as instabilidades de disruptura como sendo causadas por uma *bolha fria* movendo-se na direção do centro do plasma, como já pôde ser experimentalmente observado no perfil da temperatura eletrônica de plasma [36, 54]. De qualquer modo, mais estudos se fazem necessários antes de qualquer afirmação mais pereemptória.

Estes resultados obtidos com a rede neural, e a conseqüente discussão dos mesmos, já foram objeto de uma publicação na revista *Nuclear Fusion* [11].

5.1.2 Prevendo a descarga inteira

Uma vez que os pesos obtidos durante o treinamento e validação da rede neural referem-se aos últimos 200 *ms* das descargas, seria natural trabalhar nessa mesma faixa de tempo com os sinais de previsão. Entretanto, utilizando esses mesmos pesos numa descarga inteira de plasma, ao invés de adotar apenas os últimos 200 *ms*, conforme o treinamento, verificaram-se grandes distorções no resultado fornecido pela rede, o que não é de se surpreender, uma vez que os dados experimentais, correspondentes à região de subida dos sinais de raios-X (começo da descarga), não foram apresentados à rede, nem no treinamento, nem na validação. Por isso, concluiu-se que se deve sempre procurar trabalhar, em termos de previsão, com dados que estejam em condições semelhantes àqueles utilizados na fase de treinamento e validação da rede neural.

Como o resultado não foi satisfatório nas previsões das descargas inteiras do plasma, procurou-se modificar os parâmetros de treinamento da rede para adaptá-la à essa nova condição desejada. Desta vez, utilizaram-se os dois pulsos inteiros (133657 e 133665), no treinamento, e um terceiro (133673), também inteiro, na validação da rede. A figura 5-8 mostra os resultados obtidos da previsão, onde se observa um ajuste quase perfeito entre os resultados fornecidos pela rede e os dados experimentais (fig5-8a) para o *step* = 1. Já para o *step* = 13 (fig5-8e), a saída da rede exhibe muitas oscilações na fase de subida dos sinais de raios-X, e oscilações ainda maiores imediatamente após a disruptura menor. Entretanto, analisando-se este último resultado mais detalhadamente (fig5-9), verifica-se que a rede ainda é capaz de prever o instante da disruptura com uma antecedência de até 0,52 *ms*. Infelizmente, não se conseguiu estender o tempo de previsão além deste obtido, indicando que, talvez, os pesos mais apropriados para se trabalhar com os sinais de raios-X, para o pulso todo, ainda não tivessem sido encontrados. Conseqüentemente,

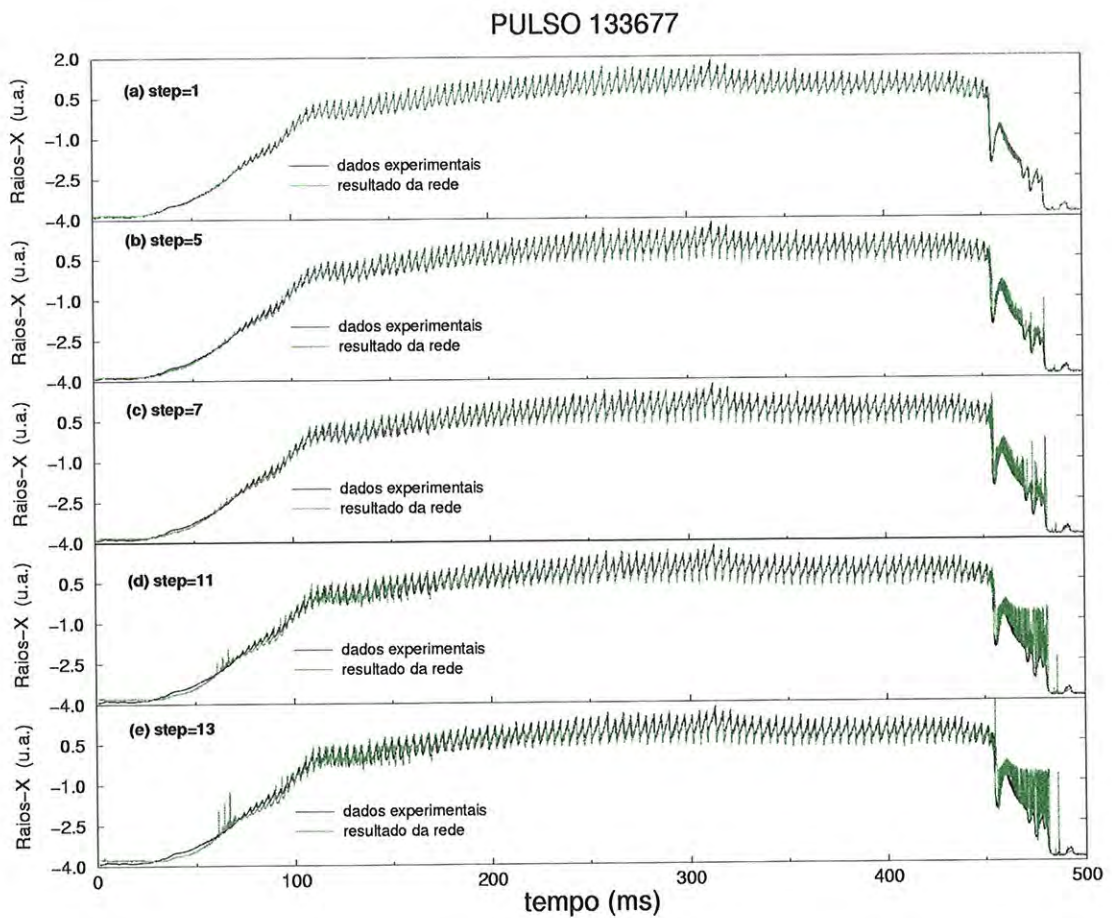


Figura 5-8: Previsão do pulso rx133677, inteiro, para vários *steps* diferentes. Como se pode observar, a previsão para *step* = 1 (a) é praticamente perfeita. Para o *step* = 13 (e), embora a rede continue prevendo o instante da disruptura de forma bastante satisfatória, a sua reprodução do sinal não chega a ser perfeita, apresentando várias oscilações no início e no fim da descarga.

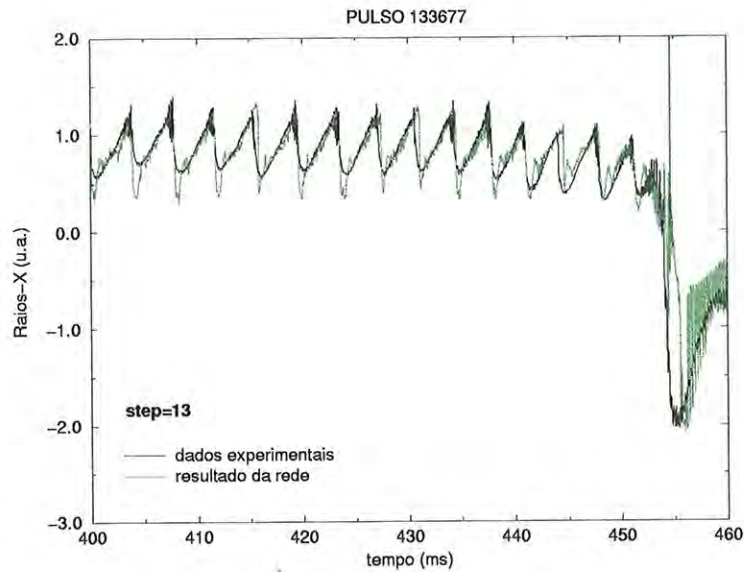


Figura 5-9: Detalhe da figura 5-8 perto da região da disruptura do sinal, na qual a previsão foi realizada com $step = 13$. Não se conseguiu estender o tempo da previsão, usando-se pulsos inteiros, com esta arquitetura 15:9:3:1 de rede.

nesta fase dos testes, concluiu-se que um número maior de dados experimentais talvez fosse necessário para treinar a rede, ou, dever-se-ia aumentar o número de épocas de treinamento, procurando-se, também, melhores parâmetros de treinamento (α , η , etc), ou, até mesmo, buscar um outro tipo de arquitetura.

5.1.3 Buscando novas arquiteturas

Diante da necessidade de se trabalhar com pulsos inteiros, uma vez que nunca se sabe, ao certo, em que momento as instabilidades de disrupturas de fato se desencadearão, procurou-se então tentar solucionar o problema da deformação dos sinais fornecidos pela rede como uma primeira tentativa de otimizar a sua performance.

Para tentar entender porque a rede deforma tanto os sinais oriundos de sua previsão, decidiu-se trabalhar com uma série bem simples. Optou-se, por exemplo, em utilizar uma função bem comportada, tipo *seno*, e com uma frequência próxima à dos sinais de raios-X, conforme a equação descrita, a seguir:

$$y(t) = a + b \sin(\omega t + c), \quad (5-1)$$

onde $a = 1$, $b = 0,5$, $c = 0$ e $w = 1,5708$.

Essa mesma série foi utilizada, tanto para treinamento e validação da rede, quanto para o processo de previsão. A figura 5-10 mostra os resultados obtidos com esse teste. Como se pode observar, a performance da rede vai degradando à medida em que se utilizam *steps* cada vez maiores para se fazer a antecipação da série, no tempo. Para o *step* 75 (figura 5-10f), por exemplo, o sinal é apresentado totalmente deformado (problema semelhante ao observado com os sinais de raios-X). Diante deste resultado curioso e inesperado, tendo em vista o fato de que a rede estava trabalhando com uma *série periódica*, pensou-se inicialmente que o problema, para tamanha deformação dos sinais, pudesse estar associado, talvez, à maneira como a rede faz as previsões. Para o caso em que se usam muitos *steps* de previsão, o que ocorre é que a rede acaba utilizando muitos pontos, que ela mesma previu, na sua própria realimentação, o que acaba gerando um acúmulo de erros.

Tomando-se, por exemplo, a antecipação em 75 *steps* de uma série temporal qualquer, dos 60 estágios de previsão que a rede de arquitetura 15:9:3:1 calcula, apenas os 15 primeiros estágios contêm dados reais da série. Nos outros 45 estágios restantes, todos os pontos utilizados para a realimentação da rede são resultados de sua própria previsão, o que compromete, em muito, a qualidade final do resultado obtido.

Uma primeira maneira de se alterar essa proporção, de forma a aumentar a quantidade de dados reais da série temporal na realimentação da rede neural, foi modificar a arquitetura dessa rede aumentando-se o número de unidades da sua camada de entrada.

Embora aparentemente simples, procurar uma outra arquitetura de trabalho não é nada trivial, pois requer uma série de tentativas que, na maioria das vezes, resultam infrutíferas. Isso porque a simples modificação de uma única unidade, em uma das camadas da rede neural, pode alterar completamente o resultado final da previsão, como pode ser observado na figura 5-11, na qual são apresentados os resultados, bastante diferentes, de duas redes neurais artificiais para o mesmo sinal de raios-X¹: uma com arquitetura 14:9:3:1 (figura 5-11a), e outra com arquitetura 15:9:3:1 (figura 5-11b). Observa-se também que os resultados são muito ruins, pois ainda não se está trabalhando com a configuração correta de rede para pulsos inteiros.

Esses dados ilustram, de maneira clara, a dificuldade com a qual se depara ao se

¹Todo o procedimento de treinamento e validação, em ambas as redes, foram mantidos os mesmos a fim de que se pudesse avaliar o efeito do número das unidades neuronais na performance de suas previsões.

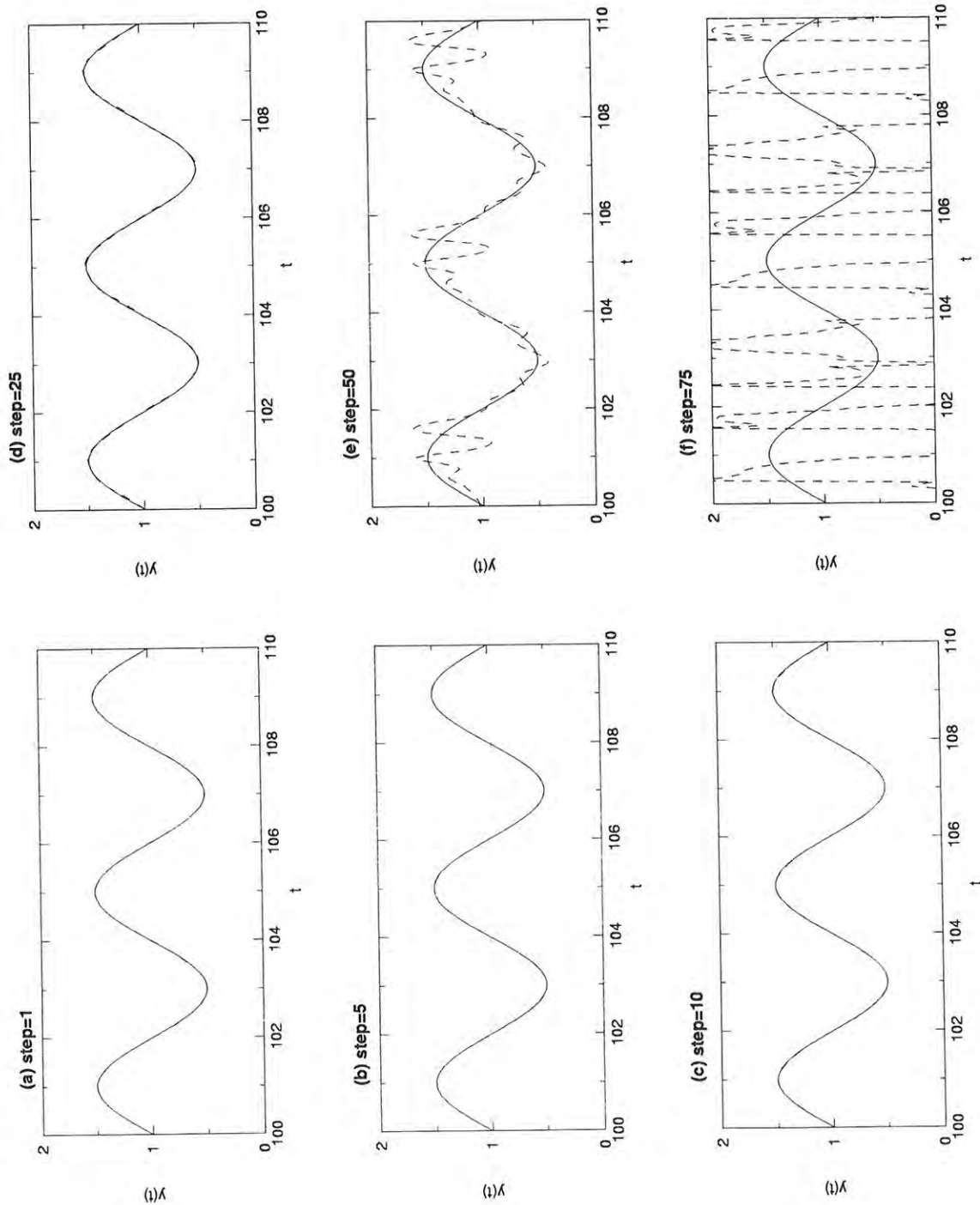


Figura 5-10: Previsão da rede neural, de arquitetura 15:9:3:1 (linha pontilhada), para a função teste $y(t) = a + b \sin(\omega t + c)$ (linha sólida), onde $a = 1$, $b = 0,5$, $c = 0$ e $\omega = 1,5708$. Observa-se que a performance da rede vai degradando à medida em que se aumentam os *steps* de previsão.

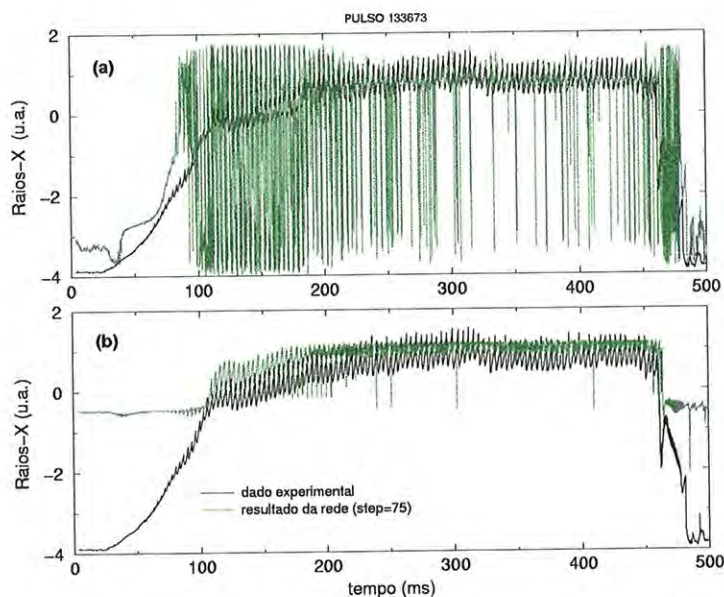


Figura 5-11: Comparação entre os resultados das previsões das redes de arquitetura (a) $14:9:3:1$ e (b) $15:9:3:1$. Como se pode observar, a simples alteração de uma única unidade neuronal, em uma das camadas da rede, modifica bastante o resultado final obtido.

trabalhar com esses tipo de sistema. E é devida à essa grande sensibilidade das redes neurais, em relação à quantidade de unidades utilizadas em suas camadas, que se procurou investigar outras possíveis arquiteturas no intuito de melhorar o desempenho desejado para as previsões de séries temporais.

Nesse contexto, após várias tentativas, chegou-se à arquitetura $33:17:33:1$, que apresentou excelentes resultados de previsão, sem nenhuma deformação do sinal, para a mesma função *seno* testada com a arquitetura $15:9:3:1$ (figura 5-12).

Embora o desempenho dessa rede, com uma nova arquitetura, tenha sido consideravelmente superior no tratamento da função teste (tipo *seno*), o mesmo não ocorreu com os dados experimentais de raios-X. Na figura 5-13a, por exemplo, observa-se um considerável atraso na tentativa de antecipação da disruptura, bem como uma grande deformação do sinal, no início da descarga.

Diante desse problema, pensou-se que uma possível explicação para esse comportamento observado poderia ser a pouca permanência, ainda, dos dados experimentais nos estágios de treinamento da rede, para *steps* grandes (~ 75), apesar de se terem aumentado os números de unidades de entrada de 15 para 33. Dessa forma, para se garantir

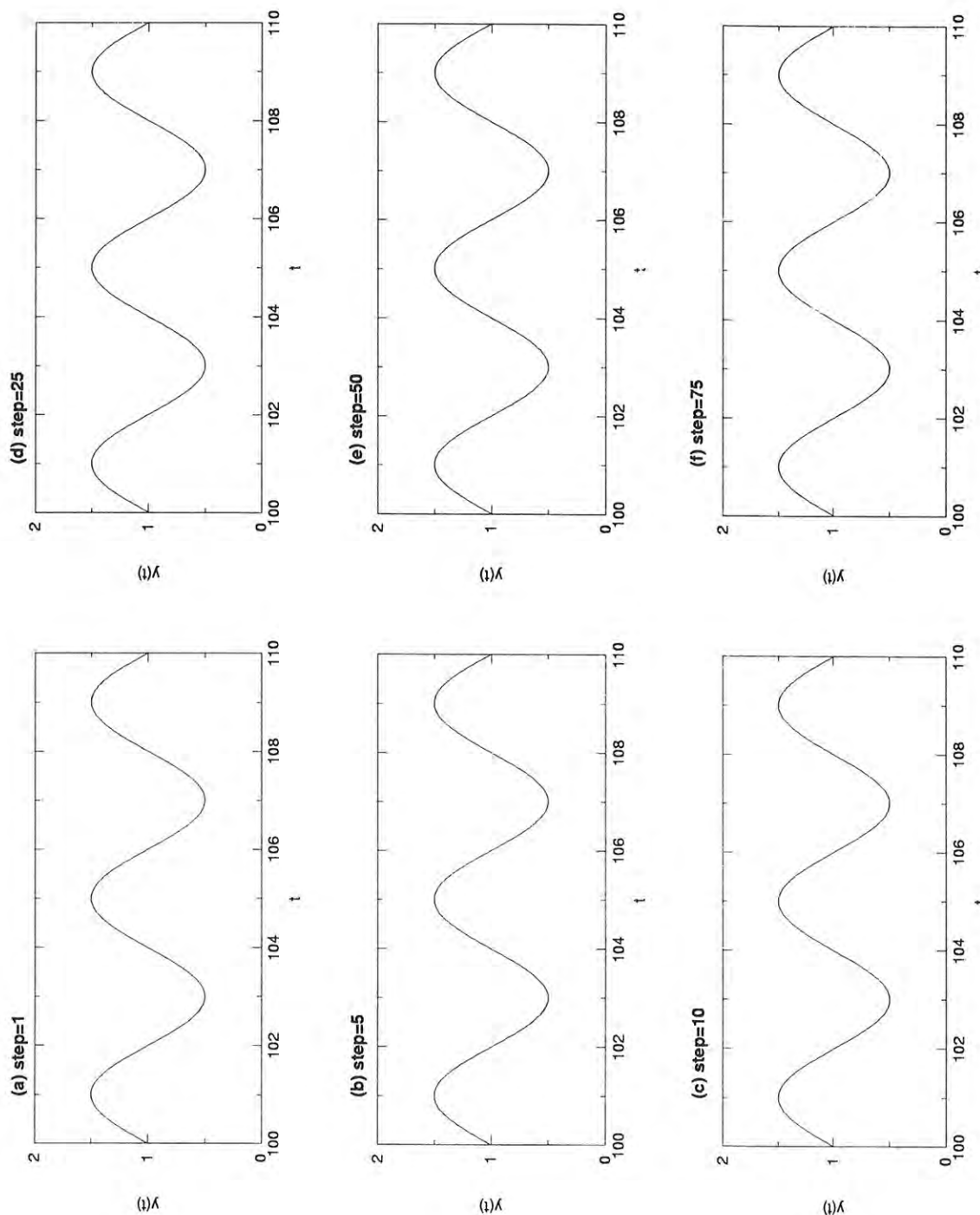


Figura 5-12: Previsão da rede neural, de arquitetura 39:17:3:1 (linha pontilhada), para a função teste $y(t) = a + b \sin(\omega t + c)$ (linha sólida), onde $a = 1$, $b = 0,5$, $c = 0$ e $\omega = 1,5708$. Observa-se que a concordância entre os resultados fornecidos pela rede e os dados da série são perfeitos, para todos os *steps*.

que houvesse uma maior predominância dos dados experimentais no treinamento da rede, em pelo menos 50% dos estágios de previsão, modificou-se a arquitetura, novamente, para $150:50:10:1$. Mesmo assim, os resultados obtidos não foram nada satisfatórios, como se pode observar na figura 5-13b, pois as deformações do sinal fornecido pela rede e os atrasos nas previsões do sinal de raios-X 133673 continuam bem evidentes. Destas figuras depreende-se também que utilizar muitas unidades numa rede neural artificial não implica diretamente na melhora de sua performance, como se poderia facilmente supor.

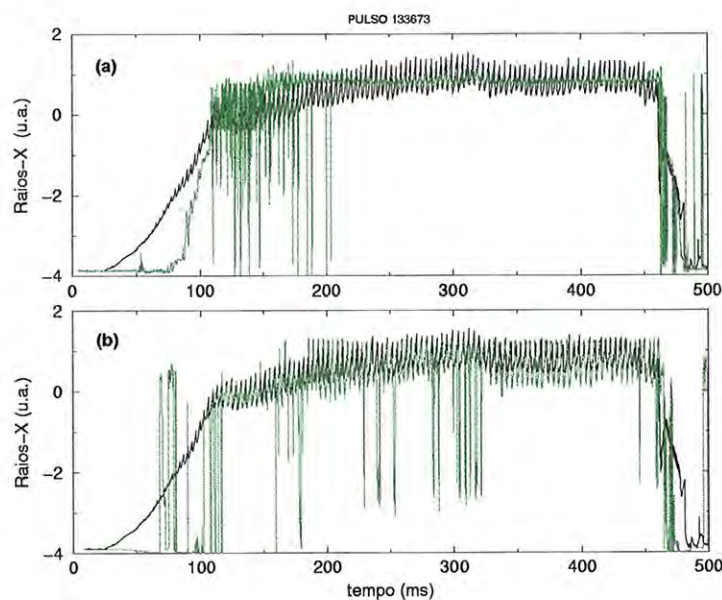


Figura 5-13: Comparação entre os resultados das previsões das redes de arquitetura (a) $33:17:3:1$ e (b) $150:50:10:1$ para o sinal de raios-X 133673, com 75 *steps*. Como se pode observar, as deformações do sinal fornecido pelas redes e o atraso de suas previsões ainda são bem evidentes, mesmo utilizando-se muitas unidades em suas camadas.

Esse comportamento das redes nos levou a pensar em duas hipóteses que poderiam explicar o problema observado:

- não é bom trabalhar com dados realimentados, ou seja, utilizar como unidades de entrada da rede neural dados que tenham sido objeto de sua própria previsão, pelo menos em casos nos quais se deseja trabalhar com valores altos de *steps*, que poderia ser o principal motivo das deformações observadas.
- a escolha do número de unidades, e sua distribuição, entre as camadas da rede neural pode não ser tão arbitrário assim, o que indica que pode haver alguma relação com

a dimensionalidade do sistema dinâmico em estudo, o que talvez pudesse explicar os atrasos observados nas previsões fornecidas pela rede.

5.2 Arquitetura 15:9:3:1, com Previsão por STEP

Em relação à primeira hipótese levantada, no final da seção anterior, procurou-se mudar a forma de se trabalhar com a rede fazendo-a antecipar o comportamento das séries temporais *direto* no *step* desejado. Desse modo, garantia-se que os dados utilizados na alimentação da rede fossem *sempre* 100% experimentais, e para qualquer *step* de previsão desejado.

Adotando-se esse novo procedimento, obteve-se um excelente resultado de previsão da série teste (tipo *seno*), que não apresentou nenhuma deformação, para qualquer *step* utilizado. E, neste caso, voltou-se a utilizar a arquitetura original 15:9:3:1 (figura 5-14), diminuindo-se, assim, a quantidade de unidades presentes nas camadas da rede neural.

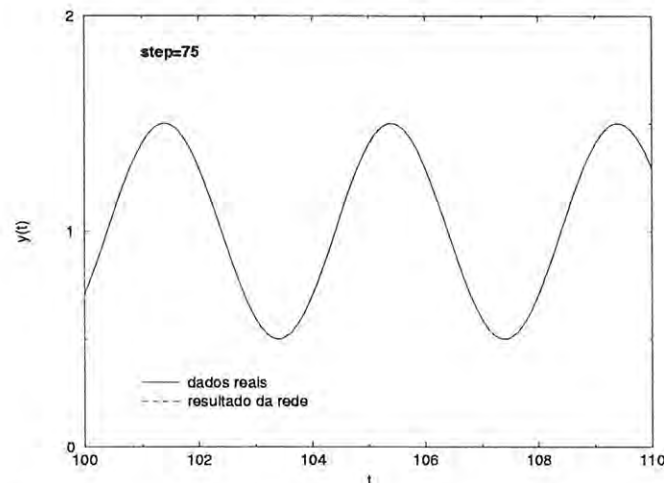


Figura 5-14: Resultado da rede neural, de arquitetura 15:9:3:1 (linha pontilhada), na previsão da função teste $y(t) = a + b \sin(\omega t + c)$ (linha sólida), onde $a = 1$, $b = 0,5$, $c = 0$ e $\omega = 1,5708$, direto no *step* desejado. Observe que a concordância entre a saída da rede e os dados da série são perfeitos. $STEP = 75$

Com os dados experimentais de raios-X, entretanto, não se conseguiu o mesmo desempenho, tão bom, para o *step* = 75. Mas, ainda assim, conseguiu-se estender o tempo de previsão, para pulsos inteiros das descargas do tokamak de 13 para 25 *steps* [22], conseguindo-se, dessa forma, melhorar a performance da rede em relação aos resultados anteriormente obtidos (figuras 5-15 e 5-16).

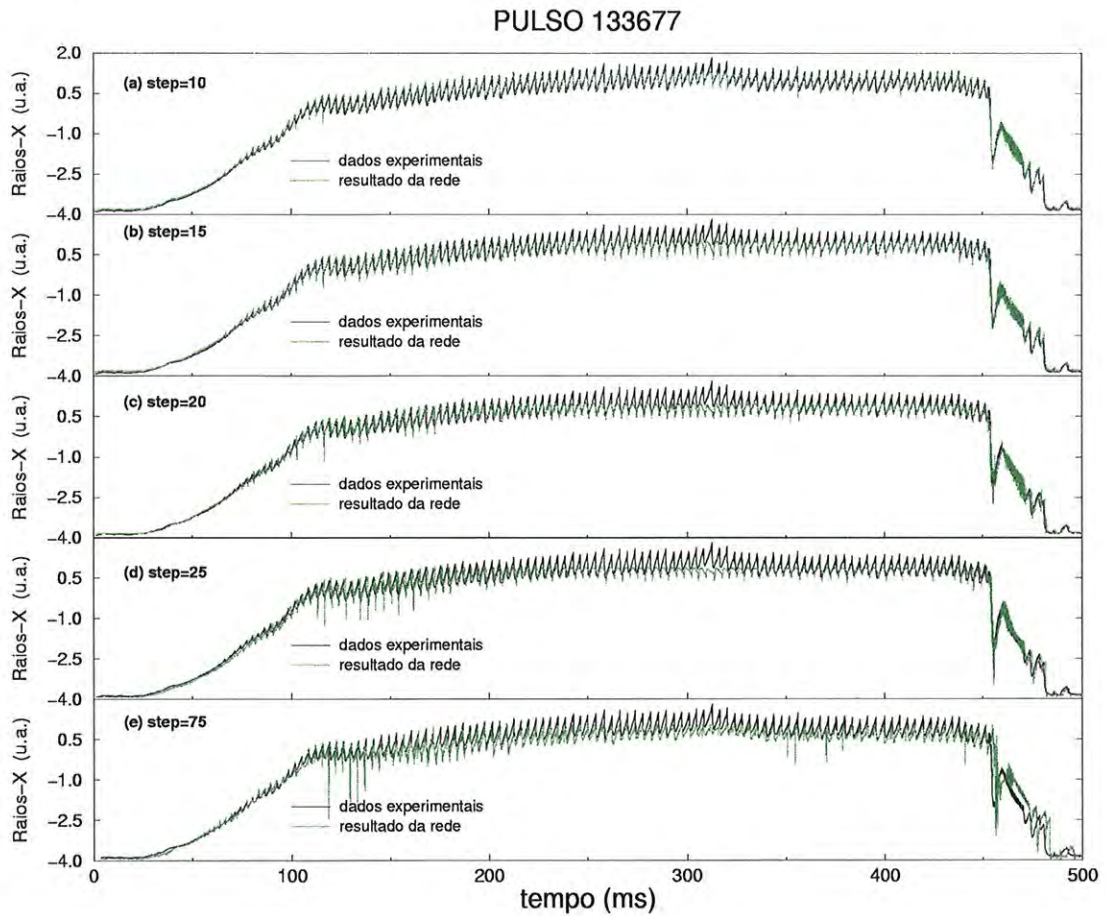


Figura 5-15: Previsão do pulso rx133677, inteiro, para vários *steps* diferentes. Como se pode observar, a previsão para *step* = 1 (a) é praticamente perfeita. Para o *step* = 25 (d), embora a rede continue prevendo o instante da disruptura de forma bastante satisfatória, a sua reprodução do sinal não chega a ser perfeita, apresentando várias oscilações no início e no fim da descarga.

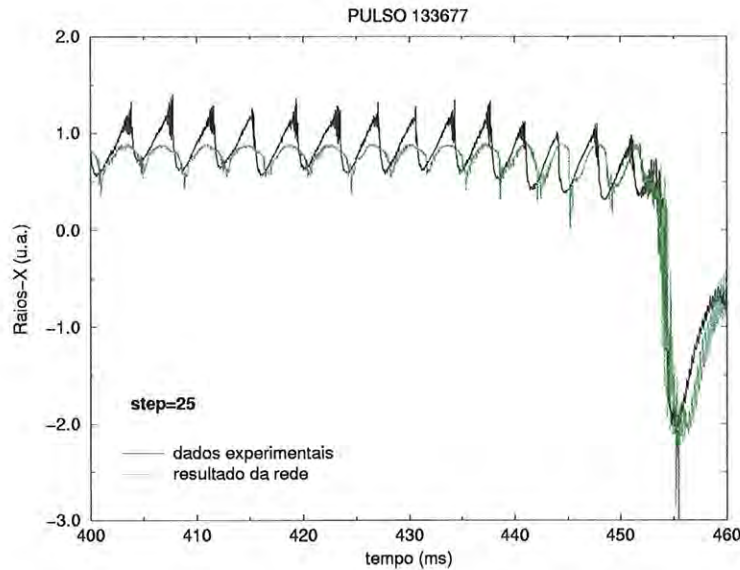


Figura 5-16: Detalhe da figura 5-15 perto da região da disruptura do sinal, na qual a previsão foi realizada com $step = 25$. Este foi o melhor resultado obtido, com pulsos inteiros, para a arquitetura $15:9:9:1$.

Com o intuito de se tentar melhorar o tempo de previsão, para pulsos inteiros do tokamak TEXT, ainda mais, e também levando-se em consideração a segunda hipótese levantada na seção anterior, a respeito do grande atraso ainda verificado nas previsões destes sinais de raios-X, começou-se a analisar a eventual relação que poderia haver entre a dimensionalidade do sistema dinâmico, em estudo, e o número de unidades (sua forma de distribuição também) utilizadas nas camadas das redes neurais.

De acordo com o estudo sobre séries caóticas, apresentado no capítulo 4, e que também foi objeto de submissão de publicação de um artigo à revista *Physica A* [53], procurou-se construir um rede neural artificial, com arquitetura $m:2m:m:1$ para tratar os sinais de raios-X moles de um pulso todo do plasma do TEXT.

5.3 Arquitetura final $15:30:15:1$

Transpondo essa nova abordagem, tratada no capítulo 4, para o tratamento dos dados experimentais de raios-X, não só se conseguiu melhorar o desempenho da rede neural, quanto à previsão das instabilidades de disruptura, trabalhando-se com os pulsos inteiros e com $steps$ de previsão em torno de 75, como também se conseguiu ampliar o tempo de

previsão para até 100 *steps*, o que equivale a dizer que é possível antever o instante da disruptura destes sinais em até ~ 4 *ms*, tempo suficiente para permitir que medidas de controle da máquina possam, talvez, ser tomadas, no sentido de evitar a ocorrência dessas indesejáveis instabilidades ou, pelo menos, atenuar os seus efeitos nocivos.

Após várias tentativas na busca da arquitetura ideal de trabalho, assim como na procura das melhores condições de treinamento e validação, chegou-se à configuração 15:30:15:1 para o conjunto de dados experimentais aqui apresentados.

As figuras 5-17 e 5-18 mostram os resultados obtidos, com essa nova rede neural, nas previsões dos pulsos 133660 (fig.5-2) e 133677 (fig.5-3) para diferentes *steps*. Os padrões de treinamento dessa rede foram extraídos de três outros sinais de raios-X: 133657, 133665 e 133673 (figuras 5-1, 5-2 e 5-3, respectivamente), cujos pontos foram espaçados no tempo de um valor $\tau = 30$. Ainda para o treinamento, adotaram-se 500 épocas, e os parâmetros usados foram $\alpha = 0,9$ e $\eta = 0,001$.

Como se pode observar, através dos sinais de raios-X expandidos no tempo (figuras 5-19 e 5-20), a rede mostrou-se capaz de prever razoavelmente bem o momento em que as disrupturas ocorreram.

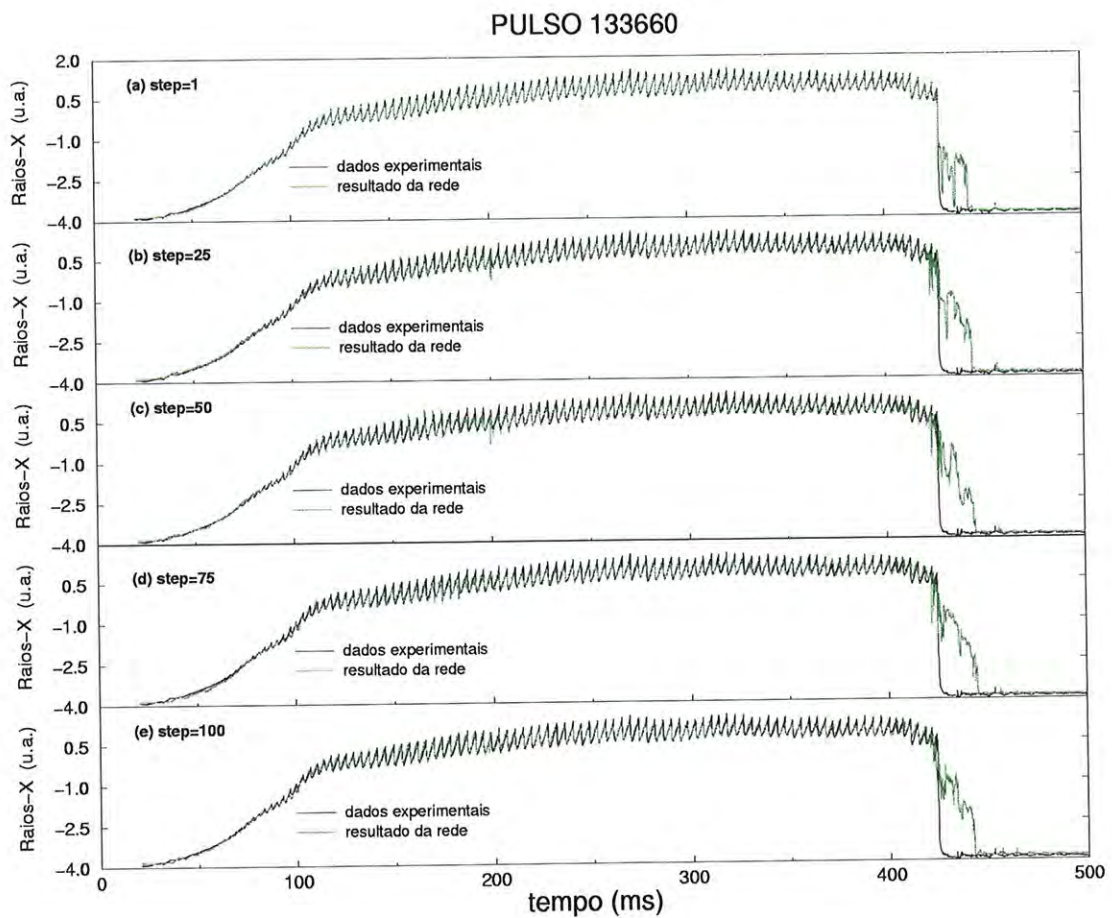


Figura 5-17: Previsão da rede neural $15:30:15:1$ para o pulso rx133660, inteiro, utilizando-se vários *steps* diferentes. Como se pode observar, a previsão para $step = 1$ (a) é praticamente perfeita. Para o demais *steps*, a rede neural continua prevendo a disruptura de forma bem satisfatória, sem deformar nenhum sinal.

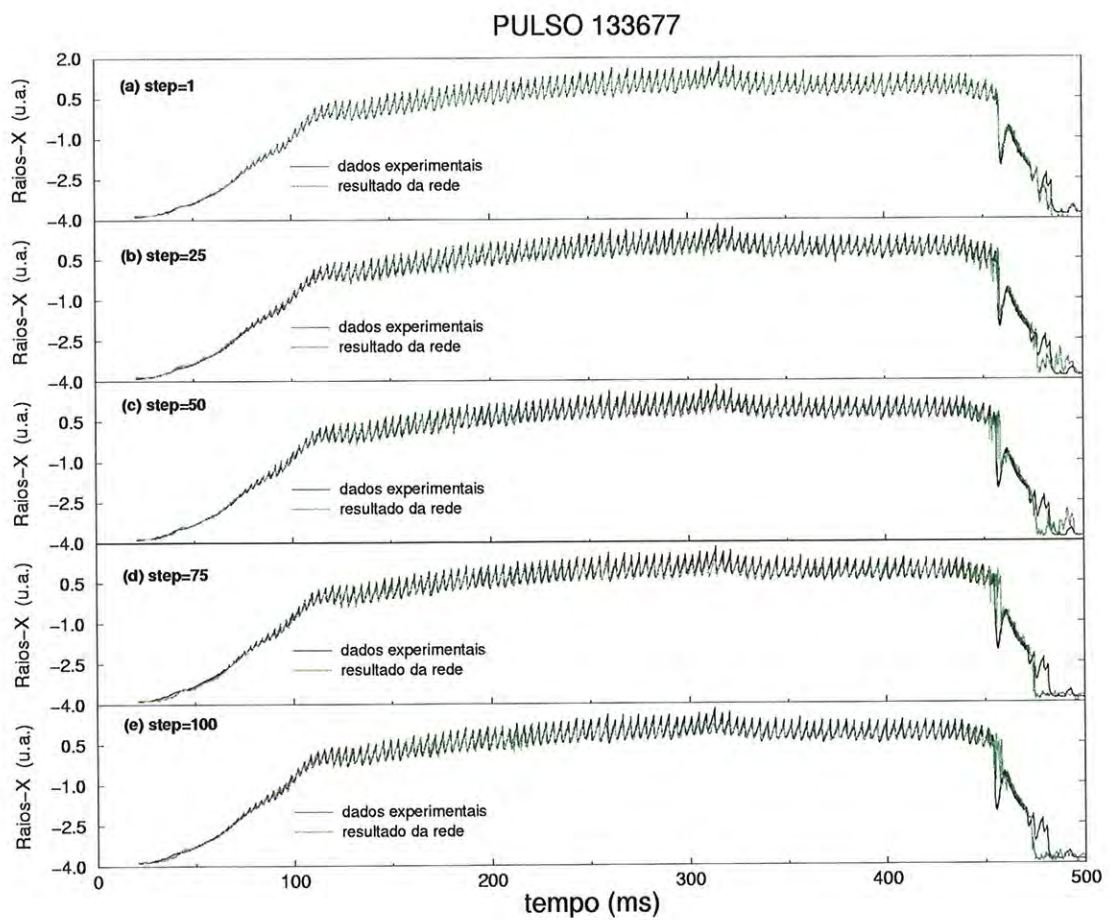


Figura 5-18: Previsão da rede neural $15:30:15:1$ para o pulso rx133667, inteiro, utilizando-se vários *steps* diferentes. Como se pode observar, a previsão para $step = 1$ (a) é praticamente perfeita. Para o demais *steps*, a rede neural continua prevendo a disruptura de forma bem satisfatória, sem deformar nenhum sinal.

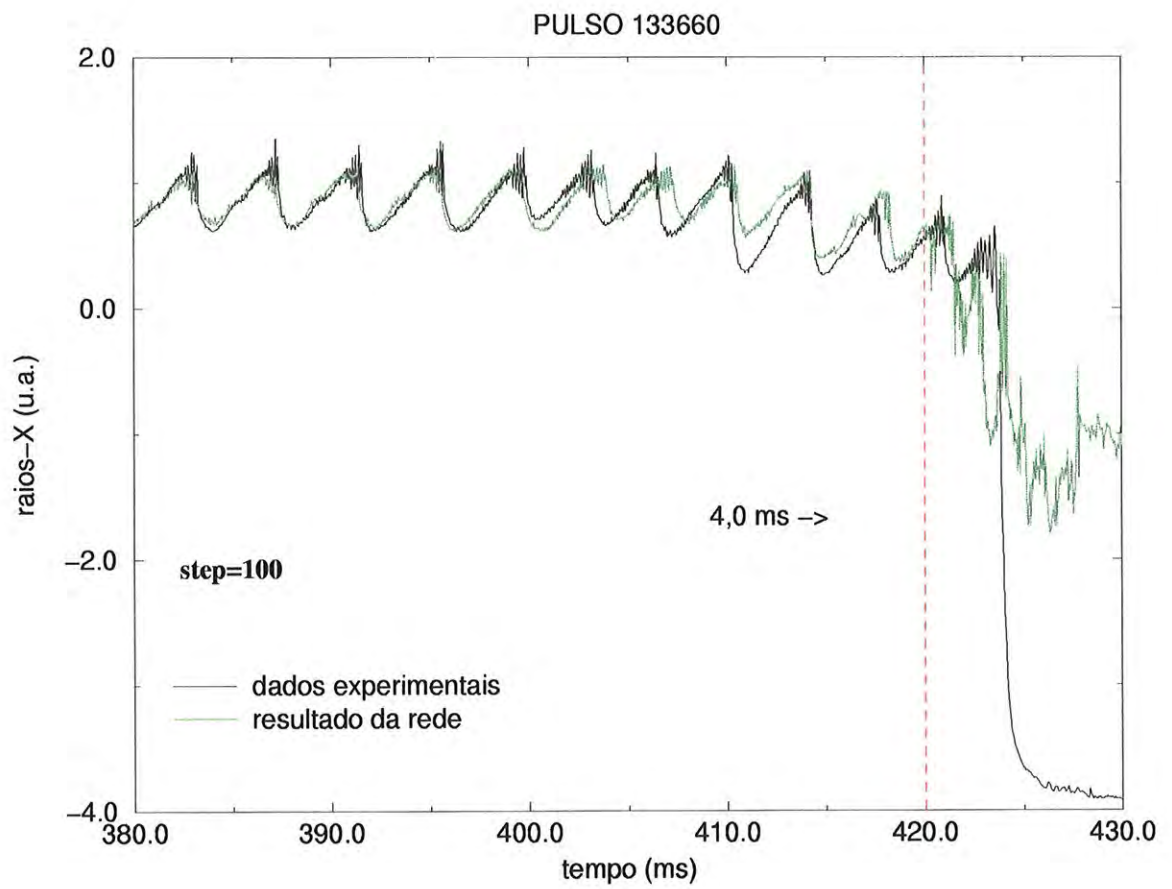


Figura 5-19: Detalhe da figura 5-17, perto da região da disruptura do sinal.

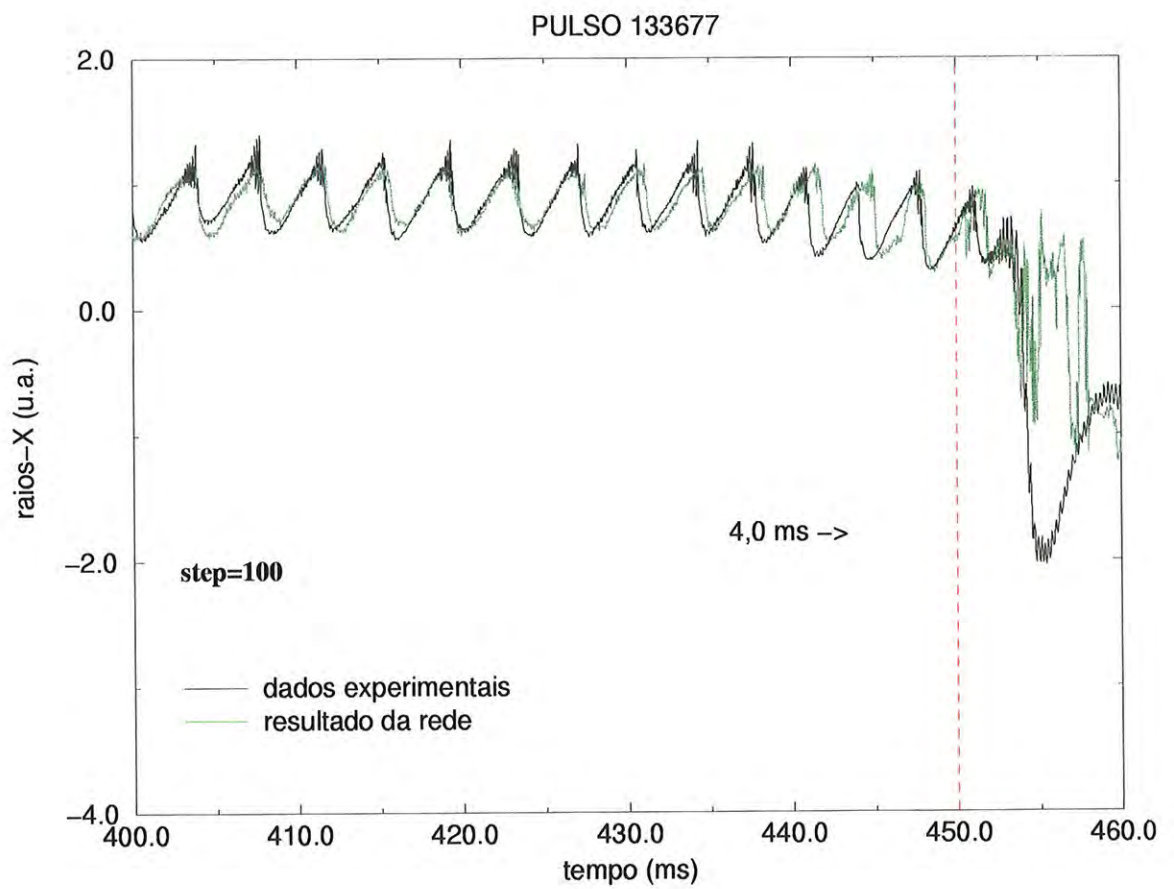


Figura 5-20: Detalhe da figura 5-18, perto da região da disruptura do sinal.

Capítulo 6

Sistema de Detecção de Raios-X (Situação Atual)

Um sistema de detecção de raios-X de baixas energias foi projetado para o tokamak TCA-BR, do Instituto de Física, da Universidade de São Paulo, para permitir estudos e investigações dos mecanismos que poderiam levar ao surgimento das instabilidades de disrupções, em plasmas magneticamente confinados, bem como tentar prever suas ocorrências através da utilização de redes neurais artificiais. Com esse sistema, pode-se também criar imagens tomográficas a partir do perfil da radiação de raios-X emitida pelo plasma, assim como também medir a temperatura eletrônica da coluna de plasma, através do método dos absorvedores [20].

No momento, o sistema, como um todo, encontra-se em fase de implementação.

6.1 Câmaras de Raios-X

Em linhas gerais, o diagnóstico de raios-X compõe-se de duas câmaras de detecção, que visualizam a mesma região do plasma, 4 conjuntos de detectores de raios-X, circuitos pré-amplificadores, sistemas de digitalização dos sinais medidos e computadores para registro e armazenamento dos dados.

As câmaras de raios-X foram produzidas em aço não magnetizável (316L) de forma que cada uma acomodasse, em seu interior, 2 conjuntos de detectores de raios-X com os correspondentes circuitos amplificadores.

Uma das câmaras foi acondicionada numa janela superior do vaso do tokamak e a outra foi instalada na janela lateral, deslocada angularmente de 90° em relação à primei-

ra, permitindo, assim, que ambas visualizem a mesma secção transversal da coluna de plasma (figura 6-1). Ambas possuem sistema de bombeamento de vácuo independente do tokamak, estando separadas deste por duas válvulas-gaveta, conforme pode ser observado na figura.

O ângulo sólido dos sensores foi definido de forma a permitir uma visualização completa de uma secção da coluna de plasma, cujo diâmetro é de 360 mm. Em cada câmara utilizam-se dois conjuntos de 38 detectores de raios-X: um conjunto fica próximo à fenda colimadora, de largura variável, e o outro, blindado, fica posicionado imediatamente atrás daquele primeiro, para medir o ruído de fundo.

Para tentar impedir que a radiação visível do espectro emitido pelo plasma seja detectada, e também para evitar a saturação dos sinais nos sistemas de medidas, anteparos absorvedores (lâminas de berílio) são tradicionalmente inseridos no caminho óptico dos detectores, o que faz com que o sistema de detecção seja preferencialmente sensível a fótons com energias correspondentes aos raios-X de baixas energias, ou seja, $0,1 \text{ keV} \leq h\nu \leq 10,0 \text{ keV}$, região de interesse para trabalhos com tomografia, medidas de temperatura eletrônica do plasma, ou uso de redes neurais.

6.2 Detectores Utilizados

Dentre os tipos de detectores que podem ser utilizados no sistema de diagnóstico de raios-X, os mais usuais são os detectores barreira-de-superfície, geralmente escolhidos pela sua alta eficiência na região dos raios-X moles, boa resolução temporal e fácil operação e calibração. Entretanto, para a finalidade de se fazer medidas tomográficas, escolheu-se os fotodiodos da Hamamatsu (C2334) que vêm em forma de "array", contendo, por exemplo, 38 sensores num espaço total 4 cm de comprimento por 0.5 cm de largura. Cada sensor possui uma área efetiva de $4 \times 1 \text{ mm}^2$, com um espaçamento de 0,1 mm entre cada um deles. Para torná-los aptos a fazerem medidas na faixa de energia desejada, retiram-se os cristais de quartzo, que os acompanham, e fazem-se as substituições por lâminas de berílio.

O sistema eletrônico, responsável pela amplificação dos sinais medidos e subtração dos ruídos de fundo, ainda está em fase de testes.

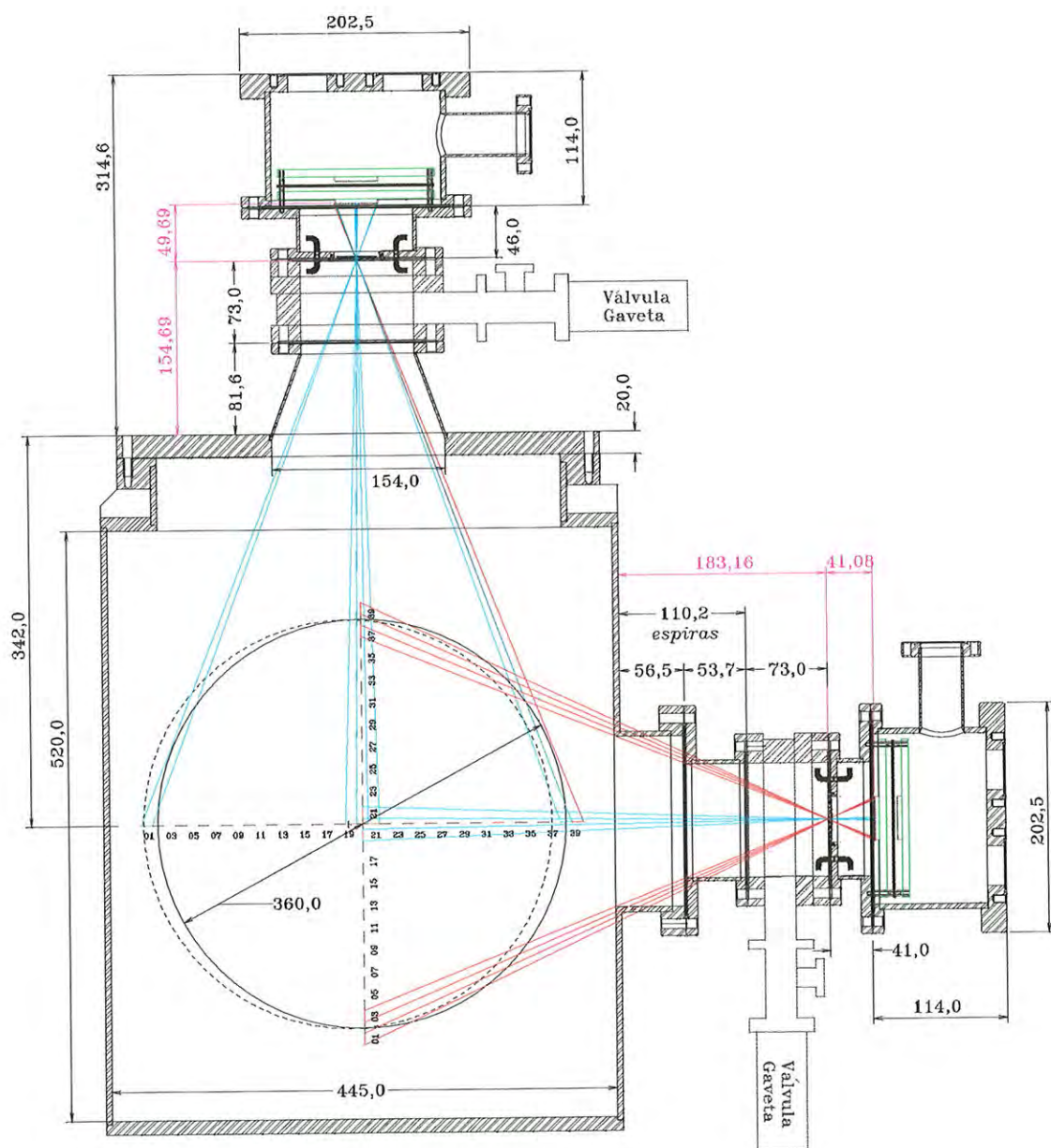


Figura 6-1: Sistema de detecção de raios-X de baixas energias projetado para o tokamak TCABR (medidas em mm).

Capítulo 7

Tomografia de Plasmas por Raios-X

Como o sistema de detecção de raios-X, projetado para o TCABR, também permite análises tomográficas do perfil da radiação emitida pelo plasma, também se fez esse estudo, no presente trabalho, concluindo-se o programa computacional responsável pela formação das imagens.

Desenvolveu-se, também, cálculos analíticos visando a substituição dos polinômios de Zernicke por funções de Bessel, na expansão dos harmônicos de Fourier da emissividade do plasma, como forma de se tentar prevenir a formação de *artefactos* na reconstrução da borda do plasma.

7.1 Reconstrução Tomográfica

O problema da reconstrução tomográfica resume-se basicamente em obter a emissividade local do plasma, $g(r, \theta)$, a partir da brilhância¹ total, $f(p, \phi)$, medida ao longo de várias cordas L (linhas de visualização) de um conjunto apropriado de detectores (figura 7-1). Como o plasma, em tokamaks, pode ser considerado opticamente fino para as suas próprias emissões de raios-X, essa transformação se torna relativamente simples. Para obtê-la, definimos inicialmente a potência irradiada por um elemento de volume do plasma (dV_{plasma}) ao longo da corda (L), como sendo (figura 7-2) [55]:

$$dP = \left[\int G(r, \theta, h\nu) \epsilon(h\nu) d(h\nu) \right] \times dV_{plasma} \times \frac{A_{det} \cos \theta_{inc}}{4\pi l^2} \quad (7-1)$$

onde,

¹Brilhância: razão entre a intensidade energética da fonte, numa direção determinada, pela projeção da área da fonte num plano perpendicular a esta direção. Mede-se, no sistema internacional, em $W/sr.m^2$.

- $h\nu \equiv$ energia dos fótons emitidos
- $G(r, \theta, h\nu) \equiv$ emissividade espectral
- $\epsilon(h\nu) \equiv$ eficiência do detector utilizado
- $\frac{A_{det} \cos \theta_{inc}}{l^2} \equiv$ ângulo sólido compreendido pelo detector
- $dV_{plasma} = dL \times \frac{A_{ap} \cos \theta'_{inc} l^2}{d^2}$

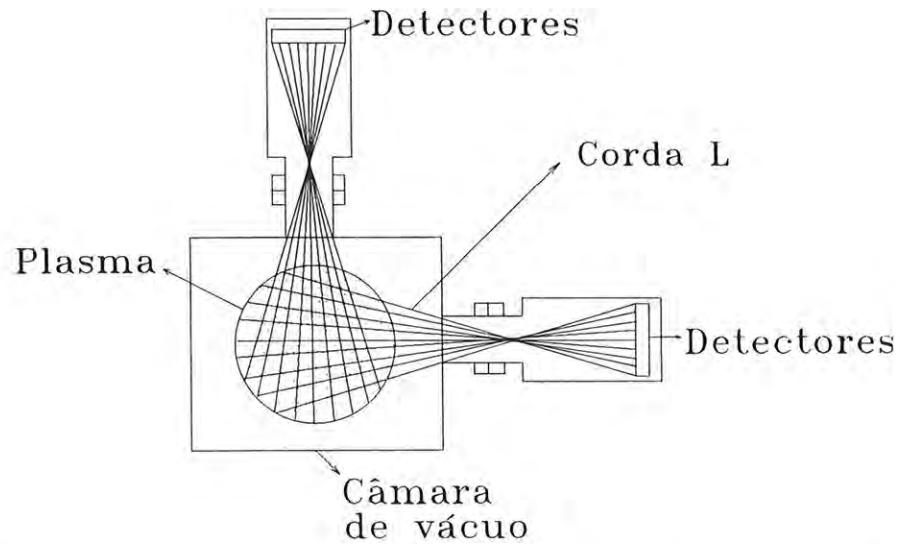


Figura 7-1: Representação esquemática do sistema de detecção de raios-X para reconstrução tomográfica.

Garantindo-se que a eficiência de todos os detectores seja a mesma, o que pode ser feito através de uma cuidadosa calibração, a integral da equação (7-1) pode ser substituída por uma função apenas da posição $g(r, \theta)$. Assim, temos:

$$dP = \left[\frac{A_{ap} A_{det} \cos \theta_{inc} \cos \theta'_{inc}}{4\pi d^2} \right] g(r, \theta) dL \quad (7-2)$$

Notando-se que os termos da equação (7-2), entre colchetes, são fatores apenas geométricos, não dependendo, portanto, de L , podemos escrever:

$$P = \left[\frac{A_{ap} A_{det} \cos \theta_{inc} \cos \theta'_{inc}}{4\pi d^2} \right] \int_{L(p, \phi)} g(r, \theta) dL \quad (7-3)$$

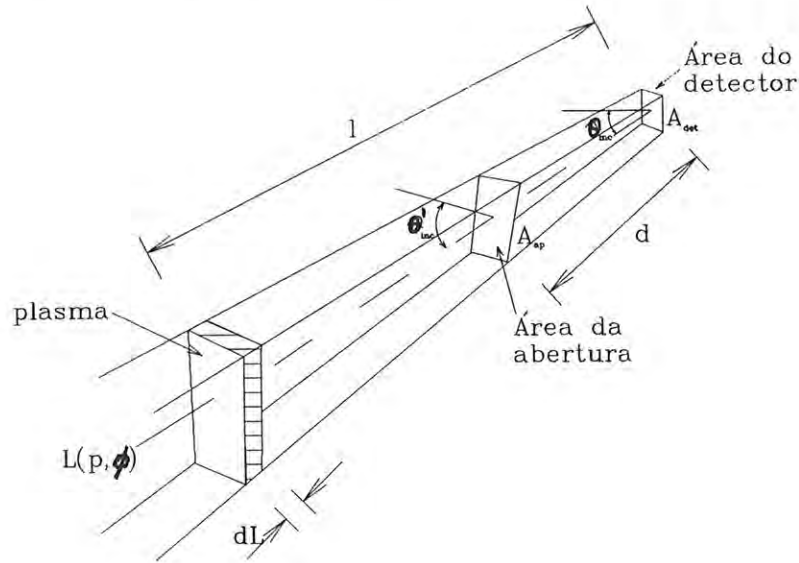


Figura 7-2: Ângulo sólido associado à corda $L(p, \phi)$ para um dado detector.

Definindo-se $f(p, \phi) \equiv P \left[\frac{A_{ap} A_{det} \cos \theta_{inc} \cos \theta'_{inc}}{4\pi d^2} \right]^{-1}$, a equação (7-3) resume-se a:

$$f(p, \phi) = \int_{L(p, \phi)} g(r, \theta) dL \quad (7-4)$$

Assim, a partir dos sinais de raios-X emitidos pelo plasma, obtemos a função $f(p, \phi)$ e, através de uma transformação apropriada, chegamos à emissividade local $g(r, \theta)$ desejada.

7.2 Técnicas de Reconstrução

Um dos métodos utilizados para se conseguir a emissividade local $g(r, \theta)$ baseia-se na utilização da técnica das soluções analíticas, por possibilitar, de uma maneira relativamente simples, a codificação das expressões matemáticas envolvidas no processo de reconstrução tomográfica, além de tornar evidentes os limites de resolução espacial do sistema de detecção adotado [55, 56]. Este método foi originalmente desenvolvido por A.M. Cormack [57, 58], para aplicações médicas, sendo apresentado a seguir:

A brilhância (df), devida a dois elementos de corda (dL), equidistantes do parâmetro de impacto (p), figura 7-3, pode ser escrita como:

$$df = g(r, \theta) dL + g(r', \theta') dL \quad (7-5)$$

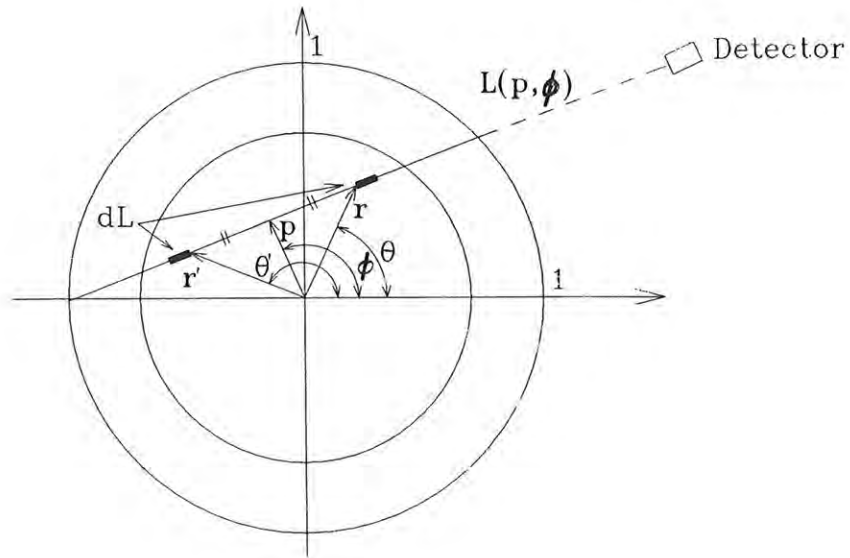


Figura 7-3: Diagrama mostrando os parâmetros que definem cada linha (corda) de visualização $(L(p, \phi))$ do detector.

Devida à simetria observada, as coordenadas r' e θ' podem ser reescritas como:

$$r' = r$$

$$\theta' = 2(\phi - \theta) + \theta = 2\phi - \theta$$

que, substituídas novamente na equação (7-5), resulta em:

$$df = [g(r, \theta) + g(r, 2\phi - \theta)] dL \quad (7-6)$$

Expandindo-se o termo $g(r, \theta)$ em séries de Fourier e substituindo-o na expressão (7-6), obtém-se:

$$df = \sum_{m=-\infty}^{\infty} [g_m(r)e^{im\theta} + g_m(r)e^{im(2\phi-\theta)}] dL \quad (7-7)$$

Fatorando-se o termo $g_m(r)e^{im\phi}$, e rearranjando-se os demais, esta expressão pode ser reescrita da seguinte forma:

$$df = 2 \sum_{m=-\infty}^{\infty} g_m(r) \cos [m(\theta - \phi)] e^{im\phi} dL \quad (7-8)$$

Mas $L^2 = r^2 - p^2$ e, portanto, $dL = \frac{r dr}{\sqrt{r^2 - p^2}}$. Além disso, pela figura 7-3, observamos que $\cos(\theta - \phi) = p/r$, o que nos permite escrever:

$$df = 2 \sum_{m=-\infty}^{\infty} \left\{ g_m(r) \frac{r dr}{\sqrt{r^2 - p^2}} \cos \left[m \cos^{-1} \left(\frac{p}{r} \right) \right] \right\} e^{im\phi} \quad (7-9)$$

Identificando-se o termo $\cos [m \cos^{-1} (p/r)]$ como sendo o polinômio de Tschebycheff, $T_m (p/r)$ [59] tem-se, após terem sido somadas as contribuições de todos os elementos que compõem a corda L :

$$f(p, \phi) = \sum_{m=-\infty}^{\infty} \left\{ 2 \int_p^1 \frac{g_m(r) T_m (p/r) r dr}{\sqrt{r^2 - p^2}} \right\} e^{im\phi} \quad (7-10)$$

Expandindo-se a expressão $f(p, \phi)$ em série de Fourier e comparando-se, termo a termo, com a série dada pela equação acima, obtém-se:

$$f_m(p) = 2 \int_p^1 \frac{g_m(r) T_m (p/r) r dr}{\sqrt{r^2 - p^2}} \quad (7-11)$$

Garantindo-se que, para valores suficientemente grandes de r , $g_m(r)$ se comporta como $(r^{m+1})^{-1}$, podemos estender o resultado da equação (7-11) para:

$$f_m(p) = 2 \int_p^{\infty} \frac{g_m(r) T_m (p/r) r dr}{\sqrt{r^2 - p^2}} \quad (7-12)$$

O problema da reconstrução tomográfica consiste agora em inverter a equação (7-12) para se obter $g_m(r)$ que, segundo Cormack [57], é dada por:

$$g_m(r) = -\frac{1}{\pi} \frac{d}{dr} \int_r^{\infty} \frac{f_m(p) T_m (p/r) r dp}{p \sqrt{p^2 - r^2}} \quad (7-13)$$

Mas esta expressão é difícil de ser implementada numericamente, pois a derivada, e a singularidade, que aparece em $p = r$, a tornam muito sensível a ruídos.

Para contornar esse problema, Cormack desenvolveu uma outra maneira de se obter a reconstrução tomográfica *sem passar pelo processo de inversão* da equação (7-12), que é basicamente a seguinte:

Inicialmente, calculam-se duas transformadas de Fourier: uma transformada de Fourier cosseno, para os termos pares, e uma transformada de Fourier seno, para os termos ímpares, da equação (7-12). Como os cálculos são totalmente análogos, desenvolveremos, aqui, apenas a transformada cosseno, obtendo-se o seguinte resultado:

$$F_c(f_{2m}(p)) = \int_0^{\infty} f_{2m}(p) \cos(pt) dp = 2 \int_0^{\infty} \cos(pt) dp \int_p^{\infty} \frac{g_{2m}(r) T_{2m} (p/r) r dr}{\sqrt{r^2 - p^2}} \quad (7-14)$$

Fazendo-se a mudança de variável: $x \equiv p/r$, obtemos:

$$F_c(f_{2m}(p)) = 2 \int_0^{\infty} g_{2m}(r) r dr \int_0^1 \frac{T_{2m}(x) \cos(rxt) dx}{\sqrt{1 - x^2}} \quad (7-15)$$

Utilizando-se a seguinte identidade matemática [60]:

$$J_{2m}(x) = (-1)^m \frac{2}{\pi} \int_0^1 \frac{T_{2m}(y) \cos(xy) dy}{\sqrt{1 - y^2}}, \quad (7-16)$$

onde $J_{2m}(x)$ é a função de Bessel de ordem $2m$, a equação (7-15) resulta em:

$$F_c(f_{2m}(p)) = (-1)^m \pi \int_0^\infty g_{2m}(r) J_{2m}(rt) r dr \quad (7-17)$$

Calculando-se, agora, a anti-transformada cosseno de Fourier, da equação acima, obtemos, para os termos pares, a seguinte expressão:

$$f_{2m}(p) = 2(-1)^m \int_0^\infty \cos(pt) dt \int_0^\infty g_{2m}(r) J_{2m}(rt) r dr \quad (7-18)$$

Para os termos ímpares, a anti-transformada é dada por [58]:

$$f_{2m+1}(p) = 2(-1)^m \int_0^\infty \text{sen}(pt) dt \int_0^\infty g_{2m+1}(r) J_{2m+1}(rt) r dr \quad (7-19)$$

A justificativa que Cormack dá para a obtenção de duas expressões aparentemente bem mais complicadas do que a equação (7-12) é a de que, expandindo-se a função $g_m(r)$ numa base apropriada, a transformada de Henkel, representada pela última integral das equações (7-18) e (7-19), pode ser realizada e, a partir disso, obterem-se as anti-transformadas de Fourier seno e cosseno, fornecendo, desta forma, uma expressão para $f_m(p)$.

7.2.1 Polinômios de Zernicke

Uma outra propriedade matemática conhecida [61]:

$$\int_0^1 R_m^l(r) J_m(rt) r dr = (-1)^l t^{-1} J_{m+2l+1}(t) \quad (7-20)$$

sugere que, se utilizarmos os Polinômios de Zernicke:

$$R_m^l = \sum_{s=0}^m \frac{(-1)^s (n+2m-s)! r^{n+2m-2s}}{s!(n+m-s)!(m-s)!} \quad (7-21)$$

como base para expandirmos $g_m(r)$, nas equações (7-18) e (7-19) e, depois disso, calcularmos as anti-transformadas seno e cosseno de Fourier, destas mesmas equações, obtém-se para $f_m(p)$ [62]:

$$\begin{aligned} f_m(p) &= \sum_{l=0}^{\infty} a_{m,l} \frac{2}{m+2l+1} \text{sen}[(m+2l+1) \cos^{-1} p], \text{ para } 0 \leq p < 1 \\ &= 0, \text{ para } p > 1 \end{aligned} \quad (7-22)$$

Somando-se os coeficientes da expansão de Fourier, acima obtidos, chegamos à seguinte expressão:

$$f(p, \phi) = \sum_{m=0}^{\infty} \sum_{l=0}^{\infty} a_{m,l}^c b_{m,l} \cos(m\phi) + \sum_{m=0}^{\infty} \sum_{l=0}^{\infty} a_{m,l}^s b_{m,l} \text{sen}(m\phi) \quad (7-23)$$

onde $b_{m,l} = \frac{2}{(m+2l+1)} \text{sen} [(m+2l+1) \cos^{-1} p]$. A equação (7-23) é a equação normalmente utilizada na implementação numérica da reconstrução tomográfica, observada em diversos trabalhos publicados, dentre eles, o trabalho desenvolvido para o tokamak TEXT, da Universidade do Texas (Austin, E.U.A.), cujo programa, em FORTRAN, foi-nos gentilmente cedido [56].

7.2.2 As Funções de Bessel

Analizando-se mais cuidadosamente os Polinômios de Zernicke, normalmente utilizados como base para a expansão dos harmônicos de Fourier da emissividade, verifica-se que eles não são muito apropriados para reconstruções de perfis de emissividade típicos. Isso porque os *nós* dos polinômios não são igualmente espaçados, apresentando distâncias cada vez menores à medida em que se aproxima da borda do plasma, o que compromete a resolução radial das medidas nesta região. Mais ainda, os Polinômios de Zernicke, como dito anteriormente, também não se anulam na borda, o que pode acabar provocando a geração de artefactos nas imagens obtidas. Já as funções de Bessel apresentam os nós quase que igualmente espaçados, melhorando bastante a resolução das medidas, além de se anularem na borda do plasma. Deste modo, fizeram-se os cálculos considerando-se esta alteração, sendo os resultados apresentados a seguir:

Até a equação 7-9 os cálculos se mantêm iguais. A mudança de procedimento ocorre daí para frente, com o rearranjo da referida equação:

$$df = \sum_{m=-\infty}^{\infty} g_m(r) \left[2T_m(p/r) \frac{rdr}{\sqrt{r^2-p^2}} \right] e^{im\phi} \quad (7-24)$$

Podemos reescrevê-la, ainda, desta forma:

$$df = \sum_{m=0}^{\infty} g_m(r) \left[2T_m(p/r) \frac{rdr}{\sqrt{r^2-p^2}} \right] e^{im\phi} + \sum_{m=1}^{\infty} g_{-m}(r) \left[2T_{-m}(p/r) \frac{rdr}{\sqrt{r^2-p^2}} \right] e^{-im\phi} \quad (7-25)$$

Para que df seja real, $g_m(r)$ e $g_{-m}(r)$ devem ser imaginários, onde são válidas as relações:

$$g_m(r) = \frac{g_m^c(r) - ig_m^s(r)}{2} \quad (7-26)$$

$$g_{-m}(r) = \frac{g_m^c(r) + ig_m^s(r)}{2} \quad (7-27)$$

Desta forma, obtemos, para a expressão de df , lembrando-se que $T_m(p/r) = T_{-m}(p/r)$:

$$df = \sum_{m=0}^{\infty} \frac{g_m^c(r) - ig_m^s(r)}{2} \left[2T_m(p/r) \frac{rdr}{\sqrt{r^2 - p^2}} \right] e^{im\phi} + \sum_{m=1}^{\infty} \frac{g_m^c(r) + ig_m^s(r)}{2} \left[2T_m(p/r) \frac{rdr}{\sqrt{r^2 - p^2}} \right] e^{-im\phi} \quad (7-28)$$

Agrupando-se os termos em senos e cossenos, temos:

$$df = \sum_{m=0}^{\infty} g_m^c(r) \left[2T_m(p/r) \frac{rdr}{\sqrt{r^2 - p^2}} \right] \cos m\phi + \sum_{m=0}^{\infty} g_m^s(r) \left[2T_m(p/r) \frac{rdr}{\sqrt{r^2 - p^2}} \right] \sin m\phi \quad (7-29)$$

o que resulta em:

$$f(p, \phi) = \sum_{m=0}^{\infty} 2 \int_p^1 g_m^c(r) T_m(p/r) \frac{rdr}{\sqrt{r^2 - p^2}} \cos m\phi + \sum_{m=0}^{\infty} 2 \int_p^1 g_m^s(r) T_m(p/r) \frac{rdr}{\sqrt{r^2 - p^2}} \sin m\phi \quad (7-30)$$

Por outro lado, sabemos que:

$$f(p, \phi) = \sum_{m=0}^{\infty} f_m^c(p) \cos m\phi + \sum_{m=0}^{\infty} f_m^s(p) \sin m\phi \quad (7-31)$$

Desta forma, comparando-se a equação (7-30) com a equação (7-31) concluímos que:

$$f_m^c(p) = 2 \int_p^1 g_m^c(r) T_m(p/r) \frac{rdr}{\sqrt{r^2 - p^2}} \quad (7-32)$$

$$f_m^s(p) = 2 \int_p^1 g_m^s(r) T_m(p/r) \frac{rdr}{\sqrt{r^2 - p^2}} \quad (7-33)$$

Garantindo-se que, para valores suficientemente grandes de r , $g_m(r)$ se comporta assintoticamente, como $(r^{m+1})^{-1}$, podemos estender os resultados acima para:

$$f_m^c(p) = 2 \int_p^{\infty} g_m^c(r) T_m(p/r) \frac{rdr}{\sqrt{r^2 - p^2}} \quad (7-34)$$

$$f_m^s(p) = 2 \int_p^{\infty} g_m^s(r) T_m(p/r) \frac{rdr}{\sqrt{r^2 - p^2}} \quad (7-35)$$

Aplicando-se a transformada de Fourier cosseno, para $m = \text{par}$, à equação (7-34), obtemos:

$$\int_0^{\infty} f_{2m}^c(p) \cos ptdp = 2 \int_0^{\infty} \cos(pt) dp \int_p^{\infty} g_{2m}^c(r) T_{2m}(p/r) \frac{rdr}{\sqrt{r^2 - p^2}} \quad (7-36)$$

Fazendo-se a mudança de variável $x = p/r$ na equação acima, obtemos:

$$\int_0^\infty f_{2m}^c(p) \cos ptdp = 2 \int_0^\infty g_{2m}^c(r) r dr \int_0^1 \frac{T_{2m}(x) \cos(rxt)}{\sqrt{(1-x^2)}} dx \quad (7-37)$$

Identificando-se a segunda integral da equação (7-37) com a função de Bessel [59]:

$$\int_0^1 \frac{T_{2m}(x) \cos(rxt)}{\sqrt{(1-x^2)}} = \frac{\pi}{2} (-1)^m J_{2m}(rt) \quad (7-38)$$

obtemos:

$$\int_0^\infty f_{2m}^c(p) \cos ptdp = (-1)^m \pi \int_0^\infty g_{2m}^c(r) J_{2m}(rt) r dr \quad (7-39)$$

Calculando-se a anti-transformada cosseno da expressão (7-39), obtemos:

$$f_{2m}^c(p) = 2(-1)^m \int_0^\infty \cos(pt) dt \int_0^\infty g_{2m}^c(r) J_{2m}(rt) r dr \quad (7-40)$$

que é um resultado análogo ao obtido por Cormack, na expressão (7-18). Aqui, ao invés de expandirmos $g_{2m}^c(r)$ em séries de polinômios de Zernicke, expandiremos em séries de funções de Bessel, da seguinte maneira [59]:

$$g_{2m}^c(r) = \sum_{l=1}^{\infty} a_{2m,l}^c J_{2m}(\lambda_{2m,l} r) \quad (7-41)$$

Substituindo-se esta expressão na equação 7-40, obtemos:

$$f_{2m}^c(p) = \sum_{l=1}^{\infty} 2a_{2m,l}^c 2(-1)^m \int_0^\infty \cos(pt) dt \int_0^\infty J_{2m}(\lambda_{2m,l} r) J_{2m}(rt) r dr \quad (7-42)$$

Usando-se a condição de ortogonalidade das funções de Bessel para $\nu > (-1/2)$ [59]:

$$\int_0^\infty J_\nu(\alpha\rho) J_\nu(\alpha'\rho) \rho d\rho = \frac{1}{\alpha} \delta(\alpha - \alpha'), \quad (7-43)$$

temos:

$$f_{2m}^c(p) = \sum_{l=1}^{\infty} 2(-1)^m a_{2m,l}^c \int_0^\infty \frac{1}{t} \cos(pt) \delta(t - \lambda_{2m,l}) dt \quad (7-44)$$

Ou seja,

$$f_{2m}^c(p) = \sum_{l=1}^{\infty} 2(-1)^m a_{2m,l}^c \frac{\cos(p\lambda_{2m,l})}{\lambda_{2m,l}} \quad (7-45)$$

Analogamente, para os termos ímpares, temos:

$$f_{2m+1}^c(p) = \sum_{l=1}^{\infty} 2(-1)^m a_{2m+1,l}^c \frac{\sin(p\lambda_{2m+1,l})}{\lambda_{2m+1,l}} \quad (7-46)$$

Aplicando-se a transformada de Fourier seno à expressão (7-35) e continuando com cálculos análogos aos efetuados para a expressão (7-34), obtemos:

$$f_{2m}^s(p) = \sum_{l=1}^{\infty} 2(-1)^m a_{2m,l}^s \frac{\cos(p\lambda_{2m,l})}{\lambda_{2m,l}} \quad (7-47)$$

Analogamente, para os termos ímpares, temos:

$$f_{2m+1}^s(p) = \sum_{l=1}^{\infty} 2(-1)^m a_{2m+1,l}^s \frac{\sin(p\lambda_{2m+1,l})}{\lambda_{2m+1,l}} \quad (7-48)$$

Podemos, então, reescrever as expressões (7-45), (7-46), (7-47) e (7-48) de uma forma mais compacta:

$$f_m^c(p) = \sum_{l=1}^{\infty} a_{m,l}^c b_{m,l} \quad (7-49)$$

e

$$f_m^s(p) = \sum_{l=1}^{\infty} a_{m,l}^s b_{m,l} \quad (7-50)$$

onde:

$$b_{m,l} = \begin{cases} 2(-1)^{m/2} \frac{\cos(p\lambda_{m,l})}{\lambda_{m,l}} & , \text{ para } m \text{ par} \\ 2(-1)^{(m-1)/2} \frac{\sin(p\lambda_{m,l})}{\lambda_{m,l}} & , \text{ para } m \text{ ímpar} \end{cases} \quad (7-51)$$

Substituindo-se as expressões (7-49) e (7-50) na equação (7-31), obtemos, finalmente:

$$f(p, \phi) = \sum_{m=0}^{\infty} \sum_{l=1}^{\infty} a_{m,l}^c b_{m,l} \cos m\phi + \sum_{m=0}^{\infty} \sum_{l=1}^{\infty} a_{m,l}^s b_{m,l} \sin m\phi \quad (7-52)$$

resultado análogo ao obtido por Cormack, onde a diferença reside apenas na expressão para os $b_{m,l}$.

7.3 Implementação Numérica

Embora seja necessária a soma de infinitos harmônicos para representar $f(p, \phi)$ de forma exata, experimentalmente isso não é possível, pois o **número de câmaras**, bem como o **número de detectores**, utilizados no sistema de detecção, limitam o número de harmônicos que podem ser determinados, sendo necessário, portanto, que a série seja truncada [55, 56].

Para N câmaras de detecção utilizadas, a brilhância pode ser medida em $2N$ valores de ângulo de impacto, ϕ . Portanto, pelo teorema de Nyquist, o maior harmônico que pode ser inteiramente resolvido pelo conjunto de câmaras é $m = N - 1$, e, para apenas umas das componentes, $m = N$ ($\sin N\theta$ ou $\cos N\theta$). Como serão utilizadas apenas duas

câmaras de detecção de raios-X, no TCABR: uma câmara vertical, que pode resolver, por si só, o modo $m = 0$ e o modo $\cos\theta$, e uma câmara horizontal, que pode resolver, por si só, também o modo $m = 0$, e o modo $\text{sen}\theta$ (com a informação proveniente de ambas as câmaras pode-se determinar tanto o modo $\cos 2\theta$, quanto o modo $\text{sen} 2\theta$, mas não ambos, simultâneamente) os limites máximos para o primeiro somatório da equação (7-23) tornam-se, desta forma, $m_{max}^c = 2$ e $m_{max}^s = 1$.

Para o segundo somatório, da mesma equação, a maior ordem para os polinômios de Zernicke é obtida de maneira análoga, com base no número de detectores existentes em cada câmara de detecção, determinando-se, assim, a resolução radial das medidas. Para que a parte radial da emissividade seja bem amostrada é necessário que o espaçamento entre os zeros do Polinômio de Zernicke de maior ordem seja maior, ou no mínimo igual, à separação entre as cordas. Na prática, porém, o valor máximo de l é obtido de uma forma semi-empírica: Adota-se um l_{max} , indicado pelo Teorema de Nyquist, e um dado **perfil teste**, de emissividade. Após a realização da reconstrução tomográfica, compara-se o resultado obtido com o perfil utilizado como teste e, a partir desta comparação, verifica-se se o l_{max} escolhido é suficientemente bom ou não.

Uma vez realizadas as medidas, para um determinado número de cordas, dado por $M \geq (l_{max} + 1)(m_{max}^c + m_{max}^s + 1)$, a expressão (7-23) torna-se um sistema de equações algébricas lineares para os coeficientes $a_{m,l}^{c,s}$, dada por:

$$\mathbf{f} = \mathbf{W} \cdot \mathbf{a} \quad (7-53)$$

Os elementos da matriz \mathbf{W} dependem somente da geometria e do número de detectores utilizados. Desta forma, uma vez definida esta matriz, ela não mais é alterada, podendo ser invertida e armazenada. Os coeficientes $a_{m,l}^{c,s}$ podem então ser calculados, em qualquer instante de tempo específico, através da simples multiplicação dada a seguir:

$$\mathbf{a} = \mathbf{W}^{-1} \cdot \mathbf{f} \quad (7-54)$$

7.3.1 Coeficientes de Zernicke e de Bessel

Estes coeficientes são então utilizados no somatório dos polinômios de Zernicke, que compõem os coeficientes de Fourier da emissividade, dados por:

$$g_m^{c,s}(r) = \sum_{l=0}^{l_{max}} a_{m,l}^{c,s} R_m^l(r) \quad (7-55)$$

Com estes coeficientes podemos finalmente realizar a soma da série de Fourier para a emissividade e obter, conseqüentemente, $g(r, \theta)$.

Este método é bastante eficiente e computacionalmente rápido. Além disso, os coeficientes $a_{m,l}^{c,s}$ são obtidos por um ajuste por mínimos quadrados, o que torna o procedimento menos sensível a ruídos. Não obstante, existem limitações que são inerentes a qualquer método numérico. Uma delas é o fato de se ter que truncar as séries que, no caso dos Polinômios de Zernicke, que não se anulam na região próxima da borda do plasma, ocasiona a geração de artefatos na referida região. Mas esta situação pode não ser muito restritiva, uma vez que a emissividade do plasma, próxima à borda, não é muito intensa.

O procedimento descrito para a obtenção dos coeficientes de Bessel é análogo ao descrito para os coeficientes de Zernicke, uma vez que as soluções são muito parecidas.

7.3.2 Programa TOMOG

O programa TOMOG, desenvolvido em FORTRAN, para fornecer imagens 2D das emissões de raios-X moles provenientes do tokamak TCABR, foi concluído a contento, e está operando de forma eficiente e rápida, condições necessárias para uma situação de análise de muitos dados tirados em um dia de operação de máquina.

O programa foi realizado para ler os dados experimentais, excluir aqueles sinais que eventualmente provierem de detectores defeituosos, quando for o caso, e apresentar uma série de gráficos na tela do computador que podem ser instantaneamente convertidos em arquivos "*postscript*", se desejado.

A título de ilustração, utilizou-se um conjunto de dados experimentais, do tokamak TEXT, da Universidade do Texas (Austin, E.U.A.), para apresentar as saídas gráficas do programa TOMOG, uma vez que ainda não se conseguiu obter dados experimentais do TCABR.

Após a leitura dos dados experimentais, o programa efetua os cálculos descritos na seção 7.2, apresentando a seguinte seqüência de gráficos:

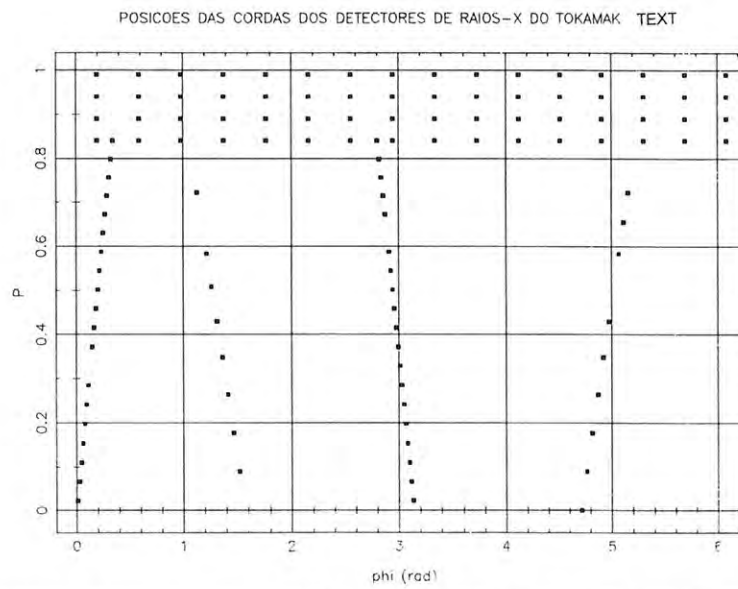


Figura 7-4: Disposição (coordenadas p, ϕ) dos detectores internos das duas câmaras de raios-X utilizadas no tokamak.

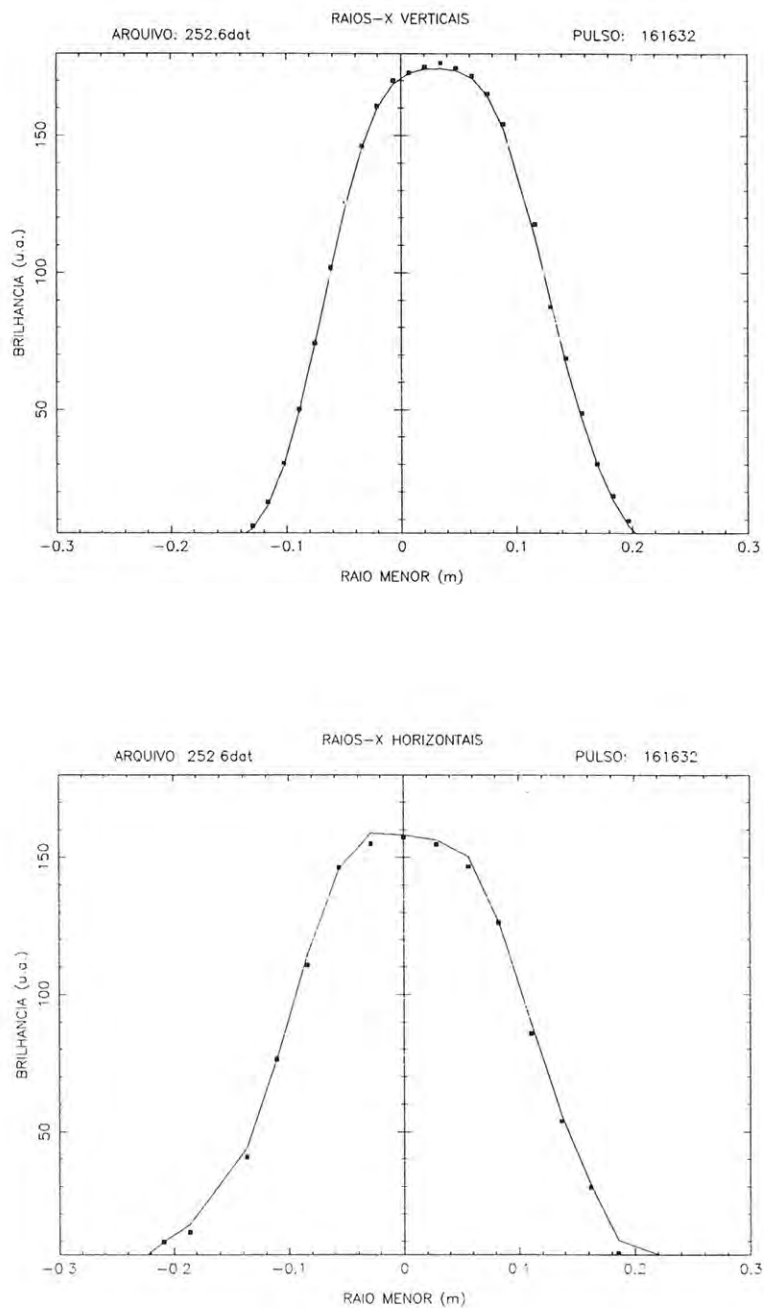


Figura 7-5: Comparação das brilhanças calculadas pelo TOMOG com os dados experimentais obtidos do TEXT.

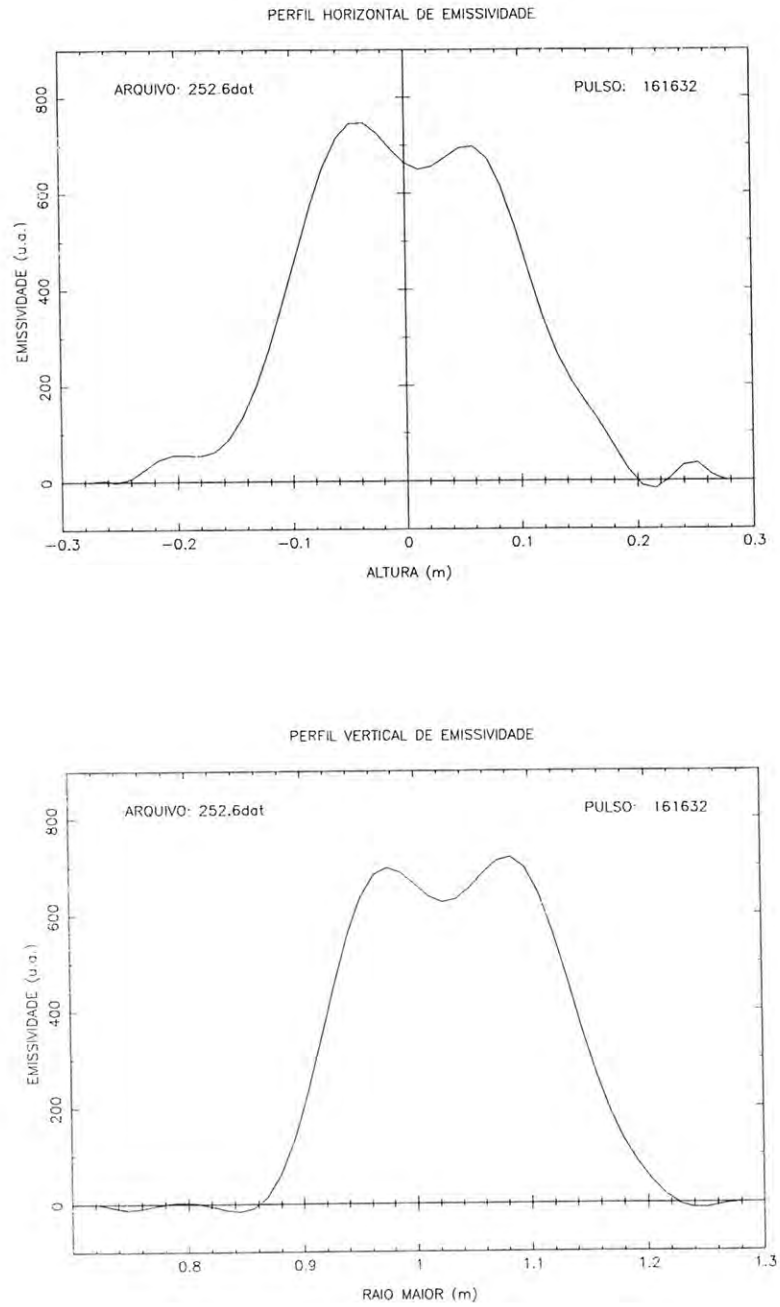


Figura 7-6: Perfis da emissividade vertical e horizontal fornecidos pelo programa TOMOG.

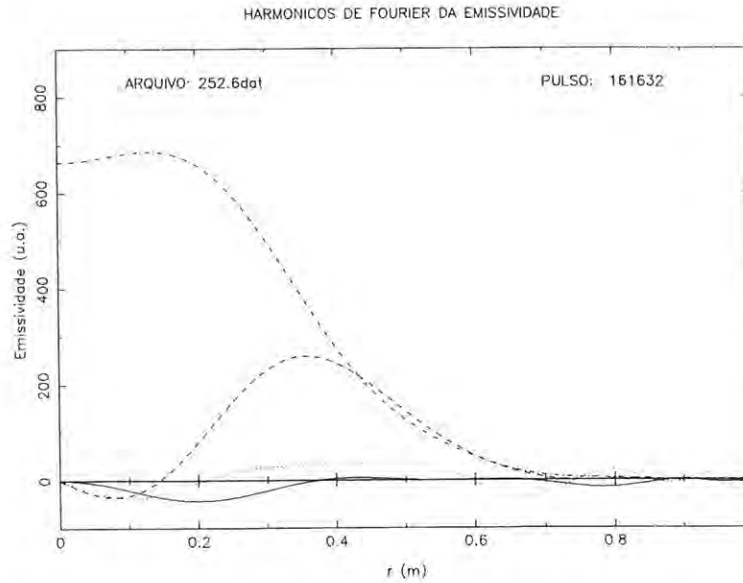
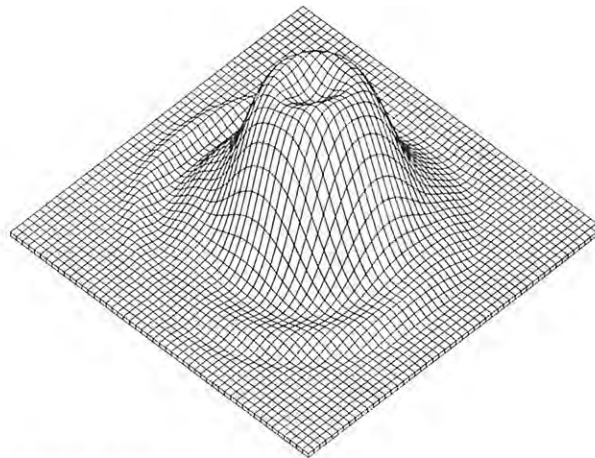


Figura 7-7: Harmônios de Fourier da emissividade obtida.

ARQUIVO: 252.6dat

PULSO: 161632



ANGULO: 50.0 graus

Figura 7-8: Imagem 3D da emissividade obtida.

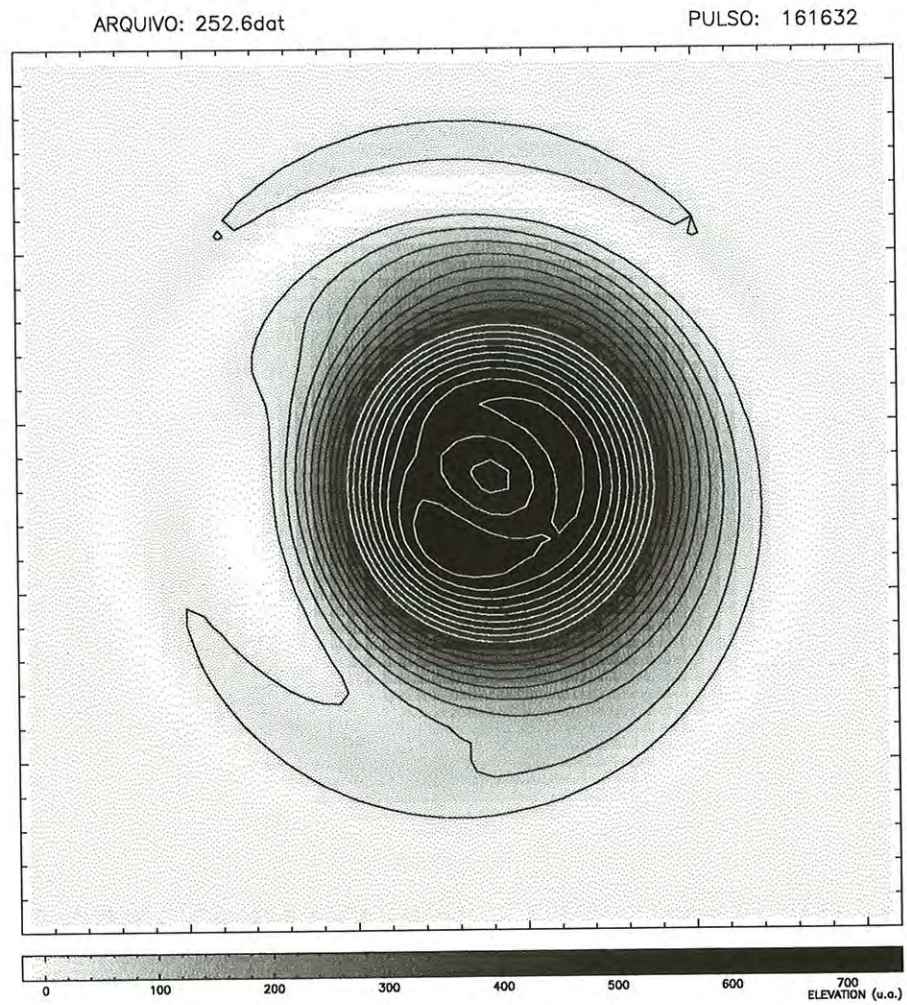


Figura 7-9: Curvas de nível da emissividade obtida.

Capítulo 8

Conclusões

Neste trabalho, verificou-se, com sucesso, que redes neurais artificiais podem ser utilizadas na previsão do comportamento futuro de sistemas dinâmicos não lineares.

Utilizando-se sistemas dinâmicos bem conhecidos, como o sistema de Lorenz, o mapa de Henon e o mapa logístico, dos quais se extraíram, aleatoriamente, várias séries temporais utilizadas no processo de previsão da rede neural adotada, verificou-se, para todos os casos, que uma arquitetura do tipo $m:2m:m:1$, onde m é a dimensão de imersão do sistema dinâmico em estudo, costuma ser um bom *chute* inicial para a construção da arquitetura de rede utilizada no sistema de previsão. De outra forma, o processo de escolha da arquitetura ideal seria bem mais difícil de se obter, uma vez que, em princípio, não existe um critério próprio para isso, sendo a sua escolha usualmente efetuada de modo empírico, tornando o trabalho árduo e bastante demorado.

Com este procedimento, obtiveram-se excelentes resultados quanto ao tempo de antecipação do comportamento futuro das séries temporais caóticas utilizadas, se comparados com resultados apresentados em trabalhos de outros autores. Em alguns casos, conseguiu-se melhorar o tempo de previsão das séries temporais do sistema de Lorenz, por exemplo, em até 8 vezes mais do que o obtido por Koga [15], que também utilizou redes neurais artificiais no processo de previsão, só que de uma maneira diferente daquela aqui descrita. O processo de previsão, utilizando-se redes neurais artificiais, é particularmente interessante porque, uma vez que a rede neural tenha aprendido o sistema dinâmico, nenhum outro treinamento se torna necessário, sendo a rede capaz de realizar, então, a previsão do comportamento futuro de qualquer outra série temporal referente àquela mesma dinâmica do sistema, o que já não ocorre com outros métodos de previsão, como o método da *aproximação local*, por exemplo [52].

Os bons resultados aqui obtidos, com respeito a séries temporais de sistemas dinâmicos já bem conhecidos, demonstram claramente que deve haver uma relação entre a escolha da arquitetura de rede, utilizada para fazer a previsão de uma determinada série temporal, e a dimensão de imersão do sistema dinâmico a ela relacionada.

Com relação ao principal objetivo deste trabalho, que envolve a utilização de redes neurais artificiais, na previsão de instabilidades de disruptura em plasmas magneticamente confinados por tokamaks, verificou-se que resultados muito melhores puderam ser obtidos ao se utilizarem sinais experimentais de raios-X de baixas energias, em vez de sinais magnéticos provenientes das bobinas de Mirnov.

De início, para a previsão das disrupturas, utilizando-se estes sinais de raios-X, adotou-se a mesma arquitetura de rede, o mesmo procedimento de treinamento, validação e previsão que havia sido utilizados com relação aos sinais de bobinas magnéticas, uma vez que os resultados, então obtidos com tal procedimento, tinham sido relativamente bons, conseguindo-se realizar uma antecipação de ~ 1 ms para o fenômeno disruptivo.

Desta forma, após um intenso esforço dispendido, conseguiu-se aumentar o tempo de previsão das instabilidades de disrupturas, utilizando-se o mesmo tipo de rede, processo de treinamento e validação anteriores, em até ~ 3 ms, muito embora o sinal fornecido pela rede tenha se mostrado deformado em relação ao sinal original [11].

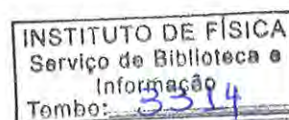
Somente quando se alterou a forma de abordar o problema da previsão, levando-se em conta a dinâmica da série temporal em questão, é que se obteve os excelentes resultados aqui apresentados, com um bom tempo de antecipação do fenômeno disruptivo (~ 4 ms), sem deformação significativa quando comparada com o sinal experimental.

Este tempo de previsão, finalmente conseguido, talvez seja suficiente para possibilitar a utilização de sistemas de defesa, tais como injeção de partículas neutras (ou "*pellets*"), aplicação de campos helicoidais ressonantes, etc, no sentido de se tentar abortar uma disruptura que estaria sendo desencadeada ou, pelo menos, atenuar os seus efeitos danosos.

Embora a rede neural não seja muito transparente, quanto à forma com que ela alcança os bons resultados de previsão das séries temporais utilizadas (essa é uma das desvantagens em se trabalhar com redes, quando se pretende entender fisicamente o problema em questão), algumas hipóteses podem ainda ser levantadas quanto aos prováveis mecanismos que poderiam desencadear os fenômenos disruptivos. Considerando-se que a rede neural foi capaz de prever as instabilidades de disrupturas utilizando sinais de raios-X, em instantes de tempo anteriores ao do crescimento das amplitudes da atividade MHD,

poder-se-ia especular, por exemplo, que talvez não seja o crescimento das ilhas magnéticas, relacionadas à superfície magnética $q = 2$, que perturbaria, conseqüentemente, a ilha interna em $q = 1$, como comumente aceito. Na verdade, algum distúrbio estaria ocorrendo primeiro no centro da coluna de plasma e, somente depois, é que a região do plasma, nas imediações de superfície em $q = 2$, seria afetada. Este mecanismo de ação já foi, inclusive, anteriormente proposto em outros trabalhos publicados, na tentativa de explicar a erosão observada no perfil da temperatura eletrônica obtido experimentalmente, considerando-se que uma “bolha fria” estaria movendo-se na direção do centro da coluna de plasma.

Finalmente, deve-se mencionar que um sistema de diagnóstico para medidas de raios-X de baixas energias no tokamak TCABR, do Instituto de Física, da USP, está em fase de implementação. Este sistema de diagnóstico, quando estiver funcionando, irá permitir o uso de redes neurais artificiais para dar continuidade ao estudo das previsões das instabilidades de ruptura, aqui iniciados, como também permitirá a realização de análises tomográficas do plasma, além de possibilitar a medida da temperatura eletrônica do plasma através do método dos absorvedores [20].



ex. 1

Bibliografia

- [1] Fiddler N.F., Nascimento I.C. - *Ciência Hoje* 7 (1988) 44.
- [2] Lasche, G.P. - *The Phys. Teacher* (1981) 234.
- [3] Weynants R.R. - *Trans. of Fus. Technol.* 25 (1994) 19.
- [4] Bateman G. - "*MHD Instabilities*", MIT Press (1978).
- [5] Vannucci A., Nascimento I.C. and Caldas I.L. - *Plasma Phys. and Control. Fus.* 31 (1989) 147.
- [6] Wesson J. - *Phys. Rev. Lett.* 44 (1980) 1215.
- [7] Westerhof E. - *Trans. of Fusion Technol.* 25 (1994) 152.
- [8] Noll P., Sonnerup L., Froger C., et al. - *Fus. Technol.* 15 (1989) 259.
- [9] Rebut P.H. - *Trans. Fus. Technol.* 27 (1995) 3.
- [10] Gruber O., Lackner K., Pautasso G., et al. - *Plasma Phys. and Control. Fus.* 35 (1993) B191.
- [11] Vannucci A., Oliveira K.A.de, Tajima T. - *Nucl. Fus.* 39 (1999) 255.
- [12] "*III Congresso Brasileiro de Redes Neurais*" - Anais, julho (1977) Florianópolis (SC) Brasil.
- [13] Rojas R. - "*Neural Networks*". Springer (1996).
- [14] Hertz J., Krogh A. and Palmer R.G. - "*Introduction to the Theory of Neural Computation*", Addison-Wesley Publishing Company (1991).
- [15] Koga J. - "*Predictions and Control of Complex Systems via Neural Networks*", Notes on Basic Science, Jaeri-Japan, 3 N^o2 July (1996) 37.

- [16] Diambra L., Plastino A. - Phys. Lett. A 216 (1996) 278.
- [17] Goeler S.von - "*Diagnostics for Fusion Experiments*", Pergamom Press (1979).
- [18] Goeler S.von, Stodiek W., et al. - Phys. Rev. Lett. 33 (1974) 1201.
- [19] Jahoda F.C. et al. - Phys. Rev. 119 (1960) 843.
- [20] Oliveira K.A.de, Vannucci A. et al. - Braz. J. of Phys. 28 (1998) 230.
- [21] Bell M.G., McGuire K.M., et al. - Nucl. Fus. 35 (1995) 1429.
- [22] Vannucci A. - "*Instabilidades de Disruptura em Tokamaks*", tese de Livre Docência apresentada ao Instituto de Física, da USP (1998).
- [23] del Bosco E. - "*Formação do Plasma no TBR-1*", dissertação de mestrado apresentada no Instituto de Física, da USP (1981).
- [24] Chen F.F. - "*Introduction to Plasma Physics*", Plenum Press (1974).
- [25] Artsimovich L.A. - Nucl. Fus. 12 (1972) 215.
- [26] Wesson J.A. - "*Tokamaks*", Clarendon Press (1987).
- [27] Caldas I.L., Vannucci A. - "*Equilíbrio e Instabilidades do Plasma em Tokamaks*", Report/IFUSP/P-523 (1985).
- [28] Caldas I.L., Tasso H. - Plasma Phys. 20 (1978) 1299.
- [29] Wesson J. A. - Nucl. Fus. 18 (1978) 87.
- [30] Rosebluth M.N., Rutherford P.H. - "*Tokamak Plasma Stability*", 1A Academic Press (1981).
- [31] Coppi B., Galvão R.M.O., Pellat R. - Sov. J. Plasma Phys. 2 (1976) 533.
- [32] Vannucci A., Gill R.D. - Nucl. Fus. 31 (1991) 1127.
- [33] Vannucci A., McCool S.C. - Nucl. Fus. 37 (1997) 1229.
- [34] Tsuji S., Nagayama Y., et al. - Nucl. Fus. 25 (1985) 305.
- [35] Greenwald M., Terry J.L., et al. - Nucl. Fus. 28 (1988) 2199.

- [36] Schüller F.C. - Plasma Phys. and Control. Fus. 37 (1995) A135.
- [37] Mirnov S. V. - Nucl. Fus. 9 (1969) 57.
- [38] Mirnov S. V., Semenov I. B. - Sov. Phys. JETP 33 (1971) 1134.
- [39] Araujo M.S.T., Vannucci A., et al. - referente ao trabalho de tese de doutoramento do Sr. Mauro Sérgio Teixeira Araújo, a ser submetido para publicação na revista *Physics of Plasmas*.
- [40] Melissinos A.C. - "*Experiments in Modern Physics*", Academic Press (1966).
- [41] Petrasso R., et al. - Rev. Sci. Instrum. 51 (1980) 585.
- [42] Nenu L., Cojocaru E., Mihailescu I.N. - "*Filter Method for Electron Temperature Determination from X-Ray Continuum Radiation Emitted by Hot Plasmas*", Bucharest Report LOP-10-80 (1980).
- [43] Kovács Z. L. - "*Redes Neurais Artificiais, Fundamentos e Aplicações*", Co-edição da Editora Collegium Cognitio e Edição Acadêmica, São Paulo, SP (1996).
- [44] Tafner M. A., et al - "*Redes Neurais Artificiais - Introdução e Princípios de Neurocomputação*", editora EKO, Blumenau, SC (1995).
- [45] Tafner M. A. - Revista Cérebro & Mente 2 N^o5 Maio (1998).
- [46] Hilton G. E. - Scientific Am., September (1992) 105.
- [47] Hernandez J. V., Vannucci A., et al - Nucl. Fus. 36 N^o8 (1996) 1009.
- [48] Internet - "<http://www-ra.informatik.uni-tuebingen.de/SNNS/>
- [49] Alligood K.T., Sauer T.D. and Yorke J.A. - "*CHAOS - An Introduction to Dynamical Systems*", Springer (1996).
- [50] Takens F. - "*Detecting Strange Attractors in Turbulence. In Dynamical Systems and Turbulence*", Lectures Notes in Mathematics, Springer-Verlag 898 (1981).
- [51] Lorenz E. - J. Atmospheric Science 20 (1963) 130.
- [52] Farmer J.D. and Sidorowich J.J. - Phys. Rev. Lett. 59 (1987) 845.

- [53] Oliveira K.A.de, Vannucci A. and Silva E.C. - “*Using Artificial Neural Network to Forecast Chaotic Time Series*”, submitted to *Physica A*.
- [54] Fredrickson E.D., et al. - *Nucl. Fus.* 33 (1993) 141.
- [55] Granetz R. S., Smeulders P.: *X-Ray Tomography on JET*, *Nuclear Fusion*, 28, n^o 3 (1988), 457.
- [56] Smith B. A.: *X-Ray Imaging of MHD Activity in the TEXT Tokamak*, Report FRCR # 384, The University of Texas (Austin, U.S.A.).
- [57] Cormack A. M.: *Representation of a Function by its Line Integrals, with some Radiological Applications*, *Journal of Applied Physics*, 34, n^o 9 (1963), 2722.
- [58] Cormack A. M.: *Representation of a Function by its Line Integrals, with some Radiological Applications. II*, *Journal of Applied Physics*, 35, n^o 10 (1964), 2908.
- [59] Arfken G.: *Mathematical Methods for Physicists*, Academic Press, Oxford (U.K.), Third Edition (1985).
- [60] Abramowitz M., Stegun I. A.: *Handbook of Mathematical Functions*, Dover Publications, New York (U.S.A.), Second Edition (1965).
- [61] Born M. and Wolf E.: *Principles of Optics*. Pergamon Press, Oxford (U.K.), Fifth Edition (1975).
- [62] Erdelyi A., Magnus W., Oberhettinger F., Tricomi F. G.: *Tables of Integral Transforms*, Mac Graw-Hill Book Company Inc., New York (U.S.A.), (1954).