

---

UNIVERSIDADE DE SÃO PAULO  
ESCOLA DE ENGENHARIA DE SÃO CARLOS  
DEPARTAMENTO DE ENGENHARIA ELÉTRICA

---

Serviço de Pós-Graduação EESC/USP

EXEMPLAR REVISADO

Data de entrada no Serviço.....25 / 08 / 04.....

Ass.:..........

**REGRAS DE OPERAÇÃO PARA SISTEMAS  
HIDROELÉTRICOS COM PREVISÃO  
POR REDES NEURAIIS**

**Rodrigo Sacchi**

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

Orientador: **Prof. Dr. Adriano Alber de França Mendes Carneiro**

DEDALUS - Acervo - EESC



31100047318

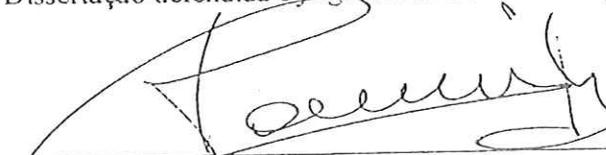


São Carlos  
Abril - 2004

FOLHA DE JULGAMENTO

Candidato: Engenheiro RODRIGO SACCHI

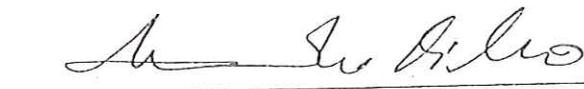
Dissertação defendida e julgada em 08-04-2004 perante a Comissão Julgadora:

  
\_\_\_\_\_  
Prof. Dr. ADRIANO ALBER DE FRANÇA MENDES CARNEIRO (Orientador)  
(Escola de Engenharia de São Carlos/USP)

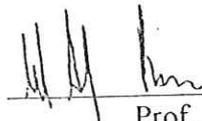
APROVADO

  
\_\_\_\_\_  
Prof. Dr. ALUÍZIO FAUSTO RIBEIRO ARAÚJO  
(Escola de Engenharia de São Carlos/USP)

APROVADO

  
\_\_\_\_\_  
Prof. Dr. MARINHO GOMES DE ANDRADE FILHO  
(Instituto de Ciências Matemáticas e de Computação/ICMC)

Aprovado

  
\_\_\_\_\_  
Prof. Assoc. MURILO ARAUJO ROMERO  
Coordenador do Programa de Pós-Graduação  
em Engenharia Elétrica

  
\_\_\_\_\_  
Profa. Tit. MARIA DO CARMO CALJURI  
Presidente da Comissão de Pós-Graduação

Dedico com amor e carinho à minha família, minha noiva e meus amigos.

## **Agradecimentos**

Ao professor Adriano Alber de França Mendes Carneiro pela oportunidade de trabalho e pela orientação, desde os tempos de graduação.

Aos professores Aluizio Fausto Ribeiro Araújo, André Carlos Ponce de Leon F. de Carvalho e Mário Oleskovicz pelas excelentes contribuições a este trabalho de pesquisa.

À minha família por sempre acreditarem no meu trabalho.

À minha noiva Maurícia, pelo carinho e companheirismo.

Aos colegas de laboratório (LSEE), pelo agradabilíssimo ambiente de trabalho.

À Coordenação de Formação de Pessoal de Nível Superior – CAPES, pela bolsa de Mestrado.

À Fundação de Amparo à Pesquisa do Estado de São Paulo – FAPESP, pelo suporte financeiro por meio de Projeto Temático.

Ao Departamento de Engenharia Elétrica da EESC-USP, por toda a infraestrutura.

“A Máquina,  
que produz em abundância,  
tem nos deixado na penúria.  
Nossos conhecimentos fizeram-nos céticos;  
nossa inteligência, empedernidos e cruéis.  
Pensamos em demasia e sentimos muito pouco.  
Mais do que máquinas, precisamos de humanidade.  
Mais do que de inteligência, precisamos de afeição e doçura.  
Sem estas virtudes, a vida será de violência e tudo estará perdido...”

*Charles Chaplin*

## Resumo

SACCHI, R. (2004). *Regras de Operação para Sistemas Hidroelétricos com Previsão por Redes Neurais*. Dissertação (Mestrado) – Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2004.

Este trabalho de pesquisa está relacionado com a aplicação de técnicas de Redes Neurais Artificiais (RNA's) na obtenção de políticas de operação conjunta de Usinas Hidroelétricas (UH's). O comportamento ótimo das Usinas Hidroelétricas depende tanto da posição relativa de cada UH ao longo da cascata como da relação entre elas. O propósito principal deste trabalho é de aplicar a rede de Mapa Auto-Organizável (SOM) de Kohonen com Modelos de Sistemas Dinâmicos, no aprendizado destas relações, objetivando seu uso para simular a operação ótima de um sistema hidroelétrico real. Este modelo é uma versão não-supervisionada do modelo NARX sendo chamado de modelo ou rede NARX Auto-Organizável (SONARX), capaz de processar padrões espaço-temporais. Desta maneira, um algoritmo de programação não-linear, especialmente desenvolvido para a realização da operação determinística de Sistemas Hidrotérmicos de Potência, é utilizado para otimizar a operação deste conjunto de usinas, considerando-se várias condições hidrológicas distintas e, tendo como função objetivo a minimização do custo com complementação térmica. Em seguida, os resultados desta otimização são então processados e utilizados no treinamento da rede SONARX; ela irá captar os comportamentos relativos de cada uma das UH's, trabalhando com afluências futuras que ela mesma irá prever. De forma geral, a partir dos resultados obtidos com um subsistema do Sistema Sudeste Brasileiro, pode-se concluir que a rede SONARX conseguiu assimilar o comportamento ótimo da operação do Sistema Teste. Os testes mostraram que o previsor neural possui uma forte tendência em seguir os resultados da otimização determinística, otimizando o uso dos recursos hídricos disponíveis para geração de energia elétrica. Posteriormente, a rede neural com Função de Base Radial (RBF) foi aplicada na interpolação das previsões feitas pela rede SONARX, uma vez que este tipo de rede é indicado para esta tarefa. Neste contexto, os centros das funções de base, da rede RBF, passam a ser definidos pela rede SONARX, mitigando o erro inerente a este tipo de modelagem baseada na técnica de MATQV.

Palavras-chave: otimização, planejamento de curto prazo, políticas de operação, redes neurais, simulação da operação, sistemas hidroelétricos

## Abstract

SACCHI, R. (2004). *Regras de Operação para Sistemas Hidroelétricos com Previsão por Redes Neurais*. M.Sc. Dissertation – Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2004.

This research work is related to the application of Artificial Neural Networks (ANN) in the obtaining of united operation policies in Hydroelectric Power plants (HP). The optimal behavior of the hydroelectric power plants depends as much on the relative position of each station throughout the cascade as on the relation between them. The main purpose of this work is to apply the Self-Organizing Map network (SOM) of Kohonen with Dynamic System Models, in the learning of these relations, objectifying its usage to simulate the optimal operation of a real hydroelectric power system. This model is an unsupervised version of the NARX model; which has been called Self-Organized model or network (SONARX), capable of processing space-time patterns. This way, a nonlinear programming algorithm, especially developed to accomplish the Hydrothermal Power System deterministic operation, is used to optimize the operation of this group of hydropower plants, considering many distinct hydrological conditions, and aiming the minimized cost with the thermal complementation. After this, the results of this optimization are processed and used in the SONARX network training; it will captivate the relative behavior of each one of the (HP), working with future affluences that it will predict. In a general way, from the acquired results from a subsystem of the Brazilian Southeast System, we may conclude that the SONARX network managed to assimilate the optimal behavior of the System Test operation. The tests showed that the neural predictor has a high tendency towards deterministic optimization results, optimizing the water resources use available for the electric energy generation. After all this, the Radial Basis Function (RBF) network was applied on the interpolation of the predictions made by the SONARX network, once this kind of network is indicated to this task. On this context, the centers of the basis functions, of RBF network, are to be defined by the SONARX network, mitigating the unrelated error to this type of modeling based on MATQV technique.

Keywords: optimization; short term planning; operation policies; neural networks; simulation of the operation; hydropower systems

## Lista de Figuras

Figura 3.1 – Sistema hidrotérmico de potência.....	19
Figura 3.2 – Função de custo com complementação térmica .....	20
Figura 3.3 – Fluxograma do Processo Iterativo do Simulador.....	24
Figura 4.1 – Topologia dos espaços de entrada e saída, e o mapeamento de características $\phi$ .....	28
Figura 4.2 – (a) Vetor de pesos de um dado neurônio $i$ . (b) Rede SOM bidimensional .....	29
Figura 5.1 - Sistema teste, parte da Bacia do Rio Paraná .....	45
Figura 5.2 - Evolução do custo de operação durante o processo de otimização.....	46
Figura 5.3 - Trajetória de mercado ao longo do período de otimização .....	46
Figura 5.4 - Trajetória dos volumes ótimos para o período de Jan/1935 até Dez/1944..	47
Figura 5.5 - Trajetória dos volumes ótimos para o período de Jan/1945 até Dez/1954..	48
Figura 5.6 - Trajetória dos volumes ótimos para o período de Jan/1955 até Dez/1964..	49
Figura 5.7 - Trajetória dos volumes ótimos para o período de Jan/1965 até Dez/1974..	50
Figura 5.8 - Trajetória dos volumes ótimos para o período de Jan/1975 até Dez/1984..	51
Figura 5.9 - Trajetória dos volumes ótimos para o período de Jan/1985 até Dez/1992..	52
Figura 5.10 – Erro de validação para previsão das Afluências de Furnas, para o modelo 1.....	56
Figura 5.11 - Erro de validação para previsão das Volumes Ótimos de Furnas, para o modelo 1.....	56
Figura 5.12 – Erro de validação para previsão das Afluências de Furnas, para o modelo 2.....	57
Figura 5.13 - Erro de validação para previsão das Volumes Ótimos de Furnas, para o modelo 2.....	57
Figura 5.14 – Erro de validação para previsão das Afluências de Furnas, para o modelo 3.....	58
Figura 5.15 - Erro de validação para previsão das Volumes Ótimos de Furnas, para o modelo 3.....	58
Figura 5.16 – Erro de validação para previsão das Afluências de Furnas, para o modelo 4.....	59
Figura 5.17 - Erro de validação para previsão das Volumes Ótimos de Furnas, para o modelo 4.....	59
Figura 5.18 – Erro de validação para previsão das Afluências de Furnas, para o modelo 5.....	60
Figura 5.19 - Erro de validação para previsão das Volumes Ótimos de Furnas, para o modelo 5.....	60
Figura 5.20 – Erro de validação para previsão das Afluências de Furnas, para o modelo 6.....	61
Figura 5.21 - Erro de validação para previsão das Volumes Ótimos de Furnas, para o modelo 6.....	61
Figura 5.22 – Erro de validação para previsão das Afluências de Furnas, para o modelo 7.....	62
Figura 5.23 - Erro de validação para previsão das Volumes Ótimos de Furnas, para o modelo 7.....	62
Figura 5.24 – Erro de validação para previsão das Afluências de Furnas, para o modelo 8.....	63
Figura 5.25 - Erro de validação para previsão das Volumes Ótimos de Furnas, para o modelo 8.....	63

Figura 5.26 – Erro de validação para previsão das Afluências de Furnas, para o modelo 9.....	64
Figura 5.27 - Erro de validação para previsão das Volumes Ótimos de Furnas, para o modelo 9.....	64
Figura 5.28 – Erro de validação para previsão das Afluências de Furnas, para o modelo 10.....	65
Figura 5.29 - Erro de validação para previsão das Volumes Ótimos de Furnas, para o modelo 10.....	65
Figura 5.30 – Erro de validação para previsão das Afluências de Furnas, para o modelo 11.....	66
Figura 5.31 - Erro de validação para previsão das Volumes Ótimos de Furnas, para o modelo 11.....	66
Figura 5.32 - Previsão das afluências para a usina de Furnas, para o período de Janeiro de 1936 até Dezembro de 1945.....	68
Figura 5.33 - Previsão das afluências para a usina de Furnas, para o período de Janeiro de 1946 até Dezembro de 1955.....	68
Figura 5.34 - Previsão das afluências para a usina de Furnas, para o período de Janeiro de 1956 até Dezembro de 1965.....	69
Figura 5.35 - Previsão das afluências para a usina de Furnas, para o período de Janeiro de 1966 até Dezembro de 1975.....	69
Figura 5.36 - Previsão das afluências para a usina de Furnas, para o período de Janeiro de 1976 até Dezembro de 1985.....	70
Figura 5.37 - Previsão das afluências para a usina de Furnas, para o período de Janeiro de 1986 até Dezembro de 1990.....	70
Figura 5.38 - Previsão dos volumes operativos ótimos para a usina de Furnas, para o período de Janeiro de 1936 até Dezembro de 1945.....	71
Figura 5.39 - Previsão dos volumes operativos ótimos para a usina de Furnas, para o período de Janeiro de 1946 até Dezembro de 1955.....	71
Figura 5.40 - Previsão dos volumes operativos ótimos para a usina de Furnas, para o período de Janeiro de 1956 até Dezembro de 1965.....	72
Figura 5.41 - Previsão dos volumes operativos ótimos para a usina de Furnas, para o período de Janeiro de 1966 até Dezembro de 1975.....	72
Figura 5.42 - Previsão dos volumes operativos ótimos para a usina de Furnas, para o período de Janeiro de 1976 até Dezembro de 1985.....	73
Figura 5.43 - Previsão dos volumes operativos ótimos para a usina de Furnas, para o período de Janeiro de 1986 até Dezembro de 1990.....	73
Figura 5.44 – Gráfico das previsões dos volumes ótimos do sistema teste, de janeiro de 1986 até dezembro de 1990.....	75
Figura 5.45 – Trajetória de Energia Armazenada Total do Sistema para o período de janeiro de 1986 até dezembro de 1990.....	76
Figura 5.46 - Número de vezes que cada dado de entrada contribuiu para a criação dos centros, durante o treinamento.....	79
Figura 5.47 - Número de vezes que cada dado de entrada contribuiu para a criação dos centros, durante o treinamento.....	80
Figura 6.1 – Três exemplos da dicotomia de separabilidade- $\phi$ para diferentes configurações de cinco pontos em duas dimensões: (a) separabilidade linear; (b) separabilidade quadrática;.....	85
Figura 6.2 – Arquitetura de uma Rede Neural com Função de Base Radial (RBF).....	88

Figura 6.3 – Comparação entre a rede SONARX em (a) e a rede SONARX-RBF em (b) na Previsão das afluências para a usina de Furnas, para o período de Janeiro de 1976 até Dezembro de 1985.....	93
Figura 6.4 – Comparação entre a rede SONARX em (a) e a rede SONARX-RBF em (b) na Previsão das afluências para a usina de Furnas, para o período de Janeiro de 1986 até Dezembro de 1990.....	94
Figura 6.5 – Comparação entre a rede SONARX em (a) e a rede SONARX-RBF em (b) na Previsão dos Volumes Operativos Ótimos para a usina de Furnas, para o período de Janeiro de 1976 até Dezembro de 1985.....	95
Figura 6.6 – Comparação entre a rede SONARX em (a) e a rede SONARX-RBF em (b) na Previsão dos Volumes Operativos Ótimos para a usina de Furnas, para o período de Janeiro de 1986 até Dezembro de 1990.....	96

## Lista de Tabelas

Tabela 1 - Tabela dos Erros de previsão das vazões afluentes para a usina de Furnas...	67
Tabela 2 - Tabela dos Erros de previsão dos volumes operativos ótimos para a usina de Furnas.....	67
Tabela 3 – Erros de Previsão dos Volumes Operativos Ótimos .....	76
Tabela 4 – Tabela comparativa entre a rede SONARX e a rede SONARX-RBF, para o período de Janeiro de 1976 até Dezembro de 1990 .....	97

## Lista de Abreviaturas e Siglas

<b>AG</b>	Algoritmos Genéticos
<b>CEPEL</b>	Centro de Pesquisas de Energia Elétrica
<b>CLFC</b>	Controle de Malha Fechada
<b>DET</b>	Despacho Econômico Térmico
<b>EESC</b>	Escola de Engenharia de São Carlos
<b>EPM</b>	Erro Percentual Médio
<b>LSEE</b>	Laboratório de Sistemas de Energia Elétrica
<b>MATQV</b>	Memória Associativa Temporal via Quantização Vetorial
<b>MIMO</b>	Múltiplas Entradas/Múltiplas Saídas ( <i>Multiple-Input/Multiple-Output</i> )
<b>MLP</b>	Perceptron de Multicamadas ( <i>MultiLayer Perceptron</i> )
<b>MRE</b>	Erro Relativo Médio ( <i>Mean Relative Error</i> )
<b>MSE</b>	Erro Médio Quadrático ( <i>Mean Squared Error</i> )
<b>NARX</b>	Modelo Auto-Regressivo Não-Linear com Variáveis Exógenas
<b>OLFC</b>	Controle de Malha Parcialmente Aberta
<b>ONS</b>	Operador Nacional do Sistema Elétrico
<b>PD</b>	Programação Dinâmica
<b>PDDE</b>	Programação Dinâmica Dual Estocástica
<b>PDE</b>	Programação Dinâmica Estocástica
<b>PNL</b>	Programação Não-Linear
<b>RBF</b>	Rede com Função de Base Radial ( <i>Radial Basis Function</i> )
<b>RNA</b>	Rede Neural Artificial
<b>ROO</b>	Regras de Operação Otimizadas
<b>ROR</b>	Regra de Operação de Reservatórios
<b>RPH<sub>i</sub></b>	Redução do Potencial Hidroelétrico da Usina <i>i</i>
<b>SISO</b>	Entrada-Única/Saída-Única ( <i>Single-Input/Single-Output</i> )
<b>SOM</b>	Mapa Auto-Organizável de Kohonen ( <i>Self-Organizing Map</i> )
<b>SONARX</b>	Modelo NARX Auto-Organizável ( <i>Self-Organizing NARX</i> )
<b>UH</b>	Usina Hidroelétrica
<b>USP</b>	Universidade de São Paulo
<b>WTA</b>	Rede competitiva <i>Winner-Take-All</i>

## Lista de Símbolos

$D(t)$	Demanda média durante o intervalo $t$ ;
$t$	intervalo de planejamento;
$MW$	Mega Watt
$\overline{MW}$	Mega Watt médio
$H(t)$	geração hidroelétrica do sistema durante o intervalo $t$ ;
$GT(t)$	geração termoelétrica complementar durante o intervalo $t$ ;
$C(t)$	função de custo com complementação térmica durante o intervalo $t$ ;
$I$	número de intervalos do horizonte de planejamento;
$j$	taxa de juros relativa a cada intervalo do horizonte de planejamento;
$ghi(t)$	geração média da usina hidroelétrica $i$ , durante o intervalo $t$ , em $\overline{MW}$ ;
$k_i$	constante que engloba aceleração da gravidade, densidade da água, rendimento turbina-gerador e fatores de conversão, da usina hidroelétrica $i$ ;
$vol_i(t)$	volume do reservatório da usina hidroelétrica $i$ , ao final do intervalo $t$ , em $hm^3$ ;
$def_i(t)$	defluência média da usina hidroelétrica $i$ , durante o intervalo $t$ , em $m^3/s$ ;
$h_{liq,i}(t)$	altura de queda líquida da usina hidroelétrica $i$ , durante o intervalo $t$ , em $m$ ;
$turb_k(t)$	turbinagem média da usina hidroelétrica $i$ , durante o intervalo $t$ , em $m^3/s$ ;
$af_{inc,i}(t)$	afluência incremental média à usina hidroelétrica $i$ , durante o intervalo $t$ , em $m^3/s$ ;
$\Omega_k$	conjunto das usinas à montante da usina hidroelétrica $i$ ;
$vol_{min,i}$	limites mínimo para o volume armazenado na usina hidroelétrica $i$ , em $hm^3$ ;
$vol_{máx,i}$	limites máximo para o volume armazenado na usina hidroelétrica $i$ , em $hm^3$ ;
$turb_{min,i}$	limites mínimo para a turbinagem da usina hidroelétrica $i$ , em $m^3/s$ ;
$turb_{máx,i}$	limites máximo para a turbinagem da usina hidroelétrica $i$ , em $m^3/s$ ;
$\lambda$	fator de acoplamento entre os reservatórios, $0 \leq \lambda \leq 1$ ;
$x(t)$	vetor de entrada da rede no instante $t$ ;
$w_i(t)$	vetor de pesos associado ao neurônio $i$ ;
$i^*$	neurônio vencedor ;
$h(i^*,i;t)$	função de vizinhança das redes SOM e SONARX;
$\alpha(t)$	é a taxa de aprendizagem de pesos para a fase de treinamento;
$u(t)$	vetor de entrada de um sistema qualquer no instante $t$ ;
$y(t)$	vetor de saída de um sistema qualquer no instante $t$ ;
$F(\cdot)$	função ou mapeamento não-linear de um dado sistema;
$P(t)$	padrão de entrada das RNAs;

$\hat{y}(t+1)$	saída da rede no instante $t$
$\overrightarrow{dist}$	distância Euclidiana;
$\  \cdot \ $	Norma Euclidiana <sup>1</sup>
$\varphi_j(r)$	Função de Base Radial da unidade neural $j$ , com relação a distância $r$ ;
$\sigma_j$	é a largura do campo receptivo da função de base radial da unidade $j$ ;
$\mu_j$	é a posição espacial do centro da função de base radial da unidade $j$ ;
$\eta$	taxa de aprendizagem para a camada de saída da rede RBF;

---

<sup>1</sup> A Norma utilizada neste trabalho de pesquisa foi a Norma Euclidiana Vetorial:  $\| \cdot \|_2$ .

## Sumário

Agradecimentos.....	ii
Resumo.....	iv
Abstract.....	v
Lista de Figuras.....	vi
Lista de Tabelas.....	ix
Lista de Abreviaturas e Siglas.....	x
Lista de Símbolos.....	xi
<b>1. Introdução.....</b>	<b>1</b>
1.1. ORGANIZAÇÃO GERAL DESTA DOCUMENTO.....	4
<b>2. Políticas de Operação.....</b>	<b>6</b>
2.1. OBJETIVOS DO TRABALHO.....	15
<b>3. Desenvolvimento da Pesquisa.....</b>	<b>17</b>
3.1. METODOLOGIA UTILIZADA.....	17
3.1.1. ALGORITMO DE OTIMIZAÇÃO.....	18
3.1.2. ALGORITMO DE SIMULAÇÃO.....	22
<b>4. Rede SONARX.....</b>	<b>25</b>
4.1. MODELAGEM MATEMÁTICA.....	25
4.1.1. A REDE SOM.....	26
4.1.2. A REDE SOM COM MODELOS DE SISTEMAS DINÂMICOS.....	32
4.1.3. ANÁLISE DA CONVERGÊNCIA DO MODELO SONARX.....	37
4.1.4. REPRESENTAÇÃO ESPAÇO DE ESTADOS DA REDE SONARX.....	39
4.1.5. POSSÍVEIS APLICAÇÕES.....	40
<b>5. Aplicação da rede SONARX em S.E.P.....</b>	<b>44</b>
5.1. SISTEMA TESTE.....	44
5.1.1. OTIMIZAÇÃO DETERMINÍSTICA.....	45
5.1.2. TREINAMENTO DA REDE SONARX.....	53
5.1.3. RESULTADOS DA PREDIÇÃO DE AFLUÊNCIA E VOLUMES ÓTIMOS.....	55
5.2. OPERAÇÃO SIMULADA PELO DECISOR NEURAL: SONARX.....	74
5.3. TESTES INVESTIGATIVOS PARA DEFINIR “QUEM” GERA OS CENTROS.....	78

<b>6. Redes com Funções de Base Radial (RBF) .....</b>	<b>82</b>
<b>6.1. TEOREMA DE COVER: SEPARABILIDADE DOS PADRÕES .....</b>	<b>83</b>
<b>6.2. PROBLEMA DE INTERPOLAÇÃO .....</b>	<b>85</b>
<b>6.3. MODELAGEM DA REDE RBF .....</b>	<b>88</b>
<b>6.4. RESULTADOS DA PREDIÇÃO COM INTERPOLAÇÃO PELA REDE RBF .....</b>	<b>93</b>
<b>7. Conclusões.....</b>	<b>98</b>
<b>8. Bibliografia .....</b>	<b>101</b>

# 1. Introdução

No planejamento da operação de sistemas hidrotérmicos de potência busca-se definir uma política de operação que atenda a demanda de energia elétrica no sistema, com confiabilidade e pelo menor custo operativo. Em outras palavras, isto equivale a determinar um cronograma ótimo de geração para cada usina do sistema, a cada intervalo, permitindo que o sistema atenda a demanda de forma confiável, sobretudo, minimizando o custo esperado de operação durante o período de planejamento (Carneiro; Kadowaki 1996).

Este processo de planejamento é muito complexo, uma vez que as afluências futuras não são conhecidas, caracterizando-o como um problema estocástico. Devido ao grande número de usinas envolvidas no processo de otimização, tem-se um grande número de variáveis, característico de problemas de grande porte. Além disso, as usinas que compõem um sistema de geração são altamente interligadas, tanto eletricamente, como hidraulicamente, o que o faz ser interconectado. Como a função de custo a ser minimizada e algumas equações dos modelos são não-lineares, temos também um problema não-linear. Por fim, as decisões tomadas em um intervalo de planejamento dependem das decisões tomadas no passado e determinam a evolução futura do sistema, o que permite classificar o problema também como dinâmico (Pereira 1985).

Devido à complexidade do planejamento e aos diferentes aspectos que devem ser abordados, o mesmo é dividido em horizonte de médio, curto e curtíssimo prazos. Nos horizontes de médio e curto prazo é tratado o aspecto energético da operação, que está relacionado com os estados dos reservatórios, em cada intervalo do horizonte de planejamento. Enquanto que no horizonte de curtíssimo prazo são tratados os aspectos elétricos do sistema, como por exemplo os limites de transmissão elétrica (Pereira 1985).

No Brasil, o planejamento de longo prazo era realizado em um horizonte de cinco anos, com discretização mensal, e devido às incertezas relacionadas às vazões

a

b

c

d

e

f

afluentes dos reservatórios, utilizava-se a Programação Dinâmica Estocástica (PDE). Apesar desta técnica preservar a estocasticidade do problema, ela apresenta a conhecida “maldição da dimensionalidade” se preservada a representação individualizada das Usinas Hidroelétricas (UH’s). Sendo assim, a representação do sistema hidroelétrico passa a ser feita através de um reservatório equivalente em energia, ao qual aflui energia, ao invés de água.

Em se tratando desta modelagem, tanto a água armazenada em cada reservatório, como as vazões afluentes devem ser transformadas em energia armazenada e energia afluente, respectivamente. Para tanto, faz-se necessária a utilização de Regras de Operação de Reservatórios (ROR), a qual é de fundamental importância para a definição do Reservatório Equivalente em Energia (Fortunato; Araripe Neto et al. 1985). Quanto mais próxima da realidade forem estas Regras, mais fiel será o Reservatório Equivalente.

Um fator importante na agregação dos reservatórios é a definição das Regras de Operação que estabelecem a maneira na qual os reservatórios devem operar. A Regra adotada no planejamento de operação do sistema brasileiro, é uma Regra de Operação em Paralelo, no qual estabelece que todos os reservatórios em um sistema devem operar com a mesma porcentagem de armazenamento dos reservatórios. Entretanto, vários estudos de operação ótima de reservatórios em geração hidroelétrica mostram que o comportamento de cada reservatório depende de sua localização na cascata (Cruz Jr.; Soares Filho 1999; da Silva; Finardi 2001).

O reservatório de cabeceira (montante), por exemplo, deve regularizar a vazão afluente utilizando seu reservatório de acumulação, enquanto que os reservatórios de jusante devem ser mantidos cheios a maior parte do tempo, operando como usinas a fio d’água. Esses estudos têm sugerido que a Regra de Operação em Paralelo subestima a capacidade de geração de sistemas hidrotérmicos.

Feito o planejamento de longo prazo, estarão definidas dentre outras coisas, as metas energéticas de operação do reservatório equivalente em energia para todo o horizonte de planejamento. Entretanto estas metas se apresentam de forma agregada, sendo necessária a realização da desagregação destas metas; o que pode ser feito através de simulações operativas. Novamente, nesta etapa de simulação são utilizadas Regras de Operação, distribuindo a demanda entre as usinas do sistema (Soares Filho; Carneiro 1993).

Recentemente, o Operador Nacional do Sistema Elétrico (ONS) adotou para o planejamento da operação em sistemas hidrotérmicos de potência, um modelo baseado em Programação Dinâmica Dual Estocástica (PDDE), o qual permite a representação de múltiplos reservatórios em reservatórios equivalentes, divididos em quatro subsistemas (Norte, Nordeste, Sudeste/Centro Oeste e Sul), procurando preservar as características das bacias hidrográficas. Esse pacote é chamado de *Newave* simula a operação do sistema ao longo do período de planejamento, para distintos cenários de seqüências hidrológicas, falhas dos componentes e variações da demanda. Calcula índices de desempenho, tais como a média dos custos de operação, dos custos marginais, o risco de déficit, os valores médios de energia não suprida, de intercâmbio de energia e de geração hidroelétrica e térmica (Rotting; Gjelsvik 1992; CEPEL 2000).

Sendo assim, as RORs são importantes em várias etapas dentro do Planejamento da Operação, principalmente com o objetivo de estudar a performance do Sistema Hidrotérmico, através de simulações operativas. Isto porque os modelos de simulação permitem que a operação seja simulada em condições próximas à real, onde afluências futuras não são conhecidas. Por outro lado, esses modelos de simulação não possuem algoritmos de otimização, realizando a operação do sistema com base em regras que especificam como as usinas devam se comportar, testando os sistemas em diversas condições, o que é de suma importância para o bom desempenho dos mesmos. (Carneiro; Kadowaki 1996).

No processo de Planejamento da Expansão uma das mais importantes etapas é o dimensionamento energético de uma usina hidroelétrica. Este dimensionamento visa especificar cinco parâmetros básicos, os quais determinam a capacidade de produção de energia de um aproveitamento: os volumes mínimo e máximo do reservatório, a potência instalada, a altura de referência e a altura de projeto das turbinas. Durante a definição destes parâmetros, alterações em seus valores devem ser consideradas sempre que os benefícios energéticos incrementais resultantes forem maiores que os respectivos acréscimos de custo. O benefício de um aproveitamento equivale à valorização econômica de seus ganhos energéticos. Estes ganhos correspondem não somente à energia que a usina é capaz de gerar, mas também aos eventuais ganhos que esta usina possa proporcionar às outras usinas do sistema no qual ela será inserida.

Assim, para avaliar os benefícios de uma usina hidroelétrica, deve-se primeiro definir um Sistema de Referência, que corresponde ao sistema existente no qual a usina será incorporada. Deve-se então proceder à simulação da operação deste sistema e

operativos ótimos, os quais podem ser utilizados para a extração de Regras de Operação baseadas em critérios ótimos.

No Capítulo 3 será descrito com mais detalhes esta metodologia de otimização determinística. Onde será apresentada a formulação do problema de planejamento da operação de um sistema hidrotérmico de potência, bem como o algoritmo de otimização utilizado. Também neste capítulo, será apresentado o algoritmo de simulação que por sua vez necessita de um “guia” para a tomada de decisões operativas a cada instante do horizonte do planejamento. Neste trabalho as RORs exercem justamente este papel de “guia” na tomada de decisões operativas.

Já o Capítulo 4 trata da modelagem matemática da rede SONARX, descrevendo de forma detalhada a rede SOM. Em seguida é introduzida a rede SONARX, uma versão não-supervisionada do modelo NARX. E finalmente, algumas aplicações possíveis para esta rede, como por exemplo a previsão de séries temporais, que faz parte do escopo deste trabalho.

O Capítulo 5 está relacionado com a parte prática, ou seja da aplicação desta Rede Neural Artificial (RNA) na obtenção da predição de afluência e volumes ótimos. Neste capítulo é apresentado o Sistema Teste, constituído por Usinas Hidroelétricas pertencentes ao Subsistema Sudeste Brasileiro e, por fim são apresentados alguns testes da rede SONARX.

No Capítulo 6 é apresentada a rede neural com Funções de Base Radial (RBF), aplicada na interpolação das previsões feitas pela rede SONARX, uma vez que este tipo de rede é indicado para esta tarefa. Neste contexto, os centros das funções de base, da rede RBF, passam a ser definidos pela rede SONARX, deixando o ajuste dos pesos de conexão com a camada de saída, da rede RBF, para um treinamento supervisionado do tipo regra Delta.

Finalmente, tem-se o Capítulo 7, referente às conclusões decorrentes dos resultados apresentados no capítulo anterior, e onde estão discutidas algumas propostas de continuidade deste trabalho de pesquisa.

## 2. Políticas de Operação

Modelos de simulação são amplamente utilizados em estudos de planejamento da operação de sistemas hidrotérmicos, sendo em geral, modelos que necessitam de regras de operação para que possam se implementados. As regras de operação usadas em simulações muitas vezes imitam decisões operacionais verdadeiras. Na prática, as usinas são operadas de acordo com regras que definem as metas de armazenamento e defluência.

A popular “Regra de Espaço”, proposta de (Bower; Hufschmidt et al. 1962), sugere que em um sistema de reservatórios em paralelo, os espaços reservados de cada reservatório devam ser proporcionais a afluência esperada para cada um deles. Como as outras diretrizes heurísticas, esta regra de espaço especifica defluências ideais e níveis de armazenamento para cada um dos reservatórios.

Uma Política de Operação para modelos de simulação de sistemas hidroelétricos pode ser definida implicitamente pela inclusão de um modelo de otimização, que determine as decisões de defluência a cada instante. Este modelo de otimização geralmente possui restrições que estabelecem tanto as descargas factíveis quanto os níveis de armazenamento ao fim do período de simulação, além da função objetivo que reflete os benefícios associados com diferentes estados finais dos reservatórios.

(Johnson; Stedinger et al. 1991) mostraram como políticas de operação eficientes para sistemas de reservatórios em paralelo podem ser obtidas através de um modelo de simulação e por um modelo de otimização de cada etapa do período de planejamento, o qual, a cada etapa, incorpora novas restrições ao modelo e usa diretrizes heurísticas para definir os níveis de armazenamento ao final do período de planejamento.

Modelos de cada etapa (mês) para períodos quando os reservatórios estão reenchendo foram desenvolvidos, os quais geram níveis de armazenamento ao final do período que satisfaz a regra de espaço, ou minimiza o vertimento esperado, quando tais

espaços estão disponíveis. Um modelo de otimização foi também desenvolvido para períodos quando os reservatórios estão sendo esvaziados.

A necessidade de funções objetivos heurísticos e restrições adicionais para abordar problemas pluri-mensais pode ser evitada, através do uso de função objetivo advinda da resolução de um modelo de PDE, o qual é capaz de capturar as características relacionadas com a dinâmica do problema (Johnson; Stedinger et al. 1991).

Como dito anteriormente, as ROR's são de fundamental importância, já que são elas que determinam quanto de energia se pode obter de um sistema e também determinam o estado do reservatório de cada usina. Estas regras exercem um papel particularmente relevante no caso do Brasil, já que na quase totalidade das usinas hidroelétricas brasileiras a altura de queda da água depende, de forma não linear, do volume do reservatório, característica esta chamada de efeito cota.

Neste caso, o estado do reservatório afeta a produtividade da usina e, portanto, a trajetória de volume das usinas em um período de operação irá afetar significativamente a energia total gerada pelo sistema. Em sistemas constituídos por usinas em cascatas, este efeito torna-se mais grave nas usinas mais à jusante, por onde passa toda a água, e onde uma pequena perda de produtividade pode significar uma grande perda de energia.

A regra básica normalmente adotada no modelo empregado no Sistema Brasileiro é a de que, por hipótese, os reservatórios operem em paralelo, isto é, os armazenamentos e deplecionamentos são feitos paralelamente em volume. Esta regra tem como maior atrativo sua simplicidade, porém, está longe de uma operação otimizada. Mesmo sua variante, a de operação em paralelo por faixas, deixa a desejar, já que persiste o problema de como determinar estas faixas de forma otimizada (Soares Filho; Carneiro 1991).

Desta maneira, (Carneiro; Kadowaki 1996) propõem uma metodologia para obtenção de ROR's mais próximas do comportamento ótimo dos sistemas hidrotérmicos, baseado no estudo do comportamento de reservatórios em cascata no planejamento ótimo da operação, realizado através de uma ferramenta determinística, não linear e à usinas individualizadas.

O padrão de comportamento diferenciado entre as usinas, em uma operação otimizada, foi a base para a busca de regras que refletissem um comportamento operativo ótimo das diversas usinas, ou seja, que reproduzissem, em média, as trajetórias obtidas pelo modelo de otimização (Soares Filho; Carneiro 1993).

Para tanto, procurou-se relacionar, na operação ótima, o estado de armazenamento de cada usina com o estado agregado de todo o sistema, de tal forma que, a cada estado do sistema como um todo correspondesse um estado particular para cada reservatório.

Assim, a obtenção das ROR's em um sistema é feita a partir da simulação da operação otimizada, sob diversas condições hidrológicas, onde se obtém, para cada reservatório, um conjunto de pontos que relacionam o estado do reservatório, através de seu volume, com o estado global do sistema, representado pela sua energia armazenada (Soares Filho; Carneiro 1993).

Esta primeira etapa gera uma nuvem de pontos para cada reservatório que devem ser usados como indicativos para uma regra de operação. Assim, para cada reservatório, faz-se o ajuste de uma curva que irá representar, na média, a operação ótima da usina em questão. Sendo que o conjunto de curvas de todas as usinas, que podem ser representadas por equações, constituem as ROR's otimizadas para um dado sistema.

As regras obtidas foram avaliadas através de sua performance na simulação da operação do Sistema Sudeste Brasileiro, obtendo-se ganhos energéticos evidentes em todas as situações hidrológicas, quando comparadas com a simulação em paralelo (Soares Filho; Carneiro 1993).

Embora as regras assim obtidas tenham apresentado boa performance nos testes realizados, por outro lado, devido ao espalhamento dos pontos, nem sempre se consegue um ajuste adequado para as curvas. As regras ficam um pouco dependentes do particular ajuste realizado. É evidente que os resultados obtidos são bem superiores aos das regras heurísticas mais comumente utilizadas, porém, este processo pôde ser aperfeiçoado.

Por esta razão, as regras apresentadas por (Carneiro; Kadowaki 1997) foram obtidas utilizando-se uma metodologia similar para a obtenção das regras otimizadas. As regras foram obtidas utilizando-se o algoritmo de otimização não linear e, posteriormente, aplicando-se técnicas de redes neurais artificiais. Assim, o estudo do comportamento de reservatórios em cascata no planejamento ótimo da operação, realizado através desta ferramenta à usinas individualizadas, forneceu um conjunto de dados que serviram para treinamento, por um algoritmo tipo *backpropagation*, de uma rede neural *feedforward*, a qual gerou as ROR's mais próximas do comportamento ótimo dos sistemas.

Consegue-se assim gerar regras mais versáteis, por um processo mais adequado à automação, e que trabalha mais facilmente com o espalhamento de pontos em algumas

usinas. Quando comparadas com a regra em paralelo, as oriundas das redes neurais apresentaram grandes ganhos energéticos, como já era de se esperar. Agora, quando comparadas com as regras obtidas por ajuste de curvas, o ganho energético não foi tão superior assim. Entretanto, o aspecto de automatização que os procedimentos por redes neurais apresentam devem ser levados em consideração.

Apesar dos bons resultados apresentados por estes métodos, algumas dificuldades foram encontradas no ajuste das Regras de Operação Otimizadas (ROO). Com o objetivo de superar tais dificuldades e melhorar a precisão dos ajustes, (Carneiro; Silva Filho 1998a) propuseram a obtenção de ROO's através da criação de uma base de conhecimento Fuzzy, onde os pontos que melhor caracterizam a relação entre as variáveis utilizadas nas regras, os chamados centróides, são obtidos pelo Método da Montanha.

A nova técnica de ajuste proposta pode ser dividida em duas partes: na primeira, é necessário que sejam encontrados pontos, os centróides, que melhor caracterizem a relação entre os dados, e na segunda, cria-se uma Base de Conhecimento Fuzzy que utiliza os centróides para estabelecer a ROO desejada.

Esta abordagem promoveu a completa automação da metodologia para a obtenção das regras otimizadas, onde as regras foram obtidas utilizando-se o algoritmo de otimização não linear e, posteriormente, aplicando-se as técnicas do Método da Montanha e Lógica Fuzzy, não havendo a necessidade de intervenção por parte do operador durante o processo de ajuste (Carneiro; Silva Filho 1998b).

Novamente diante do algoritmo de PNL, especialmente desenvolvido para a realização da operação determinística de Sistemas Hidrotérmicos de Potência, e utilizado para otimizar a operação de um conjunto de usinas, considerando-se várias condições hidrológicas distintas, percebe-se que o aspecto mais difícil de ser captado é a de como as diversas situações interferem nos estados, volumes, dos reservatórios. Este tipo de simulação é muito importante na análise do comportamento dos reservatórios, permitindo observar como eles relacionam-se entre si.

Assim, (Silva Filho; Carneiro et al. 1999b) propuseram a extração destas relações através do treinamento de uma Rede Neural Artificial com Funções de Base Radial (Rede RBF), onde os resultados destas simulações foram previamente processados.

Após treinada, a Rede RBF é integrada a um simulador da operação do sistema hidroelétrico, onde a cada instante de tempo é necessário que ela decida quanto cada

usina deve gerar para que uma demanda solicitada seja atendida. As simulações não são determinísticas e, a cada intervalo de simulação, as decisões devem ser tomadas sem nenhum conhecimento sobre o futuro.

Foi então constatado que a rede RBF conseguiu assimilar o comportamento ótimo da operação de sistemas hidroelétricos, apresentando uma excelente performance durante as simulações.

Já (Oliveira; Loucks 1997) apresentam estudos para encontrar políticas de operação pré-definidas que indiquem, em um sistema de reservatórios, as metas de defluência total e as metas de volumes armazenados individuais, como função da época do ano e do volume total armazenado no sistema. Sendo então aplicados Algoritmos Genéticos (AGs) para propor políticas de operação de sistemas de usinas hidroelétricas.

A defluência de cada reservatório é determinada na tentativa de minimizar o desvio entre os volumes armazenados atuais de cada reservatório e as metas de volumes armazenados desejados, enquanto é atendida a defluência total especificada do sistema. Estas políticas de operação definem metas de armazenamento de médio termo e longo termo, e metas de defluência total. Elas podem ser usadas para guiar os operadores em seus processos de tomadas de decisão.

Além disso, modelos de otimização podem ser usados para designar o cálculo das defluências de curto prazo que satisfazem todas as metas do sistema e deixam o sistema num estado, o mais próximo possível, daquele definido pelas metas de longo termo. Estes modelos de otimização podem ser usados como ferramenta de apoio à decisão da operação diária de sistemas de reservatórios.

Tanto as políticas pré-definidas de longo termo, como os modelos de otimização em tempo real podem ser incorporados em modelos de simulação de sistemas de reservatórios para estudar o impacto das regras de operação sob diferentes condições hidrológicas.

O AG usado naquele trabalho, empregou cromossomos de valores reais, contendo as coordenadas dos pontos que definem as funções das regras de operação lineares por partes. O esforço computacional das simulações, destas regras extraídas por AG, de sistemas médios de usinas hidroelétricas não se mostraram muito custosos.

Em outra abordagem (Lund; Guzman 1999) revisaram uma variedade de políticas de operação mono objetivas para reservatórios em série e em paralelo para o suprimento de água, controle de cheia, geração hidroelétrica, qualidade da água, e recreação. Tais regras são úteis para operação em tempo-real, estudos de simulações em

tempo-real, sazonais, e operações de longo termo, e para a compreensão do comportamento de sistemas a multireservatórios.

Com relação ao desenvolvimento e crescente uso de modelos de otimização, a vasta maioria dos planejamentos de reservatórios e estudos da operação é baseada em modelos de simulação, logo requerem especificações inteligentes das regras de operação.

Estas regras determinam as decisões de defluência e armazenamento para cada reservatório a cada instante de tempo durante a simulação e ajuda na tomada de decisão do operador de sistemas.

As regras examinadas por (Lund; Guzman 1999) são geralmente aplicadas em estudos sazonais e de médio prazo. Estudos em tempo-real, com discretização horária ou diária, geralmente têm mais detalhes de segurança, habitat e limitações de recursos que não são tão importantes para os estudos que usam formulações mais grosseiras como os de longo prazo.

Já a operação multi-objetiva requer alguma combinação das regras mono-objetivas. A premissa de muitas das regras de operação é que pode ser mais vantajoso operar um sistema de reservatórios, do que operar os reservatórios individualmente.

Regras de hidroeletricidade para reservatórios em séries varia entre estações de enchimento e de esvaziamento. Durante o período de enchimento, normalmente o problema é maximizar o nível de armazenamento de energia ao fim deste período. Durante o período de esvaziamento, o objetivo é maximizar a produção hidroelétrica para um dado armazenamento total.

O objetivo da regra de armazenamento de energia para reservatórios em série, proposta por (Lund; Guzman 1999), é de maximizar a energia total armazenada no final do período de enchimento. O período de enchimento foi definido como a estação quando as afluências do sistema excedem aquelas necessárias para o atendimento da demanda de energia elétrica. Esta regra de armazenamento de energia indica que geralmente deve-se encher primeiro os reservatórios de montante.

A água armazenada numa cota maior tem uma maior energia armazenada do que a água armazenada numa cota menor. Isto é particularmente verdade para a água armazenada em reservatórios em série, onde a água defluída do reservatório de montante pode gerar hidroeletricidade também nos reservatórios de jusante. Qualquer vertimento dos reservatórios de montante pode ser capturado pelos espaços alocados nos reservatórios de jusante.

Felizmente, as regras de armazenamento de energia e de suprimento de água para reservatórios em série são bem compatíveis para o período de enchimento, ao menos em termos de onde o armazenamento deve ser feito e na intenção de acumular a máxima quantidade de água. Entretanto, com a chegada do período de esvaziamento, regras de produção de energia são necessárias (Kelman; Damazio et al. 1989).

Reservatórios de montante geram eletricidade despachando suas turbinas, que conseqüentemente aumentam a geração de energia das usinas de jusante, ou devido ao aumento da cota de seus reservatórios, ou pelo aumento de suas turbinagens.

A regra de produção de energia de “steady-state”, apresentada por (Lund; Guzman 1999), maximiza a geração de energia em um único instante de tempo, dada uma meta total de armazenamento do sistema, prioritariamente durante a estação de esvaziamento. Em geral, a regra favorece o armazenamento nos reservatórios que proporcionem uma maior queda por unidade de volume armazenado, que tenham maior eficiência de geração, e tenham maior turbinagem, já que produção de energia é o produto da queda, da eficiência, e da turbinagem.

Dessa maneira, quando reduções no armazenamento do sistema são necessárias, elas devem ser feitas a partir dos reservatórios com menor capacidade de produzir energia. De modo contrário, se um aumento no armazenamento do sistema for necessário, ele deve ser feito nos reservatórios com maiores capacidades de produzir energia.

Existem muitas oportunidades para o uso de métodos de otimização dentro de modelos de simulação. Tais como a implementação das regras de alocação de armazenamento para reservatórios em série ou em paralelo, dentre outras, bem como as regras de enchimento ou de esvaziamento, que se aplicam em períodos distintos do ano, sendo uma importante área para estudos futuros.

Algumas regras de operação de usinas hidroelétricas foram desenvolvidas e discutidas por (Lund 2000), para reservatórios em série, em paralelo, e para reservatórios individuais para os casos onde os reservatórios re-enchem sem que atinjam o mínimo operativo, e para quando reservatórios em paralelo fossem esperados atingir seu mínimo operativo enquanto esvaziavam. Tais regras de operação apresentaram uma simplificação econômica para a alocação de armazenamento e energia em sistemas hidroelétricos a multi-reservatórios.

Os reservatórios que apresentassem as menores Reduções do Potencial Hidroelétrico ( $RPH_i$ , em inglês) deveriam ser esvaziados primeiro. Conceitualmente, o

$RPH_i$  é proporcional a perda de produção de energia, até o total re-enchimento do reservatório, por uma redução unitária do nível atual de armazenamento (Lund 2000).

Para sistemas com poucos reservatórios, com predominância para a geração de energia, tais regras podem ser os pontos de partida para estudos mais detalhados de modelagem. Para sistemas de grande porte e multi-objetivos, estudos adicionais de otimização e simulação se fazem necessários.

No contexto do Sistema Hidroelétrico Brasileiro, em que as usinas são consideradas de baixa e média queda, as usinas de jusante tendem a serem maiores, no que se refere ao nível de geração hidroelétrica, devido ao aumento das vazões afluentes. Assim, há uma tendência que estas usinas permaneçam a maior parte do tempo cheias, preservando a maior cota possível, sendo estas prioritárias no re-enchimento de seus reservatórios. Enquanto que as usinas de montante possuem os maiores reservatórios de acumulação, com a finalidade de regularizarem o máximo possível as vazões afluentes às usinas de jusante. Desta maneira estas usinas passam a ser prioritárias no esvaziamento nos períodos de seca (Carneiro 1991).

Recentemente, (Martinez; Soares Filho 2002) apresentaram a comparação entre políticas de controle de malha fechada e controle de malha parcialmente aberta, aplicadas ao planejamento hidrotérmico de médio prazo, usando um sistema brasileiro como estudo de caso.

A PDE tem sido amplamente utilizada na otimização de médio prazo em problemas de planejamento hidrotérmico devido a sua grande habilidade de lidar com a não-linearidade e estocasticidade destes problemas. Além disso, a técnica de Programação Dinâmica (PD) é uma política de Controle de Malha Fechada (CLFC) em que na saída não se tem valores numéricos das variáveis de decisão, mas sim uma regra de decisão ótima, em cada instante de tempo, em função de cada estado do sistema.

O uso desta técnica de PD, entretanto, é limitada pela conhecida “maldição da dimensionalidade”, onde o esforço computacional cresce exponencialmente com o número de variáveis de estado. Várias técnicas têm sido aplicadas para contornar este problema de dimensionalidade, inclusive a agregação de sistemas hidroelétricos através da representação a reservatório equivalente e o uso da Programação Dinâmica Dual Estocástica (PDDE), baseada na decomposição de Bender.

Uma proposta alternativa para esta otimização por malha fechada, combina a otimização determinística do problema de planejamento hidrotérmico com um modelo de previsão de vazões. Onde em cada estágio do período de planejamento deste

esquema de controle é determinada uma decisão ótima baseada na presente previsão das afluições futuras, levando em conta o estado atual do sistema. Quando uma nova previsão é feita, uma nova decisão é tomada para todo o sistema, caracterizando assim uma política de Controle de malha Parcialmente Aberto (OLFC).

A comparação foi feita através da simulação usando dados históricos de afluição, bem como séries sintéticas de afluições geradas por um modelo auto-regressivo de primeira ordem. O mesmo modelo foi usado na simulação de ambas as políticas de controle, em malha fechada para gerar funções de densidade de probabilidade condicional de afluições, e em malha parcialmente aberta para prover a previsão de afluições. A simulação foi feita considerando três usinas hidroelétricas localizadas em diferentes rios brasileiros.

Os resultados apresentados por (Martinez; Soares Filho 2002) revelaram que a política de controle de malha parcialmente aberta promove um comportamento de geração média um pouco maior, em todas as simulações. Além disso, um maior desvio padrão de geração térmica complementar, o que acarreta num maior custo total de geração.

Em períodos de estiagem, importantes para o planejamento da operação, a política de controle de malha parcialmente aberta (que combina previsão de vazões afluentes com otimização determinística) se revelou mais eficiente do que a política de controle de malha fechada (que utiliza a PDE), onde a quantidade de água disponível é bem limitada e o modelo estocástico não é capaz de prevê-lo eficientemente, ao contrário da técnica de malha parcialmente aberta que é mais eficiente para prever este tipo de cenário.

Outro benefício associado ao controle de malha parcialmente aberto é a representação individualizada no caso de sistemas de médio e grande porte, como é o caso brasileiro. E uma melhoria sensível pode ser alcançada com a utilização de modelos de previsão de afluições mais eficientes que os de primeira ordem.

## **2.1. Objetivos do Trabalho**

Este trabalho de pesquisa objetiva a obtenção das Regras de Operação de Reservatórios (ROR) bem próximas do comportamento ótimo do sistema. Para tanto, um algoritmo de Programação Não-Linear (PNL), especialmente desenvolvido para a realização da operação determinística de Sistemas Hidrotérmicos de Potência, é utilizado para otimizar a operação de um sistema hidroelétrico real com grandes UH's pertencentes ao Sistema Sudeste Brasileiro e, tendo como função objetivo a minimização do custo com complementação térmica.

Considerando as seqüências de aflúncias do histórico de vazões pode-se adotar este tipo de abordagem determinística, e com isso definir quanto cada usina deveria gerar a cada instante para que a operação fosse ótima. Apesar deste tipo de análise ser fictício, ela é muito importante na realização de uma análise do comportamento dos reservatórios, observando-se como eles relacionam-se entre si.

As Redes Neurais Artificiais (RNA) são muito boas no mapeamento do comportamento de sistemas dinâmicos complexos. Este mapeamento entrada-saída (input/output) pode ser feito do tipo caixa preta (black box), já que a descrição analítica do sistema é desconhecida. No contexto deste trabalho de pesquisa, os dados de entrada foram as vazões afluentes históricas e os estados de armazenamento dos reservatórios do sistema hidroelétrico, gerando como variáveis de saída, as previsões das futuras vazões afluentes e a previsão dos volumes operativos ótimos, caracterizando assim o comportamento otimizado do sistema de geração.

O propósito principal deste trabalho foi de aplicar a rede de Mapa Auto-Organizável (SOM, em inglês) de Kohonen com Modelos de Sistemas Dinâmicos, com o objetivo de capturar o comportamento ótimo relacionado a cada uma das UHs e simular a operação ótima deste conjunto de usinas. Sendo este modelo uma versão não-supervisionada do modelo NARX (do inglês, Nonlinear Auto-Regressive model with eXogenous input), é então chamado de modelo ou rede NARX Auto-Organizável (SONARX, em inglês), capaz de processar padrões espaço-temporais.

Além da captura do comportamento ótimo operativo, esta rede SONARX pôde ser aplicada na previsão de futuras vazões afluentes às UHs, já que este tipo de informação complementar é de fundamental importância para a tomada de decisão por parte do operador do sistema de geração.

Após seu treinamento, a Rede SONARX pode ser integrada a um simulador da operação de sistema hidroelétrico, onde a cada instante de tempo é necessário que ela decida quanto cada usina deve gerar para que uma demanda solicitada seja atendida. As simulações não são determinísticas e, a cada intervalo de simulação, as decisões devem ser tomadas sem nenhum conhecimento sobre o futuro. Como a rede foi treinada com dados provenientes de operações ótimas, é esperado que suas decisões produzam comportamentos similares.

### **3. Desenvolvimento da Pesquisa**

No capítulo anterior foi mostrada a importância das regras de operação no planejamento da operação. A seguir será discutida uma metodologia para a obtenção de regras que reflitam o comportamento ótimo dos reservatórios das usinas em cascata num sistema hidroelétrico.

As regras de operação, como já foi dito, procuram relacionar o comportamento dos reservatórios de cada usina com o estado do sistema. Assim o nosso objetivo é estabelecer uma relação entre a série de afluições às usinas do sistema, bem como a seqüência das decisões operativas já tomadas, com as futuras decisões operativas a serem tomadas para um dado conjunto de usinas. Desta forma, a aplicação das regras ótima reproduziria as trajetórias de volume obtidas por um modelo de otimização, na forma que será descrito nos próximos itens.

A proposta deste trabalho é de aplicar a rede de Mapa Auto-Organizável (SOM) de Kohonen com Modelos de Sistemas Dinâmicos, no aprendizado destas relações, objetivando seu uso para simular a operação ótima de um sistema hidroelétrico real com grandes UH's localizadas no Sistema Sudeste Brasileiro. E, em seguida, realizar a avaliação da eficiência destas novas regras, através de simulações da operação, comparando os resultados com os modelos utilizando as regras do paralelismo, dentre outras.

#### **3.1. Metodologia Utilizada**

Neste trabalho um algoritmo de PNL, especialmente desenvolvido para a realização da operação determinística de Sistemas Hidrotérmicos de Potência, é utilizado para otimizar a operação de um conjunto de usinas, considerando-se várias condições hidrológicas distintas e tendo como função objetivo a minimização do custo com complementação térmica.

Os estudos a serem realizados se localizam, como já exposto, num horizonte de médio e curto prazo, abrangendo o que pode ser chamado de planejamento energético.

Para a realização destes estudos toma-se um sistema composto por várias usinas e faz-se uma simulação da operação otimizada, para um determinado período do histórico, onde a ferramenta principal é o algoritmo determinístico de fluxo em redes, implementado pelo grupo de pesquisa do Laboratório de Sistemas de Energia Elétrica (LSEE) da EESC-USP, na linguagem FORTRAN de programação. Os detalhes sobre os algoritmos de otimização e de simulação, o sistema real utilizado e os testes realizados serão descritos nas próximas seções.

Em seguida, os resultados destas simulações são processados e utilizados no treinamento da rede SONARX, capaz de processar padrões espaço-temporais, como dito anteriormente. Este modelo de RNA foi por mim implementado também através da linguagem FORTRAN de programação. A opção de se implementar tal modelo, e não utilizar um “pacote pronto” (*software*), decorreu da necessidade de se estudar mais a fundo este campo de pesquisa tão promissor da Inteligência Artificial.

Para o presente estudo foram utilizados e desenvolvidos vários algoritmos. Nesta subseção encontram-se o algoritmo de otimização utilizado durante os testes de simulação da operação otimizada para a obtenção das regras, e o algoritmo de simulação heurística da operação desenvolvido para a realização de testes de avaliação das regras obtidas.

### **3.1.1. Algoritmo de Otimização**

No caso deste trabalho, o modelo de otimização da operação está preocupado com as decisões de médio e curto prazo de um sistema hidrotérmico de potência, aquelas relacionadas às variações de volume dos reservatórios ao longo dos anos e à quantidade de geração do sistema de usinas. O objetivo é encontrar políticas de operação que minimizem o custo de operação do sistema e que garantam níveis mínimos de confiabilidade de suprimento (Pereira 1985; Soares Filho 1987; Soares Filho; Carneiro 1991).

O modelo de otimização utilizado neste trabalho, previamente desenvolvido pelo nosso grupo de trabalho, presta-se à otimização determinística de sistemas hidrotérmicos de potência com representação individualizada para as usinas

hidroelétricas, para a maximização de benefícios em problemas que podem ser formulados como de fluxo em redes. Nos problemas de operação hidrotérmica, o algoritmo aproveita a estrutura particular da rede hidráulica, através de procedimentos que norteiam o processo de partição de variáveis em conjunto das variáveis básicas e não básicas. Estes procedimentos compõem a partição dirigida de variáveis e baseiam-se na análise da função objetivo, na análise da função de produção hidráulica, na importância de cada usina no processo otimizador e na experiência do operador. A sua principal característica é a identificação da base por um único índice.

Uma das várias utilidades deste algoritmo está no fato de que aplicado sobre o passado, isto é, com as afluências passadas conhecidas, permite identificar regras de operação para os principais reservatórios do sistema e testar os sistemas em condições hidrológicas do passado como o período crítico. Uma outra, é fazer simulações com séries sintéticas para prever futuros problemas.

O sistema de geração hidrotérmico pôde ser dividido em dois subsistemas: um subsistema hidroelétrico, e um subsistema térmico, como ilustrado na Figura 3.1, onde toda a geração é injetada à um sistema de transmissão, ilimitado e sem perdas, que irá atender toda a Demanda média  $D(t)$ , durante o intervalo  $t$ , em  $\overline{MW}$ .

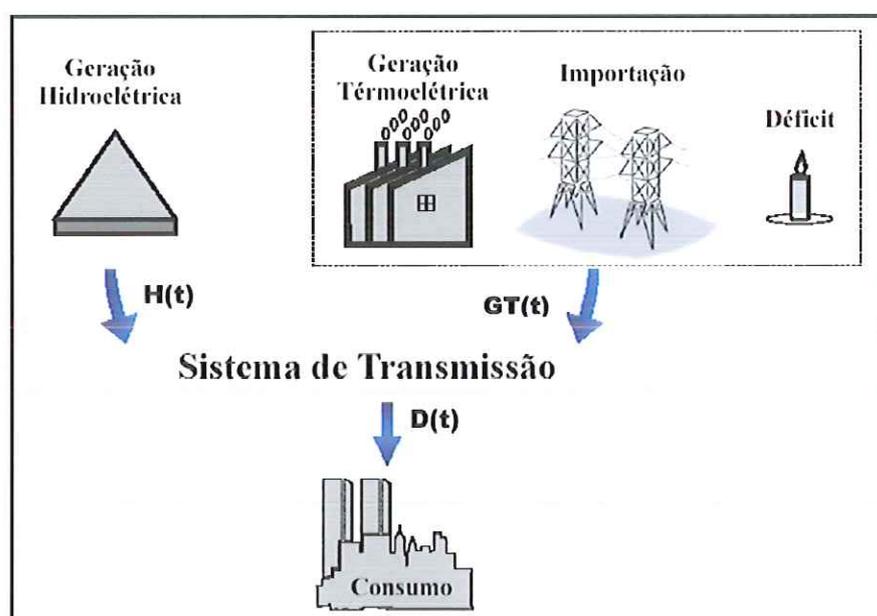


Figura 3.1 – Sistema hidrotérmico de potência

O Subsistema Hidroelétrico é constituído basicamente pelas usinas hidroelétricas, tendo seu custo operativo simplificado e considerado nulo, uma vez que o “combustível” empregado na obtenção de energia é a água. Mesmo em situações onde são cobradas taxas pelo uso da água, o custo de operação das usinas hidroelétricas é

relativamente baixo. Ele é responsável por toda a geração das hidroelétricas  $H(t)$ , durante o intervalo  $t$ , em  $\overline{MW}$ .

Já o Subsistema Termoelétrico foi representado através de uma função de custo que engloba o custo do combustível utilizado na operação das usinas termoelétricas, o custo de importação de energia de outros sistemas, e pelo custo da falta de suprimento energético, chamado de custo do déficit. Este subsistema é responsável por toda a geração não hidráulica  $GT(t)$ , chamada de complementação térmica, durante o intervalo  $t$ , em  $\overline{MW}$ .

Determinar o comportamento ótimo de um sistema de geração predominantemente hidroelétrico, como o brasileiro, implica em substituir, na medida do possível, a geração de origem não-hidráulica por geração hidráulica. Isto pode ser traduzido por encontrar as trajetórias ótimas de volume armazenado nos reservatórios, o que significa conhecer as vazões turbinadas e vertidas em cada usina, a cada intervalo de tempo (Soares Filho 1987).

Assim, para cada intervalo mensal  $t$ , o sistema hidrotérmico tem como custo de operação justamente o custo do sistema não-hidráulico complementar,  $C(t)$ , o qual depende da demanda  $D(t)$  e da geração hidráulica  $H(t)$ . Quanto maior a geração hidráulica, menor a complementação térmica  $GT(t)$  correspondente.

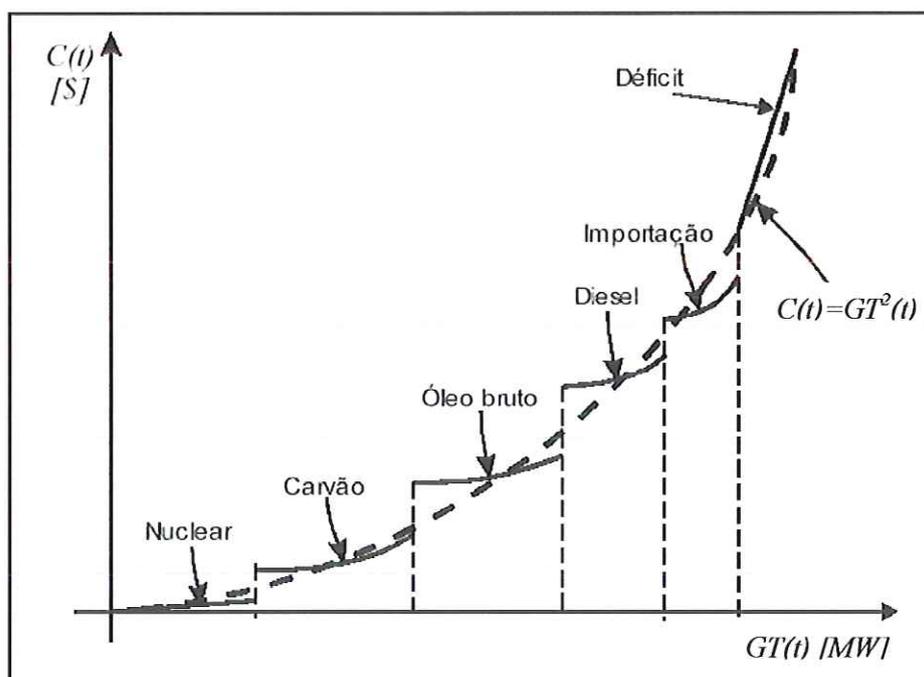


Figura 3.2 – Função de custo com complementação térmica

A função  $C(t)$ , ilustrada na Figura 3.2, é o resultado do Despacho Econômico Térmico (DET) (Lyra; Tavares et al. 1984). Observa-se que primeiramente são

utilizadas as fontes com menor custo de geração e conforme a participação destas fontes aumenta, usinas mais caras vão sendo despachadas até que sejam necessárias importações, e em último caso, cortes de carga cujos custos são representados pelo custo do déficit. Desta maneira, o problema do planejamento ótimo da operação determinística de um sistema hidrotérmico de potência pode ser formulado da seguinte forma:

$$\begin{aligned}
 & \min \sum_{t=1}^I \frac{1}{(1+j)^t} \cdot C(t) \\
 & \left\{ \begin{aligned}
 & GT(t) = D(t) - H(t) \\
 & H(t) = \sum_{i=1}^{NH} gh_i(t) \\
 & gh_i(t) = k_i \cdot h_{liq,i}(vol_i, def_i) \cdot turb_i(t) \\
 & s.a. \left\{ \begin{aligned}
 & vol_i(t+1) = vol_i(t) - turb_i(t) + af_{inc,i}(t) + \sum_{k \in \Omega_i} turb_k(t) \\
 & turb_{min,i} \leq turb_i(t) \leq turb_{máx,i} \\
 & vol_{min,i} \leq vol_i(t) \leq vol_{máx,i} \\
 & \text{dados : } vol_i(1), vol_i(I)
 \end{aligned} \right.
 \end{aligned} \right. \quad (3.1)
 \end{aligned}$$

onde:

$I$  : número de intervalos do horizonte de planejamento;

$C(t)$  : custo de operação do parque complementar térmico, dado em alguma unidade monetária;

$j$  : taxa de juros relativa a cada intervalo do horizonte de planejamento;

$GT(t)$  : geração média das fontes não-hidráulicas durante o intervalo  $t$ , em  $\overline{MW}$  ;

$D(t)$  : demanda média durante o intervalo  $t$ , em  $\overline{MW}$  ;

$H(t)$  : geração hidráulica total média durante o intervalo  $t$ , em  $\overline{MW}$  ;

$gh_i(t)$  : geração média da usina hidroelétrica  $i$ , durante o intervalo  $t$ , em  $\overline{MW}$  ;

$k_i$  : constante que engloba aceleração da gravidade, densidade da água, rendimento turbina-gerador e fatores de conversão, da usina hidroelétrica  $i$ ;

$vol_i(t)$  : volume do reservatório da usina hidroelétrica  $i$ , ao final do intervalo  $t$ , em  $hm^3$ ;

$def_i(t)$  : defluência média da usina hidroelétrica  $i$ , durante o intervalo  $t$ , em  $m^3/s$ ;

$h_{liq,i}(t)$  : altura de queda líquida da usina hidroelétrica  $i$ , durante o intervalo  $t$ , em  $m$ ;

$turb_k(t)$  : turbinagem média da usina hidroelétrica  $i$ , durante o intervalo  $t$ , em  $m^3/s$ ;

$af_{inc,i}(t)$  : afluência incremental média à usina hidroelétrica  $i$ , durante o intervalo  $t$ , em  $m^3/s$ ;

$\Omega_k$  : conjunto das usinas à montante da usina hidroelétrica  $i$ ;

$vol_{min,i}$  e  $vol_{máx,i}$  : limites mínimo e máximo para o volume armazenado na usina hidroelétrica  $i$ , em  $hm^3$ ;

$turb_{min,i}$  e  $turb_{máx,i}$  : limites mínimo e máximo para a turbinagem da usina hidroelétrica  $i$ , em  $m^3/s$ ;

A equação do custo  $C(t)$  é uma função exponencial da geração não-hidráulica  $GT(t)$ , podendo ser representada por  $C(t) = k \cdot (GT(t))^2$ , onde  $k$  é um fator de proporcionalidade.

Logo, este problema apresentado pela Equação 3.1, pode ser eficientemente resolvido como um modelo de fluxo em redes, utilizando, como no caso em questão, um algoritmo Simplex Convexo adaptado (Carvalho; Soares Filho 1987).

### 3.1.2. Algoritmo de Simulação

Os algoritmos de simulação tentam reproduzir a operação do conjunto de usinas do sistema, estabelecendo o comportamento de cada uma delas, sob condições preestabelecidas de vazão, geração e regras de operação.

Como dito, obtidas tais regras de operação, estas devem ser testadas para a avaliação da sua eficiência, comparando-se os resultados com aqueles obtidos através de outras regras. Para tanto, utiliza-se um algoritmo de simulação da operação, também desenvolvido por nosso grupo de pesquisa do LSEE, que opere com vários tipos de regras de operação, inclusive a proposta neste trabalho.

O modelo de simulação utilizado presta-se à avaliação energética de sistemas hidroelétricos de geração, com representação individualizada das usinas. Para a realização dos estudos deste projeto adotou-se uma base de tempo mensal, com sistemas estáticos e a barramento único (Kadowaki 1995).

Para estas simulações são dados de entrada, além da configuração do sistema, as afluências do período da simulação e as gerações hidráulicas totais do sistema, constituindo a meta energética de geração de todo o sistema hidroelétrico. Isto é feito através do ajuste das turbinagens, e conseqüentemente dos volumes, já que essa vazão

turbinada implica em variar o volume dos reservatórios das usinas segundo uma política de operação pré-definida.

Entende-se por política de operação um conjunto de regras de operação, uma para cada usina a reservatório do sistema hidroelétrico, que define como as usinas a reservatório vão trabalhar de forma acoplada (Silva Filho; Carneiro 1999a).

Uma das mais simples regras de operação que podem ser adotadas para um sistema hidroelétrico é a operação em paralelo dos reservatórios, a qual diz que todos os reservatórios do sistema devem manter-se com a mesma percentagem de seus volumes úteis. Por exemplo, seja  $\lambda$  um número positivo entre zero e um, responsável pelo acoplamento entre as usinas; onde para um determinado período da operação tem-se  $\lambda = 0,3$ , isto implica que todas as usinas hidroelétricas do sistema devam apresentar 30% de armazenamento dos volumes úteis em seus reservatórios de acumulação.

É importante observar que a operação em paralelo está longe do determinado pela operação ótima, a qual prevê comportamento diferenciado para cada tipo de reservatório, enquanto que a operação em paralelo determina que todos os reservatórios sejam operados mantendo um mesmo volume útil percentual.

Entretanto, alguns trabalhos propõem metodologias de obtenção de regras de operação que incorporam conceitos da operação otimizada. Essas regras também especificam o estado de cada reservatório em função do estado do sistema, porém, ao contrário do que ocorre na operação em paralelo, para um mesmo  $\lambda$  as usinas possuem diferentes volumes.

O processo iterativo pode ser entendido como o conjunto de operações realizadas em cada intervalo de tempo da simulação, de forma a fazer com que a demanda seja atendida pelo sistema hidroelétrico, utilizando-se a quantidade de água disponível e respeitando-se as restrições às usinas.

O fluxograma apresentado pela Figura 3.3 ilustra as etapas do processo iterativo de simulação. Na primeira etapa deste processo iterativo, os valores de vazões naturais afluentes e vazões incrementais são atualizados, assim como os dados de volume do mês atual e a meta de geração no próximo mês.

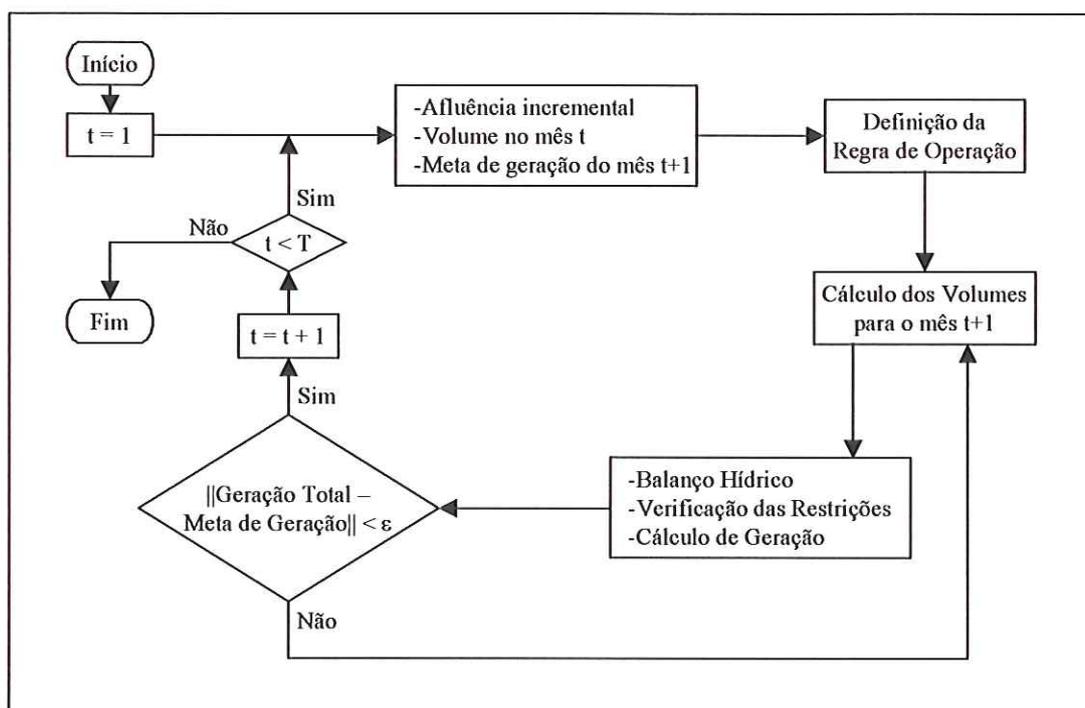


Figura 3.3 – Fluxograma do Processo Iterativo do Simulador

Na segunda etapa, são definidas as Regras de Operação que vão “guiar” as decisões operativas do simulador, segundo uma determinada Política de Operação. Estas ROR podem ser paramétricas, como discutido anteriormente, ou não paramétricas, como no caso deste trabalho de pesquisa, onde a rede SONARX exerceu o papel de “decisor operativo”, prevendo os volumes operativos ótimos de cada uma das Usinas Hidroelétricas, o que ocorre na terceira etapa do processo iterativo.

Conhecendo-se os volumes ao final do intervalo  $t+1$  e as vazões afluentes a cada reservatório durante este intervalo, determinam-se na quarta etapa os valores de vazão defluente de cada usina através da equação de balanço hídrico. Ainda nesta etapa, são verificadas as restrições operativas como limite de geração, limite de turbinagem e limite de volume operativo. Calculando-se também a geração de cada usina através de sua função de geração, e na seqüência a geração total do sistema hidroelétrico  $H(t+1)$ .

Finalmente, tem-se a comparação da geração total com a meta de geração. Se a geração total for igual à meta de geração, incrementa-se  $t$ ; caso contrário, recalculam-se os volumes operativos até que esta igualdade seja alcançada.

Após a realização da simulação em todo o horizonte de estudo, o simulador gera um conjunto de resultados que retrata como foi realizada a operação em cada usina hidroelétrica e no sistema como um todo, permitindo que se fizesse algumas análises operativas.

## 4. Rede SONARX

Neste capítulo é apresentada a Rede Neural Artificial (RNA) que foi utilizada neste trabalho de pesquisa. A arquitetura é a da rede de Mapa Auto-Organizável (SOM, em inglês) de Kohonen com Modelos de Sistemas Dinâmicos, que será apresentada em detalhes na próxima seção, correspondente a parte teórica. Juntamente com toda a formulação também serão descritas suas principais características, e inclusive seus campos de aplicação.

### 4.1. Modelagem Matemática

Este item apresenta um modelo neural não-supervisionado capaz de processar padrões espaço-temporais, baseado na rede de Mapa Auto-Organizável (rede SOM). Este modelo, passa a ser uma versão não-supervisionada do modelo NARX (do inglês Nonlinear Auto-Regressive model with eXogenous input), sendo assim chamado de modelo ou rede NARX Auto-Organizável (SONARX).

Vale ressaltar que a rede SONARX é uma instância particular de uma técnica mais geral que permite utilizar RNA's Temporais competitivas na modelagem entrada-saída de sistemas dinâmicos. Esta técnica recebe o nome de Memória Associativa Temporal via Quantização Vetorial (MATQV) (Barreto; Araújo 2001a).

Dentre as áreas de aplicação da rede SONARX, destacam-se as seguintes: identificação e controle de sistemas dinâmicos, tendo como exemplo um atuador hidráulico; a aprendizagem de trajetórias robóticas complexas, como controle preditivo; a modelagem e previsão de série temporais, que será objeto de estudo deste trabalho.

Na tentativa de representar de forma reduzida a seqüência de entrada, por meio da técnica de quantização vetorial, utiliza-se então como base do modelo aqui apresentado, a rede SOM. Nesta rede os vetores-protótipos (centróides) representam de

forma compacta os elementos da seqüência, através dos vetores de pesos associados aos neurônios da rede SOM.

Uma representação condensada de uma seqüência ou série temporal somente é possível porque a rede SONARX utiliza um processo de aprendizagem estático que precisa de várias apresentações do conjunto de treinamento, ou seja vários ciclos de treinamento, para ajustar os pesos de forma adequada.

A maior limitação da aprendizagem de vários ciclos está na ocorrência de destruição da memória adquirida à medida que novas informações vão sendo apresentadas à rede. Este fenômeno é chamada de Interferência Catastrófica (Grossberg 1986) apud (Barreto 2003).

Apesar desta limitação, foi mostrado por (Barreto 2003), que a rede SONARX tem diversas aplicações, não exigindo alterações substanciais da arquitetura da rede SOM para ser aplicada em vários domínios.

#### **4.1.1. A Rede SOM**

A habilidade humana de processar informação foi inicialmente um mistério, sendo Ross Ashby citado como o primeiro a descrever o sistema nervoso como auto-organizado. Em seu livro, (Ashby 1952) apud (Barreto 2003) estava preocupado com a noção básica de que o comportamento adaptativo não é inato, mas sim aprendido, e que através da aprendizagem o comportamento de um animal ou sistema muda, em geral para melhor.

Haja visto que em (Ashby 1947) apud (Barreto 2003) já se pensava em sistemas auto-organizáveis como um caso de sistemas orientados à meta que são capazes, sem ajuda externa explícita, melhorando sua performance ao longo do tempo, enquanto perseguem seus objetivos e que devem ser avaliados por alguma função de desempenho.

Em outras palavras, o que é usualmente referido como um comportamento auto-organizável é a formação espontânea de estruturas organizadas, como padrões ou comportamentos, a partir de condições iniciais aleatórias.

A questão básica consiste em entender como a ação de vários elementos simples (neurônios), em uma configuração inicial desordenada ou aleatória, dá margem ao surgimento espontâneo (sem regras pré-programadas) de padrões coerentes de comportamento.

A resposta reside essencialmente na seguinte observação feita por (Turing 1952) apud [Barreto, 2003 #78]: “Ordem global pode surgir espontaneamente a partir de interações locais”, observação esta que se aplica tanto ao estudo de cérebro, quanto ao de Redes Neurais Artificiais.

A organização da rede se dá em dois planos diferentes, que interagem entre si, na forma de um laço de realimentação. O primeiro deles é a chamada Ativação, caracterizada pela atividade neural, de cada unidade de processamento, produzida por uma dada rede em resposta a estímulos de entrada. E o segundo é a Conectividade, que seria o conjunto de conexões ou pesos sinápticos da rede, em modificação sob respostas aos padrões de ativação dos neurônios da rede.

Atualmente há muitas teorias que apóiam a hipótese de auto-organização do sistema nervoso, sendo possível enumerar quatro princípios que parecem estar presentes em cada uma delas:

- 1- Modificação nos pesos sinápticos tendem a se auto-amplificar;
- 2- Limitação de recursos leva à competição entre sinapses e, em consequência, à seleção das sinapses mais vigorosas, em detrimento de outras menos adequadas;
- 3- A presença de uma sinapse de valor elevado pode melhorar a adequabilidade (*fitness*) de outras sinapses;
- 4- Redundância provê conhecimento (informação).

Os quatro princípios sobre aprendizagem não-supervisionada citados anteriormente formam a base neurobiológica subjacente a um importante algoritmo neural: a rede SOM.

Uma representação econômica dos dados com todas as suas inter-relações é um dos problemas centrais da informação, e tal habilidade é também característica da operação do cérebro. Ao processar informação, há uma tendência em comprimi-la, formando representações reduzidas dos fatos mais relevantes, sem perda de conhecimento a cerca de suas inter-relações. Sendo justamente esta a característica de interesse na aplicação da rede Auto-Organizável neste trabalho.

O processamento inteligente da informação, em geral, parece ser a criação de imagens simplificadas do mundo observável em vários níveis de abstração, a partir de uma seqüência de dados de entrada.

Apenas recentemente tem sido possível obter evidências concretas sobre o cérebro formando representações neurais que podem ser consideradas imagens de várias representações sensoriais, projetadas sobre várias áreas corticais. Dá-se o nome genérico

de mapas corticais ou mapas topográficos a essas áreas de projeção. Tais mapas têm importância fundamental para os processos de aprendizagem e memória, pois definem uma representação compacta dos dados sensoriais.

Um conjunto de modelos computacionais, chamados de mapas topológicos auto-organizáveis, vem sendo propostos com o intuito de modelar diversas propriedades de mapas corticais. Estes modelos pertencem à classe das redes neurais artificiais não-supervisionadas. Nesta seção, será apresentada a teoria relativa a um desses mapas topológicos, proposto por (Kohonen 1989), conhecido como Mapa Auto-Organizável (SOM), cuja teoria apresentada envolve o estudo do algoritmo SOM e determinadas propriedades como convergência e estabilidade. A apresentação da Rede SOM nesta seção objetiva fornecer subsídios necessários à correta utilização do algoritmo.

Sem ter a ambição de modelar fielmente um neurônio real, a rede SOM foi proposta com a finalidade de modelar algumas características essenciais dos mapas corticais cerebrais. Tendo como objetivo principal transformar um sinal de entrada de dimensão qualquer em um arranjo discreto de unidades de processamento. Esta transformação deve ser realizada de forma adaptativa e mantendo relações de similaridade entre os dois espaços.

Topologicamente a rede SOM pode ser tratada da seguinte maneira. Seja  $\mathbf{X}$  um espaço contínuo de dados de entrada, sendo dele selecionado aleatoriamente um conjunto de vetores  $\mathbf{x} \in \mathbf{X}$ , sujeito a uma organização aleatória de densidade de probabilidade  $p(\mathbf{x})$ . Já a topologia do espaço discreto de saída  $\mathbf{A}$  é definida pelo arranjo geométrico do conjunto de neurônios  $i \in \mathbf{A}$ .

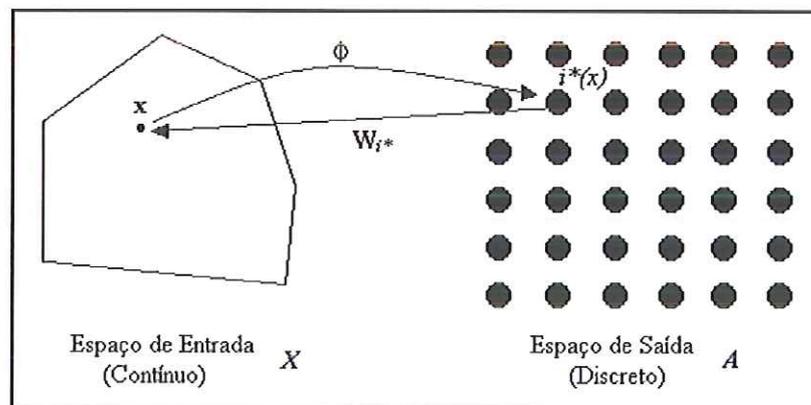


Figura 4.1 – Topologia dos espaços de entrada e saída, e o mapeamento de características  $\phi$

Desta maneira pode-se definir uma transformação não-linear  $\phi: \mathbf{X} \rightarrow \mathbf{A}$ , que leva do espaço de entrada  $\mathbf{X}$  ao espaço de saída  $\mathbf{A}$ . Tal transformação pode ser

visualizada na Figura 4.1, sendo chamada de mapeamento de características (feature map, em inglês).

Estabelecendo assim uma relação neural para tal mapeamento, o espaço de entrada  $\mathbf{X}$  pode representar o conjunto de coordenadas de receptores somato-sensoriais densamente distribuídos sobre toda a superfície do corpo. Já o espaço de saída  $\mathbf{A}$  representa o conjunto de neurônios localizados em uma certa porção do córtex cerebral que responde aos estímulos gerados por tais receptores.

Seja o espaço contínuo de entrada  $\mathbf{X}$  de dimensão  $n$ , um vetor de entrada (padrão de entrada)  $\mathbf{x} \in \mathbf{X} \subset \mathfrak{R}^n$ , pode ser representado por:

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \quad (4.1)$$

Da mesma forma, o vetor de pesos associado a cada neurônio da rede tem a mesma dimensão do vetor de entrada. A representação do vetor de pesos de um neurônio  $i$  é a seguinte:

$$\mathbf{w}_i = \begin{pmatrix} w_{i1} \\ w_{i2} \\ \vdots \\ w_{in} \end{pmatrix}, \quad i = 1, 2, \dots, M \quad (4.2)$$

onde  $M$  é o número total de neurônios da rede.

Cada neurônio  $i$  no arranjo está totalmente conectado às  $n$  unidades de entrada por meio de termos ponderadores,  $w_{ij}$ ,  $j = 1, 2, \dots, n$ , chamados pesos sinápticos, agrupados no vetor  $\mathbf{w}_i$ , como pode ser visto na Figura 4.2(a).

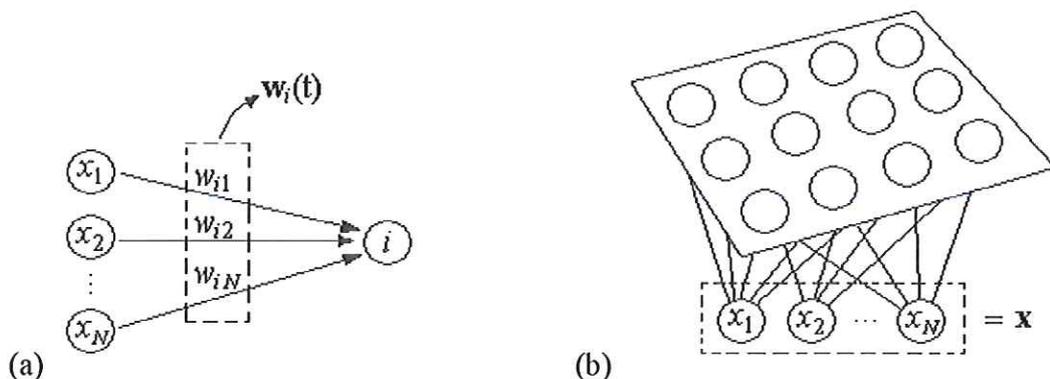


Figura 4.2 – (a) Vetor de pesos de um dado neurônio  $i$ . (b) Rede SOM bidimensional

Esta rede possui uma estrutura de propagação para frente (feedforward, em inglês) com uma única camada de processamento, constituída de  $M$  neurônios normalmente dispostos de forma retangular, no caso da topologia bidimensional. Assim a Figura 4.2(b) mostra um diagrama esquemático de um arranjo de neurônios comumente usado como espaço discreto de saída.

Outras configurações devem ser destacadas, tais como de um arranjo unidimensional, onde a camada de processamento é formada por uma única linha de neurônios; e a configuração hexagonal, a qual é bastante adequada para a visualização do mapeamento topográfico  $\phi$ .

Definida a organização da topologia da rede SOM e seus dados de entrada e saída, o algoritmo responsável pela formação do mapa auto-organizável tem como passo inicial a atribuição de valores iniciais (normalmente aleatórios) aos pesos da rede.

Em seguida se dá a Competição, a fim de se determinar dentre todos os neurônios aquele que melhor representa o padrão de entrada, sendo este definido como o neurônio vencedor. Esta etapa é muito importante tanto na fase de treinamento quanto de operação da rede SOM. A Cooperação representa a interação entre o neurônio vencedor e seus vizinhos, mais intensa a medida que se trata de um neurônio próximo do vencedor. Já o Ajuste dos pesos sinápticos da rede se dá somente durante a fase de treinamento da rede, onde o neurônio vencedor, dado um padrão de entrada, recebe um ajuste mais intenso, enquanto que seus vizinhos são ajustados em menor intensidade.

Para encontrar o neurônio vencedor, o vetor de pesos de todos os neurônios da rede SOM é comparado, um a um, com o vetor de entrada. O neurônio cujo vetor de pesos está mais próximo do vetor de entrada é considerado o vencedor  $i^*$ . A distância euclidiana, definida pela Equação 4.3, pode ser implementada para tal finalidade, como na Equação 4.4.

$$\|x(t) - w_i(t)\| = \sqrt{\sum_{j=1}^n [x_j(t) - w_{ij}(t)]^2} \quad (4.3)$$

$$i^*(x(t)) = \arg \min_{i \in A} \|x(t) - w_i(t)\| \quad (4.4)$$

onde  $i^*(x)$  é o índice que representa o neurônio vencedor para o padrão de entrada  $x(t)$ .

O neurônio vencedor  $i^*(x)$  determina o centro espacial de um grupo dentre seus vizinhos. Sendo que o neurônio vencedor e sua vizinhança interagem lateralmente de forma cooperativa. A intensidade desta interação lateral entre o vencedor  $i^*$  e um neurônio  $i$  qualquer é, em geral, descrita matematicamente na forma de uma função vizinhança  $h(i^*, i; t)$ . Esta função define o que se chama de vizinhança topológica centrada no neurônio vencedor  $i^*(x)$ .

A função vizinhança deve satisfazer os seguintes requisitos:

- alcançar seu valor máximo para uma distância lateral nula;
- ser simétrica em relação ao neurônio vencedor;
- sua amplitude decai monotonicamente com o aumento da distância lateral.

Uma escolha comum para  $h(i^*, i; t)$  que satisfaz os requisitos acima é a função gaussiana, definida por:

$$h(i^*, i; t) = \exp\left(-\frac{\|x_i - w_{i^*}\|^2}{2\sigma^2(t)}\right) \quad (4.5)$$

onde o parâmetro  $\sigma(t)$  define a “largura efetiva” da vizinhança topológica.

A partir deste tipo de função vizinhança, pode-se derivar uma rede competitiva simples, conhecida como rede WTA (do inglês, Winner-Take-All), em que apenas o neurônio vencedor tem seu vetor de pesos atualizado a cada iteração do algoritmo.

O ajuste dos pesos sinápticos é o último passo na formação de um mapa auto-organizável. Como a rede SOM é não-supervisionada, o vetor de pesos de um dado neurônio  $i$  deve ser modificado como função do estímulo de entrada apenas. Durante a adaptação sináptica o objetivo é extrair algum tipo de regularidade presente no espaço de entrada  $X$ . A partir das suposições feitas por [Hebb, 1949 #107] apud [Barreto, 2003 #78] na tentativa de relacionar estímulos de entrada à uma estrutura neural com suas sinapses neurais, Kohonen propôs uma representação matemática para o ajuste dos pesos para a rede SOM :

$$w_i(t+1) = w_i(t) + \alpha(t) \cdot h(i^*, i; t) \cdot [x(t) - w_i(t)] \quad (4.6)$$

onde  $\alpha(t)$  define a taxa de aprendizagem.

A taxa de aprendizagem controla a velocidade com que os pesos sinápticos são modificados, ou seja, adquire novo conhecimento, devendo cair monotonicamente com o treinamento, para garantir tanto a convergência como a estabilidade da rede SOM.

Desta maneira, o mapeamento  $\phi$  resultante da aplicação deste algoritmo SOM, representado pelo conjunto de vetores de pesos sinápticos  $\{w_i\}$  que compõem  $A$ , provê uma aproximação discreta do espaço de entrada  $X$ . Cada neurônio da rede é responsável por uma determinada região do espaço de entrada que define a sua região de influência ou seu campo receptivo. A região de influência de um dado neurônio pode indicar que todos os padrões de treinamento que ativam tal neurônio compõem um grupo ou categoria. Assim, padrões pertencentes a um mesmo grupo guardam fortes semelhanças entre si, enquanto padrões pertencentes a grupos diferentes guardam semelhanças menores. Dessa forma, é correto afirmar que uma propriedade intrínseca do algoritmo SOM é a categorização dos padrões de entrada.

Após rotular adequadamente cada grupo com base nos vetores de treinamento, a rede pode ser utilizada para classificar padrões que venham a ser apresentados. O vetor de pesos do neurônio responsável por uma dada categoria é considerado o “protótipo” daquela categoria com o qual o novo padrão de entrada é comparado para determinar se ele pertence àquela categoria. Diante disso, os vetores de pesos dos neurônios da rede podem também receber outras denominações tais como vetores protótipos ou centróides.

Resumidamente, na operação desta rede um dado vetor de entrada  $x \in X$  é fornecido a rede, o algoritmo SOM primeiro identifica um determinado neurônio vencedor  $i^*(x)$  no espaço de saída  $A$  em concordância com o mapeamento característico  $\phi$ . O vetor de pesos  $w_{i^*}$  associado ao neurônio  $i^*$  pode ser visto como um “ponteiro” para o espaço de entrada  $X$ ; ou seja, os elementos do vetor  $w_i$  podem ser vistos como as coordenadas da imagem do neurônio  $i$  projetada no espaço de entrada.

#### 4.1.2. A Rede SOM com Modelos de Sistemas Dinâmicos

Conforme apresentado no item 4.1.1, a rede SOM é uma das mais populares arquiteturas de RNAs não-supervisionadas. Sua principal área de aplicação envolve tarefas de reconhecimento de padrões, tais como formação de categorias e agrupamentos (clusters, em inglês), classificação e visualização de dados (Kaski; Kangas et al. 1998).

Isto se deve, em grande parte, à simplicidade do seu algoritmo de aprendizagem, que pode ser resumido em dois passos principais:

1. Busca pelo neurônio vencedor,  $i^*(t)$ , para o vetor de entrada  $\mathbf{x}(t)$ :

$$i^*(t) = \arg \min_{i \in A} \|\mathbf{x}(t) - \mathbf{w}_i(t)\| \quad (4.7)$$

2. Atualização dos vetores de pesos,  $\mathbf{w}_i(t)$ , da rede:

$$\mathbf{w}_i(t+1) = \mathbf{w}_i(t) + \alpha(t) \cdot h(i^*, i; t) \cdot [\mathbf{x}(t) - \mathbf{w}_i(t)] \quad (4.8)$$

onde  $\alpha(t)$  é a taxa de aprendizagem e  $h(i^*, i; t)$  é uma função vizinhança do tipo gaussiana:

$$h(i^*, i; t) = \exp\left(-\frac{\|r_i(t) - r_{i^*}(t)\|^2}{\sigma^2(t)}\right) \quad (4.9)$$

onde  $\sigma(t)$  define a largura da vizinhança, enquanto  $r_i(t)$  e  $r_{i^*}(t)$  são, respectivamente, as posições dos neurônios  $i$  e  $i^*$  no arranjo geométrico  $\mathbf{A}$ , ao qual eles pertencem. As variáveis  $0 < \alpha(t), \sigma(t) < 1$  decaem exponencialmente com o passar do tempo:

$$\alpha(t) = \alpha_0 \left(\frac{\alpha_T}{\alpha_0}\right)^{(t/T)} \quad \text{e} \quad \sigma(t) = \sigma_0 \left(\frac{\sigma_T}{\sigma_0}\right)^{(t/T)} \quad (4.10)$$

onde  $\alpha_0$  e  $\sigma_0$  são os valores iniciais e  $\alpha_T$  e  $\sigma_T$  os finais de  $\alpha(t)$  e  $\sigma(t)$ , transcorridas  $T$  iterações de treinamento.

Esta rede SOM possui uma importante extensão de seu modelo, permitindo que esta rede aprenda o mapeamento entrada-saída estático. Para isso, a saída atual pode ser mapeada somente em função da entrada atual (Walter; Ritter 1996):

$$\mathbf{y}(t) = G(\mathbf{u}(t)) \quad (4.11)$$

onde  $\mathbf{u}(t) \in \mathfrak{R}^n$  e  $\mathbf{y}(t) \in \mathfrak{R}^m$  denotam, respectivamente, o vetor de entrada e o vetor de saída do sistema. O mapeamento  $G(\cdot)$  é chamado de mapeamento direto (entrada  $\rightarrow$  saída), sendo em geral, desconhecido e não-linear.

Por outro lado, em muitos problemas de sistemas dinâmicos deseja-se saber o mapeamento inverso  $G^{-1}(\cdot)$ . Porém, este pode ser de difícil aproximação, já que normalmente é definido como “mal-colocado” no sentido de *Hadamard*.

A formulação proposta no mapeamento estático consiste em modificar o vetor de entrada da rede SOM,  $\mathbf{x}(t)$ , de modo que ele passe a ter duas partes:

- $\mathbf{x}^{in}(t)$  : que corresponde à informação de entrada do mapeamento que se quer aproximar.
- $\mathbf{x}^{out}(t)$  : que corresponde à informação da saída deste mesmo mapeamento.

Como conseqüência, os vetores dos neurônios  $i$  têm também suas dimensões aumentadas. Matematicamente, tem-se que:

$$\mathbf{x}(t) = \begin{pmatrix} \mathbf{x}^{in}(t) \\ \mathbf{x}^{out}(t) \end{pmatrix} \quad (4.12)$$

Assim, pode-se utilizar a rede SOM para aprender o mapeamento mostrado na Equação 4.11, definindo os vetores  $\mathbf{x}^{in}(t)$  e  $\mathbf{x}^{out}(t)$  da seguinte maneira:

$$\begin{pmatrix} \mathbf{x}^{in}(t) \\ \mathbf{x}^{out}(t) \end{pmatrix} = \begin{pmatrix} \mathbf{u}(t) \\ \mathbf{y}(t) \end{pmatrix} \quad (4.13)$$

Entretanto, o maior interesse deste trabalho é utilizar uma formulação para mapeamentos dinâmicos, o que justifica a utilização da formulação proposta por (Walter; Ritter 1996), onde mostraram que esta formulação pôde ser estendida para aproximar também mapeamentos dinâmicos. E este tipo de mapeamento passa a ser genericamente descrito como:

$$\mathbf{y}(t+1) = F[\mathbf{y}(t), \dots, \mathbf{y}(t-n_y+1); \mathbf{u}(t), \dots, \mathbf{u}(t-n_u+1)] \quad (4.14)$$

onde  $n_u$  e  $n_y$  definem as ordens da memória de entrada e de saída, respectivamente, e o mapeamento  $F(\cdot)$  é não-linear e desconhecida. De acordo com a Equação 4.14, a saída do sistema no instante  $t+1$ , depende dos  $n_y$  valores passados da saída e dos  $n_u$  valores passados da entrada. A representação matemática de sistemas dinâmicos mostrada na Equação 4.14 é genericamente conhecida como modelo NARX (do inglês, Nonlinear AutoRegressive model with eXogenous variables) (Sjöberg; Zhang et al. 1995; Aguirre 2000; Norgaard; Ravn et al. 2000).

Para a rede SOM ser capaz de aprender mapeamentos dinâmicos, ela deve possuir algum tipo de mecanismo de memória. Ou seja, ela deve ser capaz de armazenar informação passada sobre os vetores de entrada e saída do sistema. Estes mecanismos de memória permitem que a rede SOM seja capaz de processar as séries temporais das variáveis de entrada e saída,  $\{\mathbf{u}(t), \mathbf{y}(t)\}$ ,  $t = 1, \dots, N$ , que serão usados para construir o modelo. Dentre os possíveis mecanismos de memória que vêm sendo utilizados para construir versões temporais da rede SOM, apenas linhas de atraso (*delay line*, em inglês) serão utilizados neste trabalho.

A questão que se coloca neste instante passa então a ser como utilizar a rede SOM para obter uma aproximação do mapeamento  $F(\cdot)$  de posse apenas das séries temporais das variáveis de entrada e saída.

Uma possível solução para esta questão utiliza um raciocínio semelhante ao do caso estático, ou seja, escolher adequadamente as variáveis que compõem os vetores  $\mathbf{x}^{in}(t)$  e  $\mathbf{x}^{out}(t)$ . Desta maneira, na busca pela aproximação do mapeamento representado na Equação 4.14, define-se:

$$\mathbf{x}^{in}(t) = [\mathbf{y}(t), \dots, \mathbf{y}(t - n_y + 1); \mathbf{u}(t), \dots, \mathbf{u}(t - n_u + 1)] \quad (4.15)$$

$$\mathbf{x}^{out}(t) = \mathbf{y}(t + 1) \quad (4.16)$$

Apesar de parecer uma simples extensão da formulação estática de (Walter; Ritter 1996), é importante ressaltar um aspecto que torna esta abordagem mais complexa. Como consequência direta do uso de atrasadores, acrescentou-se ao problema de aproximação dos mapeamentos dois outros graus de liberdade relativos à determinação das ordens  $n_y$  e  $n_u$  do modelo. Estas variáveis influenciam diretamente na precisão da aproximação e na complexidade (número de pesos) do modelo.

Desta maneira, denotando por  $P(t)$  o padrão de entrada da rede SOM no instante  $t$ , tem-se as seguintes definições:

$$P(t) = \begin{pmatrix} \mathbf{x}^{in}(t) \\ \mathbf{x}^{out}(t) \end{pmatrix} = \begin{pmatrix} \mathbf{y}(t), \dots, \mathbf{y}(t - n_y + 1); \mathbf{u}(t), \dots, \mathbf{u}(t - n_u + 1) \\ \mathbf{y}(t + 1) \end{pmatrix} \quad (4.17)$$

Os padrões  $P(t)$  assim gerados podem ser apresentados sequencialmente ou aleatoriamente para treinar a rede SOM.

Durante o treinamento, o neurônio vencedor,  $i^*(t)$ , é encontrado usando apenas a porção correspondente a  $\mathbf{x}^{in}(t)$ :

$$i^*(t) = \arg \min_{i \in A} \|\mathbf{x}^{in}(t) - \mathbf{w}_i^{in}(t)\| \quad (4.18)$$

Na atualização dos pesos, ambos  $\mathbf{w}_i^{in}(t)$  e  $\mathbf{w}_i^{out}(t)$  são utilizados:

$$\mathbf{w}_i^{in}(t + 1) = \mathbf{w}_i^{in}(t) + \alpha(t) \cdot h(i^*, i; t) \cdot [\mathbf{x}^{in}(t) - \mathbf{w}_i^{in}(t)] \quad (4.19)$$

$$\mathbf{w}_i^{out}(t + 1) = \mathbf{w}_i^{out}(t) + \alpha(t) \cdot h(i^*, i; t) \cdot [\mathbf{x}^{out}(t) - \mathbf{w}_i^{out}(t)] \quad (4.20)$$

Ainda durante a fase de treinamento, a rede SOM aprende a associar a saída do mapeamento  $\mathbf{x}^{out}(t)$ , com a entrada correspondente  $\mathbf{x}^{in}(t)$ , além de realizar também a quantização vetorial do espaço de entrada e do espaço de saída.

E é justamente a quantização vetorial que permite uma representação compacta dos dados. Por isso, esta técnica é chamada de Memória Associativa Temporal por Quantização Vetorial (MATQV).

O modelo neural utilizado neste trabalho, por ser baseado na rede SOM e implementar uma versão não-supervisionada do modelo NARX, é então chamado de rede SONARX.

Após o treinamento, a rede SONARX é utilizada para obter estimativas dos valores de saída dos mapeamentos, a partir do vetor de pesos  $\mathbf{w}_{i^*}^{out}(t)$  do neurônio vencedor. Definindo a saída da rede no instante  $t$  por  $\hat{y}(t+1)$ , onde o neurônio vencedor  $i^*(t)$  é determinado segundo a Equação 4.18, tem-se:

$$\hat{y}(t+1) \cong \mathbf{w}_{i^*}^{out}(t) \quad (4.21)$$

Assim, este processo de estimação se repete até que a rede convirja para o estado de melhor aproximação do mapeamento em questão, sendo importante ressaltar ao menos dois aspectos importantes do método MATQV.

Em primeiro lugar, o neurônio vencedor  $i^*(t)$  é o elemento responsável por fazer a ligação entre a os dados de entrada  $\mathbf{x}^{in}(t)$ , com os dados de saída  $\mathbf{x}^{out}(t)$ , do mapeamento. Esta associação fica codificada nos respectivos vetores de pesos,  $\mathbf{w}_{i^*}^{in}(t)$  e  $\mathbf{w}_{i^*}^{out}(t)$ .

E em segundo lugar, existe uma diferença fundamental da estratégia de treinamento associativa do método MATQV em relação àquela comumente usada no treinamento de redes supervisionadas (MLP ou RBF), que é baseada na redução do erro de aproximação. Nestas redes, o vetor  $\mathbf{x}^{in}(t)$  é utilizado na entrada da rede, enquanto o vetor  $\mathbf{x}^{out}(t)$  é utilizado na saída para calcular o erro de aproximação usado para guiar o ajuste dos pesos da rede. Por outro lado, quando se usa a técnica MATQV, o vetor  $\mathbf{x}^{out}(t)$  é apresentado na entrada da rede juntamente com o vetor  $\mathbf{x}^{in}(t)$ , sem o cálculo explícito do erro de aproximação.

Porém, o método MATQV é também baseado na redução do erro de aproximação, mas de forma implícita, ou seja, o erro é usado indiretamente para o ajustar os pesos de saída,  $\mathbf{w}_i^{out}(t)$ .

### 4.1.3. Análise da Convergência do Modelo SONARX

Nesta seção deseja-se saber se a rede SONARX, descrita como sendo a aplicação da técnica MATQV sobre o modelo de rede SOM, realmente aproxima de forma assintótica um determinado mapeamento dinâmico. Esta verificação será feita através de uma análise teórica do processo de convergência da rede SONARX.

Esta análise será feita com base no método de aproximação estocástica de (Robbins; Monro 1951) apud [Barreto, 2003 #78]. Ela visa mostrar que o erro de estimação decresce assintoticamente à medida que o treinamento da rede evolui, convergindo para um estado de equilíbrio estável, considerando que a função a ser mapeada é a definida pela Equação 4.14 e que a única fonte de informação disponível consiste em pares entrada-saída observados  $\{\mathbf{u}(t), \mathbf{y}(t)\}, t = 1, \dots, N$ .

De acordo com as definições da seção anterior, a rede SONARX pode ser usada para prover uma aproximação,  $\hat{F}(\cdot)$ , da função não-linear  $F(\cdot)$  usando mecanismos de memória associativa. Neste caso, a rede deve fornecer uma estimativa da saída do sistema,  $\hat{\mathbf{y}}(t+1)$ , dado o vetor de entrada  $\mathbf{x}^{in}(t)$ . Sendo representado matematicamente da seguinte forma:

$$\hat{\mathbf{y}}(t+1) \cong \mathbf{w}_{i^*}^{out}(t) = \hat{F}[\mathbf{x}^{in}(t); i^*(t)] \quad (4.22)$$

Podendo ser definida uma medida para o valor absoluto do erro de previsão da saída  $\mathbf{y}(t+1)$  futura do sistema:

$$\epsilon(t) = \|\mathbf{y}(t+1) - \hat{\mathbf{y}}(t+1)\| = \|\mathbf{x}^{out}(t) - \mathbf{w}_{i^*}^{out}(t)\| \quad (4.23)$$

Através desta definição pode-se agora provar que o método não-supervisionado proposto para aproximação de função é, na verdade, um procedimento “auto-supervisionado” (*self-supervised*) de aprendizagem baseada no erro de aproximação.

A análise que se segue é válida apenas para a fase de convergência do processo de aprendizado, fase esta em que a função vizinhança é, em geral, invariante no tempo. Para levar em conta, a interação entre o neurônio vencedor e seus vizinhos durante o processo de convergência, define-se a seguinte função  $J$ :

$$J(t) = \sum_{i=1}^L h(i^*, i, t) \epsilon^2(t) = \sum_{i=1}^L h(i^*, i, t) \|\mathbf{x}^{out}(t) - \mathbf{w}_i^{out}(t)\|^2 \quad (4.24)$$

onde  $L$  é o número de neurônios da rede.

A convergência da rede pode ser avaliada através do valor esperado do erro médio quadrático de aproximação, dado por:

$$\tilde{J} = E \left\{ \sum_{i \in L} h(i^*, i) \in^2 \right\} = \int \sum_{i \in L} h(i^*, i) \|\mathbf{x}^{out} - \mathbf{w}_i^{out}\|^2 p(\mathbf{X}) d\mathbf{X} \quad (4.25)$$

onde se assume que o valor esperado é tomado sobre uma seqüência infinita,  $\mathbf{X} = \{\mathbf{x}^{out}(t) ; t = 1, 2, \dots, \infty\}$  de amostras estocásticas sobre a qual o estimador é formado. O termo  $p(\mathbf{X})$  refere-se à função densidade de probabilidade da seqüência  $\mathbf{X}$ , enquanto o termo  $d\mathbf{X}$  é um “volume infinitesimal” no espaço em que  $\mathbf{X}$  está definida. O índice  $t$  é omitido para simplificar a notação, pois ele já está implícito na ordem temporal da seqüência  $\mathbf{X}$ .

Visto que a função densidade de probabilidade  $p(\mathbf{X})$  é desconhecida e que a seqüência  $\mathbf{X}$  é finita para os casos reais, a técnica de aproximação estocástica de Robins-Monro será utilizada para a minimização de  $\tilde{J}$  visando a obtenção de um valor ótimo para o vetor de pesos  $\mathbf{w}_i^{out}$ . A idéia básica desta técnica é tentar diminuir o valor da função  $J(t)$  a cada novo passo  $t$ , caminhando na direção oposta ao gradiente desta função em relação ao vetor  $\mathbf{w}_i^{out}(t)$  atual. A formulação recursiva desta idéia é dada por:

$$\mathbf{w}_i^{out}(t+1) = \mathbf{w}_i^{out}(t) - \frac{1}{2} \alpha(t) \frac{\partial J(t)}{\partial \mathbf{w}_i^{out}(t)} \quad (4.26)$$

A derivada parcial de  $J(t)$  na Eq. 4.24 é dada por:

$$\frac{\partial J(t)}{\partial \mathbf{w}_i^{out}(t)} = -2 \cdot h(i^*, i, t) [\mathbf{x}^{out}(t) - \mathbf{w}_i^{out}(t)] \quad (4.27)$$

Logo, a Equação 4.26 passa a ser escrita como:

$$\mathbf{w}_i^{out}(t+1) = \mathbf{w}_i^{out}(t) - \alpha(t) h(i^*, i, t) [\mathbf{x}^{out}(t) - \mathbf{w}_i^{out}(t)] \quad (4.28)$$

Nota-se que esta é exatamente a regra de aprendizagem mostrada na Equação 4.20. Logo, conclui-se que a partir de condições iniciais arbitrárias, a seqüência de valores do vetor  $\mathbf{w}_i^{out}(t)$  converge para a vizinhança do valor ótimo  $\tilde{\mathbf{w}}_i^{out}(t)$ , correspondente ao mínimo de  $\tilde{J}$ .

É importante ressaltar que esta técnica tende a convergir para mínimos locais de  $J$ , dependendo das condições iniciais de  $\mathbf{w}_i^{out}(t=0)$ . Entretanto, para a rede SOM unidimensional, foi mostrado que existe apenas um único estado de equilíbrio estável,

representando um mínimo global (Cottrell; Fort et al. 1998), podendo este resultado ser aplicado diretamente à rede SONARX.

#### 4.1.4. Representação Espaço de Estados da Rede SONARX

Para caracterizar a rede SONARX em termos da representação Espaço de Estado, é necessário definir um vetor de estados, que leve em conta o tipo de mecanismo de memória usado pelo modelo para processar padrões dinâmicos. Neste caso, para a rede SONARX, há duas fontes de memória que estão representadas pelo vetor de entrada,  $\mathbf{x}^{in}(t)$ , e pelo vetor de saída,  $\mathbf{x}^{out}(t)$ .

Para o caso em que se aproxima o modelo do sistema, pode-se definir o seguinte vetor de estados:

$$s(t) = \mathbf{x}^{in}(t) = \{\mathbf{y}(t) \oplus \dots \oplus \mathbf{y}(t - n_y + 1)\} \oplus \{\mathbf{u}(t) \oplus \dots \oplus \mathbf{u}(t - n_u + 1)\} \quad (4.29)$$

Sendo também necessária a definição do conjunto de parâmetros adaptáveis do modelo, que incluem as matrizes de pesos  $\mathbf{w}_i^{in}(t)$  e  $\mathbf{w}_i^{out}(t)$ :

$$\mathbf{W}(t) = \begin{pmatrix} \mathbf{w}_i^{in}(t) \\ \mathbf{w}_i^{out}(t) \end{pmatrix}, \quad i = 1, \dots, N \quad (4.30)$$

Conforme mostrado na Equação 4.21, a saída da rede SONARX é uma estimativa da próxima saída do sistema:

$$y(t) = f_y(s(t), \mathbf{W}) \cong \mathbf{w}_i^{out}(t) \quad (4.31)$$

Os valores iniciais dos parâmetros são dados por:

$$\mathbf{W}(0) = \begin{pmatrix} \mathbf{w}_i^{in}(0) \\ \mathbf{w}_i^{out}(0) \end{pmatrix} = \begin{pmatrix} \text{rand}(0,1) \\ \text{rand}(0,1) \end{pmatrix}, \quad \forall i \quad (4.32)$$

em que  $\text{rand}(0, 1)$  é uma função que atribui aos componentes de  $\mathbf{x}$  números aleatórios extraídos de uma distribuição uniforme entre 0 e 1. Por fim, a função de ajuste de  $\mathbf{W}$  é dada por:

$$\Delta \mathbf{W}(t) = \begin{pmatrix} \Delta \mathbf{w}_i^{in}(t) \\ \Delta \mathbf{w}_i^{out}(t) \end{pmatrix} = \begin{pmatrix} \eta(t)h(i^*, i, t)[\mathbf{x}^{in}(t) - \mathbf{w}_i^{in}(t)] \\ \eta(t)h(i^*, i, t)[\mathbf{x}^{out}(t) - \mathbf{w}_i^{out}(t)] \end{pmatrix} \quad (4.33)$$

conforme calculados pelas Equações 4.19 e 4.20.

Na próxima seção se discutirá como a rede SONARX pode ser usada em uma ampla gama de aplicações, tais como identificação de sistemas não-lineares, aprendizagem de trajetórias complexas, controle preditivo não-linear e principalmente na previsão de séries temporais, que é o enfoque deste trabalho de pesquisa.

#### 4.1.5. Possíveis Aplicações

O foco da presente seção está na breve discussão das possíveis aplicações da rede SONARX. A primeira aplicação a ser analisada é a identificação de sistemas não-lineares, onde se está interessado na aplicação da rede SONARX na modelagem de sistemas dinâmicos do tipo entrada-saída. Por hora, será tratado apenas o caso em que as séries de entrada e saída são univariadas, ou seja, o sistema é do tipo SISO (Single-Input/Single-Output).

Assumindo que  $u \in \mathfrak{R}$  denota a variável de entrada e  $y \in \mathfrak{R}$  simboliza a variável de saída para o sistema de interesse, uma forma bastante comum de se representar um sistema dinâmico SISO é através do modelo NARX:

$$y(t+1) = f[y(t), \dots, y(t-n_y+1); u(t), \dots, u(t-n_u+1)] \quad (4.34)$$

onde  $n_u$  e  $n_y$  definem as ordens da memória de entrada e de saída, respectivamente.

Para prosseguir com o processo de identificação, assume-se ainda que se está de posse de pares de amostras entrada-saída,  $\{u(t), y(t)\}, t = 1, \dots, N$ , coletadas experimentalmente.

A utilização da rede SONARX na modelagem de sistemas SISO é bastante simples, bastando para isso escolher os vetores  $\mathbf{x}^{in}(t)$  e  $\mathbf{x}^{out}(t)$  da seguinte forma:

$$\mathbf{x}^{in}(t) = [y(t), \dots, y(t-n_y+1); u(t), \dots, u(t-n_u+1)] \quad (4.35)$$

$$\mathbf{x}^{out}(t) = y(t+1) \quad (4.36)$$

A determinação do neurônio vencedor obedece a Equação 4.18 e o ajuste dos pesos da rede SONARX segue as Equações 4.19 e 4.20. Após o treinamento, as saídas da rede SONARX podem ser dadas por:

$$\hat{y}(t+1) \cong w_{i^*}^{out}(t) \quad (4.37)$$

Para exemplificar esta aplicação, a identificação de um sistema não-linear do tipo SISO, temos o estudo de caso apresentado por (Barreto 2003). Escolheu-se a tarefa de identificação de um atuador de uma estrutura mecânica hidráulica (grua), mais precisamente na identificação do atuador do braço, visto que este é o mais afetado pelo comportamento oscilatório da estrutura.

Já a segunda aplicação estaria direcionada à identificação de sistemas do tipo Múltiplas Entradas / Múltiplas Saídas (MIMO), nos quais existem mais de uma variável de entrada e mais de uma variável de saída. Este caso, representado pela Equação 4.14, é o mais geral e difícil sistema de ser modelado.

Tendo em vista esta finalidade, (Barreto 2003) escolheu a tarefa de aprendizagem e reprodução de trajetórias robóticas complexas. O objetivo é fazer com que a rede SONARX aprenda a estimar as próximas posições angulares do braço do robô PUMA 560 durante a execução de uma trajetória contendo posturas repetidas.

Para isso, assume-se como variáveis de entrada, as posições cartesianas do efetuador do robô,  $\mathbf{x}(t) \in \mathfrak{R}^3$ , e como variáveis de saída, os seis ângulos das juntas,  $\theta(t) \in \mathfrak{R}^6$ :

$$\mathbf{u}(t) = \mathbf{x}(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{pmatrix} \quad (4.38)$$

$$\mathbf{y}(t) = \boldsymbol{\theta}(t) = \begin{pmatrix} \theta_1(t) \\ \vdots \\ \theta_6(t) \end{pmatrix} \quad (4.39)$$

Assim, para este caso a Equação 4.14 é representada da seguinte forma:

$$\theta(t) = f[\theta(t), \dots, \theta(t - n_\theta + 1); \mathbf{x}(t), \dots, \mathbf{x}(t - n_x + 1)] \quad (4.40)$$

Logo, a partir desta equação, os vetores  $\mathbf{x}^{in}(t)$  e  $\mathbf{x}^{out}(t)$  da SONARX são definidos:

$$\mathbf{x}^{in}(t) = [\boldsymbol{\theta}(t), \dots, \boldsymbol{\theta}(t - n_\theta + 1); \mathbf{x}(t), \dots, \mathbf{x}(t - n_x + 1)] \quad (4.41)$$

$$\mathbf{x}^{out}(t) = \boldsymbol{\theta}(t + 1) \quad (4.42)$$

Assim o neurônio vencedor é determinado pela Equação 4.18, e o ajuste dos pesos segue as Equações 4.19 e 4.20. Finalmente tem-se a estimativa do próximo vetor de ângulos das juntas dada por:

$$\hat{\theta}(t+1) \cong \mathbf{w}_{j^*}^{out}(t) \quad (4.43)$$

Desta maneira, são então feitas as previsões dos ângulos da junta do robô PUMA 560, e esta estimativa pode ser usada no controle do manipulador deste robô.

Finalmente será discutida a aplicação da rede SONARX na modelagem e previsão de séries temporais, que é o enfoque principal deste trabalho. Busca-se por esta aplicação a capacidade de prever o comportamento futuro de uma série particular de eventos a partir do conhecimento do seu presente e do seu passado, ou seja, remover possíveis redundâncias nos dados e, por conseguinte, descobrir regularidades estatísticas ou dinâmicas presentes nas séries (Barreto; Araújo 2001b).

De uma forma geral, a previsão de séries temporais é um problema de processamento de sinais em que se tem uma seqüência de  $N$  amostras de uma determinada variável escalar,  $\{y(t), y(t-1), \dots, y(t-N+1)\}$ , uniformemente espaçadas no tempo, e cujo objetivo é obter uma estimativa,  $\hat{y}(t+1)$ , para o próximo elemento da série. Uma possível formulação matemática para este problema supõe que a série observada possui uma evolução temporal e pode ser descrita por um modelo auto-regressivo não-linear:

$$y(t+1) = f[y(t), y(t-1), \dots, y(t-n_y+1)] + \varepsilon(t) \quad (4.44)$$

onde  $1 \leq n_y \ll N$  é a ordem do modelo,  $\varepsilon(t)$  é uma variável aleatória, de média zero e variância  $\sigma_c^2$ , representando as imprecisões de medida e/ou de modelagem. A função  $f(\cdot)$  é o mapeamento matemático que relaciona o passado e o presente da série com o seu futuro.

Para efeito de simplificação, a abordagem mais comum assume que  $\varepsilon(t)$  é uma seqüência de ruído branco gaussiano de média zero e variância  $\sigma_c^2$ . Assim, tem-se a seguinte representação:

$$\hat{y}(t+1) = \hat{f}[y(t), y(t-1), \dots, y(t-n_y+1); \mathbf{W}] \quad (4.45)$$

onde  $\hat{f}(\cdot)$  define uma aproximação da função  $f(\cdot)$  e  $\mathbf{W}$  é a matriz de pesos deste modelo matemático a ser ajustado à série.

Assim, define-se o erro de previsão como sendo a diferença entre o valor realmente observado para a próxima amostra da série e a estimativa  $\hat{y}(t+1)$  produzida

pelo modelo, sendo isto válido para todo instante  $t$  do horizonte de previsão. Desta maneira tem-se:

$$e(t) = |y(t) - \hat{y}(t)| \quad (4.46)$$

Sendo ainda necessária uma avaliação global deste tipo de modelagem, no que se refere aos erros de predição, define-se então o Erro Percentual Médio (EPM), o Erro Quadrático Médio (MSE, em inglês) e o Erro Relativo Médio (MRE, em inglês), dados pelas seguintes expressões:

$$EPM = \frac{100}{N} \sum_{t=1}^N \frac{|y(t) - \hat{y}(t)|}{y(t)} \quad (4.47)$$

$$MSE = \frac{1}{N} \sum_{t=1}^N [y(t) - \hat{y}(t)]^2 \quad (4.48)$$

$$MRE = \frac{1}{N} \sum_{t=1}^N \left[ \frac{|y(t) - \hat{y}(t)|}{y(t)} \right]^2 \quad (4.49)$$

onde  $N$  é o tamanho da seqüência a ser prevista pelo modelo da rede SONARX.

## **5. Aplicação da rede SONARX em S.E.P.**

Como já dito anteriormente, dependendo de como as usinas hidroelétricas forem operadas, isto é, de como a água dos seus reservatórios for utilizada, têm-se diferentes quantidades de energia gerada. Isto porque a produtividade das usinas depende significativamente do volume dos reservatórios, pois variando-se o volume, varia-se a altura de queda e conseqüentemente a sua produtividade.

Além disso, usinas hidroelétricas situadas em um sistema em cascata sofrem a influência da sua localização na cascata. Assim, a energia obtida pela água armazenada no reservatório de uma dada usina depende não só das condições deste reservatório, mas também do volume dos reservatórios das usinas localizadas à jusante, pois esta água será turbinada em todas as usinas rio abaixo.

Como são as regras de operação que determinam os estados dos reservatórios das usinas que compõem o sistema, o objetivo deste capítulo é justamente o de aprender o comportamento de um sistema composto por várias usinas em cascata operando de forma otimizada, e com isso, determinar regras de operação que consigam representar esta operação ótima.

A ferramenta utilizada para a realização destes estudos é o algoritmo determinístico de fluxo em redes já apresentado no item 3.1.1. Sendo a seguir apresentado o sistema teste utilizado neste trabalho de pesquisa.

### **5.1. Sistema Teste**

O sistema teste utilizado é composto por 7 (sete) importantes usinas do Sistema Sudeste Brasileiro: as usinas de Emborcação, Itumbiara e São Simão, localizadas ao longo do rio Paranaíba; enquanto as usinas de Furnas, Marimondo e Água Vermelha, localizadas ao longo do rio Grande; e finalmente a usina de Ilha Solteira, localizada no

rio Paraná. Este sistema está interligado em cascata formando um Y como mostra a Figura 5.1.

Este sistema, com uma potência instalada de 12.572 MW, apresenta uma configuração muito interessante para o estudo em questão, tendo em vista a interligação hidráulica, com conseqüente complexidade operativa das usinas quando analisadas em conjunto.

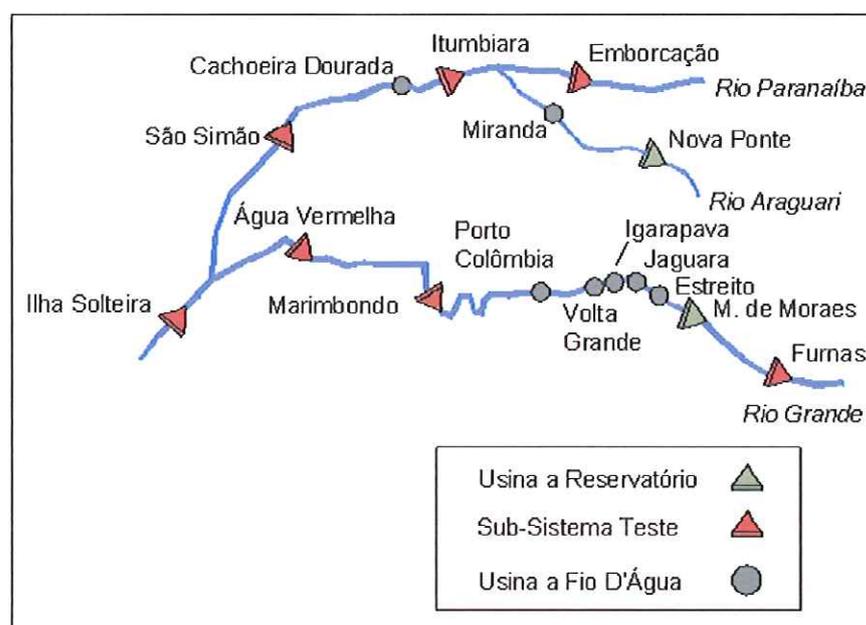


Figura 5.1 - Sistema teste, parte da Bacia do Rio Paraná

### 5.1.1. Otimização Determinística

Durante a otimização determinística utilizou-se uma demanda de 12.570 MW, que corresponde a 100% da potência instalada deste parque hidroelétrico. Este montante de demanda foi assim definido para que o parque termoeletrico complementar pudesse atuar gerando uma certa quantidade de energia, a qual deveria ser minimizada neste processo de otimização, como pode ser visto na Figura 5.3, enquanto na Figura 5.2 pode ser observada a queda do custo de operação durante este processo.

A otimização deste sistema define de forma determinística a operação ótima do sistema, já que são conhecidas as vazões afluentes desde janeiro de 1931 até dezembro de 1994.

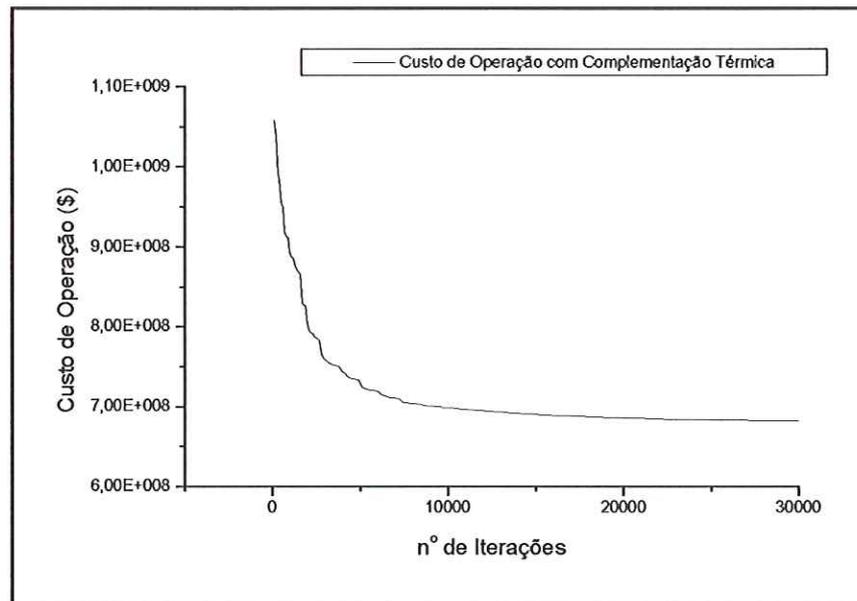


Figura 5.2 - Evolução do custo de operação durante o processo de otimização

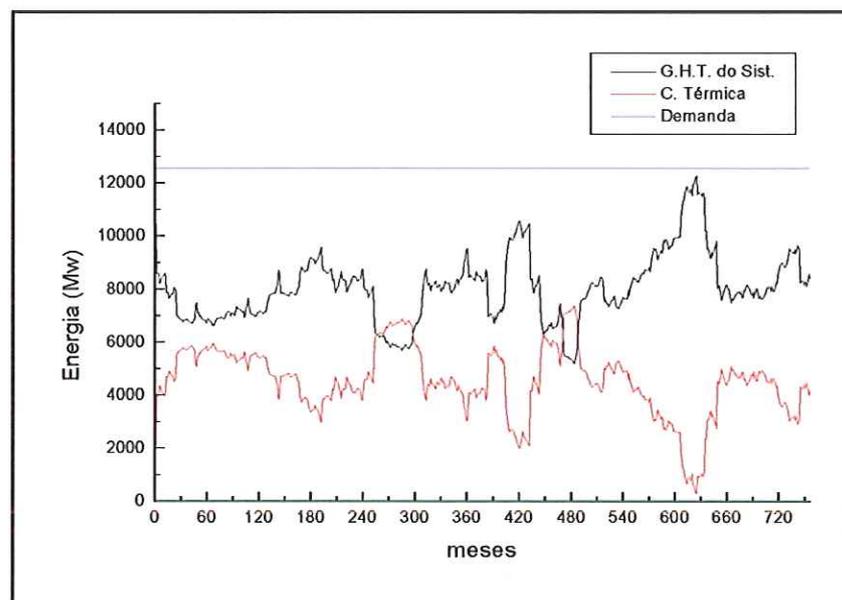


Figura 5.3 - Trajetória de mercado ao longo do período de otimização

Visto que o ano hidrológico, definido como o período compreendido entre o início da estação seca até o final da estação chuvosa, na bacia hidrográfica do rio Paraná começa em maio e termina em abril do ano seguinte, tivemos que fazer a otimização, levando em consideração somente o período compreendido entre maio de 1931 até abril de 1994, totalizando 63 anos de otimização, o que corresponde a 756 meses. Posteriormente, este resultado foi pré-processado e serviu como conjunto de padrões para o treinamento da rede SONARX.

Como a rede SONARX depende de dados do passado para fazer suas previsões futuras, houve a necessidade da exclusão de alguns anos iniciais e finais deste conjunto de dados. Desta maneira, tivemos um conjunto de padrões para treinamento e teste, correspondente ao período histórico de janeiro de 1935 até dezembro de 1992, totalizando 58 anos de otimização, o que corresponde a 696 meses. Como pode ser observado da Figura 5.4 até a Figura 5.9.

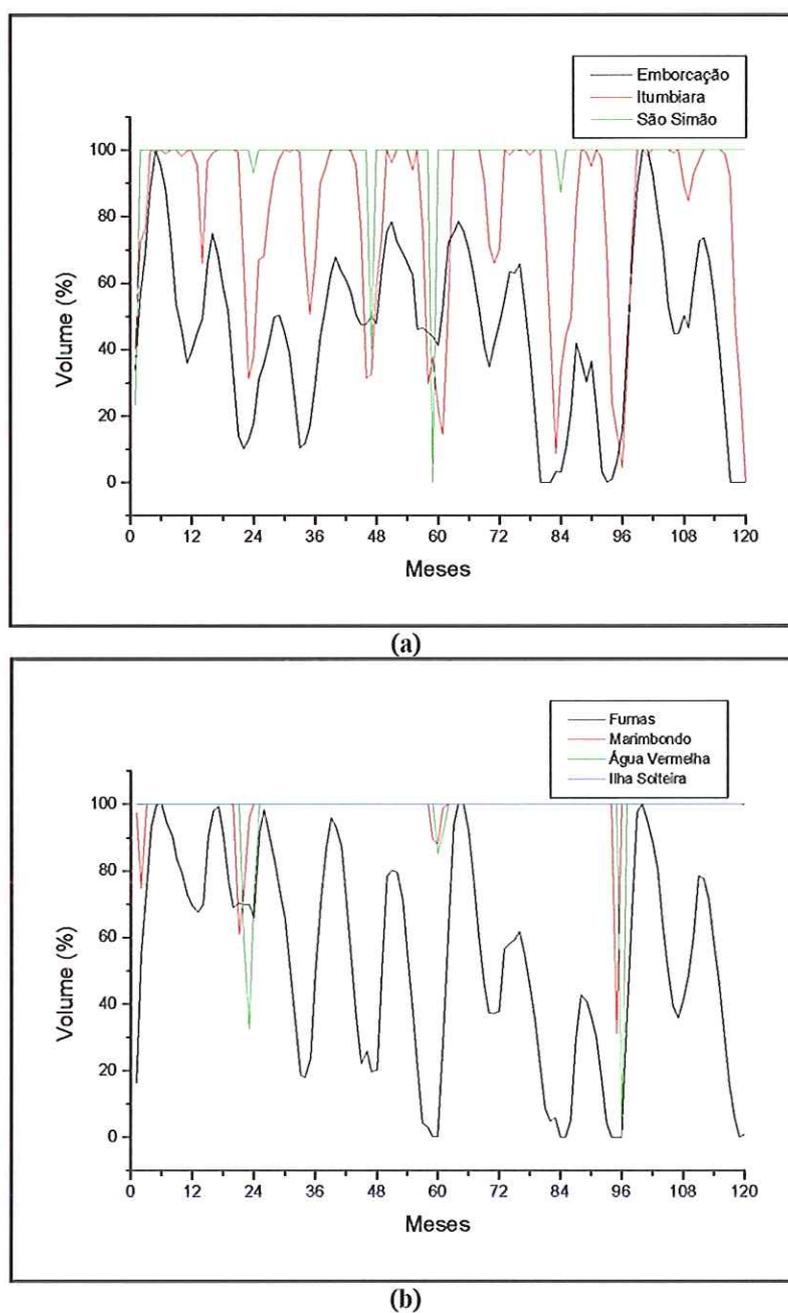
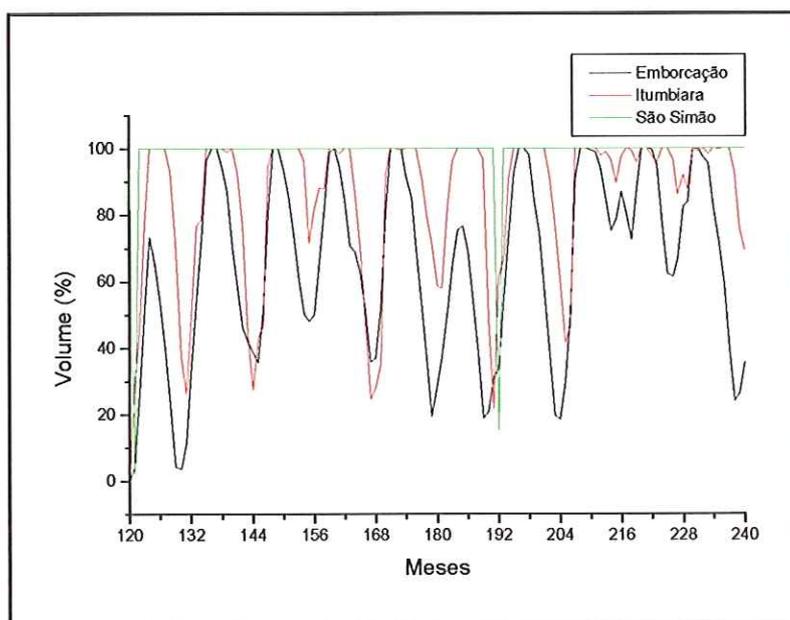
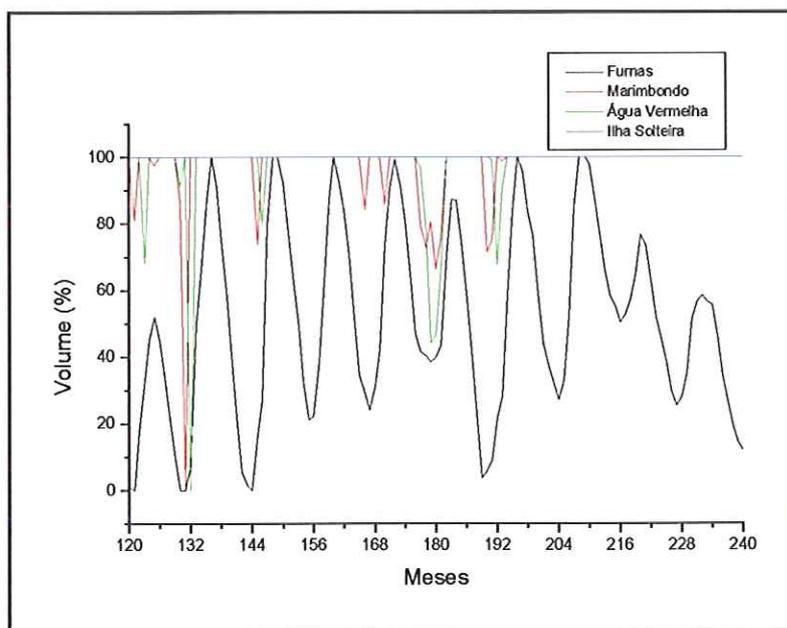


Figura 5.4 - Trajetória dos volumes ótimos para o período de Jan/1935 até Dez/1944

As usinas em cascata têm a tendência de valorizar a água, de tal forma que o processo de deplecionamento é feito de montante à jusante, e o processo de enchimento de jusante à montante. Com isso, as usinas de jusante tendem a permanecer como usinas a fio d'água, sem alterações significativas nos seus volumes, cabendo aos reservatórios de cabeceira a regulação das vazões afluentes ao sistema, o que faz com que, a produtividade das usinas permaneça a maior possível.

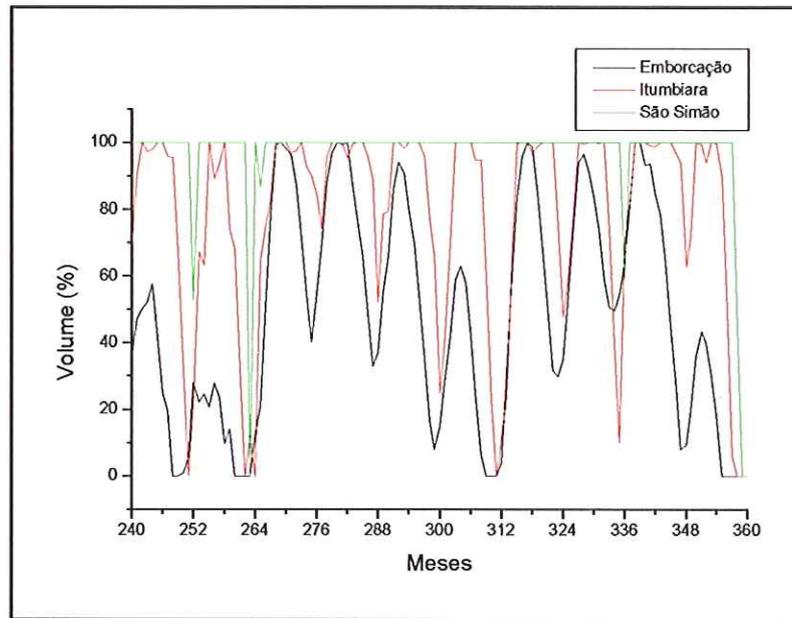


(a)

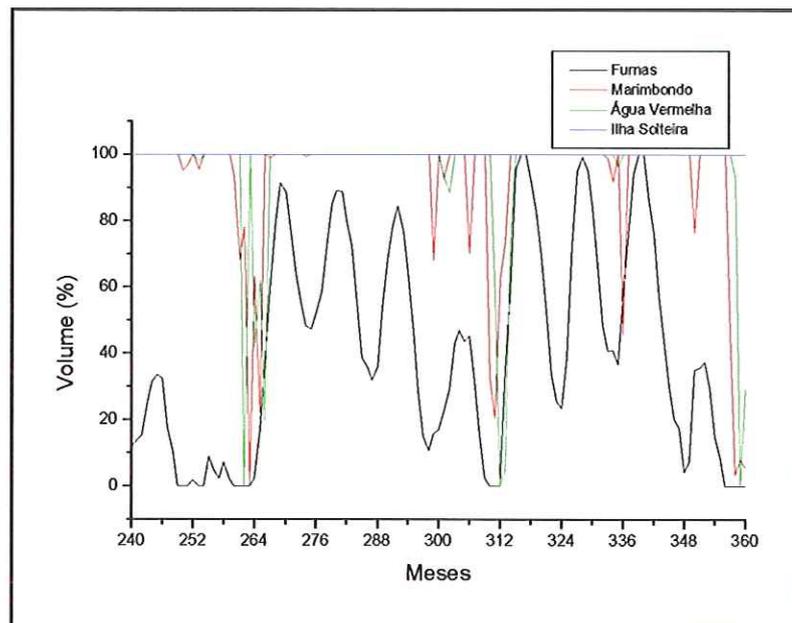


(b)

Figura 5.5 - Trajetória dos volumes ótimos para o período de Jan/1945 até Dez/1954



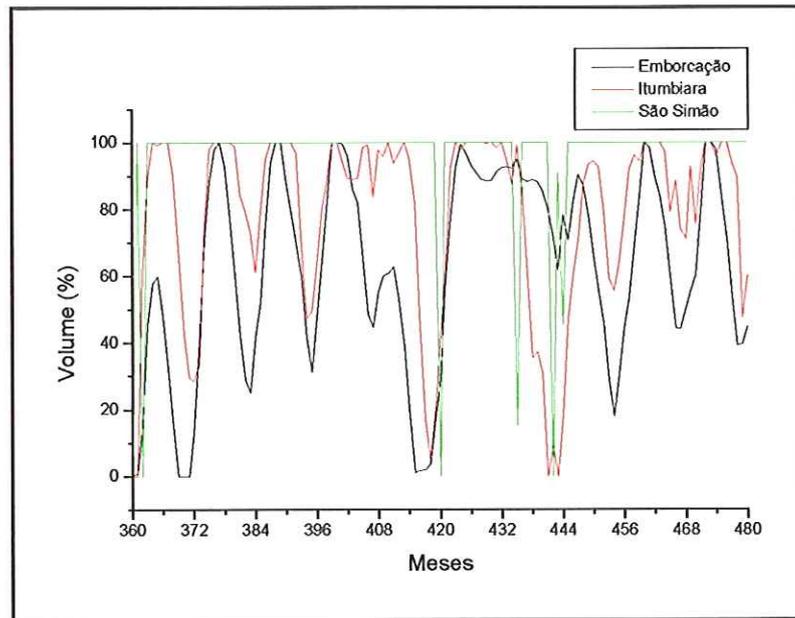
(a)



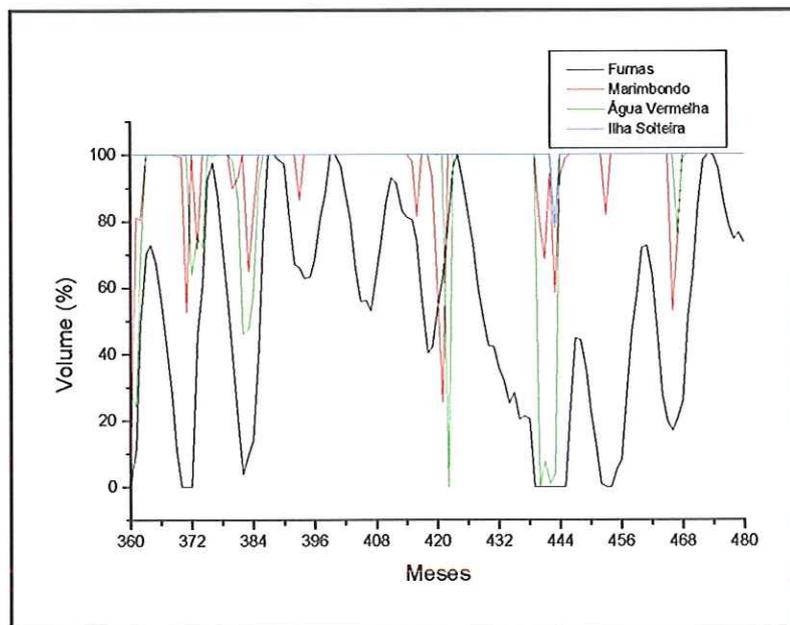
(b)

Figura 5.6 - Trajetória dos volumes ótimos para o período de Jan/1955 até Dez/1964

Como pode ser visto nestas figuras, que descrevem as trajetórias operativas ótimas, no caso da usina de Emborcação, definida como a trajetória de cor preta em (a), seu reservatório é praticamente todo utilizado ano após anos, na tarefa de regularizar as vazões afluentes à usina de jusante Itumbiara. Mas como o seu reservatório não é grande o suficiente para isto, a usina de Itumbiara é “requisitada” a auxiliar neste papel regulador. Observando a trajetória de cor vermelha em (a), nota-se que o reservatório de Itumbiara depleciona seu volume em praticamente todos os anos, tão logo o reservatório de Emborcação esteja no mínimo operativo ou quase no mínimo.



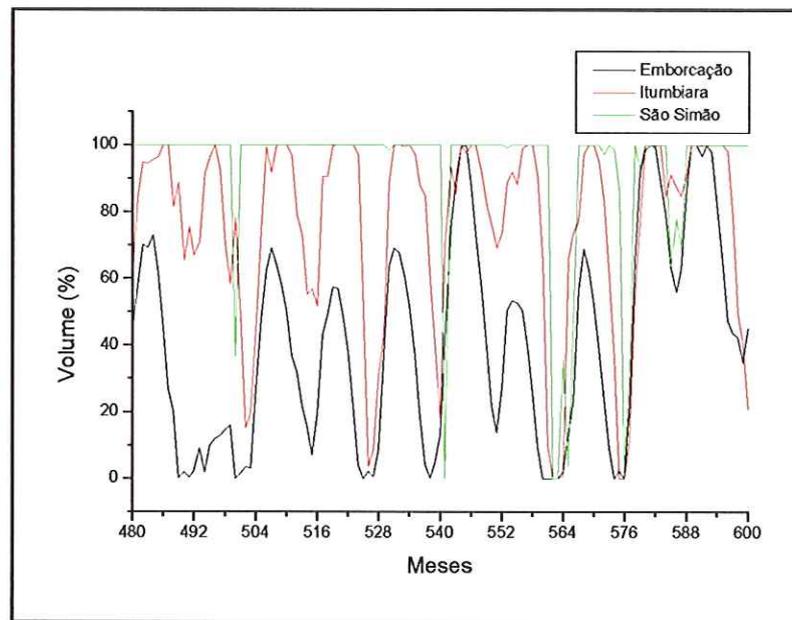
(a)



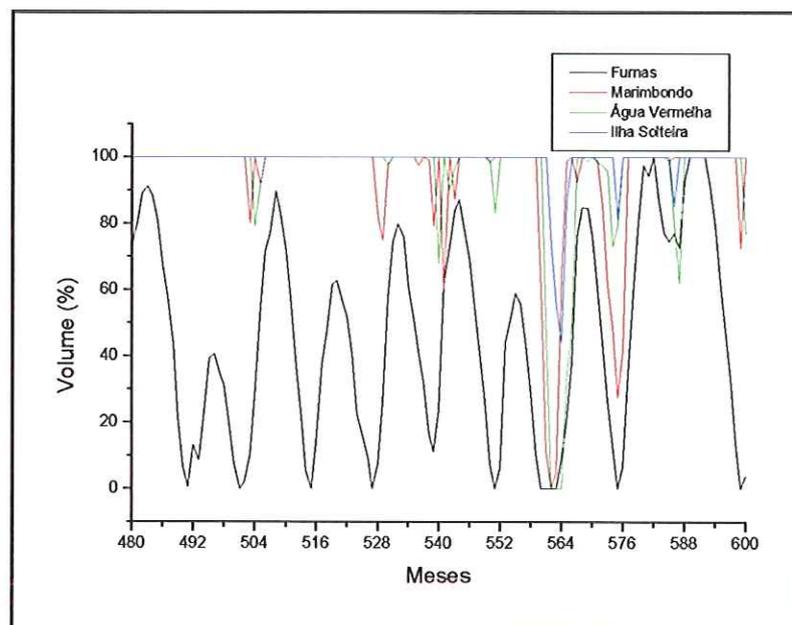
(b)

**Figura 5.7 - Trajetória dos volumes ótimos para o período de Jan/1965 até Dez/1974**

Ainda com relação às 3 (três) usinas do rio Paranaíba, logo abaixo tem-se a usina de São Simão, que permaneceu quase que o tempo todo operando a fio d'água, como pode ser observado na trajetória de cor verde em (a). Esta decisão operativa, como já foi discutido, busca preservar altura de queda desta usina, mantendo uma maior produtividade, o que corresponde a uma eficiência sistêmica, já que toda a água defluída por Emborcação e Itumbiara invariavelmente passa por São Simão. Entretanto, nas ocasiões que São Simão deplecionou seu reservatório, ela o fez antevendo um período chuvoso à sua frente, evitando assim futuros vertimentos, o que corresponde a um



(a)

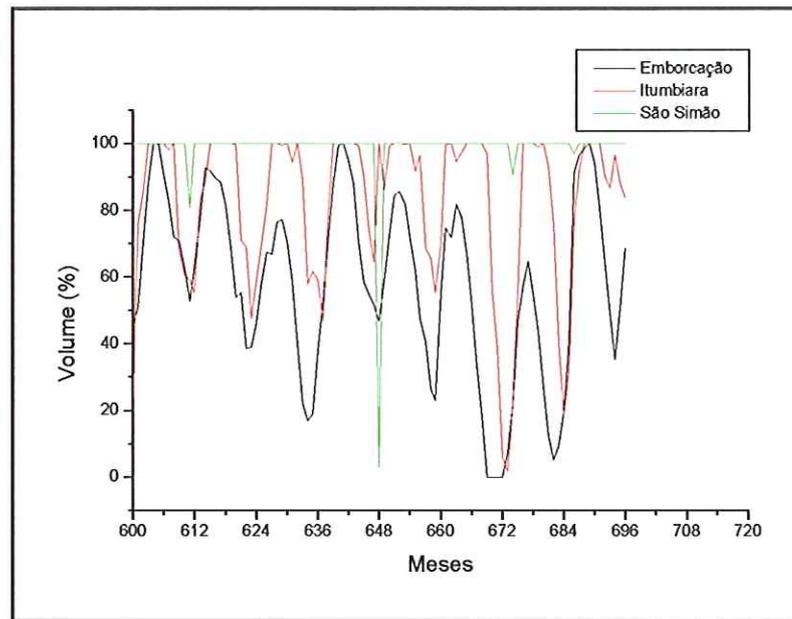


(b)

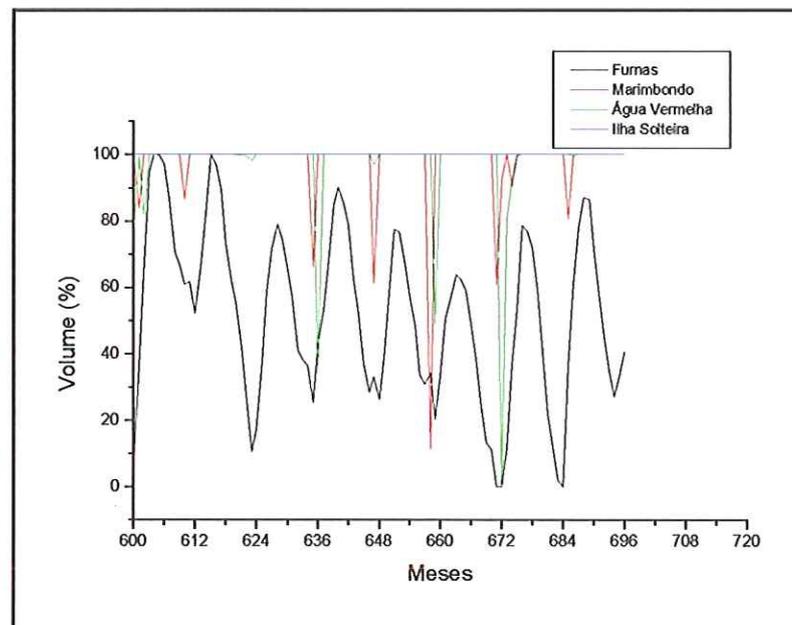
Figura 5.8 - Trajetória dos volumes ótimos para o período de Jan/1975 até Dez/1984

desperdício energético muito grande.

Com relação às 3 (três) usinas pertencentes ao rio Grande, a usina de cabeceira, Furnas, também exerce o papel regulador desta cascata. Como o seu reservatório é bem grande, ela consegue regularizar boa parte das vazões afluentes, ficando para Marimbondo deplecionar e regularizar as vazões em poucos meses do período de otimização, como pode ser visto nas trajetórias de cor preta e vermelha em (b), respectivamente. Já a usina de Água Vermelha, permanece quase que o tempo todo operando a fio d'água, assim como São Simão, segundo a trajetória de cor verde em (b).



(a)



(b)

**Figura 5.9 - Trajetória dos volumes ótimos para o período de Jan/1985 até Dez/1992**

Finalmente temos a usina de Ilha Solteira, pertencente ao rio Paraná, tida com a usina mais a jusante do sistema teste, por isso e pelas razões já comentadas de eficiência energética, Ilha Solteira teve uma trajetória de operação ótima, quase que totalmente “flat”, com pouquíssimos meses em que isto não ocorreu.

Uma outra característica observada foi a maneira como as usinas se comportam para manterem a máxima produtividade possível. As usinas utilizam somente a parcela de água dos reservatórios possível de ser recuperada nos próximos períodos chuvosos, e ainda evitando a ocorrência de vertimentos. Assim, nos períodos de menor vazão

afluente, menor é o deplecionamento dos reservatórios das usinas, isto porque, com o conhecimento da pouca afluência futura, o sistema prefere a utilização das térmicas, poupando água para o enchimento futuro dos reservatórios para a obtenção de produtividade.

Assim, a partir destes resultados, observou-se certos comportamentos para os reservatórios das usinas de acordo com a sua posição na cascata, comprovando o que já foi dito, de que os reservatórios das usinas possuem comportamentos diferenciados e estes devem ser levados em consideração durante a operação.

Os resultados confirmam o papel que cada tipo de usina possui no sistema, de acordo com a sua localização e sua importância. As usinas de montante com incumbência de absorver a sazonalidade das afluências e da demanda, são as que mais oscilam, variando o seu volume. As usinas de jusante, com o papel de manter a máxima produtividade, oscilam o menos possível o seu volume. E as usinas intermediárias possuem um comportamento misto entre os dois tipos citados anteriormente.

### **5.1.2. Treinamento da rede SONARX**

Houve a necessidade de um pré-processamento de todos os dados envolvidos neste processo de treinamento e teste de arquitetura da rede proposta. Como esta rede neural trabalha com dados de natureza contínua, optou-se por uma normalização destes dados, a fim de garantir a estabilidade da rede através da utilização de dados limitados entre 0 e 1. Esta normalização foi feita tanto para os dados de afluências como para os dados de volumes operativos de cada usina hidroelétrica.

No caso dos dados de afluência, tomou-se como limite superior o recorde histórico já registrado para cada uma das usinas. De posse desta vazão afluente histórica máxima, dividiu-se todo o histórico de vazões por este pico máximo, mês a mês, e usina por usina. Gerando um “histórico” de afluências normalizado pelo pico histórico de cada usina hidroelétrica.

Já os dados relativos aos volumes operativos ótimos, tomou-se como parâmetro de normalização, o próprio volume útil operativo de cada usina hidroelétrica. Sendo o volume útil operativo aquele referente a diferença entre o máximo volume operativo e o mínimo volume operativo, onde as correspondências se definem entre o volume mínimo operativo e o dado igual a 0, e o volume máximo operativo e o dado igual a 1.

Para uma melhor compreensão dos resultados apresentados a seguir, é importante fazer as seguintes observações. A usina de interesse, ou seja aquela para a qual se pretende realizar as previsões é a usina de Furnas. Para isso foram utilizados os dados de afluência e volume desta e de mais seis usinas, formando assim um pequeno sistema teste de usinas pertencentes ao Subsistema Sudeste, como pode ser visto na Figura 5.1.

Os dados de entrada estão dispostos da seguinte maneira: de um total de 84 entradas do passado, a primeira metade (1 até 42) se refere aos dados de afluência enquanto que a outra metade (43 até 84) aos dados de volume. Os primeiros dados são referentes à usina de Emborcação, mais precisamente os dados de afluência deslocados no tempo de  $t$ ,  $t-1$ ,  $t-2$ ,  $t-11$ ,  $t-12$  e  $t-13$ , constituindo os 6 primeiros dados de entrada (1 até 6), considerando isso sucessivamente para todas as outras usinas Itumbiara (7-12), São Simão (13-18), Furnas (19-24), Marimbondo (25-30), Água Vermelha (31-36) e Ilha Solteira (37-42). O mesmo ocorre para os dados de volume, totalizando assim as 84 entradas. Essa janela do passado foi assim definida com a intenção de fornecer ao modelo condições de distinguir e capturar as sazonalidades mensais, e anuais das séries.

Já os dados de saída são fornecidos por quatro unidades, sendo que a primeira está relacionada aos dados de previsão um passo a frente, ou seja  $t+1$ , a segunda unidade de saída com  $t+3$ , a terceira com  $t+6$  e finalmente a quarta com  $t+12$ .

Para o treinamento da rede SONARX, todo o histórico de vazões afluentes e todo o histórico de volumes operativos ótimos, ambos normalizados, foram divididos em 11 períodos de 5 anos cada. Sendo o primeiro correspondente ao período de janeiro de 1936 até dezembro de 1940, o segundo referente ao período de janeiro de 1941 até dezembro de 1945, e assim por diante até dezembro de 1990.

A rede foi treinada tendo como conjunto de padrões de treinamento apenas dez dentre os onze conjuntos disponíveis, deixando um conjunto de 5 anos somente para o teste da rede. Ou seja, cada período de teste foi retirado do conjunto total de dados disponíveis; a partir disso a rede tem seus parâmetros ajustados, de forma que todos os padrões do conjunto de treinamento sejam apresentados de forma aleatória à entrada da rede. A seguir, o conjunto de teste é utilizado para verificar o desempenho da rede já treinada, e este processo é repetido por um número finito de vezes, sendo que neste trabalho a rede passou por 100 épocas de treinamento em cada teste. Finalmente, este procedimento é aplicado para cada conjunto de teste, resultando para cada conjunto de

treinamento um determinado número de centros na estrutura da rede, o que pode ser chamado de Validação Cruzada.

A seguir serão apresentados os resultados obtidos, lembrando que a rede SONARX teve como base de dados para treinamento as séries históricas de vazões afluentes e decisões operativas ótimas para todas as usinas envolvidas no problema de otimização energético. Como já mencionado, objetivando a previsão tanto das vazões afluentes como dos estados do volume do reservatório da Usina Hidroelétrica de Furnas.

### **5.1.3. Resultados da Predição de Afluência e Volumes Ótimos**

Os testes se referem à previsão tanto das vazões afluentes como dos estados do volume do reservatório da Usina Hidroelétrica de Furnas, tendo como base de dados para o treinamento da rede SONARX as séries de vazões afluentes históricas às sete usinas hidroelétricas, bem como os volumes operativos ótimos, determinados pelo algoritmo de otimização determinística, destas usinas.

Para tanto, esta seção está sub-dividida segundo algumas modificações propostas ao modelo original da rede, e que serão discutidas posteriormente. Tais modificações foram implementadas com o intuito de melhorar a criação do conjunto de centros (ou neurônios) ainda na fase de treinamento, além da implementação de uma modificação na fase de testes, que visava uma melhor recuperação do padrão absorvido pela rede, tudo isso, objetivando melhorar o desempenho deste modelo de rede SONARX.

Além disso, os gráficos de barras estão dispostos de maneira que se possa comparar os erros obtidos na validação dos modelos propostos. Na fase de treinamento a rede foi avaliada na recuperação dos próprios padrões treinados, enquanto na fase de teste a rede foi avaliada na recuperação de padrões novos e desconhecidos.

Dentre os 11 (onze) modelos propostos, os quatro primeiros utilizam somente os dados do passado na sua fase de treinamento, enquanto que do quinto modelo até o oitavo modelo passaram a utilizar além dos dados do passado, os dados do futuro na fase de treinamento.

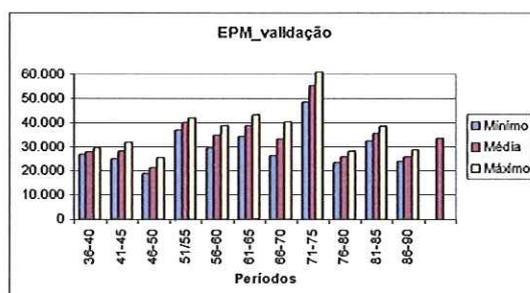
Os primeiros resultados apresentados são os referentes ao modelo original da rede SONARX, com apenas uma modificação em sua proposição original, a maneira como os centros são criados, durante o próprio treinamento. Assim, como já foi mencionado na Seção 4.1.2, o processo de treinamento pode ser resumido em 3 etapas:

criação dos centros, definida pela Equação 5.1, dada a seguir; busca pelo neurônio vencedor, descrito pela Equação 4.7; e atualização dos vetores de pesos, dada pela Equação 4.8.

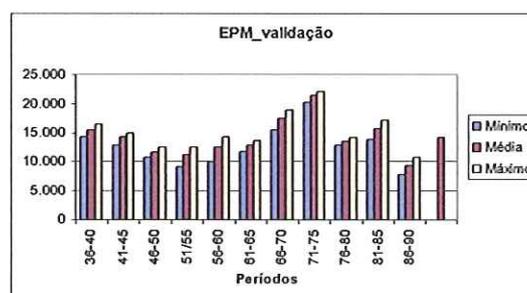
$$\overrightarrow{dist} = |x_{in} - w_{in}| > \varepsilon \quad (5.1)$$

onde  $\varepsilon$  é uma dada tolerância definida pelo usuário, e um novo centro só seria criado caso esta Equação fosse satisfeita durante o treinamento, por um dado novo padrão de treinamento.

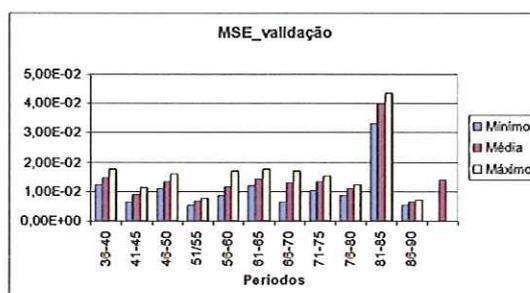
A seguir estão os erros de previsão das vazões afluentes, e a dos volumes operativos à usina de Furnas, dada pelas Figuras 5.10 e 5.11, onde somente os dados do passado (já conhecido) são considerados para a criação dos centros.



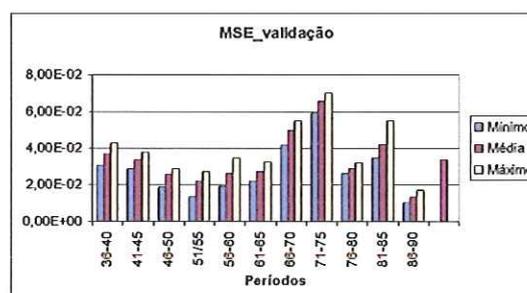
(a)



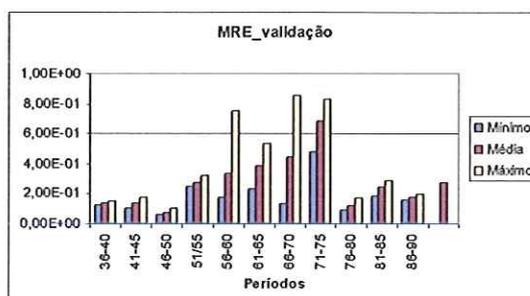
(a)



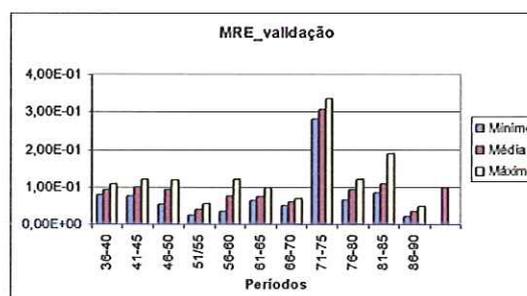
(b)



(b)



(c)

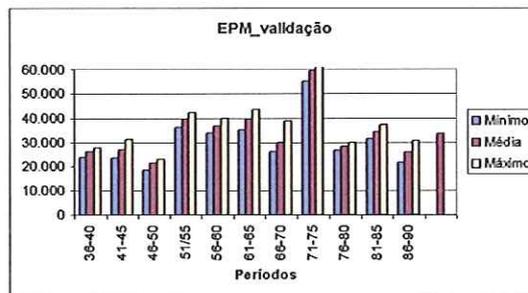


(c)

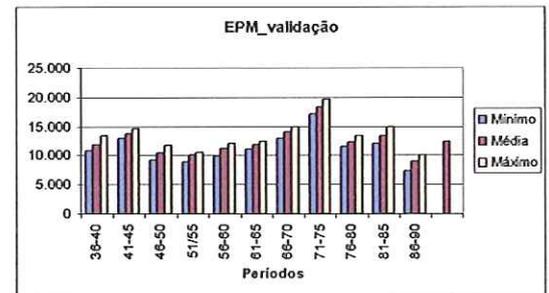
Figura 5.10 – Erro de validação para previsão das Afluências de Furnas, para o modelo 1

Figura 5.11 - Erro de validação para previsão das Vazões Ótimas de Furnas, para o modelo 1

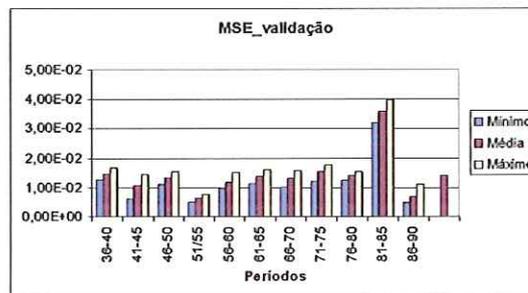
O segundo conjunto de resultados, apresentado pelas Figuras 5.12 e 5.13, tem uma modificação na recuperação do neurônio vencedor, já na fase de testes. Esta modificação se constitui em definir tal neurônio vencedor não mais pela Equação 4.7, mas sim pela distância Euclidiana composta por 3 sub-vetores definidos como os dados referente às usinas, não pertencentes a cascata da usina de interesse, no caso a usina de Furnas; o segundo sub-vetor seria o conjunto de dados referentes às usinas pertencentes a mesma cascata da usina de interesse; e finalmente o terceiro sub-vetor seria o conjunto de dados referentes à própria usina de interesse.



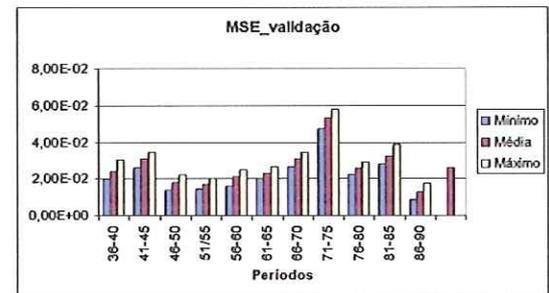
(a)



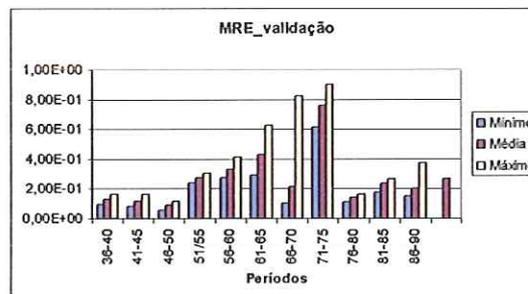
(a)



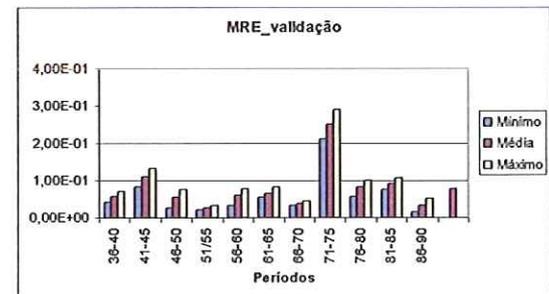
(b)



(b)



(c)



(c)

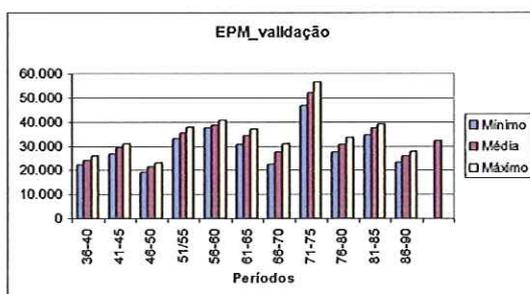
Figura 5.12 – Erro de validação para previsão das Afluências de Furnas, para o modelo 2

Figura 5.13 - Erro de validação para previsão das Volumes Ótimos de Furnas, para o modelo 2

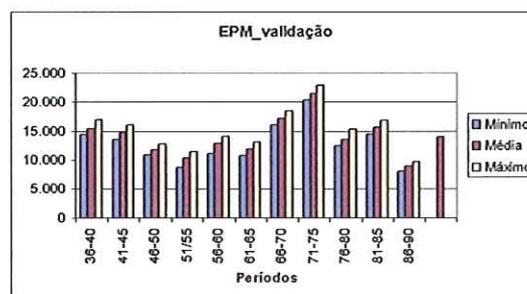
O terceiro conjunto de resultados, apresentado pelas Figuras 5.14 e 5.15, tem uma modificação no treinamento da rede, onde neste caso os centros também poderão ser criados caso a restrição da Equação 5.2 seja atendida, violando uma tolerância individual, devendo-se criar um novo centro. Com relação à recuperação do centro vencedor, aplicou-se a formulação original, através da distância Euclidiana do vetor completo dos dados de entrada.

$$|x_i - w_i| > \begin{cases} \delta_{afluencia} & , i = 1, \dots, 42 \\ \delta_{volume} & , i = 43, \dots, 84 \end{cases} \quad (5.2)$$

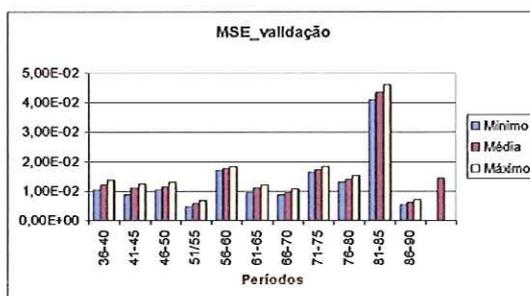
onde:  $\delta_{afluencia}$  e  $\delta_{volume}$  - tolerância para os dados de afluência e volume respectivamente.



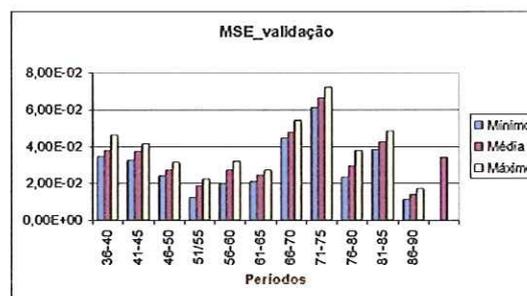
(a)



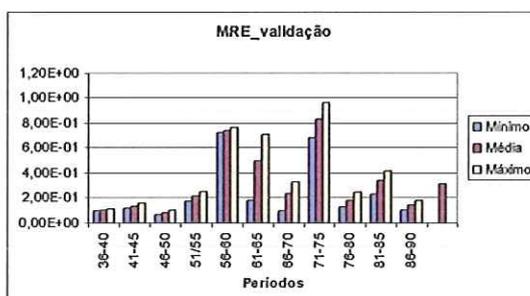
(a)



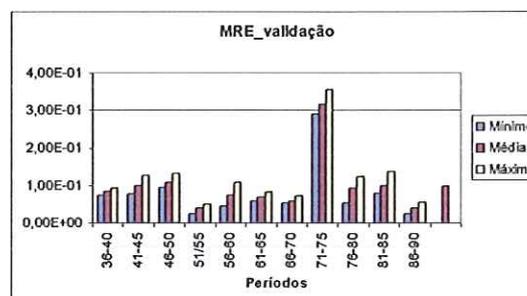
(b)



(b)



(c)

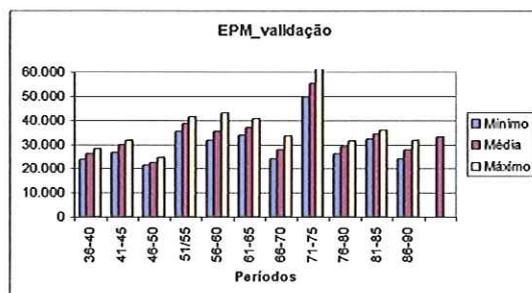


(c)

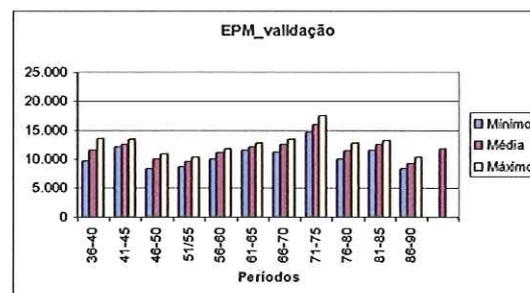
**Figura 5.14 – Erro de validação para previsão das Afluências de Furnas, para o modelo 3**

**Figura 5.15 - Erro de validação para previsão das Volumes Ótimos de Furnas, para o modelo 3**

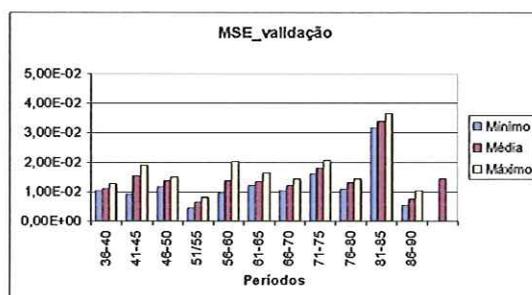
O quarto conjunto de resultados, apresentado pelas Figuras 5.16 e 5.17, ainda apresenta um treinamento com base nas definições propostas pelo caso anterior, porém com uma modificação na fase de teste, tendo em vista a modificação proposta no caso 2, onde se define o neurônio vencedor não mais pela Equação 4.7, mas sim pela distância Euclidiana composta por 3 sub-vetores.



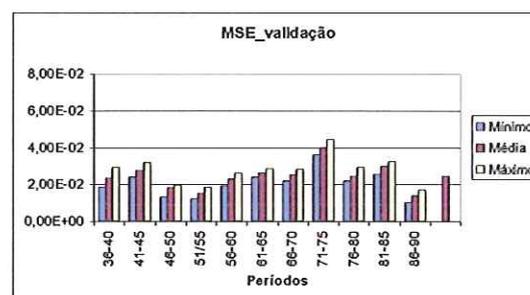
(a)



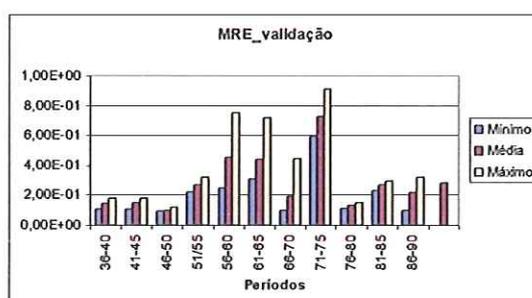
(a)



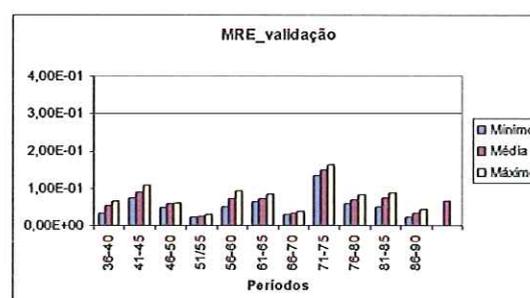
(b)



(b)



(c)



(c)

Figura 5.16 – Erro de validação para previsão das Afluências de Furnas, para o modelo 4

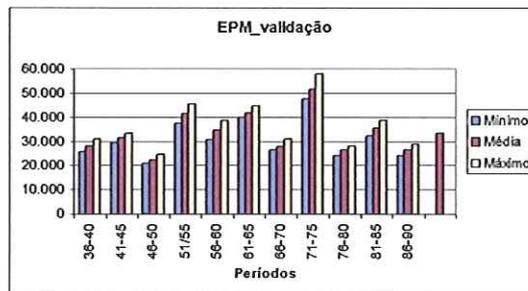
Figura 5.17 - Erro de validação para previsão das Volumes Ótimos de Furnas, para o modelo 4

Os próximos resultados se referem aos modelos 5, 6, 7 e 8, repetindo as proposições feitas nos quatro primeiros, no que se refere às modificações das modelagens propostas. A diferença é que de agora em diante e para todos os futuros modelos apresentados ter-se-á que os dados relevantes para o treinamento não mais contarão apenas com os dados do passado, mas também com os dados do futuro, afim

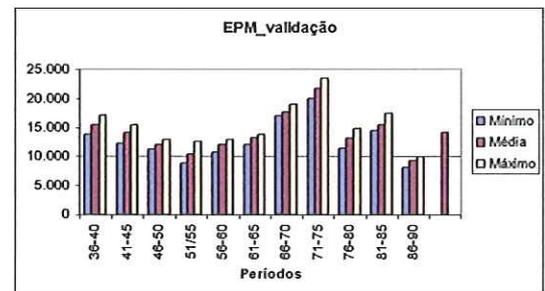
de caracterizar algumas situações não tão bem captadas só com base nos dados do passado.

Isso porque ao se levar em consideração somente os dados do passado, podem ocorrer de duas ou mais trajetórias do passado serem suficientemente parecidas, o que poderia levar a situações distintas no futuro.

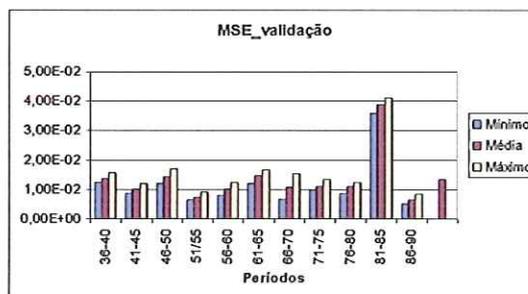
Este tipo de situação, não estaria sendo captada pelo modelo SONARX original, dificultando uma boa previsão. Logo, decidiu-se inserir dentre os dados de treinamento, os dados do futuro conhecido, com a finalidade de fornecer subsídios à rede, na fase de treinamento, para a criação de centros distintos mesmo numa situação de dados do passado razoavelmente semelhantes.



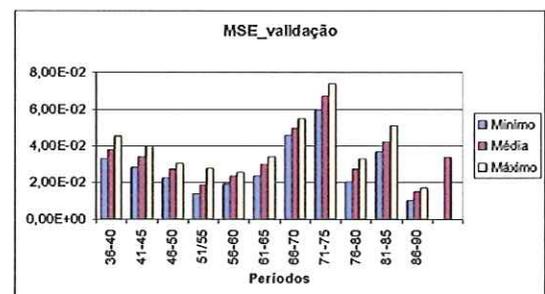
(a)



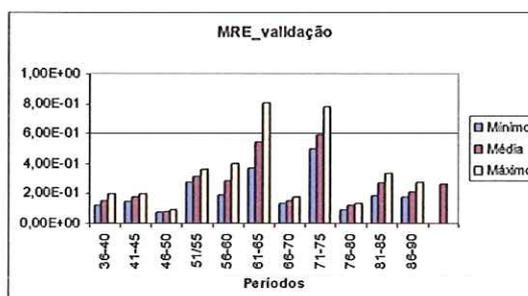
(a)



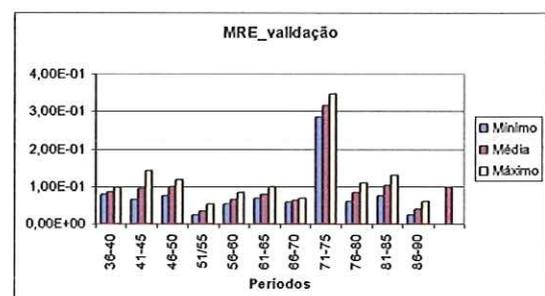
(b)



(b)



(c)



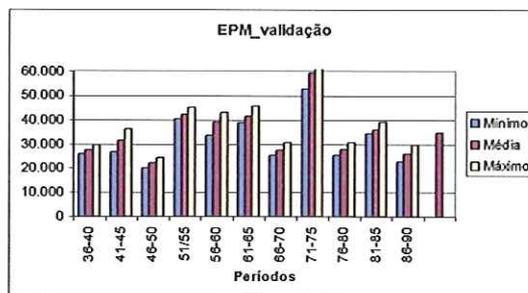
(c)

Figura 5.18 – Erro de validação para previsão das Afluências de Furnas, para o modelo 5

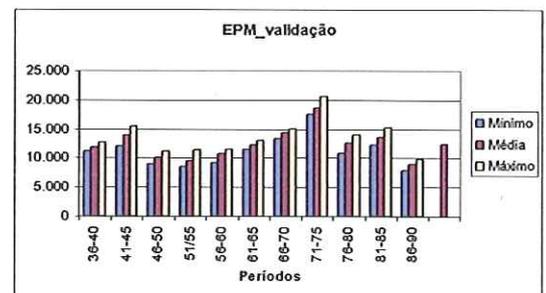
Figura 5.19 - Erro de validação para previsão das Volumes Ótimos de Furnas, para o modelo 5

Os resultados apresentados nas Figuras 5.18 e 5.19, referentes ao modelo 5, obtiveram uma ligeira melhora de desempenho para alguns trechos do histórico, se comparado aos resultados apresentados pelo modelo 1, já que a partir deste modelo levou-se em consideração dados do futuro para a criação dos centros.

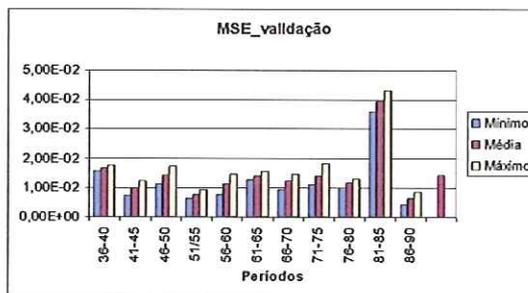
Fenômeno similar aparece nos resultados apresentados pelo modelo 6, que é similar ao modelo 2, com a adição das informações do futuro nos padrões de treinamento. Os erros apresentados por este modelo podem ser vistos nas Figuras 5.20 e 5.21, para a previsão de Vazões Afluentes e Volumes Operativos Ótimos, respectivamente.



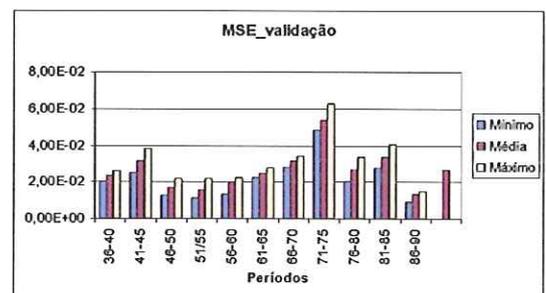
(a)



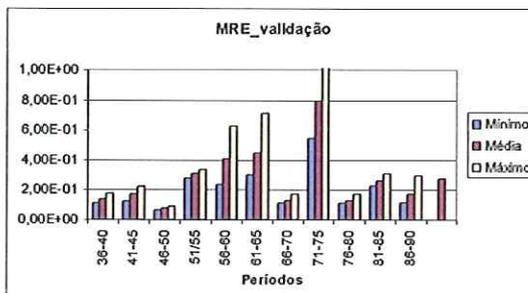
(a)



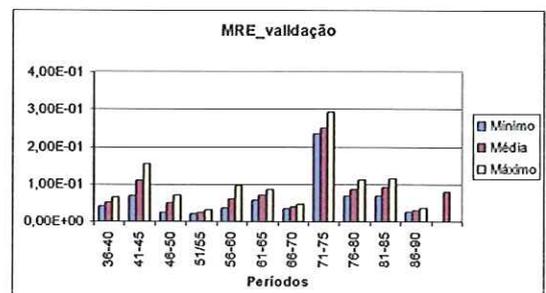
(b)



(b)



(c)

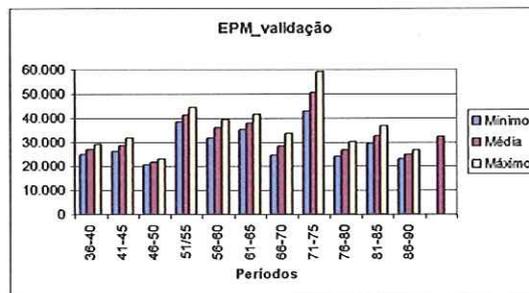


(c)

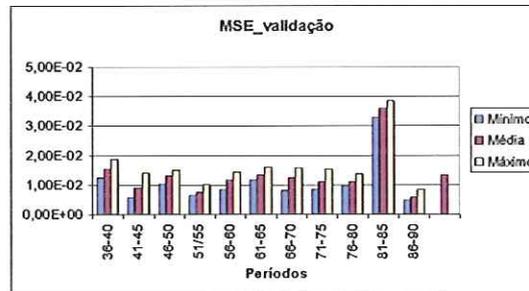
**Figura 5.20 – Erro de validação para previsão das Afluências de Furnas, para o modelo 6**

**Figura 5.21 - Erro de validação para previsão das Volumes Ótimos de Furnas, para o modelo 6**

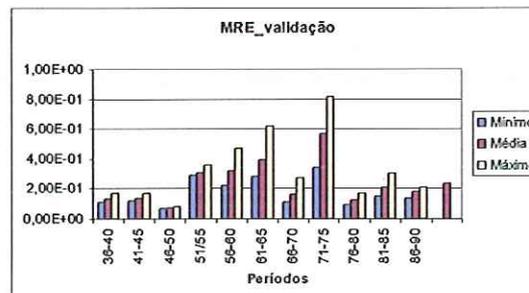
Novamente, no caso do modelo 7, não foi encontrada uma melhora significativa no desempenho de todos os períodos do histórico. Este modelo 7 possui as mesmas características do modelo 3, pois apresenta uma modificação no treinamento da rede, onde neste caso os centros também poderão ser criados caso a restrição da Equação 5.2 seja atendida, violando uma tolerância individual, devendo-se criar um novo centro. E com relação à recuperação do centro vencedor, aplicou-se a formulação original, através da distância Euclidiana do vetor completo dos dados de entrada, com a diferença de levar em consideração os dados futuros conhecidos na criação dos centros. Os erros encontrados por este modelo 7, podem ser observado nas Figuras 5.22 e 5.23.



(a)

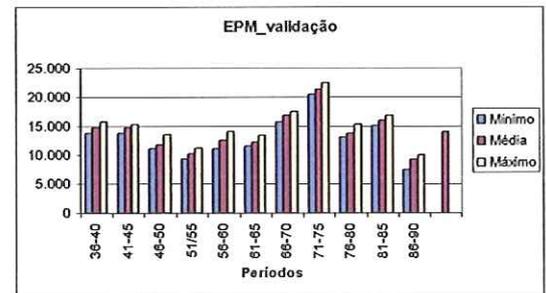


(b)

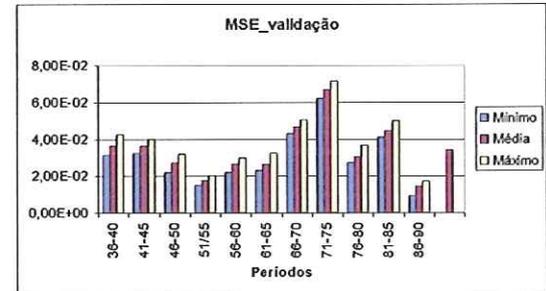


(c)

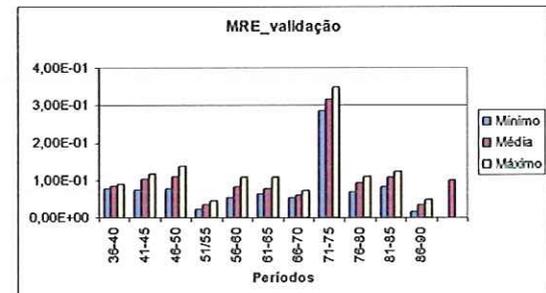
Figura 5.22 – Erro de validação para previsão das Afluências de Furnas, para o modelo 7



(a)



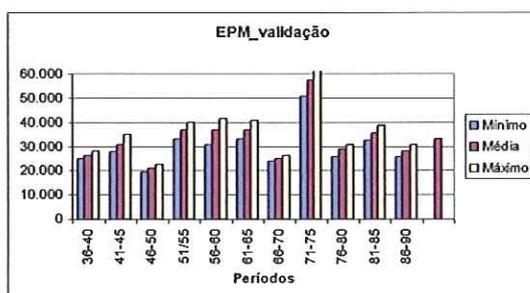
(b)



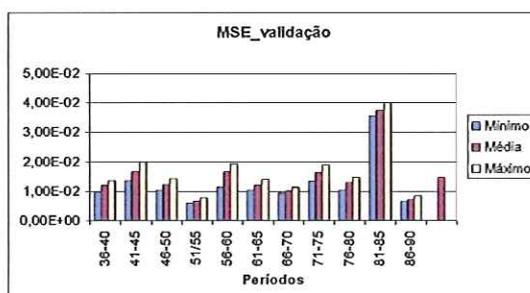
(c)

Figura 5.23 - Erro de validação para previsão das Volumes Ótimos de Furnas, para o modelo 7

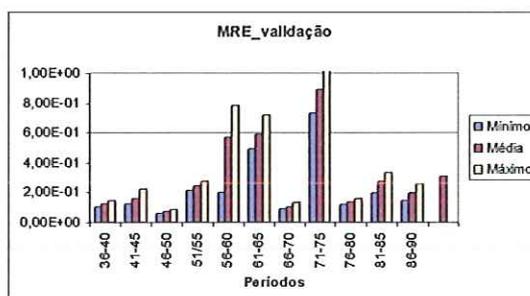
Concluindo a avaliação investigativa decorrente da introdução das informações do futuro conhecido na criação de um novo centro, tem-se os resultados apresentados nas Figuras 5.24 e 5.25, correspondente ao modelo 8, proveniente desta modificação no modelo 4, que ainda apresenta um treinamento com base nas definições propostas pelo caso anterior, porém a modificação na fase de teste, define o neurônio vencedor não mais pela Equação 4.7, mas sim pela distância Euclidiana composta por 3 sub-vetores.



(a)

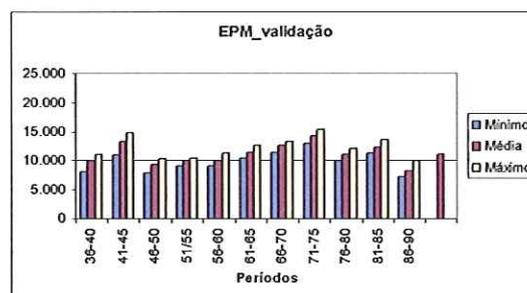


(b)

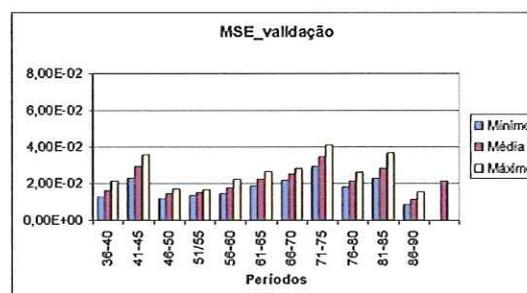


(c)

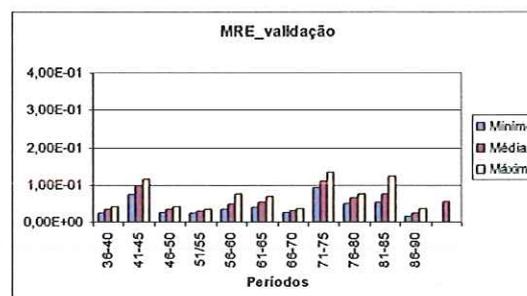
**Figura 5.24 – Erro de validação para previsão das Afluências de Furnas, para o modelo 8**



(a)



(b)



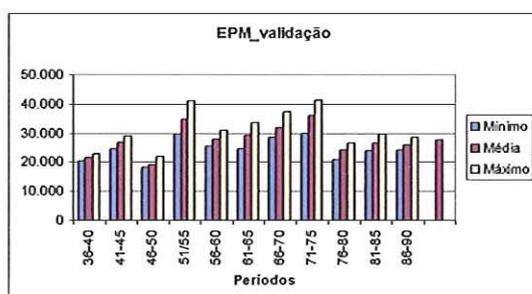
(c)

**Figura 5.25 - Erro de validação para previsão das Volumes Ótimos de Furnas, para o modelo 8**

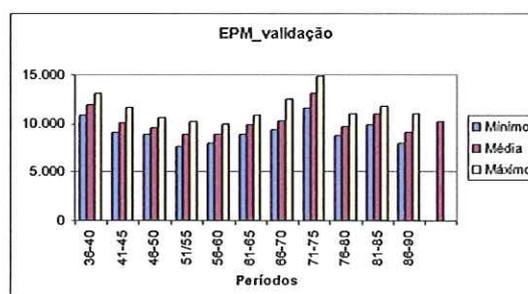
Já o nono, décimo e décimo primeiro conjuntos de resultados, apresentados pelas Figuras 5.26 à 5.31, têm uma modificação no treinamento da rede, onde é feito um deslocamento temporal de 2 meses. Neste deslocamento todos os dados de 2 meses

antes do mês a ser previsto são retirados do treinamento e tomados como desconhecidos pela rede SONARX.

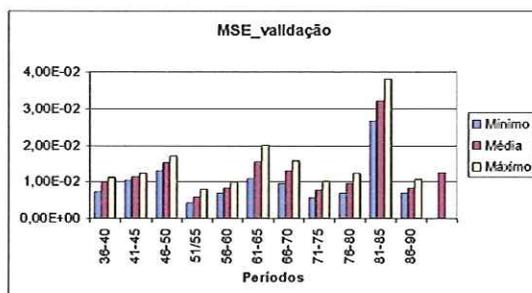
Com isso, num primeiro momento a rede deverá aprender a prever  $t-2$ ,  $t-1$  e  $t$ . Posteriormente, já na fase de teste, a rede passa a fornecer um conjunto de neurônios candidatos a vencedores, com base no passado. Em seguida, as previsões de  $t-2$  e  $t-1$  são comparadas em conjunto para determinar o neurônio vencedor dentre estes candidatos, já que os dados de  $t-2$  e  $t-1$  são de fato conhecidos.



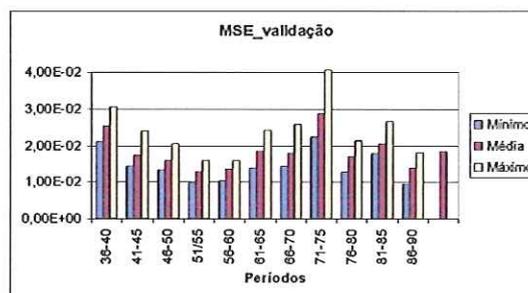
(a)



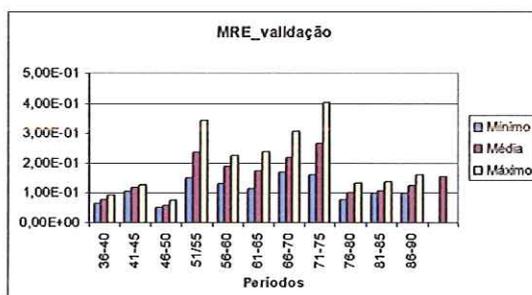
(a)



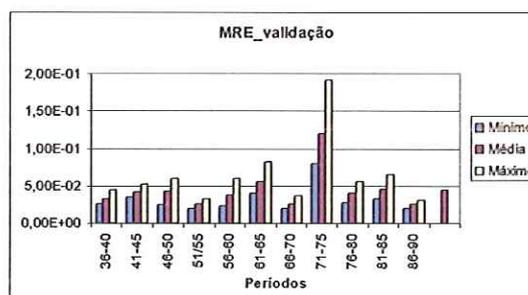
(b)



(b)



(c)



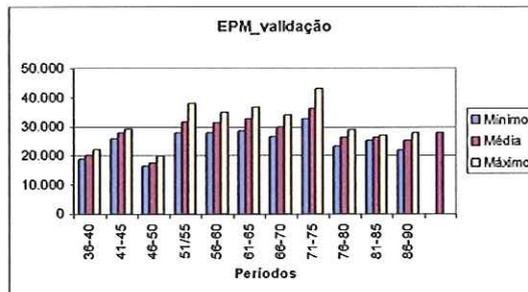
(c)

Figura 5.26 – Erro de validação para previsão das Afluências de Furnas, para o modelo 9

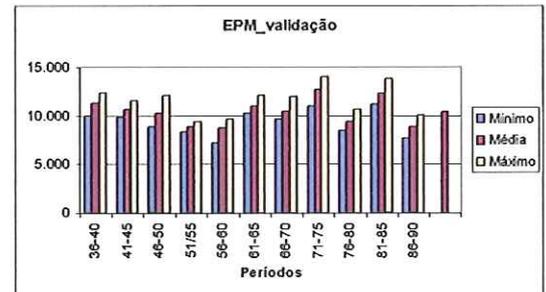
Figura 5.27 - Erro de validação para previsão das Volumes Ótimos de Furnas, para o modelo 9

O décimo conjunto de resultados, apresentado pela Figuras 5.28 e 5.29, tem uma modificação no treinamento da rede, onde neste caso os centros também poderão ser criados caso a restrição da Equação 5.2 seja atendida, violando uma tolerância individual, devendo-se criar um novo centro. E com relação a recuperação do centro

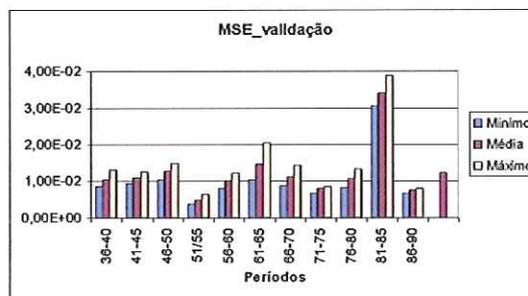
vencedor, aplicou-se a formulação original, através da distância Euclidiana do vetor completo dos dados de entrada.



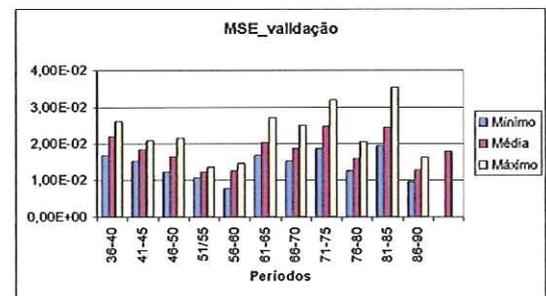
(a)



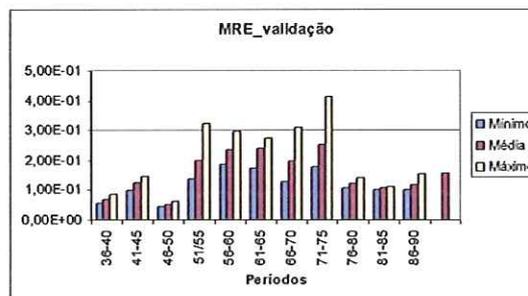
(a)



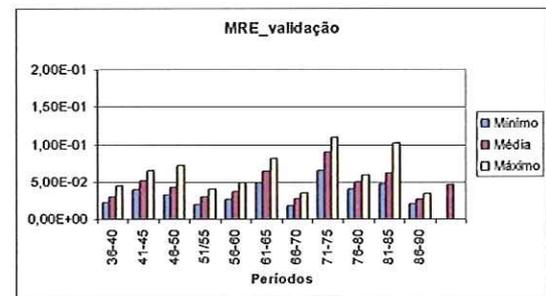
(b)



(b)



(c)



(c)

**Figura 5.28 – Erro de validação para previsão das Afluências de Furnas, para o modelo 10**

**Figura 5.29 - Erro de validação para previsão das Volumes Ótimos de Furnas, para o modelo 10**

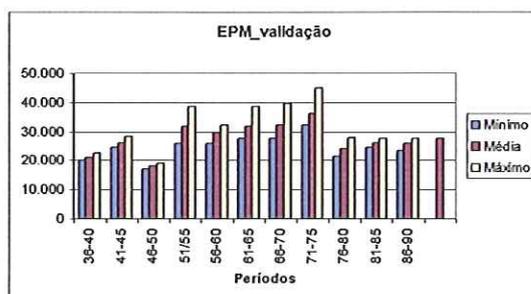
E finalmente temos o décimo primeiro conjunto de resultados apresentados pelas Figuras 5.30 e 5.31. Neste caso, os centros poderiam ser criados na fase de treinamento com base nos dados do passado, conforme a proposição original da rede, e também pela violação de uma tolerância entre os dados futuros reais e os dados definidos pelo neurônio vencedor.

De maneira geral, dentre as variações propostas sobre o modelo neural SONARX, os resultados se mostraram satisfatórios tanto para a previsão das vazões afluentes quanto para a previsão dos volumes operativos ótimos. Entretanto, estes

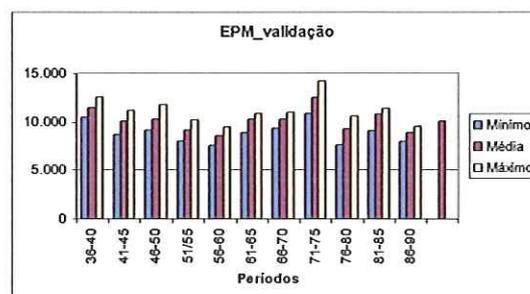
modelos encontraram dificuldades de acerto em períodos do histórico, pois possuem padrões tipicamente distintos dos demais períodos.

Analisando a previsão das vazões afluentes para o período de janeiro de 1981-1985, por exemplo, que é caracterizado como um período cheio, tendo como parâmetro de comparação o erro MSE dado por (b), a rede SONARX não teve um bom desempenho, isto porque o modelo não conseguiu uma boa extrapolação (ou generalização) para este período.

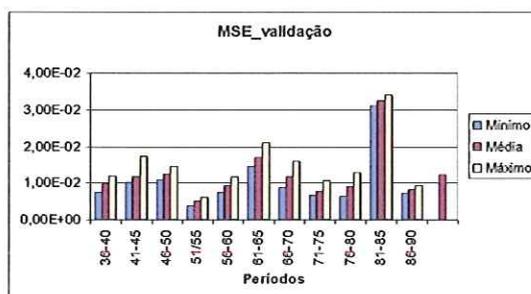
Com relação à previsão dos volumes operativos ótimos, observou-se um desempenho mais homogêneo entre os períodos testados. Isto porque a rede SONARX conseguiu extrapolar de forma satisfatória os vários padrões distintos de comportamento, evidenciando a importância de todos os dados de entrada.



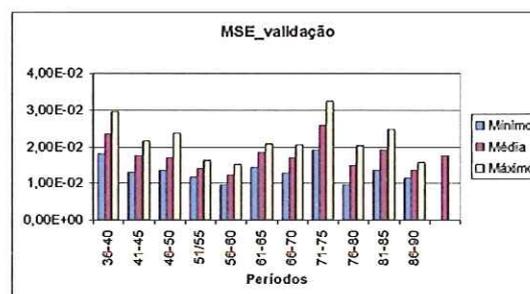
(a)



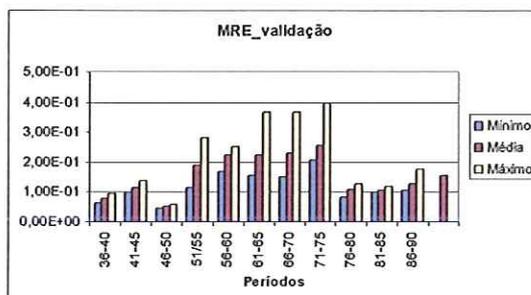
(a)



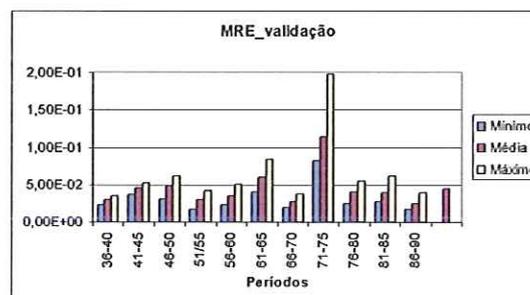
(b)



(b)



(c)



(c)

**Figura 5.30 – Erro de validação para previsão das Afluências de Furnas, para o modelo 11**

**Figura 5.31 - Erro de validação para previsão das Volumes Ótimos de Furnas, para o modelo 11**

Estes resultados apresentados anteriormente, podem ser melhor analisados através das Tabelas 1 e 2, mostrando de forma comparativa a tendência de melhora no desempenho da rede SONARX na previsão tanto dos dados de vazões afluentes, como dos volumes operativos ótimos.

**Tabela 1 - Tabela dos Erros de previsão das vazões afluentes para a usina de Furnas**

	EPM (%)		MSE ( $10^{-3}$ )		MRE ( $10^{-1}$ )	
	treino	teste	treino	teste	treino	teste
modelo 1	27,65	33,25	8,15	13,80	1,79	2,74
modelo 2	27,93	33,47	8,39	14,10	1,88	2,66
modelo 3	12,24	32,34	1,80	14,30	0,44	3,16
modelo 4	17,73	33,07	3,67	14,20	0,76	2,80
modelo 5	30,68	33,51	10,20	13,40	2,12	2,62
modelo 6	31,18	34,52	10,80	14,20	2,18	2,75
modelo 7	28,45	32,27	8,58	13,20	1,84	2,37
modelo 8	15,01	32,99	2,55	14,40	0,53	3,06
modelo 9	30,57	27,56	10,40	12,40	1,52	2,22
modelo 10	27,96	27,70	8,76	12,20	1,86	1,56
modelo 11	29,31	27,54	9,56	12,20	2,11	1,54

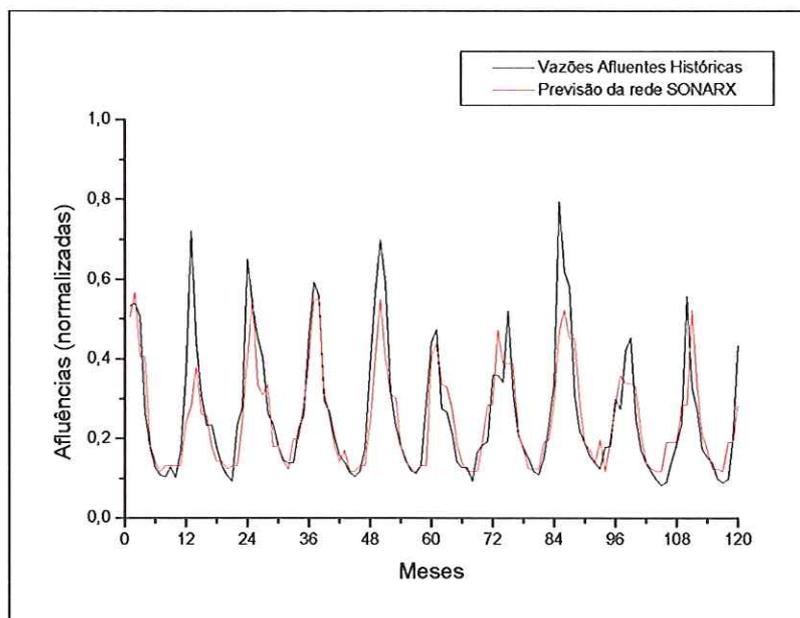
No caso da previsão das vazões afluentes, as melhoras de desempenho comparando-se o modelo 1 (formulação original) e o modelo 11 (modificado) foram de 17,2 % para o EPM, de 11,6 % para o MSE, e de 43,8 % para o MRE.

**Tabela 2 - Tabela dos Erros de previsão dos volumes operativos ótimos para a usina de Furnas**

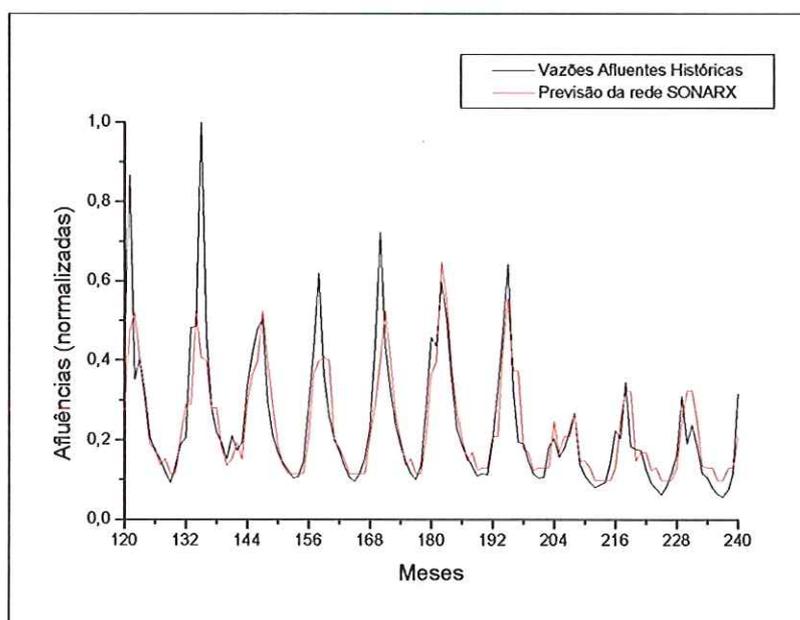
	EPM (%)		MSE ( $10^{-2}$ )		MRE ( $10^{-2}$ )	
	treino	teste	treino	teste	treino	teste
modelo 1	8,80	14,06	1,46	3,34	3,25	9,69
modelo 2	8,52	12,39	1,37	2,61	3,07	7,96
modelo 3	7,70	13,98	1,17	3,38	2,51	9,78
modelo 4	5,13	11,66	0,65	2,41	1,37	6,52
modelo 5	7,79	14,04	1,15	3,35	2,47	9,63
modelo 6	7,79	12,43	1,15	2,62	2,49	7,89
modelo 7	7,25	13,95	1,01	3,38	2,15	9,95
modelo 8	6,05	11,07	0,78	2,13	1,60	5,42
modelo 9	6,66	10,20	0,87	1,82	1,94	4,51
modelo 10	9,03	10,40	1,40	1,79	3,22	4,64
modelo 11	7,50	10,11	0,99	1,74	2,23	4,49

Já na previsão dos volumes operativos ótimos, as melhoras de desempenho foram de 28,1 % para o EPM, de 47,9 % para o MSE, e de 53,7 % para o MRE.

Finalmente tem-se as Figuras 5.32 até 5.37 mostrando as trajetórias de Vazões Afluentes à usina de Furnas registradas para o período do histórico de janeiro de 1936 até dezembro de 1990, em preto, que aqui foram divididos em 6 gráficos de 10 anos cada.



**Figura 5.32 - Previsão das aflúncias para a usina de Furnas, para o período de Janeiro de 1936 até Dezembro de 1945**

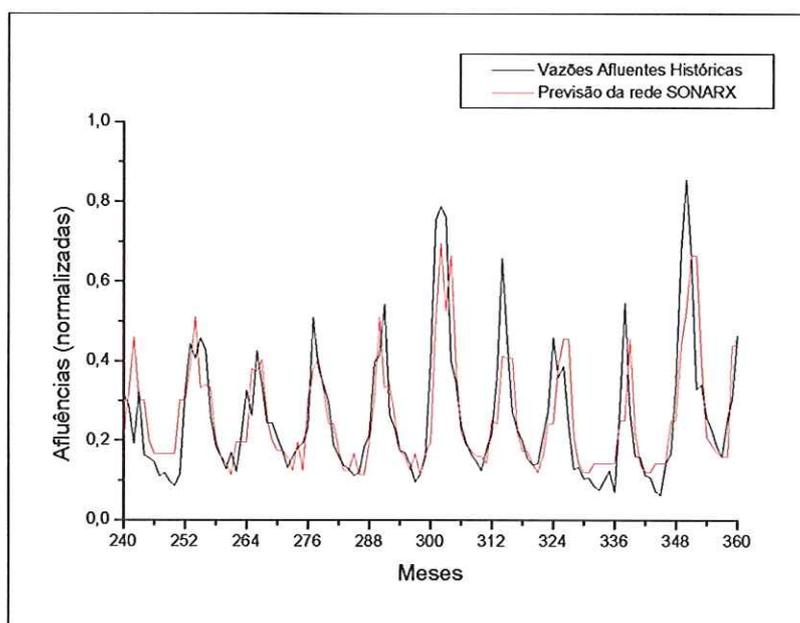


**Figura 5.33 - Previsão das aflúncias para a usina de Furnas, para o período de Janeiro de 1946 até Dezembro de 1955**

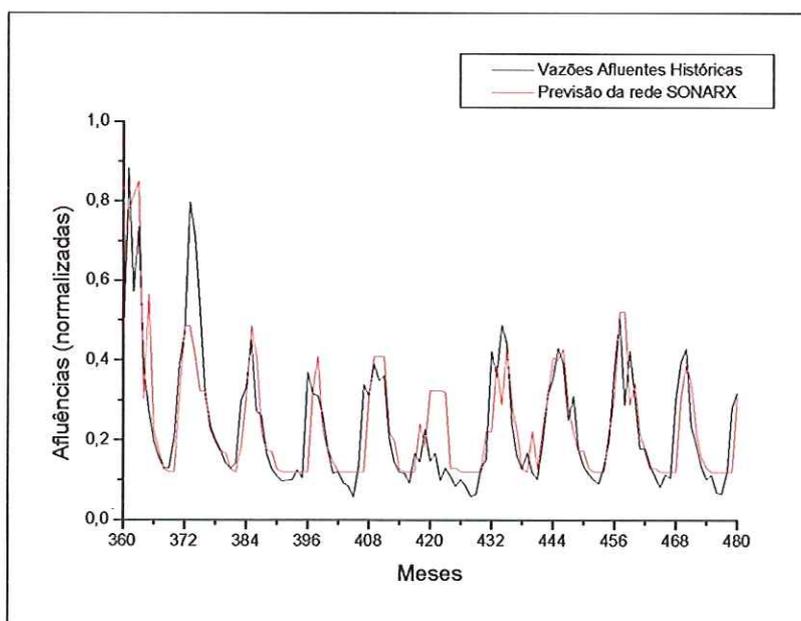
Juntamente com estas vazões afluentes históricas estão também as vazões afluentes previstas pela rede SONARX, em vermelho, sendo que estas previsões são sempre de um passo a frente, ou seja, do mês seguinte ao mês atual.

Como pode ser observado nas Figuras 5.32 até 5.37, esta rede neural teve um excelente desempenho na previsão dos meses de seca, para quase que todos os anos

avaliados; sendo o período de estiagem definido de maio a novembro de cada ano. Isto ocorre porque há um menor desvio padrão nas medições das vazões afluentes deste período, se comparado com o período chuvoso, definido de dezembro de um ano até abril do ano seguinte.

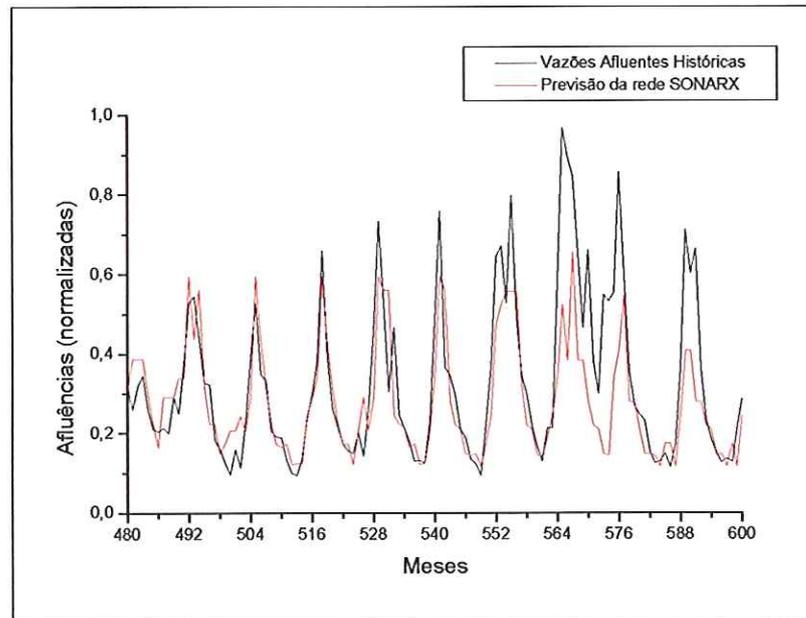


**Figura 5.34 - Previsão das aflúncias para a usina de Furnas, para o período de Janeiro de 1956 até Dezembro de 1965**

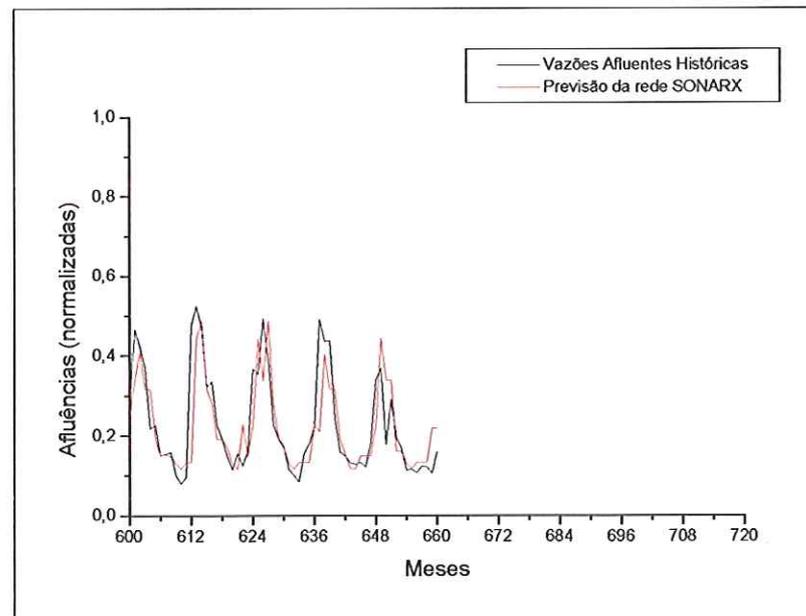


**Figura 5.35 - Previsão das aflúncias para a usina de Furnas, para o período de Janeiro de 1966 até Dezembro de 1975**

Neste sentido, pode se concluir que para os picos de vazão afluente a rede SONARX conseguiu assimilar e prever com uma boa aproximação boa parte deles, mas para alguns destes picos, a rede não apresentou um bom desempenho.

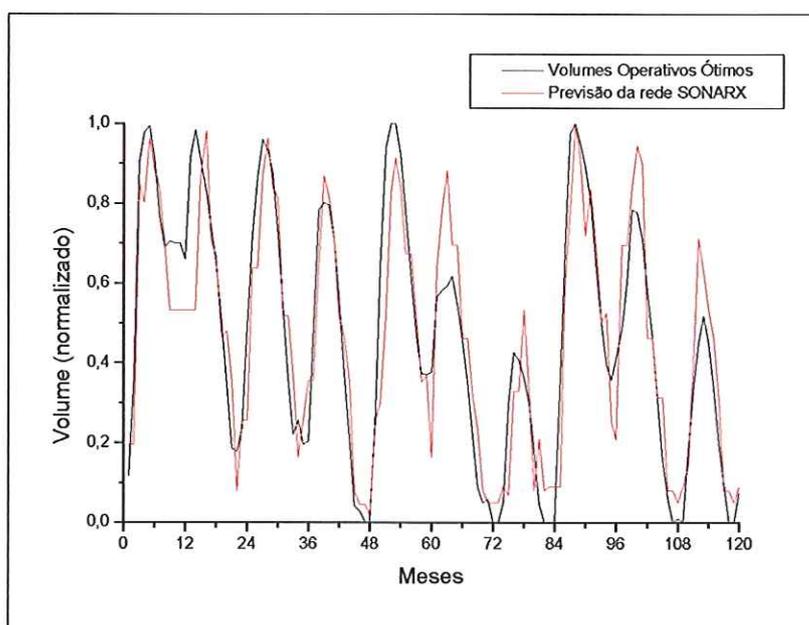


**Figura 5.36 - Previsão das aflúncias para a usina de Furnas, para o período de Janeiro de 1976 até Dezembro de 1985**

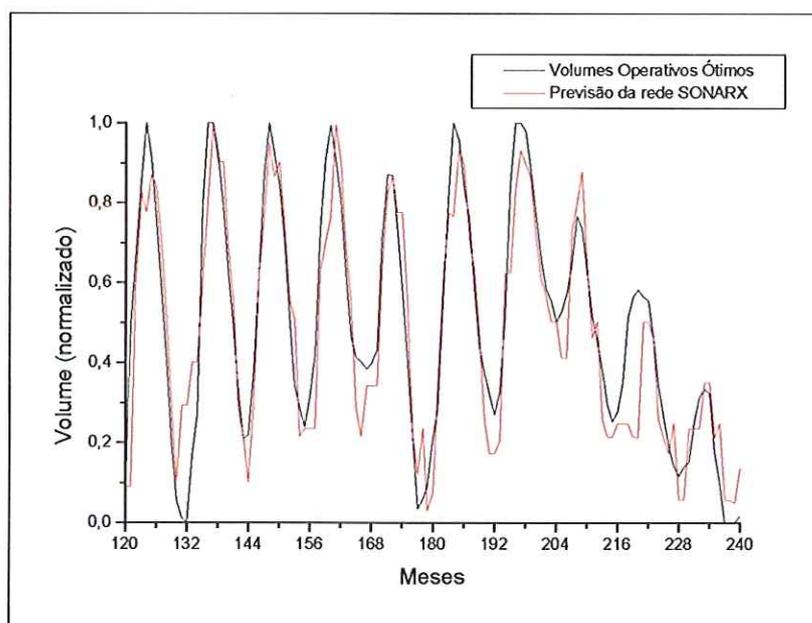


**Figura 5.37 - Previsão das aflúncias para a usina de Furnas, para o período de Janeiro de 1986 até Dezembro de 1990**

As Figuras 5.38 até 5.43 mostram as trajetórias dos Volumes Operativos Ótimos da usina de Furnas determinados pelo otimizador não-linear determinístico para o período do histórico de janeiro de 1936 até dezembro de 1990, em preto, e que novamente foram divididos em 6 gráficos de 10 anos cada.



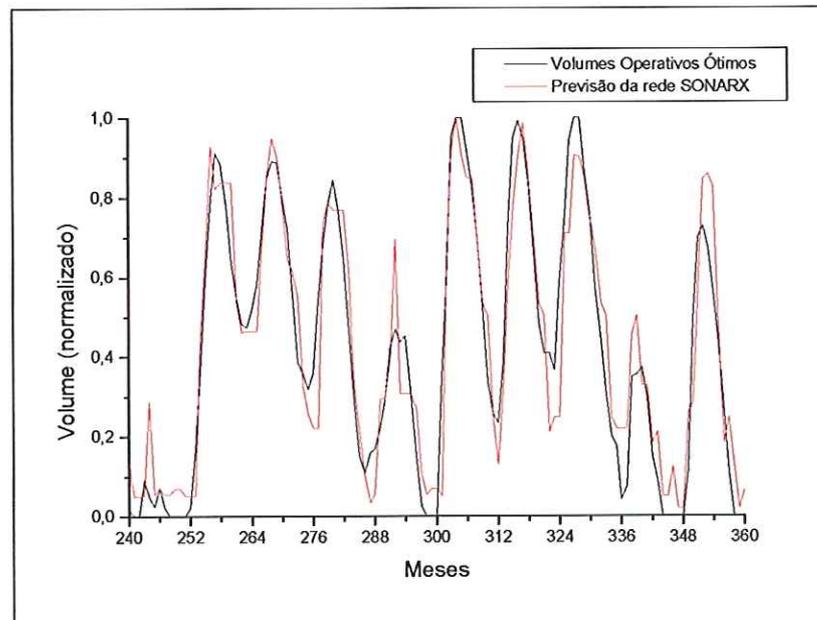
**Figura 5.38 - Previsão dos volumes operativos ótimos para a usina de Furnas, para o período de Janeiro de 1936 até Dezembro de 1945**



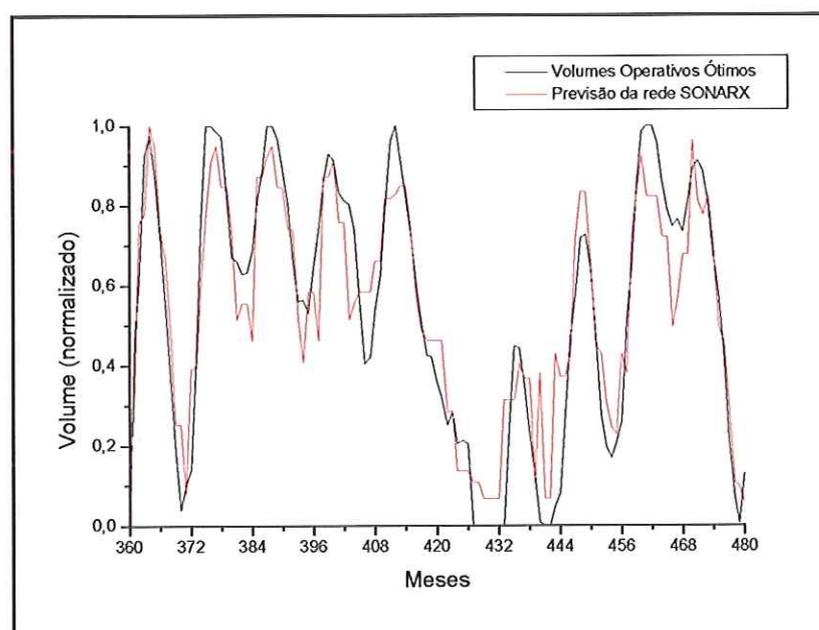
**Figura 5.39 - Previsão dos volumes operativos ótimos para a usina de Furnas, para o período de Janeiro de 1946 até Dezembro de 1955**

Juntamente com estes volumes operativos ótimos estão também os volumes previstos pela rede SONARX, em vermelho, sendo que estas previsões também são sempre de um passo a frente, ou seja, do mês seguinte ao mês atual.

Como pode ser observado, a trajetória das previsões destes volumes foi bastante aderente à trajetória ótima operativa em todos os meses analisados. Inclusive se tomado

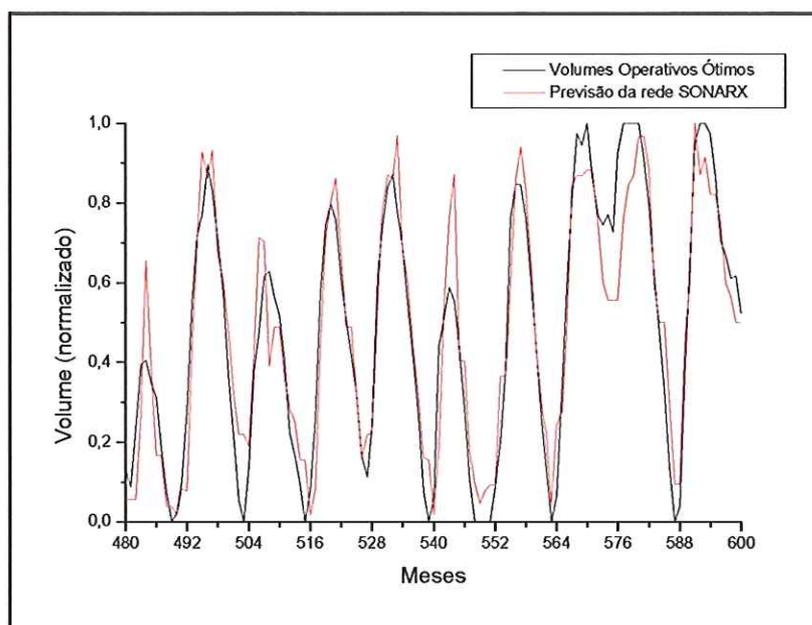


**Figura 5.40 - Previsão dos volumes operativos ótimos para a usina de Furnas, para o período de Janeiro de 1956 até Dezembro de 1965**

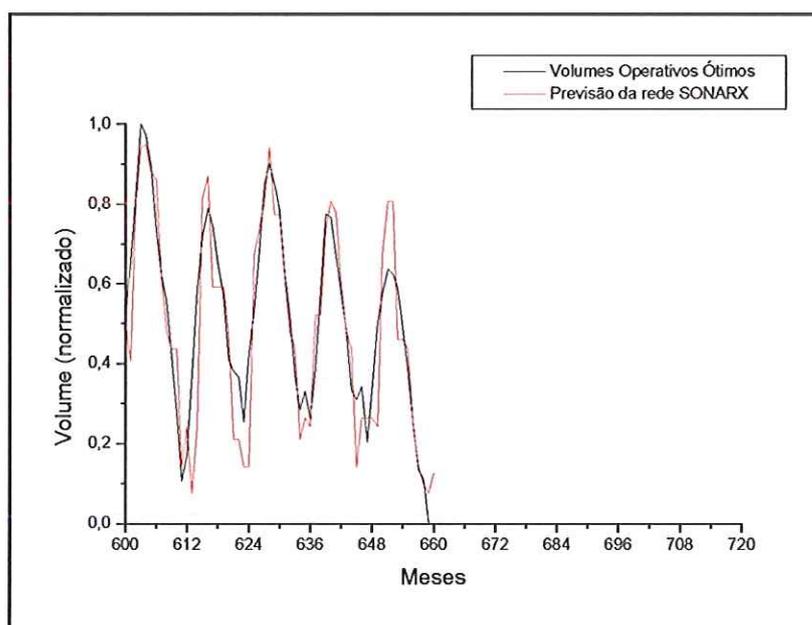


**Figura 5.41 - Previsão dos volumes operativos ótimos para a usina de Furnas, para o período de Janeiro de 1966 até Dezembro de 1975**

o período de janeiro de 1952 até dezembro de 1956 (meses de 193 até 252), por exemplo, a rede foi capaz de prever que viria alguns anos seco pela frente, seguido de uma estação razoavelmente chuvosa, o que ocasionou num total deplecionamento deste reservatório neste período.



**Figura 5.42 - Previsão dos volumes operativos ótimos para a usina de Furnas, para o período de Janeiro de 1976 até Dezembro de 1985**



**Figura 5.43 - Previsão dos volumes operativos ótimos para a usina de Furnas, para o período de Janeiro de 1986 até Dezembro de 1990**

## **5.2. Operação Simulada pelo Decisor Neural: SONARX**

O Planejamento da Operação de Sistemas Hidroelétricos tem como principal objetivo coordenar a operação das usinas pertencentes a um dado parque hidroelétrico. Esta coordenação da operação define quanto cada usina deve gerar a cada instante, a fim de que o sistema consiga atender à carga que lhe é solicitada, a um mínimo custo, segundo alguns critérios de confiabilidade.

Esta tomada de decisão não é uma tarefa fácil, principalmente porque estes sistemas normalmente são altamente interligados hidráulicamente. Dessa maneira, as características individuais de cada usina hidroelétrica e suas posições relativas dentro da cascata exigem a adoção de Regras de Operação individuais por parte dos operadores.

Agregada a estas dificuldades, a operação de um sistema hidroelétrico é um problema dinâmico, uma vez que decisões tomadas no presente implicam diretamente em futuras condições de armazenamento de todo o sistema.

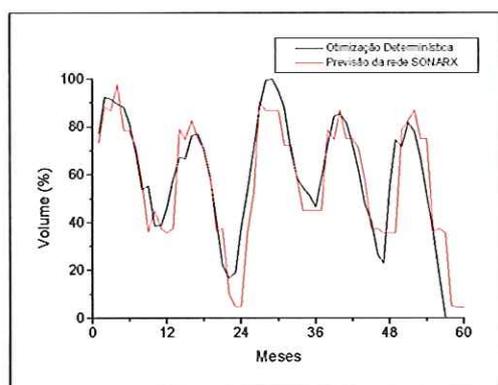
Como dito anteriormente, a rede SONARX foi aplicada justamente na extração das relações existentes entre as usinas, que caracterizam o comportamento ótimo operativo, para ser utilizada em simulações da operação de um subsistema teste real.

Depois de treinada, a rede SONARX foi integrada a um simulador da operação do sistema hidroelétrico, onde a cada instante de tempo é necessário que ela decida o estado de armazenamento de cada reservatório, e conseqüentemente quanto cada usina deve gerar a cada mês subsequente ao mês atual.

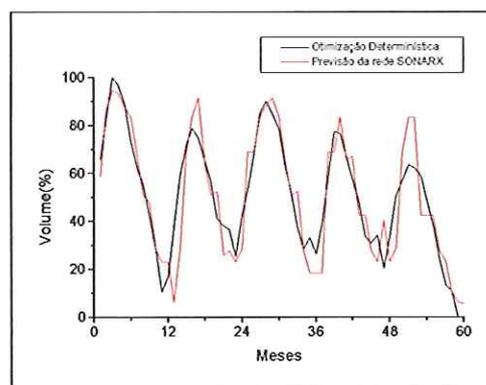
As simulações não são determinísticas e, a cada intervalo de simulação, as decisões devem ser tomadas sem nenhum conhecimento sobre o futuro. Como a rede foi treinada com dados provenientes de operações ótimas, observou-se que suas decisões produziram comportamentos similares à operação ótima.

As previsões propriamente ditas são apresentados na Figura 5.44. Sendo que as usinas estão relacionadas de (a) a (g), e de acordo com suas posições relativas na cascata do sistema teste, tem-se um sistema com usinas em série e em paralelo, como já foi dito anteriormente.

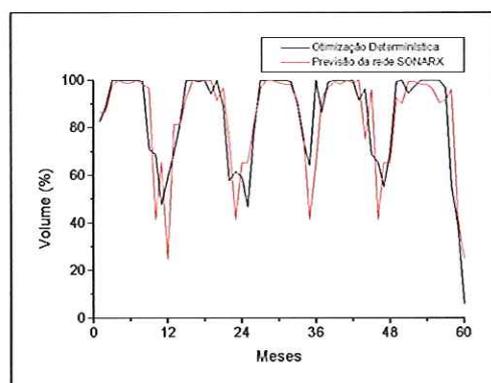
Como pôde ser observado nesta figura, a rede SONARX foi capaz de acompanhar a trajetória otimizada do volume operativo das 7 usinas hidroelétricas deste sistema teste. Mesmo para as usinas de montante, que tendem a variar muito o volume dos seus reservatórios, afim de regularizar as vazões afluentes às usinas de jusante,



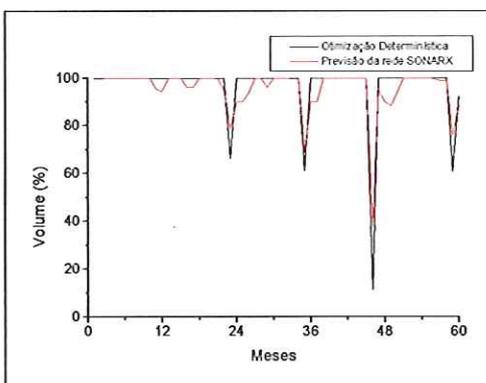
(a) Emborcação



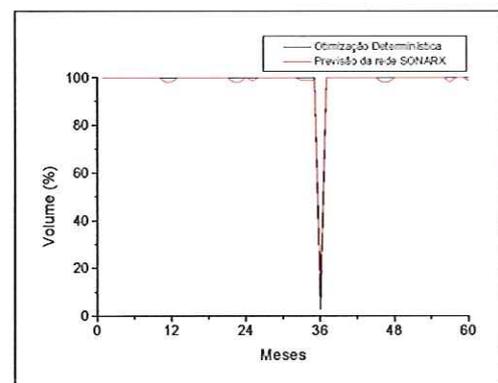
(d) Furnas



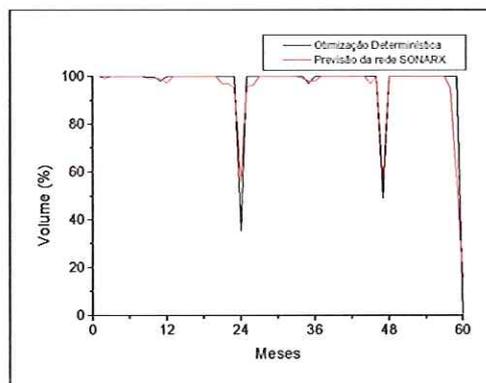
(b) Itumbiara



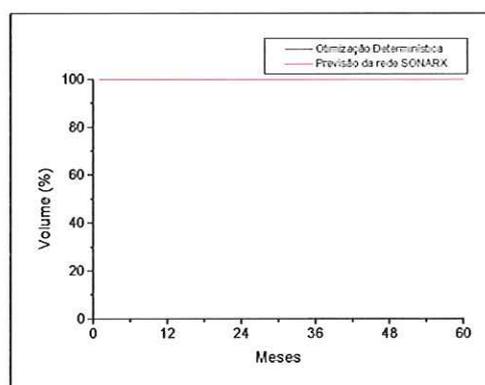
(e) Marimbondo



(c) São Simão



(f) Água Vermelha



(g) Ilha Solteira

Figura 5.44 – Gráfico das previsões dos volumes ótimos do sistema teste, de janeiro de 1986 até dezembro de 1990

como são os casos de Emborcação, Itumbiara e Furnas. Assim, as usinas de jusante permanecem cheias, com alta produtividade para todo o sistema.

A Tabela 3 apresenta os Erros Percentuais Médios (EPM), os Erros Médios Quadráticos (MSE, em inglês), e os Erros Relativos Médios (MRE, em inglês) definidos pelas equações 4.47, 4.48 e 4.49, respectivamente.

Tabela 3 – Erros de Previsão dos Volumes Operativos Ótimos

Usina Hidroelétrica	EPM (%)	MSE ( $10^{-2}$ )	MRE ( $10^{-2}$ )
Emborcação	9,64	1,63	4,20
Itumbiara	8,86	1,86	3,07
São Simão	0,73	0,10	0,24
Furnas	9,06	1,35	2,53
Marimbondo	5,13	1,76	1,52
Água Vermelha	4,18	1,68	1,19
Ilha Solteira	0,00	0,00	0,00

As trajetórias de Energia Armazenada Total do Sistema (E.A.T. do Sist.) produzidas pelo decisor neural junto ao simulador, em vermelho, são comparadas às produzidas pelo programa otimizador, em preto, e pela operação segundo a regra de operação em paralelo, em azul. Quanto mais próxima a trajetória de energia da operação pela rede SONARX estiver da trajetória de energia determinada pelo programa otimizador, mais eficiente será a operação neural, como pode ser visto na Figura 5.45.

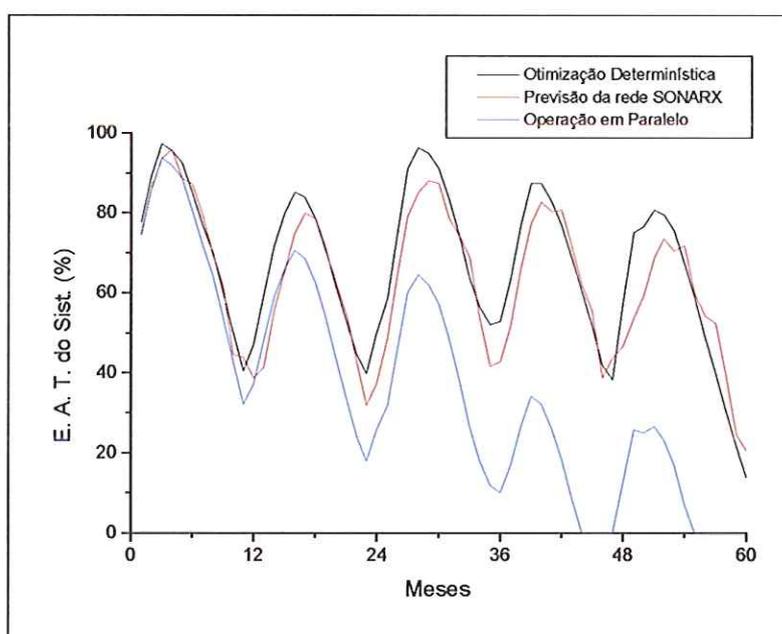


Figura 5.45 – Trajetória de Energia Armazenada Total do Sistema para o período de janeiro de 1986 até dezembro de 1990

Pelos resultados obtidos constata-se que a coordenação das usinas hidroelétricas através da rede SONARX foi bastante eficiente, no sentido de conseguir chegar ao final do ano hidrológico (final de abril, para o sistema teste) com energias armazenadas elevadas.

Além disso, a trajetória de Energia Armazenada produzida pelo decisor neural, ao longo do horizonte de operação, mostrou uma tendência em seguir a trajetória ótima determinística, considerada como referencial ideal de operação.

### 5.3. Testes Investigativos para definir “quem” gera os centros

Estes casos teste se referem ao questionamento de qual(is) variável(is) seria(m) responsável(is) pela criação dos centros, tanto através do treinamento baseado somente no vetor completo (composto somente pelo passado e a variável de interesse do futuro) para a criação dos centros, analisando dimensão por dimensão (variável por variável de entrada) para criação de um novo centro.

A contagem descrita a seguir, se refere ao somatório das avaliações de treinamento da rede SONARX, para a predição de vazões afluentes e/ou da trajetória do volume ótimo da Usina de Furnas. Isto feito através da divisão das séries históricas em 11 períodos consecutivos de 5 anos, desde 1936 até 1990, onde num primeiro momento o primeiro período, correspondente à 1936-1940, foi avaliado como desconhecido, tendo como padrões conhecidos os 10 períodos restantes, os quais foram usados como padrões de treinamento da rede. Em seguida, tomou-se como desconhecido o segundo período, correspondente à 1941-1945, com todos os outros 10 períodos restantes conhecidos, e assim sucessivamente.

Primeiramente serão apresentados os gráficos referentes ao treinamento pelo vetor completo. Neste caso, serão computadas quantas vezes cada entrada contribui de forma significativa para que a distância entre o vetor de entrada  $x_{in}$  e o vetor de pesos  $w_{in}$  seja grande o suficiente para a criação de um novo centro ( $\overrightarrow{dist} = |x_{in} - w_{in}| > \varepsilon$ ), ou seja, serão computadas as contribuições indiretas de cada entrada para a criação de um novo centro. Este computo é feito para todas as variáveis de entrada  $i$  que ultrapassarem o seguinte limiar de tolerância:

$$|x_i - w_i| > cte, \quad \text{onde } cte = 0.1$$

Desta maneira temos a Figura 5.45 ilustrando o computo de todas as vezes que cada variável de entrada  $i$  contribuiu significativamente para a criação do centro. Haja visto que temos neste caso 84 variáveis de entrada, representadas no eixo das abscissas, onde as 42 primeiras referem-se aos dados de entrada de vazões afluentes do passado e as 42 restantes aos dados de entrada de volumes operativos otimizados no passado, enquanto que no eixo das ordenadas aparecem as contribuições.

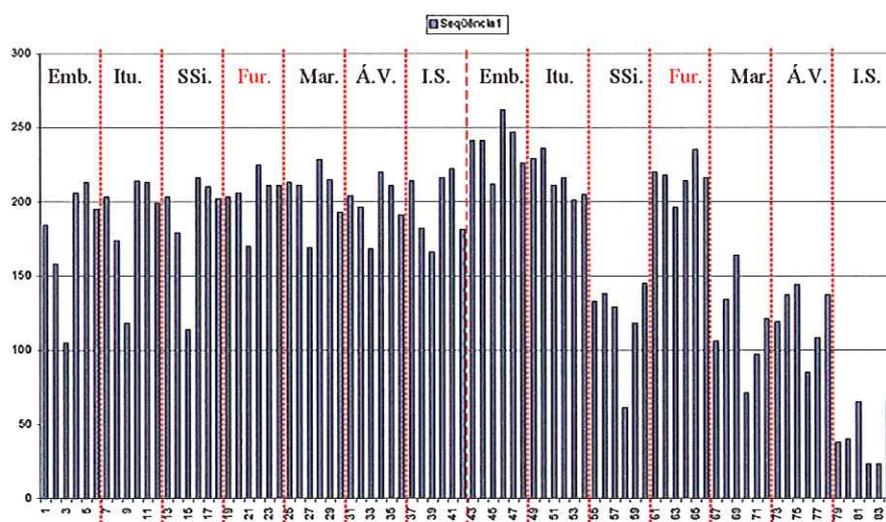


Figura 5.46 - Número de vezes que cada dado de entrada contribuiu para a criação dos centros, durante o treinamento

Fazendo uma análise de forma global, ou seja, observando as características entre as usinas, das entradas relativas à grandeza volume operativo, nota-se que as usinas de montante Emborcação, Itumbiara e Furnas, representadas pelas entradas 43-48, 49-54 e 61-66, contribuem de forma significativa na composição do vetor distância  $\overrightarrow{dist}$ . Sendo que as usinas a jusante São Simão, Marimbondo, Água Vermelha e Ilha Solteira, representadas pelas entradas 55-60, 67-72, 73-78 e 79-84, têm sua contribuição diminuída gradativamente ao longo da cascata. Este fenômeno se dá por causa da operação otimizada que determina que as usinas de montante regularizem a cascata, variando muito seus reservatórios, enquanto que as usinas mais a jusante devem conservar suas alturas de queda, deixando seus reservatórios “flats”, na maior parte do tempo, garantindo assim uma maior produtividade para a cascata. Desta maneira os reservatórios que oscilam mais, tendem a contribuir mais significativamente com diferentes padrões de entrada, enquanto que os de jusante contribuem também de forma importante, porém em menor número.

Diferentemente do volume, para a afluência praticamente todas as usinas contribuem, de forma até semelhantes, só havendo diferenças significativas nos dados temporais de uma mesma usina.

Fazendo uma análise mais detalhada dos dados de entrada referentes à grandeza afluência, nota-se que no primeiro ano do passado, os dados mais próximos do presente contribuem mais significativamente. Basta observar as entradas 1, 2 e 3 por exemplo, que representam as entradas de afluência da usina de Emborcação, deslocadas no tempo

de  $t-1$ ,  $t-2$  e  $t-3$  respectivamente, onde as contribuições de  $t-1$  são bem maiores do que de  $t-3$ . Este fenômeno se repete para todas as outras usinas do sistema.

Porém vale ressaltar que as contribuições referentes às entradas do ano anterior são bastante significativas para a criação de novos centros, uma vez que são computadas em grande quantidade. Já que a rede tende a caracterizar as diferenciações anuais, de um ano para o outro, no que se refere a anos secos, médios ou úmidos. Observa-se também que, embora as diferenças entre  $t-12$ ,  $t-13$  e  $t-14$  sejam pequenas, há uma queda das contribuições, nesta ordem.

Entretanto agora serão apresentados os gráficos referentes ao treinamento pelo critério de dimensão por dimensão. Neste caso, serão computadas quantas vezes cada entrada contribui diretamente para a criação de um novo centro. Este computo é feito quando os centros são criados para todas as variáveis de entrada  $i$  que ultrapassem o seguinte limiar de tolerância:

$$|x_i - w_i| > \begin{cases} \delta_{afluencia} & , i = 1, \dots, 42 \\ \delta_{volume} & , i = 43, \dots, 84 \end{cases}$$

onde:  $\delta_{afluencia}$  - tolerância para as entradas de afluência

$\delta_{volume}$  - tolerância para as entradas de volume

Assim, como pode ser observado na Figura 5.46, este treinamento só capta os “picos de erros”, ou seja, somente as maiores diferenciações entre o padrão de entrada e o conjunto de pesos do neurônio vencedor são considerados nesta contagem, já que seriam justamente estas “grandes” diferenças, as responsáveis pela criação de um novo centro.

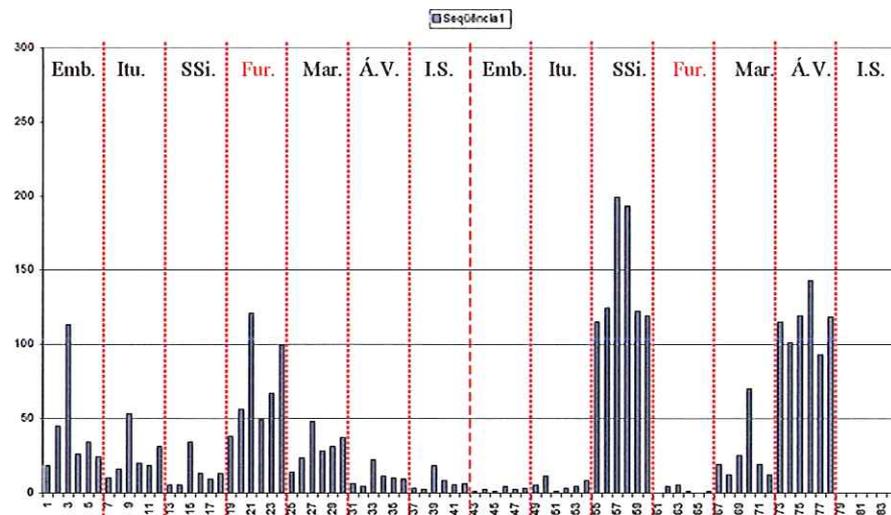


Figura 5.47 - Número de vezes que cada dado de entrada contribuiu para a criação dos centros, durante o treinamento

Novamente fazendo uma análise global das entradas relativas à grandeza volume operativo, nota-se desta vez que as usinas que mais contribuem para a criação de novos centros são as usinas de São Simão e Água Vermelha (representadas pelas entradas 55-60 e 73-78, respectivamente). Desta maneira, as usinas de jusante seriam as principais responsáveis pela criação dos centros, e essa importância decairia gradativamente a medida em que se analisa as entradas referentes às usinas mais a montante. Isto porque no resultado da operação otimizada determinística as usinas a jusante tendem a ficar “flat”, como já foi dito. Somente em poucas situações estas usinas têm seus reservatórios deplecionados, e são justamente estes deplecionamentos que são captados por esta técnica para a criação de um novo centro. Conseqüentemente, os reservatórios de montante passam a criar poucos centros, já que seus reservatórios variam mais ao longo de todo o ano, porém com baixas magnitudes de variação, não “sensibilizando” a tolerância de criação de centros.

Fato que ocorre de forma inversa para as entradas de afluência, uma vez que as usinas mais importantes para a criação de centros são as usinas de montante Emborcação e Furnas (representadas pelas entradas 01-06 e 19-24, respectivamente). Já as usinas intermediárias e de jusante contribuem menos.

Agora, fazendo uma análise mais detalhada dos dados de entrada referente à grandeza afluência, nota-se que no primeiro ano do passado, os dados mais próximos do presente, já não contribuem significativamente para a criação dos centros. Basta observar as entradas 1, 2 e 3 por exemplo, que representam as entradas de afluência da usina de Emborcação, deslocadas no tempo de  $t-1$ ,  $t-2$  e  $t-3$  respectivamente. Fazendo com que os dados referentes à três meses passados influencie de forma significativa no processo. Fenômeno este que se repete para todas as outras usinas do sistema.

## 6. Redes com Funções de Base Radial (RBF)

O modelo de uma rede neural supervisionada pode seguir diferentes linhas. O algoritmo “back-propagation” para o modelo de rede multicamada MLP (sob supervisão) pode ser visto como uma aplicação de um método de otimização, conhecido em estatística como aproximação estocástica. Neste capítulo nós tomamos uma abordagem diferente observando o modelo de uma rede neural como um “problema de ajuste de curva” (aproximação) em um espaço de alta dimensão. De acordo com este ponto de vista, aprender é equivalente a encontrar uma superfície em um espaço multidimensional que define o melhor ajuste para os dados de treinamento, cujo critério de “melhor ajuste” é medido por algum parâmetro estatístico. De forma correspondente, a generalização é equivalente ao uso desta superfície multidimensional para interpolar os padrões do conjunto teste de dados. Este ponto de vista é de fato a motivação por trás do método de funções de base radial o que conduz o trabalho de pesquisa de interpolação tradicional em espaços multidimensionais. No contexto de uma rede neural, as unidades escondidas fornecem um ajuste de funções que constituem uma base arbitrária para os padrões de entrada (vetores) quando eles expandidos no espaço de unidade-escondida; estas funções são chamadas de *funções de base radial*. As funções de base radial foram introduzidas pela primeira vez na interpolação de problemas reais multivariados. O primeiro trabalho neste contexto foi investigada por (Powell 1985), sendo hoje um dos principais campos da pesquisa em análise numérica.

(Broomhead; Lowe 1988) foram os primeiros a explorar o uso de funções de base radial na construção de redes neurais. Dentre outras grandes contribuições para a teoria, modelagem e aplicação de redes com funções de base radial estão trabalhos como (Moody e Darken, 1989), (Renals, 1989), e (Poggio e Firosi, 1990a). Este último trabalho enfatiza o uso da teoria da regularização aplicada a estas classes de redes neurais como um método de generalização de dados novos.

A arquitetura básica de uma rede com Função de Base Radial (RBF) possui apenas três diferentes camadas. A camada de entrada é constituída por nós fontes (unidades sensoriais). A segunda camada é a camada escondida de dimensão suficientemente grande, a qual executa um papel diferente daquela da rede Perceptron de Multicamadas (MLP). A camada de saída é responsável pela resposta da rede para os padrões de ativação aplicados à camada de entrada. A transformação a partir do espaço de entrada para o espaço de uma das unidades escondidas é não-linear, enquanto que a transformação do espaço da camada escondida para o espaço de saída é linear. Uma justificativa matemática racional pode ser verificada no trabalho de (Cover 1965). Em particular, este trabalho mostrou que o problema de classificação de padrões num espaço de grandes dimensões é melhor linearmente separável do que em problemas de baixas dimensões; sendo esta a razão da construção de uma rede RBF com muitas unidades na camada intermediária. Entretanto, através de uma modelagem cuidadosa, é possível reduzir a dimensão do espaço da camada escondida, especialmente se os centros das unidades escondidas forem ajustados de forma adaptativa.

### **6.1. Teorema de Cover: Separabilidade dos Padrões**

Quando uma rede com função de base radial (RBF) padrão é usada em tarefas de classificação de padrões complexos, o problema é basicamente resolvido pela transformação deste dentro de um espaço de grande dimensão de maneira não-linear. O teorema de Cover, sobre a separabilidade dos padrões, é a justificativa básica para isto, o qual estabelece que um problema complexo de classificação de padrões levado a um espaço de alta dimensão é melhor linearmente separável do que num espaço de baixa dimensão (Cover 1965). Em outras palavras, havendo padrões linearmente separáveis, o problema de classificação será relativamente fácil de resolver.

Dessa maneira, (Haykin 1994) desenvolveu uma enorme investigação na operação de uma rede RBF como um classificador através do estudo da separabilidade de padrões.

Considere uma família de superfícies, em que cada uma delas divide o espaço de entradas em duas regiões. Denotando  $X$  como um conjunto de  $N$  padrões (pontos)  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N$ , onde cada um deles é separado dentre as duas classes  $X^+$  e  $X^-$ . Esta *dicotomia* (partição binária) dos pontos diz para separá-los com respeito à família de superfícies,

desde que exista uma família de superfície que separe os pontos da classe  $X^+$  dos da classe  $X^-$ . Para cada padrão  $\mathbf{x} \in X$ , define-se um vetor composto por um conjunto de funções de valores reais  $\{\varphi_i(\mathbf{x}) | i = 1, 2, \dots, M\}$ , como mostrado por:

$$\varphi(\mathbf{x}) = [\varphi_1(\mathbf{x}), \varphi_2(\mathbf{x}), \dots, \varphi_M(\mathbf{x})]^T \quad (6.1)$$

seja que o padrão  $\mathbf{x}$  é um vetor  $p$ -dimensional do espaço de entrada. O vetor  $\varphi(\mathbf{x})$  então mapeia os pontos do espaço de entrada  $p$ -dimensional aos pontos correspondentes do novo espaço de dimensão  $M$ . Refere-se então à  $\varphi_i(\mathbf{x})$  como uma *função escondida*, já que ela atua de forma similar a uma unidade escondida de uma rede neural feedforward.

A dicotomia  $\{X^+, X^-\}$  de  $X$  é dita  *$\varphi$ -separável* se existir um vetor  $m$ -dimensional  $\mathbf{w}$  que permita escrever:

$$\mathbf{w}^T \varphi(\mathbf{x}) \geq 0, \quad \mathbf{x} \in X^+ \quad (6.2)$$

e

$$\mathbf{w}^T \varphi(\mathbf{x}) < 0, \quad \mathbf{x} \in X^- \quad (6.3)$$

O hiperplano definido pela equação:

$$\mathbf{w}^T \varphi(\mathbf{x}) = 0 \quad (6.4)$$

define a superfície de separação do espaço de entrada.

Considere uma classe natural de mapeamentos obtidos pelo uso de uma combinação linear dos elementos do vetor de padrões. As superfícies de separação correspondentes a cada mapeamento são referenciadas como as *variedades racionais de  $r$ -ésima ordem*. Uma variedade racional de ordem  $r$ , do espaço de dimensão  $p$ , é descrita por uma equação homogênea de grau  $r$  dos elementos do vetor de entrada  $\mathbf{x}$ , como mostrado por:

$$\sum_{0 \leq i_1 < i_2 < \dots < i_r \leq p} a_{i_1 i_2 \dots i_r} x_{i_1} x_{i_2} \dots x_{i_r} = 0 \quad (6.5)$$

onde  $x_i$  é o  $i$ -ésimo componente do vetor de entrada  $\mathbf{x}$ , e  $x_0$  é ajustado à unidade afim de expressar a equação na forma homogênea. Exemplos de superfícies de separação deste tipo são os *hiperplanos* (variedades racionais de primeira ordem), *parabolóides* (variedades de segunda ordem), e *hiperesferas* (variedades de segunda ordem com certas restrições lineares nos coeficientes). Estes exemplos são mostrados na Figura 6.1 para uma configuração de cinco pontos em um espaço de entrada bi-dimensional. Vale ressaltar que, em geral, os casos de separabilidade linear implicam tanto em

separabilidade esférica, como em separabilidade quadrática; entretanto, o inverso não é necessariamente verdadeiro.

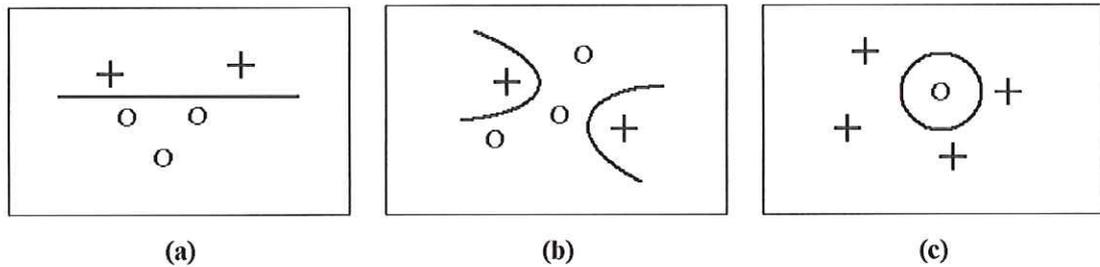


Figura 6.1 – Três exemplos da dicotomia de separabilidade- $\varphi$  para diferentes configurações de cinco pontos em duas dimensões: (a) separabilidade linear; (b) separabilidade quadrática; (c) separabilidade esférica.

*Separabilidade polinomial*, descrita anteriormente, pode ser vista como uma generalização natural da separabilidade linear. O ponto importante a ser notado aqui é que, dado um conjunto de padrões  $\mathbf{x}$  do espaço de entrada de dimensão arbitrária  $p$ , pode-se encontrar um mapeamento não-linear  $\varphi(\mathbf{x})$  de dimensão  $M$ , o qual possui separabilidade linear no espaço  $\varphi$ .

## 6.2. Problema de Interpolação

O ponto importante que surge do teorema de Cover sobre a separabilidade de padrões é que, resolvendo um problema de classificação de padrões não-linearmente separável, em geral, há benefício prático a ser ganho no mapeamento do espaço de entrada em um novo espaço de dimensão suficientemente grande. Basicamente, um mapeamento não-linear é usado para transformar um problema de classificação não-linearmente separável em um linearmente separável. Da mesma forma, pode-se usar um mapeamento não-linear para transformar um complexo problema de filtragem não-linear em um problema simples que envolve filtragem linear.

Considere uma rede neural de três camadas, com um neurônio na camada de saída. Ela foi modelada para fazer um mapeamento não-linear entre um espaço de entrada e o espaço da camada oculta (intermediária), seguida de um mapeamento linear a partir deste espaço da camada oculta para o espaço da camada de saída. Denotando  $p$  como a dimensão do espaço de entrada, esta rede passa a representar um mapeamento

entre o espaço de entrada  $p$ -dimensional e o espaço de saída unidimensional, escrito como:

$$s : \mathfrak{R}^p \rightarrow \mathfrak{R}^1 \quad (6.6)$$

Este mapeamento  $s$  pode ser pensado como um hiperplano  $\Gamma \subset \mathfrak{R}^{p+1}$ , e visto como um gráfico multidimensional da saída como função da entrada. Na prática, o hiperplano  $\Gamma$  é desconhecido e os dados de treinamento são normalmente contaminados por ruídos. Assim, as fases de treinamento e de generalização do processo de aprendizagem podem ser vistas respectivamente como:

- A fase de treinamento consiste na otimização do procedimento de ajuste da superfície  $\Gamma$ , com base nos dados conhecidos e apresentados à rede na forma de dados entrada-saída (padrões).
- A fase de generalização é o mesmo que interpolar entre pontos dos dados treinados, ou seja, o mesmo que interpolar a partir do hiperplano gerado no processo de ajuste de maneira a aproximar ao máximo o tal hiperplano  $\Gamma$  desconhecido.

O problema de interpolação pode ser definido matematicamente como:

dado um conjunto de  $N$  diferentes pontos  $\{\mathbf{x}_i \in \mathfrak{R}^p \mid i = 1, 2, \dots, N\}$  e um correspondente conjunto de  $N$  números reais  $\{d_i \in \mathfrak{R}^1 \mid i = 1, 2, \dots, N\}$ , encontrar uma função  $F: \mathfrak{R}^p \rightarrow \mathfrak{R}^1$  que satisfaça a condição de interpolação:

$$F(\mathbf{x}_i) = d_i, \quad i = 1, 2, \dots, N \quad (6.7)$$

Para que se obtenha uma boa interpolação, a função  $F$  (ou o hiperplano  $\Gamma$ ) deve ser construída de forma a passar por todos os pontos de dados do treinamento.

A técnica de Funções de Base Radial (RBF) consiste na escolha de uma função  $F$  que tenha a seguinte forma (Powell 1988):

$$F(\mathbf{x}) = \sum_{i=1}^N w_i \varphi(\|\mathbf{x} - \mathbf{x}_i\|) \quad (6.8)$$

onde  $\{\varphi(\|\mathbf{x} - \mathbf{x}_i\|) \mid i = 1, 2, \dots, N\}$  é um conjunto de funções quaisquer (geralmente não-lineares), conhecidas como *funções de base radial*, e  $\|\cdot\|$  denota a *norma Euclidiana*, normalmente a mais utilizada. Os pontos de dados conhecidos  $\mathbf{x}_i \in \mathfrak{R}^p \mid i = 1, 2, \dots, N$  são tomados como os *centros* das funções de base radial.

Inserindo as condições de interpolação da Equação 6.7 na Equação 6.8, obtém-se o seguinte conjunto de equações lineares para os coeficientes (pesos) desconhecidos da expansão  $\{w_i\}$ :

$$\begin{bmatrix} \varphi_{11} & \varphi_{12} & \cdots & \varphi_{1N} \\ \varphi_{21} & \varphi_{22} & \cdots & \varphi_{2N} \\ \vdots & \vdots & \vdots & \vdots \\ \varphi_{N1} & \varphi_{N2} & \cdots & \varphi_{NN} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_N \end{bmatrix} = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_N \end{bmatrix} \quad (6.9)$$

onde

$$\varphi_{ji} = \varphi(\|\mathbf{x}_j - \mathbf{x}_i\|) \quad j, i = 1, 2, \dots, N \quad (6.10)$$

sendo

$$\mathbf{d} = [d_1, d_2, \dots, d_N]^T \quad (6.11)$$

$$\mathbf{w} = [w_1, w_2, \dots, w_N]^T \quad (6.12)$$

Os vetores,  $N$  por  $1$ ,  $\mathbf{d}$  e  $\mathbf{w}$  representam o *vetor de resposta desejada* e o *vetor de pesos lineares*, respectivamente. Sendo ainda que  $\Phi$  denota a matriz  $N$  por  $N$  com elementos  $\varphi_{ji}$ :

$$\Phi = \{\varphi_{ji} \mid j, i = 1, 2, \dots, N\} \quad (6.13)$$

chamada de *matriz de interpolação*.

Assim, pode-se reescrever a Equação 6.9 de forma compacta:

$$\Phi \mathbf{w} = \mathbf{d} \quad (6.14)$$

Existe uma classe de funções de base radial que seguem estas propriedades, segundo o teorema (Light 1992):

“Sejam  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N$  pontos distintos em  $\mathfrak{R}^p$ . Então a matriz de interpolação  $\Phi$ ,  $N$  por  $N$ , cujo  $j$ -ésimo elemento  $\varphi_{ji} = \varphi(\|\mathbf{x}_j - \mathbf{x}_i\|)$ , é definido positivo.”

Este teorema se aplica às seguintes funções:

#### 1. Multiquadrática inversa

$$\varphi(r) = \frac{1}{(r^2 + c^2)^{1/2}} \quad \text{para algum } c > 0, \text{ e } r \geq 0 \quad (6.15)$$

#### 2. Funções Gaussianas

$$\varphi(r) = \exp\left(-\frac{r^2}{2\sigma^2}\right) \quad \text{para } \sigma > 0, \text{ e } r \geq 0 \quad (6.16)$$

Investigações teóricas e resultados práticos, entretanto, mostram que o tipo de não-linearidade de  $\varphi(\cdot)$  não é crucial para a performance da rede RBF (Powell 1988).

Retornando ao teorema de Light, nota-se que a matriz de interpolação  $\Phi$  é positiva definida, e então pode-se resolver a Equação 5.14 para a obtenção dos pesos  $\mathbf{w}$ , da seguinte maneira:

$$\mathbf{w} = \Phi^{-1} \mathbf{d} \quad (6.17)$$

onde  $\Phi^{-1}$  é a inversa da matriz de interpolação  $\Phi$ .

### 6.3. Modelagem da rede RBF

Esta secção descreve um modelo de rede neural artificial inspirado na resposta localmente sintonizada (locally tuned response) observada em neurônios biológicos. Neurônios com esta característica podem ser encontrados em muitos sistemas nervosos. Estas células nervosas têm uma resposta característica, que é seletiva para um espectro finito do espaço do sinal de entrada.

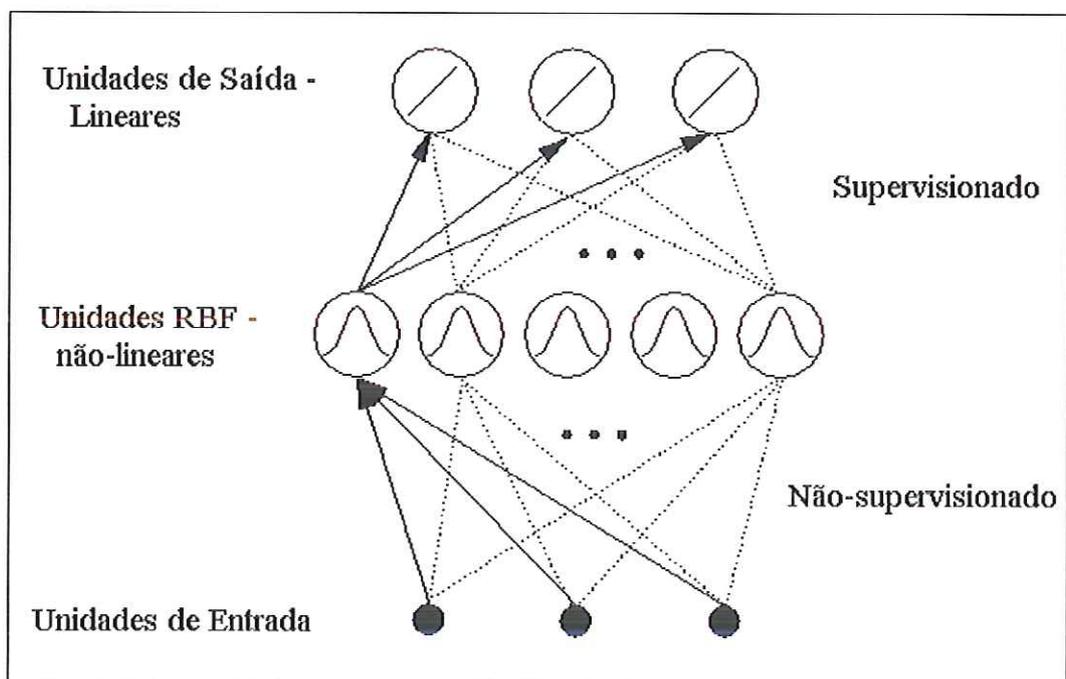


Figura 6.2 – Arquitetura de uma Rede Neural com Função de Base Radial (RBF)

A rede RBF possui uma estrutura feedforward constituída de uma única camada escondida de  $J$  unidades de sintonia local, as quais são totalmente conectadas à camada

de saída com  $L$  unidades lineares, como na Figura 6.2. Todas as unidades escondidas recebem simultaneamente o vetor de entrada  $\mathbf{x}$   $n$ -dimensional. Assim, as saídas  $\varphi_j$  de cada unidade escondida são obtidas em função da proximidade entre a entrada  $\mathbf{x}$  e o parâmetro  $\mu_j$  (definida como o centro de cada função radial). Esta proximidade é avaliada por uma função radial, como a Equação 6.16, função Gaussiana:

$$\varphi_j(\mathbf{x}; \mu_j; \sigma_j) = \exp\left(-\frac{\|\mathbf{x} - \mu_j\|^2}{2\sigma_j^2}\right) \quad (6.18)$$

onde o parâmetro  $\sigma_j$  é a largura do campo receptivo do espaço de entrada da unidade  $j$ . Isto implica que  $\varphi_j$  apresenta um valor considerável somente quando a distância  $\|\mathbf{x} - \mu_j\|$  for menor que a largura  $\sigma_j$ . Dado um vetor de entrada  $\mathbf{x}$ , a saída da rede RBF passa então a ser definida por:

$$\mathbf{y}(\mathbf{x}) = \sum_{j=1}^J w_{lj} \varphi_j(\mathbf{x}) \quad (6.19)$$

Redes RBF são muito boas para aproximar mapeamentos contínuos, ou contínuos por parte, do tipo  $f: \mathcal{R}^n \rightarrow \mathcal{R}^L$ , quando  $n$  é suficientemente pequeno. De acordo com as Equações 6.18 e 6.19, a rede RBF pode ser vista como uma aproximação de uma função desejada  $f(\mathbf{x})$  pela superposição de funções de base radial não-ortogonais, podendo ser classificada como um aproximador universal. O grau de precisão pode ser controlado por três parâmetros: o número de funções de base, sua localização, e sua largura.

Durante a fase de treinamento procura-se reduzir o erro na saída da rede, que neste trabalho foi considerado o MSE, a valores aceitáveis através da adaptação dos parâmetros livres na rede RBF: os centros dos campos receptivos  $\mu_j$ , as larguras dos campos receptivos  $\sigma_j$  e os pesos entre a camada escondida e a camada de saída  $w_{lj}$ . Este treinamento pode utilizar a aprendizagem supervisionada, não-supervisionada ou híbrida, sendo esta última a mais empregada e, inclusive, utilizada neste trabalho.

O treinamento híbrido combina aprendizagem não-supervisionada com supervisionada. A primeira treina a camada escondida, definindo os parâmetros desta camada, como localização dos centros e larguras dos campos receptivos. A segunda etapa define os valores dos pesos entre as camadas escondidas e de saída. Nesta segunda fase, os parâmetros definidos na primeira fase não se modificam.

Na etapa de treinamento não-supervisionado para a camada escondida são então determinados os centros das funções de base radial e as larguras dos campos receptivos. Dentre os métodos para determinação dos centros das funções de base radial, temos:

- Seleção Aleatória

Os centros são vetores de entrada aleatoriamente selecionados. Esta técnica demanda que os padrões de treinamento representem precisamente todo o espaço de soluções. Este método é simples e direto, no entanto pode exigir grande número de unidades intermediárias, e escolher centros muito próximos uns dos outros pode acarretar mau funcionamento da rede.

- Fixação de grade regular

Neste caso, os centros são fixados em uma grade regular, cobrindo todo o espaço dos vetores de entrada. Em geral, este método exige muitas unidades intermediárias para vetores de entrada com muitas dimensões, o que provoca um crescimento exponencial do número das unidades escondidas, fenômeno conhecido como maldição da dimensionalidade.

- Técnicas de agrupamento

Os centros são definidos por técnicas de agrupamento. Entre as quais se destacam os algoritmos de k-médias e de mapas auto-organizáveis (SOM). Os mapas auto-organizáveis caracterizam-se por agrupar padrões espacialmente próximos que compartilhem micro-características. O centro que apresentar maior produto escalar com um dado vetor de entrada, adiciona este vetor de entrada ao seu grupo. Esta foi a técnica empregada neste trabalho de pesquisa.

A técnica dos mapas auto-organizáveis se mostra bastante eficiente e define os centros receptivos das funções de base radial de forma automatizada, durante esta etapa de treinamento não-supervisionado.

Neste trabalho os centros foram então definidos pela rede SONARX, caracterizada com uma rede de Mapa Auto-Organizável (SOM) de Kohonen com Modelos de Sistemas Dinâmicos, em outras palavras, uma versão não-supervisionada do modelo NARX.

Ainda na etapa de treinamento não-supervisionado da rede RBF, precisa-se definir uma heurística para a determinação da largura do campo receptivo  $\sigma_j$ , que pode ser único para todas as unidades ou pode ser diferente para cada unidade escolhida. Algumas das principais heurísticas são:

- Utilização da distância euclidiana média entre centros:

$$\sigma = \frac{1}{ng} \sum_{j=1}^{ng} \|\mu_j - \mu_{j(mprox)}\| \quad (6.20)$$

onde  $ng$  é o número de grupos formados e  $\mu_{j(mprox)}$  é o centro com menor distância euclidiana com relação ao centro  $\mu_j$ . Este método produz um único valor de raio.

- Utilização da distância euclidiana entre centro e vetor de entrada:

$$\sigma_j^2 = \frac{1}{nv} \sum_{\mathbf{x} \in \psi_j} \|\mu_j - \mathbf{x}\|^2 \quad (6.21)$$

onde  $\psi_j$  é o conjunto dos  $nv$  vetores de entrada com menor distância euclidiana para o centro  $\mu_j$ .

- Utilização da distância euclidiana entre os centros determinados pelo método k-médias:

$$\sigma_j^2 = \frac{1}{nv} \sum_{\mathbf{x} \in S_j} \|\mu_j - \mathbf{x}\|^2 \quad (6.22)$$

onde  $S_j$  é o agrupamento com  $nv$  vetores de entrada pertencentes a ele.

- Utilização da distância euclidiana entre centros:

$$\sigma_j = \alpha \|\mu_j - \mu_{j(mprox)}\| \quad (6.23)$$

onde  $\mu_{j(mprox)}$  é o centro com menor distância euclidiana com relação ao centro  $\mu_j$  e o parâmetro geralmente se situa  $1,0 \leq \alpha \leq 1,5$ .

Devido ao espalhamento não uniforme dos centros, definidos pela rede SONARX no espaço de entrada, a heurística utilizada neste trabalho foi a que utiliza a distância euclidiana entre centros (apresentada pela Equação 6.23), já que este método permite definir larguras distintas dos campos receptivos para cada centro,

conseqüentemente para cada função de base radial, caracterizando de forma suave todo o espaço de entrada.

Já na etapa do treinamento supervisionado para a camada de saída passam a ser determinados os pesos entre a camada escondida e a de saída. A primeira etapa neste processo é o cálculo do erro. Dentre as maneiras de se calcular o erro, optou-se por:

- Erro quadrático médio (MSE):

$$MSE = \frac{1}{N} \sum_{t=1}^N [y(t) - \hat{y}(t)]^2 \quad (6.24)$$

onde  $y(t)$  e  $\hat{y}(t)$  são respectivamente os resultados desejado e obtidos pela rede RBF, e  $N$  é o número total de padrões.

Calculado o erro, este deve ser minimizado por algum procedimento que altere e defina valores “ótimos” para os pesos  $w_{ij}$ ; dentre os quais destacam-se: o método dos mínimos quadrados, a matriz pseudo-inversa (ambos classificados como métodos lineares), e a regra delta, a qual foi aqui adotada.

A regra delta, proposta por (Widrow; Jr. 1960), busca obter o ponto de mínimo através de um processo de iteração local, utilizando um padrão do conjunto de treinamento por vez. Isto pode ser traduzido no seguinte algoritmo iterativo para o cálculo do vetor de pesos  $\mathbf{w}(t)$  sucessivamente:

$$w_{ij}(t+1) = w_{ij}(t) + \Delta w_{ij}(t) \quad (6.25)$$

onde  $\Delta w_{ij}(t)$  representa a parcela de atualização dos pesos, descrita por :

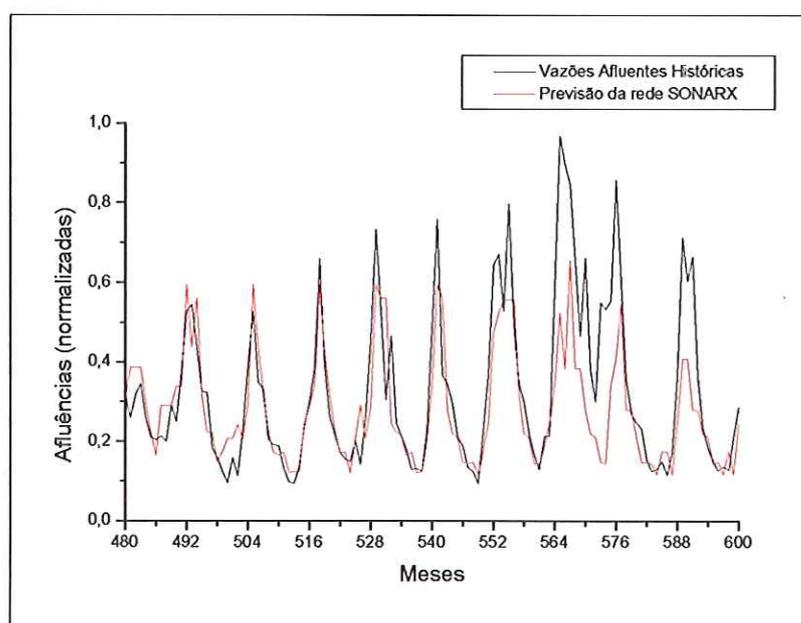
$$\Delta w_{ij}(t) = \eta \cdot (y_j(t) - \hat{y}_j(t)) \cdot \varphi_j \quad (6.26)$$

sendo  $\eta$  a taxa de aprendizagem e  $\varphi_j$  a função de ativação do neurônio  $j$  da camada escondida.

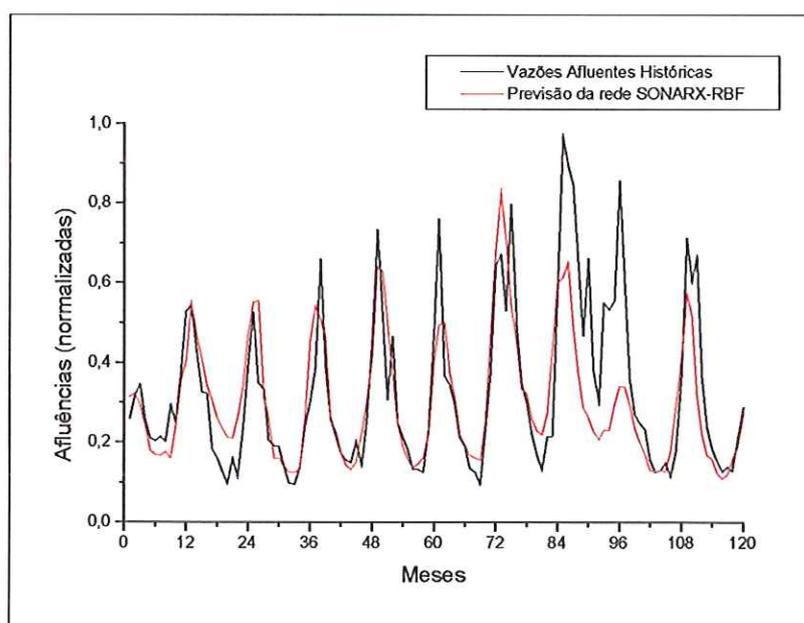
Naturalmente, existe uma dificuldade inerente a este tipo de abordagem, uma vez que a definição desta taxa de aprendizagem  $\eta$  deve ser feita pelo usuário. Uma escolha correta desta taxa depende muito da sua experiência na aplicação desta metodologia a problemas reais.

#### 6.4. Resultados da Predição com Interpolação pela rede RBF

Nesta seção é apresentada a implementação de um modelo de interpolação acoplado a rede SONARX, que teve o propósito de mitigar o erro inerente a este tipo de modelagem baseada na técnica de MATQV. A rede utilizada foi a rede com Função de Base Radial (RBF), escolhida pela sua eficácia na interpolação dos padrões de entrada,



(a)

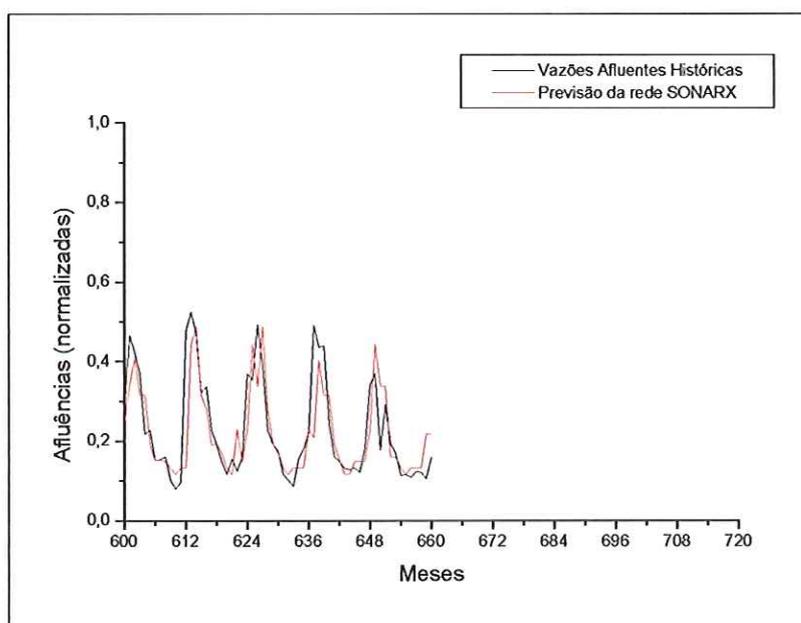


(b)

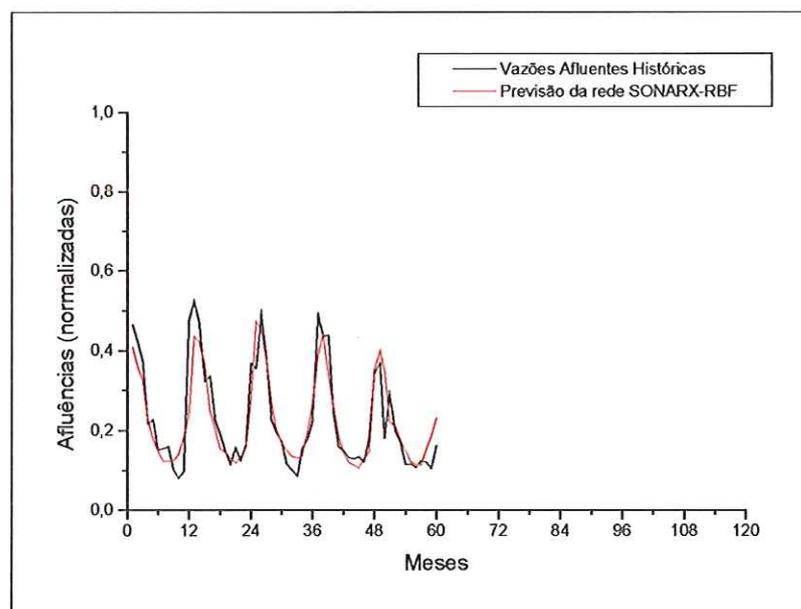
Figura 6.3 – Comparação entre a rede SONARX em (a) e a rede SONARX-RBF em (b) na Predição das aflúncias para a usina de Furnas, para o período de Janeiro de 1976 até Dezembro de 1985

como foi mostrado no item 6.2.

Pelo mesmo motivo que a rede SONARX foi implementada usando a linguagem de programação FORTRAN, resolveu-se implementar também a rede RBF na mesma linguagem. Com isso, chegou-se a um modelo híbrido SONARX-RBF, no qual a rede SONARX ficou responsável por definir a localização dos centros receptivos, e a rede RBF pela tarefa de interpolação dos padrões de entrada.



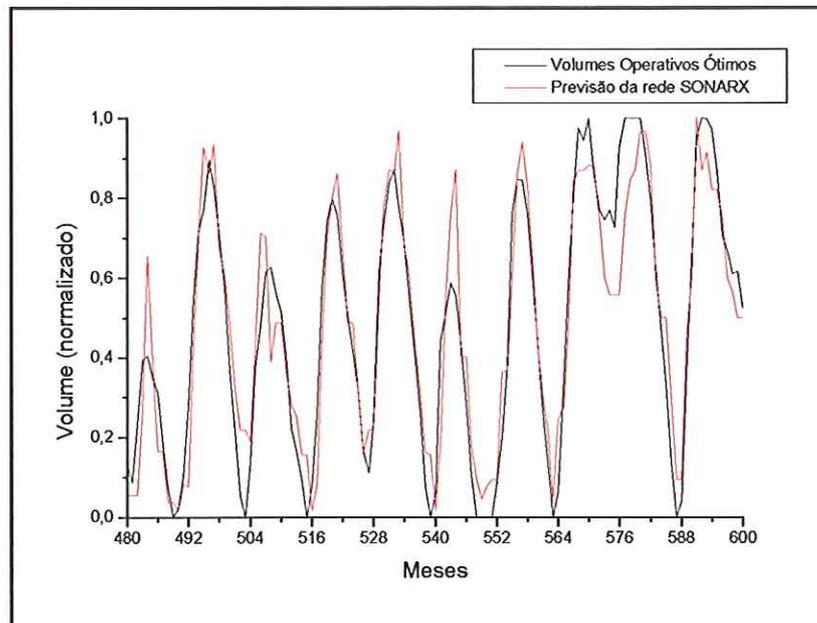
(a)



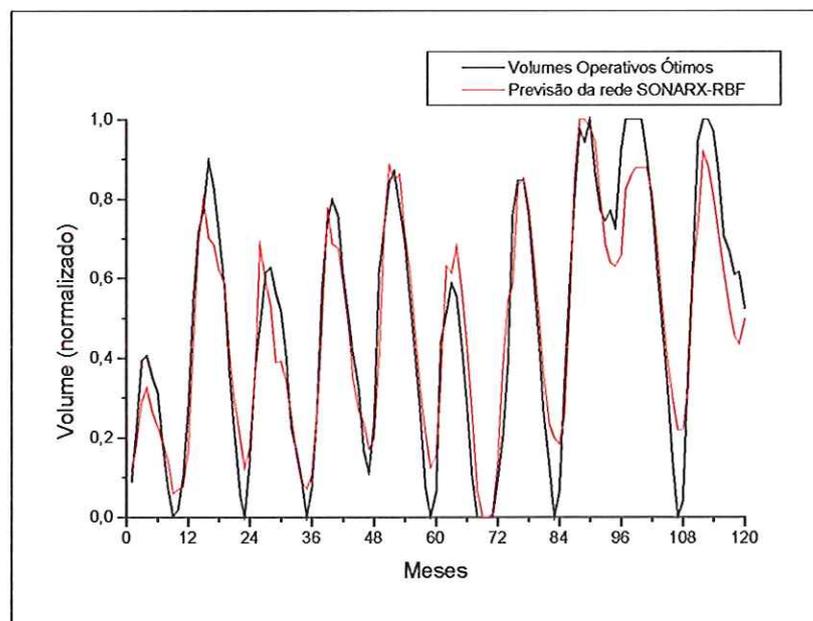
(b)

**Figura 6.4 – Comparação entre a rede SONARX em (a) e a rede SONARX-RBF em (b) na Previsão das aflúncias para a usina de Furnas, para o período de Janeiro de 1986 até Dezembro de 1990**

A previsão das vazões afluentes através da rede SONARX, sem interpolação, é apresentada nas Figuras 6.3(a) e 6.4(a), enquanto que a previsão das vazões usando a rede SONARX-RBF, com interpolação, está nas Figuras 6.3(b) e 6.4(b).



(a)

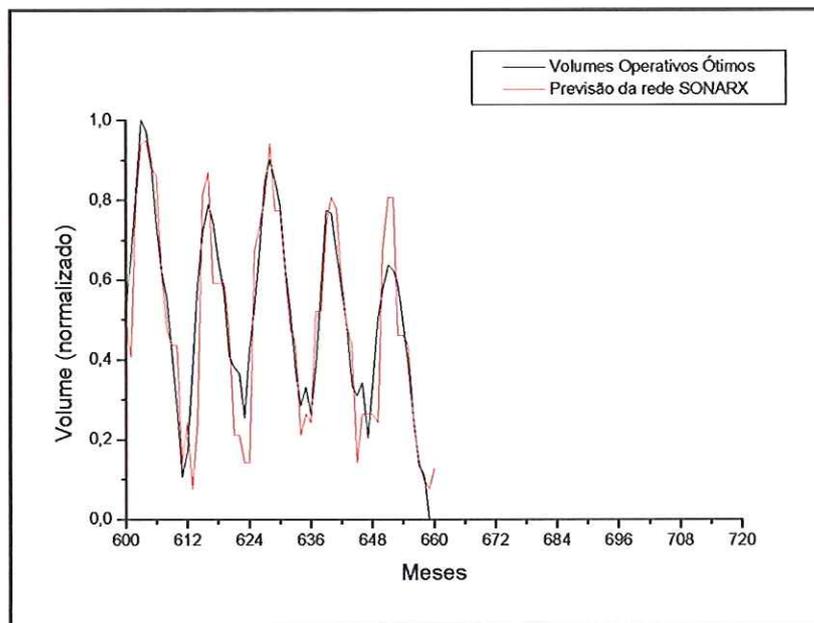


(b)

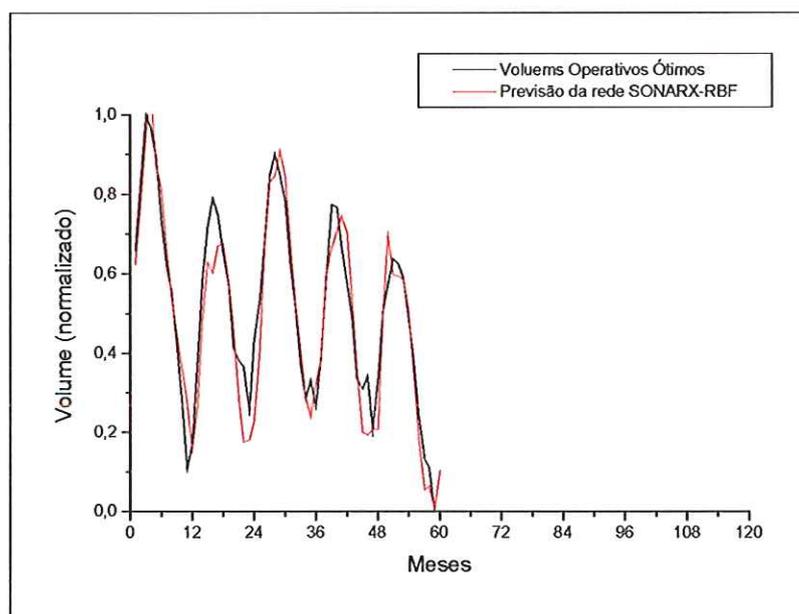
**Figura 6.5 – Comparação entre a rede SONARX em (a) e a rede SONARX-RBF em (b) na Previsão dos Volumes Operativos Ótimos para a usina de Furnas, para o período de Janeiro de 1976 até Dezembro de 1985**

Analisando as trajetórias de ambas as previsões de vazões em (a) e (b), para o período de janeiro de 1976 até dezembro de 1990, nota-se uma suavização da trajetória

de previsões definida pela rede SONARX-RBF em relação à gerada pela rede SONARX, provocada justamente pela interpolação dos padrões.



(a)



(b)

**Figura 6.6 – Comparação entre a rede SONARX em (a) e a rede SONARX-RBF em (b) na Previsão dos Volumes Operativos Ótimos para a usina de Furnas, para o período de Janeiro de 1986 até Dezembro de 1990**

Esta suavização na trajetória de previsões provocou uma melhoria no desempenho do previsor neural, como pode ser observado na Tabela 4, onde são comparados os Erros Médios Quadráticos (MSE), definido pela Equação 4.48, em

destaque na cor vermelha, já que este foi o erro escolhido para ser minimizado na etapa de treinamento supervisionado da rede SONARX-RBF.

Com relação as previsões dos volumes operativos ótimos, apresentados nas Figuras 6.5 e 6.6, cabem as mesmas considerações feitas para as previsões das vazões afluentes. Observou-se uma melhoria de desempenho ainda maior nas previsões dos volumes ótimos pelo uso da rede SONARX-RBF, através da diminuição dos Erros Médios Quadráticos (MSE), como pode ser visto também na Tabela 4.

Tabela 4 – Tabela comparativa entre a rede SONARX e a rede SONARX-RBF, para o período de Janeiro de 1976 até Dezembro de 1990

		Vazões Afluentes		Volumes Operativos Ótimos	
		SONARX	SONARX-RBF	SONARX	SONARX-RBF
76-80	EPM (%)	24,05	24,50	9,25	7,09
	MSE ( $10^{-3}$ )	9,08	5,93	14,90	7,61
	MRE ( $10^{-2}$ )	10,80	13,10	4,00	2,10
81-85	EPM (%)	26,16	28,31	10,75	10,05
	MSE ( $10^{-3}$ )	32,40	28,50	19,00	13,70
	MRE ( $10^{-2}$ )	10,60	12,30	4,00	3,70
86-90	EPM (%)	25,79	21,97	8,84	6,44
	MSE ( $10^{-3}$ )	8,28	4,20	13,40	6,69
	MRE ( $10^{-2}$ )	12,70	11,10	2,50	1,40

Fazendo uma análise comparativa de desempenho para as previsões de vazões, compreendido entre o período de 1976 a 1980, observou-se uma diminuição de **34,7%** do erro MSE. Para o período de 1981 a 1985 obteve-se uma melhoria de **12,0%**, e para o período de 1986 a 1990, uma melhoria de **49,3%**.

Na análise da previsão dos volumes operativos ótimos, obteve-se uma melhoria de **48,9%**, **27,9%** e **50,1%** do erro MSE, respectivamente para os períodos de 1976 a 1980, 1981 a 1985 e 1986 a 1990.

## 7. Conclusões

No presente trabalho obteve-se um conjunto de Regras de Operação de Reservatórios (ROR), individuais para cada Usina Hidroelétrica (UH), bem próximas do comportamento ótimo do sistema, pois é inviável definir uma ROR geral para simular a operação de um dado sistema, visto que diferentes posições relativas dentro das cascatas e diferentes características construtivas são peculiaridades de cada UH.

Este comportamento diferenciado entre as usinas pode ser observado e analisado na seção 5.1.1., definindo-se quanto cada usina deveria gerar a cada instante para que a operação fosse ótima. Apesar deste tipo de análise ser fictício, ela é muito importante na realização de uma análise do comportamento dos reservatórios, observando-se como eles relacionam-se entre si.

Visto que a descrição analítica deste problema é desconhecida, optou-se pelas Redes Neurais Artificiais (RNA) para fazer o mapeamento segundo a modelagem do tipo caixa preta (black box). Para a tarefa de capturar o comportamento ótimo de cada uma das UH e simular a operação ótima de um subsistema do Sistema Sudeste Brasileiro, formado por 7 (sete) usinas reais, aplicou-se as redes neurais SONARX e SONARX-RBF.

Neste contexto, os dados de entrada foram as vazões afluentes históricas e os estados de armazenamento dos reservatórios do sistema hidroelétrico, gerando como variáveis de saída, as previsões das futuras vazões afluentes e a previsão dos volumes operativos ótimos, informações importantíssimas para a tomada de decisão por parte do operador do sistema de geração.

A partir dos resultados obtidos com o sistema teste, pode-se concluir que a rede SONARX conseguiu assimilar o comportamento ótimo da operação. Os testes mostraram que o previsor neural possui uma forte tendência em seguir os resultados da otimização determinística, otimizando o uso dos recursos hídricos disponíveis para geração de energia elétrica. Além de uma boa previsão das futuras vazões afluentes.

A rede SONARX foi capaz de acompanhar a trajetória otimizada do volume operativo das 7 usinas hidroelétricas deste sistema teste. Mesmo para as usinas de montante, que tendem a variar muito seus reservatórios, buscando regularizar as vazões afluentes às usinas de jusante, como são os casos de Emborcação, Itumbiara e Furnas.

Como estas previsões são informações importantes para a tomada de decisão, este decisor neural foi integrado a um programa simulador da operação de sistemas hidroelétricos. Na comparação realizada entre este decisor, e uma regra heurística de operação em paralelo, o decisor neural se mostrou muito mais eficiente.

Como a rede SONARX é baseada na técnica de MATQV, implementou-se um modelo de interpolação acoplado a esta rede, na tentativa de mitigar o erro inerente a esta técnica. A rede escolhida foi a rede RBF, dando origem ao modelo neural SONARX-RBF. Com base nos resultados apresentados na seção 6.4, observou-se uma diminuição considerável no erro de previsão, tanto para as vazões afluentes quanto para os volumes operativos ótimos, decorrente da suavização das trajetórias provocada justamente por esta interpolação.

### **Contribuições do Trabalho**

No decorrer deste trabalho, tive a oportunidade de passar por um excelente aprendizado de RNA pertencente ao amplo espectro do conhecimento de Inteligência Artificial, aplicado em um problema de Sistemas Elétricos de Potência. Além disto, tive um profundo amadurecimento do problema de Planejamento da Operação.

Para a consolidação de um grupo de pesquisa, é de fundamental importância que os membros deste grupo adquiram novas competências, podendo com isso, solucionar os problemas complexos ainda existentes no campo da Engenharia Elétrica. Neste sentido, este trabalho contribuiu para a consolidação das “ferramentas” matemático-computacionais em nosso grupo do LSEE.

Este trabalho ainda teve uma excelente contribuição ao meio científico, ao abordar o problema de Planejamento da Operação sob um novo enfoque, estabelecendo como uma nova variável de previsão os Volumes Operativos Ótimos, através de modelos de RNA. Dessa maneira, fica evidente a abertura para trabalhos futuros, que será discutido logo a seguir.

### **Desenvolvimentos Futuros**

Dentre os possíveis futuros encaminhamentos deste trabalho de pesquisa, destacam-se dois aspectos. O primeiro seria um estudo e avaliação da previsão das vazões afluentes considerando somente os dados históricos das vazões afluentes da usina de interesse, diminuindo assim o número de variáveis de entrada, uma possível causa da deterioração da informação. O que poderia ser feito tanto para o modelo de rede SONARX como para a rede SONARX-RBF.

O primeiro seria o treinamento e teste da rede SONARX-RBF para os outros períodos do histórico que não foram avaliados por este modelo. Posteriormente, poderia se pensar em obter novas ROR usando esta modelagem. Regras estas que como visto neste trabalho de pesquisa, não se faz uso da previsão das futuras vazões afluentes, para definir sua estratégia de operação.

Finalmente, poder-se-ia simular a operação hidrotérmica, para todo o histórico conhecido, podendo assim avaliar seu comportamento frente as mais variadas condições hidrológicas. Desta maneira, poderiam ser feitas análises comparativas entre esta Política de Operação e diversas outras, como por exemplo a Programação Dinâmica Dual Estocástica (base do modelo "Newave"), a qual vem sendo aplicada no Setor Elétrico Brasileiro.



## 8. Bibliografia

- Aguirre, L. A. (2000). Introdução à Identificação de Sistemas. Belo Horizonte - MG, Editora UFMG.
- Ashby, R. (1947). "Principles of the Self-Organizing Dynamic System." Journal of General Psychology(37): 125-128.
- Ashby, R. (1952). Design for a Brain: The Origin of Adaptive Behavior. New York, Wiley.
- Barreto, G. d. A. (2003). Redes Neurais Não-Supervisionadas Temporais para Identificação e Controle de Sistemas Dinâmicos. Engenharia Elétrica. São Carlos-SP, Universidade de São Paulo: 237.
- Barreto, G. d. A.; A. F. R. Araújo (2001a). "Time in self-organizing maps: An overview of models." International Journal of Computer Research, Special Issue on Neural Networks: Past, Present and Future 10(2): 139-179.
- Barreto, G. d. A.; A. F. R. Araújo (2001b). A self-organizing NARX network and its application to prediction of chaotic time series. International Joint Conference on Neural Networks - IJCNN '01, Washington D.C., USA.
- Bower, B. T.; M. M. Hufschmidt, et al. (1962). Operating Procedures: Their Role in the Design and Implementation of Water Resource Systems by Simulation Analysis. Design of Water Resource Systems. H. University. Cambridge, Mass.: Cap. 11, 443-458.
- Broomhead, D. S.; D. Lowe (1988). Multivariable functional interpolation and adaptive networks. Complex Systems. 2: 321-355.

- Carneiro, A. A. F. M. (1991). Contribuição ao Planejamento da Operação Energética de Sistema Hidrotérmica de Potência. Departamento de Engenharia Elétrica. Campinas, Universidade Estadual de Campinas.
- Carneiro, A. A. F. M.; M. Kadowaki (1996). Regras de Operação para Grandes Sistemas Hidroelétricos em Cascata. XI Congresso Brasileiro de Automática, São Paulo-SP, Brazil.
- Carneiro, A. A. F. M.; M. Kadowaki (1997). Operação Simulada de Sistemas Hidroelétricos com Regras por Redes Neurais. XIV SNPTEE - Seminário Nacional de Produção e Transmissão de Energia Elétrica, Belém, Brazil.
- Carneiro, A. A. F. M.; D. Silva Filho (1998a). Regras Otimizadas de Operação de Sistemas Hidroelétricos Obtidas por Lógica Fuzzy. 12 CONGRESSO BRASILEIRO DE AUTOMÁTICA, Uberlândia, Brazil.
- Carneiro, A. A. F. M.; D. Silva Filho (1998b). Fuzzy logic applied to operation rules for large hydrothermal power systems. POWERCON '98 - International Conference on Power System Technology, Beijing, China.
- Carvalho, M. F. H.; S. Soares Filho (1987). "An Efficient Hydrothermal Scheduling Algorithm." IEEE Transactions on Power System PWRS-2(3): 537-542.
- CEPEL (2000). Modelo NEWAVE - Manual de Referência. Rio de Janeiro - RJ, Brasil, CEPEL: 103.
- Cottrell, M.; J. C. Fort, et al. (1998). "Theoretical Aspects of the SOM Algorithm." Neurocomputing 21: 119-138.
- Cover, T. M. (1965). "Geometrical and statistical properties of systems of linear inequalities with applications in pattern recognition." IEEE Transactions on Electronic Computers 14: 326-334.

- Cruz Jr., G. d.; S. Soares Filho (1999). General Composite Representation of Hydroelectric Systems. 21st IEEE International Conference of Power Industry Computer Applications.
- da Silva, E. L.; E. C. Finardi (2001). Planning of Hydrothermal Systems Using a power Plant Individualistic Representation. IEEE Power Tech Conference 2001, Porto, Portugal.
- Fortunato, L. A. M.; T. A. Araripe Neto, et al. (1985). Operating Planning Studies of the Brazilian Generation System. IFAC Symposium on Planning and Operation of Electric Energy System - Preprints, Rio de Janeiro.
- Grossberg, S. (1986). Adaptive self-organization of serial order in behavior: speech, language, and motor Control. Pattern Recognition by Humans and Machines: Speech Perception. E. SchwabH. Nusbaum, Academic Press. 1: 187-294.
- Haykin, S. (1994). Neural Networks: A Comprehensive Foundation. NY, Macmillan Publishing Company.
- Johnson, S. A.; J. R. Stedinger, et al. (1991). "Heuristic Operating Policies for Reservoir System Simulation." Water Resources Research 27(5): 673-685.
- Kadowaki, M. (1995). Simulação da Operação de Sistemas Hidroelétricos de Potência a Usinas Individualizadas com Regras Otimizadas. Departamento de Engenharia Elétrica. São Carlos, Universidade de São Paulo: 131.
- Kaski, S.; J. Kangas, et al. (1998). "Bibliografy of the Self-Organizing Map (SOM) Papers: 1981-1997." Neural Computing Surveys 1: 102-350.
- Kelman, J.; J. M. Damazio, et al. (1989). "The determination of flood control volumes in a multireservoir system." Water Resources Research 25(3): 337-344.
- Kohonen, T. (1989). Self-Organization and Associative Memory. Berlin, Springer-Verlag.

- Light, W. A. (1992). Some aspects of radial basis function approximation. Approximation Theory, Spline Functions and Applications, Boston, MA.
- Lund, J. R. (2000). "Derived Power Production and Energy Drawdown Rules for Reservoirs." Journal of Water Resources Planning and Management: 108-111.
- Lund, J. R.; J. Guzman (1999). "Derived Operating Rules for Reservoirs in Series or in Parallel." Journal of Water Resources Planning and Management **125**(3): 143-153.
- Lyra, C.; H. Tavares, et al. (1984). "Modelling and optimization of hydrothermal generation scheduling." IEEE Transactions on Power Apparatus and System **103**(8): 2126-2133.
- Martinez, L.; S. Soares Filho (2002). "Comparison Between Closed-Loop and Partial Open-Loop Feedback Control Policies in Long Term Hydrothermal Scheduling." IEEE Transactions on Power System **17**(2): 330-336.
- Norgaard, M.; O. Ravn, et al. (2000). Neural Networks for Modelling and Control of Dynamic Systems, Springer-Verlag.
- Oliveira, R.; D. P. Loucks (1997). "Operating Rules for Multireservoir Systems." Water Resources Research **33**(4): 839-852.
- Pereira, M. V. F. (1985). Optimal Scheduling of Hydrothermal System - An Overview. IFAC - Electric Energy Systems, Rio de Janeiro, Brazil.
- Powell, M. J. D. (1985). Radial basis functions for multivariable interpolation: A review. IMA Conference on Algorithms for the Approximation of Functions and Data, Shrivenham, UK.
- Powell, M. J. D. (1988). Radial basis function approximations to polynomials. Numerical Analysis, Dundee, UK.

- Robbins, H.; S. Monro (1951). "A Stochastic Approximation Method." Annals of Mathematical Statistics 22: 400-407.
- Rotting, T. A.; A. Gjelsvik (1992). "Stochastic Dual Dynamic Programming for Seasonal Scheduling in the Norwegian power System." IEEE Transactions on Power System 7(1): 273-279.
- Silva Filho, D.; A. A. F. M. Carneiro (1999a). Sistema de apoio à decisão para simulação da operação de sistemas hidroelétricos de geração. 20 IBERIAN LATIN-AMERICAN CONGRESS ON COMPUTACIONAL METHODS IN ENGINEERING / BRAZILIAN CONGRESS ON COMPUTACIONAL MECHANICS, São Paulo, Brazil.
- Silva Filho, D.; A. A. F. M. Carneiro, et al. (1999b). Redes RBF Aplicadas à Simulação da Operação de Usinas Hidroelétricas. IV Congresso Brasileiro de Redes Neurais, São José dos Campos-SP, Brazil.
- Silva Filho, D.; A. A. F. M. Carneiro, et al. (2000). Influência das Regras de Operação Sobre Parâmetros de Dimensionamento de Usinas Hidroelétricas. VII SEPOPE - Symposium of Specialists in Electric Operational and Expansion Planning, Curitiba, Brazil.
- Sjöberg, J.; Q. Zhang, et al. (1995). "Nonlinear Black-Box Modeling in System Identification: A Unified Overview." Automatica 31(12): 1691-1724.
- Soares Filho, S. (1987). "Planejamento da Operação de Sistemas Hidrotérmicos." SBA - Controle e Automação 1(2): 122-131.
- Soares Filho, S.; A. A. F. M. Carneiro (1991). "Optimal Operation of Reservoirs for Electric Generation." IEEE Transactions on Power Delivery 6(3): 1101-1107.

- Soares Filho, S.; A. A. F. M. Carneiro (1993). Reservoir Operation Rules for Hydroelectric Power System Optimization. Athens Power Tech - IEEE/NTVA, Athens, Greece.
- Turing, A. (1952). The Chemical Basis of Morphogenesis. Phil. Trans. of the Royal Soc.
- Walter, J.; H. Ritter (1996). "Rapid Learning with Parametrized Self-Organizing Maps." Neurocomputing 12: 131-153.
- Widrow, B.; M. E. H. Jr. (1960). Adaptive switching circuits. IRE WESCON Convention Record.
- Wood, A. J.; B. F. Wollenberg (1984). Power Generation, Operation and Control. New York, John Wiley and Sons.