



UNIVERSIDADE ESTADUAL DE CAMPINAS

FACULDADE DE ENGENHARIA ELÉTRICA

DEPARTAMENTO DE ENGENHARIA DE SISTEMAS

Previsão de Carga no Período de Demanda de Ponta Utilizando Redes Neurais Artificiais

Este exemplar corresponde à redação final da tese
defendida por Wagner da Silva Lima
e aprovada pela Comissão
Julgadora em 07/09/96
Takaaki Ohishi
Orientador

por: Eng^o Wagner da Silva Lima
orientador: Prof^o Dr. Takaaki Ohishi

Dissertação submetida à
Faculdade de Engenharia
Elétrica da Universidade
Estadual de Campinas, para
preenchimento dos pré-
requisitos parciais para
obtenção do Título de Mestre
em Engenharia Elétrica.

Junho de 1996

UNIDADE BC
N.º CHAMADA:
TUNICAMP
L 628 p
V. EF.
T. REJO 67/29.009
PRCO. 667/96
C D
PREÇO R\$ 11,00
DATA 05/11/96
N.º CPD

CM.00099467-8

FICHA CATALOGRÁFICA ELABORADA PELA
BIBLIOTECA DA ÁREA DE ENGENHARIA - BAE - UNICAMP

L628p

Lima, Wagner da Silva

Previsão de carga no período de demanda de ponta
utilizando redes neurais artificiais / Wagner da Silva
Lima.--Campinas, SP: [s.n.], 1996.

Orientador: Takaaki Ohishi.

Dissertação (mestrado) - Universidade Estadual de
Campinas, Faculdade de Engenharia Elétrica e de
Computação.

1. Sistema de energia elétrica. 2. Redes neurais
(Computação). I. Ohishi, Takaaki. II. Universidade
Estadual de Campinas. Faculdade de Engenharia Elétrica e
de Computação. III. Título.

Ao Mestre dos Mestres, **Jesus**, ao qual
o amanhã pertence, dedico este trabalho.

Esta pesquisa teve o suporte financeiro da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (**CAPES**) e da Fundação de Amparo à Pesquisa do Estado de São Paulo (**FAPESP**). Nossos agradecimentos ao Instituto Agrônomo de Campinas (**IAC**) e ao Centro de Ensino e Pesquisa em Agricultura (**CEPAGRI/UNICAMP**).

Agradecimentos

Ao prof. Takaaki pela atenção, amizade, paciência e persistência no desenvolver das etapas que resultaram neste trabalho.

À todos os colegas do DENSIS que conviveram comigo durante a minha estada em Campinas, o Gélson, o Cássio, o Makoto, a Suze, o Walcir, o Inácio, o Vítor, a Márcia e a Bete.

A CAPES que financiou os dois primeiros anos do mestrado e aos professores da Escola de Engenharia Elétrica da UFG que me ajudaram na fase final.

Aos meus pais, que mesmo longe, me deram muita força e incentivo e viabilizaram a minha ida à Campinas.

Aos meus irmãos em Cristo da República da Bença, Alexandre, Andréa, Márcio, Iniza, Gilberto, Claudião, Christiano e Eliene pela maravilhosa companhia, apoio e orações para que eu acabasse esta tese.

Ao pessoal do Kerygma pela imensa força e apoio que deram, em especial ao Cláudio, Karina, Vaneska, Luiz, Ilka e Denilson.

Aos irmãos de Goiânia e Campinas que colaboraram de alguma forma com orações, apoio, hospitalidade e amizade.

Ao Eterno Deus, Excelso e Supremo por sua bênção e por me fazer andar por caminhos maravilhosos demais para mim, toda a minha gratidão. A Ele toda a Honra, Glória e Louvor. Amém!

Resumo

A habilidade de prever precisamente a carga do sistema é vital ao planejamento e operação eficiente, econômica e segura de um sistema de potência. Este trabalho investiga a utilização de redes neurais artificiais para previsão de carga no período de demanda de ponta a curto e curtíssimo prazos.

Dois algoritmos de previsão são testados e avaliados com relação a precisão e esforço computacional. Uma análise da influência de dados climáticos sobre a carga é realizada. Procurou-se encontrar uma arquitetura compacta e robusta que pudesse levar em consideração a sazonalidade da carga anual, sem comprometer a precisão da previsão.

O primeiro algoritmo (*MWS*) utiliza os dados dos últimos dez dias típicos para previsão do perfil de maneira estática e dinâmica. O segundo algoritmo (*AAS*) utiliza os dados históricos do ano anterior para previsão do ano vigente (previsão estática e dinâmica).

O algoritmo *MWS* com previsão dinâmica obteve os melhores resultados para os horizontes de dez minutos (curtíssimo prazo) à frente, uma e 24 horas à frente. Várias dificuldades foram encontradas para considerar a entrada e saída do horário de verão. Apenas a variável temperatura máxima foi a mais significativa em termos de variáveis climáticas. A escassez de dados climáticos mais consistentes no final da tarde impediram uma avaliação mais completa da influência das condições climáticas na previsão.

Os resultados obtidos demonstraram um bom desempenho das redes neurais com erro médio percentual absoluto em torno de 2% para os três horizontes previstos.

Abstract

The ability to accurately predict the system load is vital to the efficient, economic, and secure operation and planning of a power system. This work investigates the use of artificial neural networks for short and very short-term load peak demand forecasting.

Two forecasting algorithms are tested and evaluated based on their precision and computational load. The influence of weather conditions on load demand is investigated. We sought a robust and compact topology which considers annual load sazonality, in order to preserve the forecast precision.

The algorithm (*MWS*) uses data from the last 10 typical days to forecast the load peak demand profile with static and dynamic methods. The second algorithm (*AAS*) uses historical data from the previous year's load and weather database to forecast current year using static and dynamical methods.

The *MWS* algorithm with dynamic forecasting yields the best load peak demand forecasting results for 10 minutes (very short-term forecasting), 1 and 24 hours ahead. The maximum temperature is the most significant weather variable. Scarce consistent evening weather data prevent a more complete evaluation of the influence of weather conditions on load forecasting.

The results show good performance of neural networks with around 2% mean percent absolute error for forecasts on the three horizons evaluated.

Conteúdo

Agradecimentos	iii
Resumo	iv
Abstract	v
Conteúdo	vi
Capítulo 1 - Introdução	1
1.1- Redes neurais artificiais aplicadas à previsão de carga.....	1
1.2- Motivação do trabalho.....	3
1.3- Objetivo.....	4
1.4- Descrição do trabalho.....	4
1.5- Organização da dissertação.....	5
Capítulo 2 - Previsão de Carga	7
2.1- Introdução.....	7
2.2- Os diversos horizontes de previsão de carga.....	8
2.3- A importância da previsão de carga a curto-prazo.....	10
2.4- Modelo trivial para previsão de carga.....	11
2.5- Previsão de carga no Brasil.....	14
2.6- Características da carga do sistema.....	15
2.6.1- Caracterização da carga e influência dos fatores operativos e externos da mesma.....	15
2.6.2- Fatores que influenciam a carga.....	16
2.6.3- Outra classificação.....	19
2.7- Evolução histórica dos procedimentos de previsão de carga.....	20
2.8- Classificação dos modelos de previsão de carga a curto prazo.....	22
2.8.1- Modelos pico de carga.....	22
2.8.2- Modelos perfil de carga.....	23
2.9- Técnicas de previsão.....	24

2.9.1- Regressão linear múltipla.....	24
2.9.2- Alisamento exponencial.....	25
2.9.3- Séries temporais estocásticas.....	26
2.9.4- Método de espaço de estado (ou Filtro de Kalman).....	27
2.9.5- Sistemas especialistas.....	28
2.9.6- Redes neurais artificiais.....	28
2.10- Considerações práticas.....	30
Capítulo 3 - Redes Neurais Artificiais	33
3.1- Fundamentos das redes neurais artificias.....	33
3.2- Histórico.....	34
3.3- O neurônio artificial.....	36
3.4- Função de ativação.....	38
3.5- Topologia das redes neurais artificiais.....	40
3.6- Tipos de treinamento.....	42
3.7- Redes neurais multicamadas.....	43
3.8- Algoritmo de retropropagação do erro (BP).....	45
3.8.1- Problemas com o algoritmo BP.....	50
Capítulo 4 - Previsão de Carga a Curto Prazo Utilizando RNA's	52
4.1- Aplicação de RNA's à previsão de carga a curto.....	52
4.2- Estado da arte.....	52
4.3- Previsão do período de demanda de ponta.....	58
4.4- Algoritmos propostos para previsão da ponta via RNA.....	61
4.4.1- Algoritmo <i>Moving Window (MWS)</i>	62
4.4.2- Algoritmo <i>Ano-Ano (AAS)</i>	63
4.5- Comentários adicionais.....	68
Capítulo 5 - Estudo de Caso	69
5.1- Introdução.....	69
5.2- Período de demanda de ponta para o sistema CESP.....	69

5.2.1- Consideração de efeitos especiais para previsão na hora da ponta.....	71
5.2.1.a- Efeito da hora do pôr-do-sol.....	72
5.2.1.b- Efeito do dia da semana.....	73
5.2.1.c- Efeito dos feriados.....	74
5.2.1.d- Efeito do horário de verão.....	75
5.2.1.e- Comentários adicionais.....	76
5.3- Análise dos dados da CESP.....	77
5.3.1- Dados de carga.....	77
5.3.2- Dados climáticos.....	78
5.3.3- Outros dados.....	78
5.4- Considerações no projeto de RNA's.....	79
5.4.1- Fatores determinantes no desempenho das RNA's.....	79
5.4.2- Controle dos fatores determinantes no desempenho das RNA's..	80
5.4.2.a- Padronização.....	81
5.4.2.b- Normalização.....	81
5.4.2.c- Funções de ativação.....	82
5.4.2.d- Funções de custo.....	83
5.4.2.e- Parâmetros de controle do gradiente descendente e do tempo de treinamento.....	83
5.4.2.f- Dimensionamento da RNA.....	86
5.5- Seleção das variáveis de entrada.....	87
5.6- Testes realizados.....	88
Capítulo 6 - Análise dos Resultados	91
6.1- Introdução.....	91
6.1.1- Método de avaliação dos resultados.....	91
6.1.1.a- Comparação gráfica.....	92
6.1.1.b- Diagramas de dispersão ou <i>scattergrams</i>	93
6.1.1.c- Mean percent error (<i>MPE</i>).....	93
6.1.1.d- Mean absolute percent error (<i>MAPE</i>).....	94

6.1.1.e- Standard deviation absolute percent error (<i>SDAPE</i>).....	95
6.1.1.f- Maximum absolute percent error (<i>MAXAPE</i>).....	96
6.2- Resultados obtidos.....	95
6.3- Comparação com o modelo da CESP.....	102
6.4- Considerações finais.....	103
Capítulo 7 - Conclusão	105
7.1- Conclusões gerais.....	105
7.2- Trabalhos futuros.....	107
Referências Bibliográficas	109
Apêndice A	119

Capítulo 1

INTRODUÇÃO

Este capítulo introduz brevemente a computação neural aplicada à previsão de carga, fornece a motivação e uma descrição geral do trabalho, destaca os objetivos da dissertação e descreve como ela está organizada.

1.1 - Redes Neurais Artificiais Aplicadas à Previsão de Carga

O paradigma neural é uma forma alternativa de computação que procura imitar o funcionamento do cérebro humano na solução de problemas de reconhecimento de padrões. O interesse dos pesquisadores numa forma de computação que se inspirasse no cérebro veio do fato deste último possuir características altamente desejáveis para um sistema artificial, tais como: robustez e tolerância a falhas; flexibilidade; capacidade para lidar com informações probabilísticas, ruidosas ou inconsistentes; processamento paralelo; arquitetura compacta e com pouca dissipação de energia. Além destas características, a capacidade de aprendizado, generalização e associação, motivou o grande interesse por este tipo de computação alternativa [1].

Uma rede neural artificial (RNA) é caracterizada pelo emprego de simples processadores em paralelo e pelo conhecimento armazenado em sua própria topologia final. Além disso, a menos de algumas exceções, muito deste conhecimento é aprendido por exemplos de maneira que estas redes possam adquirir a habilidade de resolver um problema específico. Em outras palavras, RNA's não exigem um algoritmo pré-definido para resolver problemas, mas elas necessitam de exemplos das situações que envolvem tais problemas [2].

RNA's são, como seu próprio nome declara, redes artificiais de elementos neurais. Os elementos neurais, *neurônios*, são simples processadores nem rápidos e nem poderosos quando analisados individualmente. O poder destas estruturas vem de sua utilização de inúmeros destes processadores simples colocados numa topologia de rede

como pode ser visto na figura.1.1. Estas redes podem coletivamente apresentar algumas propriedades interessantes e, em muitos casos, ser modeladas adaptativamente de maneira a aprender os exemplos. Contudo, uma vez treinada, uma RNA pode responder rapidamente e ser muito útil em aplicações em tempo-real [2].

Até o momento RNA's têm sido aplicadas consistentemente em problemas de Engenharia Elétrica [3]. Três dos principais usos desta abordagem são derivados de sua habilidade de previsão, otimização e classificação. Previsão de carga para uma concessionária de energia é um dos exemplos de sua capacidade de predição.

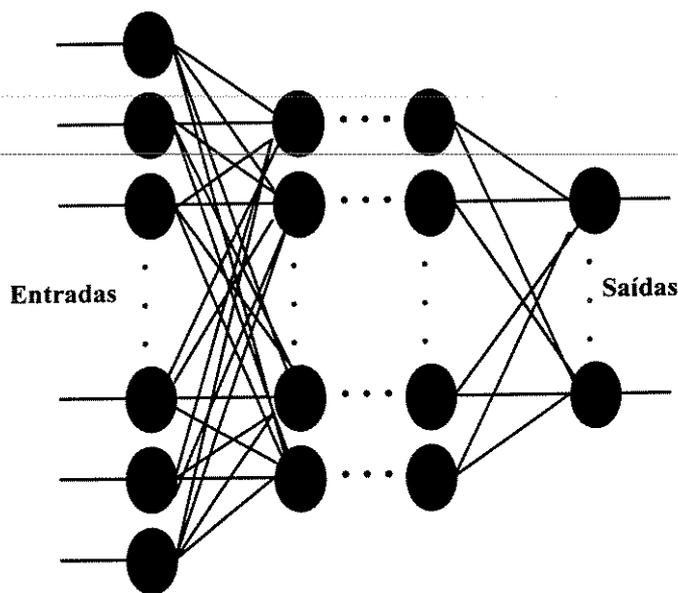


Figura.1.1: Exemplo de uma Rede Neural Artificial Multicamada

As aplicações de RNA's em sistemas de potência podem ser divididas dentro das seguintes categorias [4]:

- previsão de carga
- análise de segurança
- análise de contingência
- diagnóstico e processamento de alarmes
- análise de controle e observabilidade
- modelagem e identificação

Dentre os problemas citados acima, um tem recebido especial atenção dos pesquisadores nos últimos anos: a previsão de séries temporais (previsão de carga a curto-prazo). Afinal, prever o futuro, em especial do comportamento de séries que variam ao longo do tempo, é fundamental em análises e tomadas de decisão e, até hoje, continua sendo um desafio para a estatística e a computação. As RNA's têm se mostrado uma poderosa ferramenta na previsão de séries temporais, em especial em previsão de carga a curto-prazo. Sua habilidade em extrair complicadas relações não-lineares a partir de dados de entrada ruidosos tem produzido resultados surpreendentes, na maioria das vezes melhores que os obtidos por procedimentos estatísticos convencionais [5-7]. Exemplos de aplicações bem sucedidas das RNA's na previsão de carga são encontradas em várias referências. O capítulo 3 apresenta uma breve análise de algumas destas aplicações.

1.2- Motivação do Trabalho

Tradicionalmente, a previsão de carga a curto prazo é feita em base horária, ou seja, a previsão fornece a estimativa da carga média para cada hora do próximo dia. Entretanto, nos intervalos em que há grande variação na demanda de carga, principalmente nos períodos de entrada e saída do período de ponta, a demanda média pode ser muito diferente do valor instantâneo. Para se trabalhar com uma curva de carga mais realista, as empresas estão adotando uma discretização menor nos intervalos de grande variação de demanda, de modo a se obter uma curva mais fiel ao comportamento da carga.

A escassez na literatura em previsão de carga a curto e curtíssimo prazo com discretização menor que uma hora, o grande sucesso do uso de RNA's para a solução de tais problemas e a necessidade de desenvolvimento de algoritmos de previsão de carga com novas metodologias específicas para problemas de previsão no Brasil, nos motivou a estudar a performance do uso de RNA's para a solução da previsão de carga própria no período de ponta com discretização menor que uma hora.

1.3- Objetivo

O objetivo principal desta dissertação é estudar o uso de RNA's na previsão de carga com horizontes de curto e curtíssimo prazo. O trabalho de pesquisa envolve estudos nas áreas de RNA's, análise de variáveis que influenciam a carga e uma análise de quatro algoritmos propostos para prever o período de demanda de ponta com discretização menor que uma hora.

1.4- Descrição do Trabalho

Esta seção apresenta as cinco partes principais da dissertação: o estudo sobre previsão de carga a curto e curtíssimo prazo; redes neurais artificiais; a modelagem das RNA's para a previsão de carga a curto e curtíssimo prazo; a análise de dados reais de uma concessionária de energia elétrica; e o estudo de caso.

O estudo sobre previsão de carga e RNA's compreendeu um resumo dos principais conceitos nas duas áreas, seguido de um levantamento das aplicações mais importantes em previsão de carga e principalmente de RNA's na previsão de carga a curto prazo.

Além de buscar um embasamento teórico necessário ao desenvolvimento de um modelo de previsão de carga, procurou-se a identificação de considerações práticas visando aplicá-las na modelagem de RNA's para previsão de carga a curto e curtíssimo prazo para sua implementação do estudo de caso. Dentre as metodologias analisadas constatou-se a predominância do uso do algoritmo de retro-propagação do erro ou *back-propagation* para o treinamento das redes.

Dentre os vários fatores indicadores de desempenho das RNA's foram considerados três: convergência, generalização e escalabilidade. Para controlar estes fatores utilizou-se mecanismos como: escolha da função de ativação mais adequada ao problema, tendo sido testadas as funções sigmóide e tangente hiperbólica; manipulação

dos parâmetros de controle do gradiente descendente e do número de épocas apresentadas à rede, tais como taxa de aprendizagem e termo de momento; e escolha da arquitetura mais apropriada.

O ambiente de simulação, desenvolvido em MATLAB, teve o objetivo apenas de testar o desempenho dos algoritmos e das redes implementadas não visando fazer um sistema de apoio à previsão ou um sistema previsor em si. Tal *software* pode ser desenvolvido com base nos algoritmos apresentados neste trabalho.

No estudo de caso, modelou-se as RNA's para executar previsões da carga própria do sistema CESP no período de demanda de ponta (para a CESP este período vai das 17 às 22 horas), com horizontes de 10 minutos, 1 hora e 24 horas à frente. Os históricos utilizados contém dados de carga com base horária, carga com discretização de 10 em 10 minutos e dados climáticos.

Os resultados obtidos foram avaliados pelas seguintes métricas de desempenho: comparação gráfica, diagramas de dispersão ou "*scattergrams*", erro médio percentual (*MPE*), erro médio percentual absoluto (*MAPE*), percentagem de previsões entre 0 e 4% e máximo erro percentual (*MAXAPE*).

1.5- Organização da Dissertação

Esta dissertação está dividida em sete capítulos que cobrem: previsão de carga, redes neurais artificiais, modelagem e algoritmos de previsão utilizando RNA's, estudo de caso, avaliação dos resultados e conclusão.

No capítulo 2 é feito um levantamento dos principais conceitos de previsão de carga. O capítulo apresenta os diversos horizontes de previsão de carga e sua importância do ponto de vista do planejamento da operação de um sistema de potência, apresenta um modelo trivial de previsão de carga a curto-prazo, analisa a situação da previsão de carga no Brasil. Em seguida, faz-se uma caracterização da carga e a

influência dos fatores operativos e externos da mesma, além de uma descrição dos modelos e técnicas mais usadas em previsão de carga a curto e curtíssimo prazo.

No capítulo 3 é feito um estudo dos principais conceitos na área de redes neurais artificiais. O capítulo apresenta um breve histórico dos acontecimentos mais marcantes na evolução das RNA's, descreve o funcionamento de um neurônio artificial, das funções de ativação, das topologias e dos tipos de treinamento existentes. Em seguida, faz-se uma descrição mais detalhada do tipo de rede utilizada na dissertação, com ênfase para o algoritmo de treinamento utilizado, o de retro-propagação do erro.

No capítulo 4, apresentamos o estado da arte das aplicações de RNA's a previsão de carga a curto prazo, analisa-se o problema da previsão de carga no horário de demanda de ponta e apresenta os algoritmos de previsão que serão empregados para solucionar o problema citado acima.

No capítulo 5, faremos o estudo da previsão do horário de ponta de carga própria do sistema CESP, onde é feita uma análise dos dados históricos.

No capítulo 6, é feita uma análise dos resultados segundo algumas métricas estatísticas de análise do desempenho da previsão para os horizontes de curto e curtíssimo prazos.

No capítulo final resumimos os resultados obtidos e apresentamos algumas conclusões e sugestões de possíveis trabalhos futuros.

Capítulo 2

PREVISÃO DE CARGA

Este capítulo introduz e discute brevemente as funções e objetivos da previsão de carga, fornece uma análise geral da situação brasileira no setor, destaca os principais modelos clássicos e técnicas de previsão a curto e curtíssimo prazo e avalia algumas questões práticas na modelagem de um sistema de previsão.

2.1 - Introdução

Etimologicamente, a palavra **previsão**, *prae e videre*, sugere que se quer ver uma coisa antes que ela exista. Alguns autores preferem a palavra **predição**, para indicar algo que deverá existir no futuro. Ainda outros utilizam o termo **projeção**. Neste trabalho iremos usar consistentemente a palavra **previsão** com o sentido indicado acima.

É importante salientar que a previsão constitui apenas um meio de fornecer informações para uma conseqüente tomada de decisões, visando a determinados objetivos [8].

O interesse pela previsão se justifica na medida em que a previsão é um subsídio relevante à tomada de decisões. Como decisões criteriosas geralmente proporcionam economias consideráveis, a análise de séries temporais objetivando métodos eficientes de previsão se reveste de grande importância [9].

O acompanhamento da carga de um sistema de potência pela geração em todos os instantes é a exigência básica na operação do mesmo [10]. Para fornecer com boa qualidade a energia elétrica, isto é, de uma maneira segura e econômica, uma companhia elétrica tem que resolver diversos problemas técnicos e econômicos no planejamento e controle da operação de um sistema de energia elétrica. Com o propósito de planejamento e controle ótimos desses sistemas novas técnicas de otimização são

desenvolvidas e aplicadas com uma considerável redução dos custos; em concordância com este objetivo o conhecimento da carga futura desse sistema é o primeiro pré-requisito [10-11].

2.2- Os Diversos Horizontes de Previsão de Carga

Num sistema de potência, a previsão de demanda de energia elétrica visa atender este mercado de maneira segura e econômica.

Os diversos níveis de planejamento da operação, e sua execução, definem o horizonte de interesse da previsão, o qual é usualmente enquadrado nas seguintes categorias:

Longo Prazo: consiste em analisar o nível de consumo do mercado para um horizonte de até 5 anos à frente visando fornecer subsídios para determinar o nível de complementação térmica de modo a minimizar o custo esperado. Podem ser utilizadas técnicas de cenário ou outras para se fazer esta previsão;

Médio Prazo: abrange um horizonte de até um ano à frente, coincidente com a fase de planejamento da operação energética; necessário para determinar uma política de operação dos reservatórios, fixação de cronogramas de manutenção de equipamentos de geração e transmissão, bem como a determinação de contratos de suprimentos de energia e potência entre empresas, de maneira a se atender a demanda prevista com a capacidade disponível, e também estabelecer os preços de médio prazo;

Curto Prazo: é a previsão feita com antecedência desde algumas horas até algumas semanas; utilizada na fase de programação da operação, onde se procura atender as metas estabelecidas no médio prazo, levando-se em conta as variações elétricas e energéticas não esperadas. Nessa fase, é realizada a análise de segurança do sistema nos aspectos elétrico, energético e hidráulico, adequando o cronograma de manutenção elaborado no médio prazo.

A informação de carga para a programação da operação hidráulica do sistema, alocação de unidades, coordenação hidro-térmica e a avaliação de intercâmbios é obtida a partir da previsão de carga a curto-prazo do sistema.

Curtíssimo Prazo: num horizonte de poucos minutos a uma hora à frente, estaremos na fase de operação em tempo-real, e a previsão de carga a curtíssimo prazo é utilizada para detectar condições perigosas de operação e para a análise *on-line* de contingências, possibilitando intervenções corretivas a tempo.

A previsão de carga à curtíssimo prazo é uma informação básica para a função de despacho econômico com horizonte de previsão na ordem de segundos e minutos.

A observação e análise do passado da demanda crescem de importância à medida que o horizonte decresce. Num prazo mais curto, as variações cíclicas do comportamento da carga é mais observável e mais significativa. No entanto, quando chegamos ao curto e curtíssimo prazo, o requisito fundamental passa a ser a adaptabilidade da previsão, considerando-se os impactos climáticos que atuam sobre a carga ou monitorando-se continuamente a variação da carga observada em relação à prevista, de modo a possibilitar ações corretivas.

Quanto mais curto o horizonte de previsão, em geral, maior a necessidade de precisão e detalhe. Outro requisito, nesse horizonte, é a robustez do modelo adotado para a previsão, já que ele deve se recuperar rapidamente de situações não usuais.

Neste trabalho o problema de se determinar o perfil de carga no horário de ponta se situa nos horizontes de curto e curtíssimo prazo que será a partir de agora o foco de nossas discussões. Ao longo da tese nós usaremos ao termo “curto” para indicar horizonte de previsão de uma a 24 horas à frente. E “curtíssimo” para indicar horizonte de previsão de poucos minutos a meia hora à frente.

2.3- A Importância da Previsão a Curto Prazo

Podemos estar interessados em diversos tipos de previsão a curto prazo dependendo de sua utilidade para a programação da operação. Logo, o interesse da previsão de carga a curto-prazo (PCCP) é, tipicamente:

- pico de carga diário do sistema.
- valor da carga numa certa hora do dia.
- curva de carga diária ou semanal (horário ou intervalo menor) do sistema.

A PCCP tem um papel importante na formulação da operação econômica, confiável, segura e estratégica de um sistema de potência. O principal objetivo da PCCP é fornecer uma previsão de carga para:

- a programação básica da geração.
- análise de segurança do sistema de potência em qualquer instante.
- informação contínua ao operador.

A aplicação primária da PCCP é fornecer informações necessárias às funções de programação que determinam a alocação de geração mais econômica consistente com as exigências de confiabilidade, restrições operacionais, políticas, ambientais e limitações do equipamento.

Uma segunda aplicação da PCCP é a avaliação preditiva da segurança do sistema. A previsão de carga é um dado importante para a análise de contingência para a detecção de futuras condições sobre as quais o sistema possa se tornar vulnerável. Esta informação permite aos operadores preparar a ação corretiva necessária para operar o sistema de potência de forma segura.

Uma terceira aplicação da PCCP é fornecer aos operadores do sistema, informações periódicas, isto é, a mais recente previsão de carga, com a última previsão de tempo, sendo que a característica de seu comportamento aleatório deve ser levado em

conta. Os operadores necessitam desta informação para operar o sistema de maneira econômica e confiável. A figura.2.1 resume as maiores aplicações da PCCP.

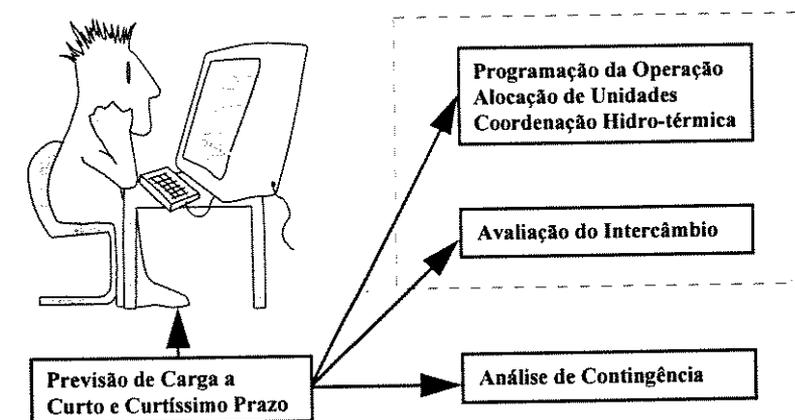


Figura.2.1: Usos da PCCP

2.4- Modelo Trivial para Previsão de Carga

A previsão de carga feita manualmente pelos operadores tem sido trocada por pacotes de *softwares* nos modernos centros de operação de sistema de potência (COS). Os maiores componentes da PCCP do sistema são os modelos de PCCP, as fontes de dados e a interface homem-máquina (IHM).

Os modelos de PCCP implementam a representação do sistema de cargas e os algoritmos de previsão. As fontes de dados são o histórico de carga, as bases de dados sobre as condições climáticas, os parâmetros da base de dados, os dados entrados de forma manual pelos operadores, os dados de tempo-real obtidos pelo COS e a conexão de dados do serviço de previsão do tempo.

Os dados que entram manualmente podem incluir atualizações dos dados do tempo, parâmetros de previsão de carga ou comandos de execução. Em geral, os modelos de PCCP usam dados de cargas integralizadas (MWh ou MWh/h).

A base de dados da telemetria em tempo-real é usada pelo CAG para determinar a carga “medida” que é, tipicamente, integralizada (e conseqüentemente ajustada) antes

de serem usados por um modelo de PCCP. As saídas da PCCP são fornecidas aos terminais de operações e informações dos operadores e à outros setores do COS que necessitam destas informações. Podemos observar melhor este fluxo de informações através da figura.2.2.

A figura.2.2 mostra o esquema de um sistema de PCCP para um COS moderno. Embora possa ser sugerido o sistema de previsão de carga mostrado na figura acima, a preocupação específica deste trabalho está nos algoritmos e na modelagem da carga que farão a previsão, avaliando o problema da previsão do perfil de carga no período de ponta.

Para cada horizonte de previsão, esta irá fornecer subsídios para um tipo de função da programação da operação. Portanto, para cada função podemos ter um horizonte de previsão diferente.

A nível de COS a previsão a curtíssimo prazo tem muita importância, pois fornece uma estimativa para o operador, na operação em tempo-real, tomar alguma decisão que seja importante para operação confiável do sistema, ou avaliar se poderá, nos próximos instantes, haver um nível considerado “perigoso” para a operação do sistema.

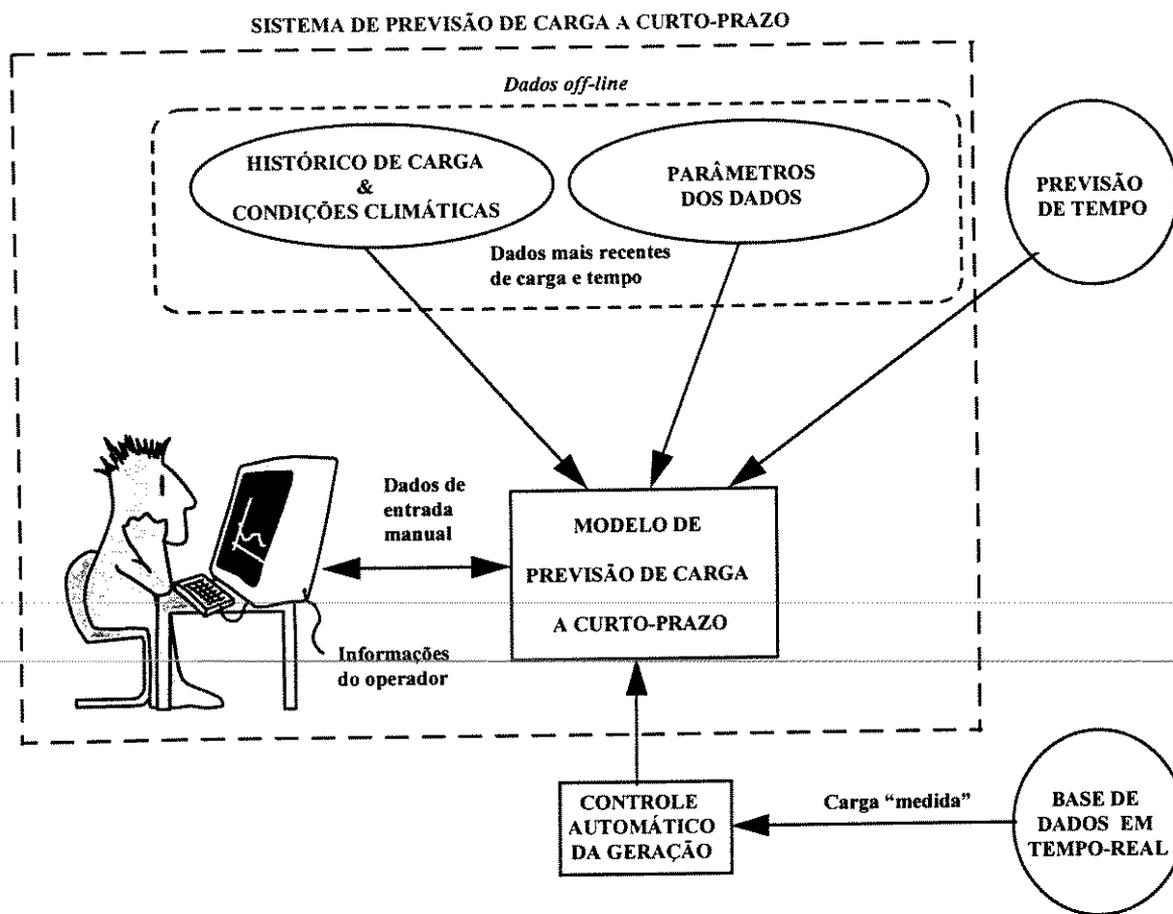


Figura.2.2: Sistema de PCCP

A precisão da previsão de carga tem efeitos significativos sobre a operação do sistema e nos custos de produção. Os operadores precisam antecipar os padrões de carga do sistema tanto quanto ter geração suficiente para satisfazer a demanda. Ao mesmo tempo, níveis suficientes de reserva girante e reserva quente são necessárias para mitigar os impactos da incerteza inerente na previsão e na disponibilidade das unidades de geração.

A capacidade de reserva girante e quente estão colocadas em níveis ditados pela medida de segurança e confiabilidade desejada para a operação do sistema de potência. Assim, pela redução do erro de previsão, os níveis de reserva podem ser reduzidos sem afetar a confiabilidade e confiança do sistema. Desta maneira o custo de produção é reduzido.

Uma previsão pessimista de carga resulta em falha para fornecer a reserva necessária que resultam em custos mais altos devido ao uso de unidades de ponta que são caras. Uma previsão otimista da carga, por outro lado, envolve a partida de muitas unidades resultando num acréscimo desnecessário e daqui um aumento nos custos de operação. No ano de 1985, por exemplo, para o sistema da Inglaterra que é predominantemente térmico, foi estimado que 1% de aumento no erro da previsão foi associado com um aumento nos custos de operação de 10 milhões de libras por ano [12].

2.5- Previsão de Carga no Brasil

O novo cenário do setor de energia elétrica do Brasil apresenta desafios que as empresas terão de superar com rapidez e qualidade, tanto do ponto de vista técnico quanto econômico-financeiro.

A realidade tarifária é apenas o início de um processo que visa modernizar as relações entre os prestadores de um serviço, que é a energia elétrica, e seu mercado consumidor. A partir daí, o conhecimento preciso e antecipado desse mercado se torna fundamental para a garantia de um bom desempenho das concessionárias. Novas tecnologias demandam tempo para absorção, e na área de previsão de carga já se constata um alto nível de carência técnica.

Não investir hoje na modernização e na real capacitação técnica de sistemas de previsão de carga, poderá deixar o setor elétrico, num futuro próximo, muito frágil diante de justificativas dos desvios dessas previsões. Tal fato acarretará erros nos contratos de energia e demanda e conseqüente comprometimento orçamentário das empresas de energia elétrica, além da avaliação incorreta quanto aos níveis de riscos operativos do Sistema Interligado Brasileiro [13].

Recentemente, no Brasil, tem-se feito um esforço de desenvolvimento metodológico e tecnológico do setor de energia elétrica do país. Este esforço tem sido alocado, mais intensamente, do lado da oferta de energia elétrica.

Conseqüentemente, os recursos humanos e técnicos explorados têm sido direcionado para essa área do Planejamento da Operação (PO), em detrimento da área da demanda onde o esforço foi significativamente menor. Nesta área alguns estudos de mercado através das análises macro-econômicas e mais recentemente técnicas de cenários enriqueceram a previsão nessas áreas. Logo, verifica-se que existe um hiato na área de previsão de carga no setor elétrico brasileiro [14].

As atividades relacionadas com a operação de um sistema de energia elétrica são caracterizada pelo encadeamento temporal do planejamento da operação e da operação em tempo-real, bem como suas funções e inter-relações. O objetivo de qualquer uma dessas atividades é atender a carga própria das empresas, ao mínimo custo total da operação para o sistema interligado (N/NE e S/SE), no caso brasileiro.

No Plano da Operação, destaca-se a importância da previsão de carga:

1. nos balanços de energia e demanda que definem os contratos de suprimento entre empresas;
2. na avaliação das disponibilidades de energia e potência;
3. no cálculo das estratégias ótimas de operação;
4. nas análises da segurança da operação do sistema.

2.6- Características da Carga do Sistema

2.6.1- Caracterização da Carga e Influência dos Fatores Operativos e Externos

A curva de carga de uma concessionária espelha, em verdade, toda a sua razão de ser, que é o requisito de energia do seu mercado consumidor. Estudos técnicos relacionando os vários aspectos da empresa com a sua curva de carga têm sido

desenvolvidos há décadas, mas o atual estágio de conhecimento da carga no Brasil, ainda é insatisfatório.

O reconhecimento dos fatores que afetam o perfil da curva de carga é apenas o passo inicial para o aprimoramento de sua análise. Uma vez que o previsor elabore sistematicamente análises das curvas de carga de sua empresa, a partir da classificação das ocorrências abaixo relacionadas, ele poderá iniciar um histórico devidamente segmentado que lhe será extremamente útil em futuras previsões, bem como futuros modelos matemáticos.

A carga do sistema é a soma de todas as demandas individuais em todos os nós do sistema de potência. Em princípio poderia-se determinar o padrão de carga do sistema se cada padrão de consumo individual fosse conhecido. Entretanto, a demanda ou o padrão de uso de uma carga individual (aparelho) ou consumidor é completamente randômico e altamente imprevisível. Também, há uma enorme diversidade de padrões de uso individuais numa certa região. Estes fatores tornam impossível prever o nível de demanda do sistema extrapolando o padrão individual estimado de uso. Felizmente, entretanto, a totalidade das cargas individuais resultam em padrões de consumo distintos que podem ser modelados matematicamente.

2.6.2- Fatores que Influenciam a Carga

O comportamento da carga do sistema é influenciado por um certo número de fatores. Podemos classificar estes fatores dentro de quatro grandes categorias:

- econômico
- temporal
- climático
- efeitos aleatórios

Para modelar a carga do sistema, é necessário entender o impacto de cada classe de fatores nos padrões de consumo de eletricidade. Nós, agora, discutiremos brevemente os efeitos de cada classe.

Fatores econômicos: o ambiente econômico em que a concessionária de energia elétrica opera, tem um efeito claro sobre a demanda de consumo de eletricidade. Fatores tais como, serviço de área demográfica, níveis de atividade industrial, mudanças no setor agrícola, a natureza e o nível de penetração/saturação do aumento da população, desenvolvimentos na regulação de tendências econômicas têm impactos significativos sobre o crescimento/queda da tendência da carga. Além disso, programas de governo também influenciam a carga como o plano real e o plano cruzado, no caso brasileiro.

Tipicamente estes fatores econômicos operam com constante de tempo consideravelmente maior do que uma semana. É importante considerar para estes fatores na atualização dos modelos de previsão de um ano para o próximo ou possivelmente de uma estação para outra. Os fatores econômicos não são, entretanto, explicitamente representados nos modelos de PCCP devido às maiores escalas temporais associadas à eles.

Fatores temporais: três fatores temporais principais (efeitos sazonais, ciclos semanais ou diários e feriados civis e religiosos) têm um importante papel que influencia os padrões de carga. Mudanças sazonais determinam o pico de carga de verão ou inverno. Certas mudanças no padrão de carga ocorrem gradualmente como resposta a variações sazonais tais como, o número de horas de luz do dia e as mudanças na temperatura. Por outro lado, existem eventos sazonais que trazem de forma abrupta modificações estruturais importantes nos padrões de carga. Estes deslocamentos são devido à mudanças na taxa de consumo (hora do dia ou demanda sazonal), início do ano escolar e reduções significativas das atividades durante o período de férias (período Natal-Ano Novo).

A periodicidade diária e semanal da carga é consequência de um padrão de descanso-trabalho do serviço numa área populacional. Há um padrão de carga bem definido para semanas “típicas”.

A existência de feriados tem em geral, o efeito de diminuição significativa dos valores de carga para níveis bem abaixo do normal. Além disso, nos dias que antecedem ou sucedem os feriados, modificações nos padrões de uso são observados devido a tendência de longos finais de semana. Outro fator importante a ser considerado é que o perfil de carga para cada feriado é diferente.

Fatores climáticos: condições meteorológicas são responsáveis por variações significativas nos padrões de carga. Isto é verdade pois muitas concessionárias tem grandes componentes de carga sensíveis ao tempo, tais como, aquelas devido ao aquecimento interno, ar-condicionado e irrigação agrícola.

Em muitos sistemas, a temperatura é a mais importante variável climática em termos de seus efeitos sobre a carga. Para qualquer dia, o desvio da variável de temperatura de um valor normal pode causar mudanças significativas na carga tal que necessite de uma reprogramação no plano da operação inicialmente estabelecido. Além disso, temperaturas passadas também afetam o perfil de carga.

Por exemplo, uma série de dias quentes podem resultar em um aumento de carga que resultará num novo pico de carga do sistema. Para um sistema com clima e geografia não uniformes, muitas variáveis de temperatura ou muitas áreas podem ser consideradas para variações na carga do sistema.

Umidade é um fator que pode afetar a carga do sistema de maneira similar à temperatura, particularmente em áreas quentes e úmidas. Tempestades também tem um forte efeito sobre a carga devido à mudança na temperatura que elas induzem. Outros fatores que tem impacto sobre o comportamento da carga são a velocidade do vento, precipitação e cobertura de nuvens/luminosidade.

Distúrbios aleatórios: agrupamos sobre esta classificação uma variedade de eventos aleatórios que causam variações nos padrões de carga que não podem ser explicados em termos dos fatores discutidos anteriormente.

Um sistema de potência é continuamente sujeito a distúrbios aleatórios refletindo o fato de que a carga do sistema é composta de um grande número de demandas individuais. Além do grande número de distúrbios muito pequenos, existem grandes cargas (mineradoras, grandes industriais, siderúrgicas), cuja operação pode causar grandes variações no uso de eletricidade. O horário de operação destes equipamentos são usualmente desconhecidos para os operadores do sistema, portanto eles representam grandes distúrbios imprevisíveis. Há também certos eventos tais como grandes greves e programas especiais na televisão cuja ocorrência é conhecida *a priori*, mas cujo efeito sobre a carga é incerto.

2.6.3- Outra Classificação

Na literatura podemos encontrar outro tipo de classificação dos fatores que influenciam a carga tal como vemos abaixo. Portanto, podemos ainda segmentar o histórico de carga analisando o perfil de carga segundo a seguinte classificação:

- origem
- tipo
- duração
- efeito

Quanto a origem, podemos classificar em:

- Sócio-cultural (jogos de copa do mundo, programas de televisão)
- Econômica (o plano real, tarifação de energia)
- Climática (temperatura, umidade, luminosidade)
- Política legal (horário de verão, propaganda eleitoral gratuita e outros)

Quanto ao tipo, podemos classificar em:

- Localizado - afeta apenas parte da curva de carga (desligamento de 3 horas).
- Disperso (geral) - afeta toda a curva de carga, ou a sua maior parte (plano real).

Quanto à duração, podemos classificar em:

- Efetivo (fornecimento horo-sazonal)
- Esporádico (propaganda eleitoral gratuita e entrada de frente fria)
- Periódico (horário de verão)
- Único (desligamento)

Quanto ao efeito (modulação), podemos classificar em:

- Transportador - apenas modula a curva (propaganda eleitoral gratuita)
- Redutor - reduz a energia total (greve de um grande consumidor)
- Incrementador - aumenta o consumo (plano real)

Embora, como podemos observar acima, muitos fatores possam influenciar a carga e determinar o nível de consumo de energia elétrica, a escolha das variáveis que mais sensibilizam a carga é que vai determinar o sucesso da previsão dentro dos limites exigidos.

2.7- Evolução Histórica dos Procedimentos de Previsão de Carga

Historicamente, métodos de previsão de carga têm sido utilizados desde que se constituíram companhias de energia elétrica para atender a demanda de consumidores.

Desde então percebeu-se que algumas variáveis climáticas exercem influência sobre a carga, algumas vezes de forma acentuada. Estudos de regressão carga-clima [15, 16] foram então implementados e utilizados, até que se notou que um modelo estático, como era o caso desses modelos de regressão, não assegura que no futuro a qualidade da previsão é mantida.

Em busca de adaptabilidade, passou-se a trabalhar com abordagens de séries temporais, extrapolando-se para o futuro uma descrição dos ciclos sazonais da série de carga, tais como, por exemplo, funções características, análise harmônica e alisamento exponencial [17-20]. Variáveis climáticas, nesse enfoque, não são consideradas.

Posteriormente, surgiram trabalhos que identificam, na carga observada, uma curva padrão de características cíclicas acrescida de um resíduo que sofre influência climática [8, 16, 19, 21]. Passou-se, simultaneamente com a abordagem de séries temporais, a associar à previsão uma medida de credibilidade, fornecendo a margem esperada de dispersão da carga observada futura em relação à prevista.

Mais recentemente, têm sido aplicados métodos de séries temporais baseados nos modelos ARIMA [9, 22-25], e a modelagem de Função de Transferência [23], ambas desenvolvidas por BOX & JENKINS [26], bem como as técnicas de espaços de estados usando Filtro de Kalman para estimação recursiva do estado [14, 27, 28]. Essas abordagens têm sido, via de regra, aplicadas ao resíduo resultante da identificação da carga padrão. Grande parte dos trabalhos publicados nos últimos vinte anos procura levar em consideração tanto as características da série histórica de carga quanto a influência de variáveis climáticas.

A maioria dos métodos baseia a previsão numa análise retrospectiva dos dados, não levando em conta o fato presente e suas conseqüências futuras. Previsões cujo modelo de carga foi ajustado num histórico demasiadamente grande podem ser prejudicadas, em determinado instante, pois muitas características da carga passada podem deixar de ser significativas para a previsão da carga atual.

Os efeitos aleatórios, difíceis de expressar quantitativamente, necessitam de distribuições de probabilidade para descrever sua incerteza.

Um modelo que opera somente com a série histórica, transformando mecanicamente dados de entrada em informações de saída, não pode prever

acontecimentos desta natureza e corre o risco de, a partir de determinado instante, se tornar inadequado. Surge, então, a necessidade de métodos mais flexíveis que permitam incorporar à série histórica informações transmitidas pelo analista e/ou possuam uma estrutura adaptativa que possibilite uma percepção rápida de mudanças externas, expressando-as convenientemente em termos quantitativos [9].

Com o intuito de se considerar as variáveis climáticas, algumas tentativas já foram feitas utilizando sistemas especialistas e redes neurais artificiais. Estas novas metodologias são capazes de associar uma modelagem não-linear com regras extraídas de um especialista (operador) para compor o modelo de previsão.

2.8- Classificação dos Modelos de Previsão de Carga a Curto-Prazo

Podemos classificar os modelos de previsão de carga de diversas formas. Basicamente, na literatura há dois tipos básicos de modelos de previsão a curto-prazo. Os *modelos pico de carga* e os *modelos perfil de carga*.

2.8.1- Modelos Pico de Carga

Nos modelos pico de carga apenas o pico diário e semanal é modelado, usualmente como uma função das variáveis climáticas. O tempo não tem parte em tais modelos que são tipicamente da forma: **pico de carga = carga base + componente dependente do clima**, onde a carga base é uma média da componente de carga insensível ao clima e a outra componente é sensível ao clima. As variáveis climáticas podem incluir a temperatura na hora do pico de carga ou uma combinação de temperaturas previstas e históricas. Umidade, intensidade luminosa, velocidade do vento e precipitação também são considerados nestes modelos. Exemplos de modelos de pico de carga podem ser encontrados em [17, 20, 29, 30, 31].

As vantagens do modelo de pico de carga estão na simplicidade de sua estrutura e a sua pouca demanda de dados para inicializar e atualizar os parâmetros do modelo.

Estes são estimados através de regressão linear ou não-linear. As desvantagens de tais modelos são tais que eles não definem a hora que o pico de carga ocorre, nem fornecem qualquer informação sobre o perfil da curva de carga. Desde que estes modelos são essencialmente estáticos, fenômenos dinâmicos como a correlação entre os períodos não podem ser previstos.

2.8.2- Modelos Perfil de Carga

Tais modelos descrevem a carga como uma série temporal discreta sobre os intervalos de previsão. O intervalo de tempo de carga amostrada é tipicamente uma hora ou meia-hora para o curto prazo ou menos de meia-hora para o curtíssimo prazo (10 minutos no nosso caso), enquanto que a quantidade medida é geralmente a energia consumida sobre o intervalo de amostragem em MWh ou MWh/h.

Muitas técnicas de previsão de carga descrevem o perfil de carga incluindo o pico de carga. Entretanto, o pico de carga é difícil de prever com muita precisão. Assim combinar modelos de perfil de carga e pico de carga pode ser desejável.

Basicamente existem dois tipos de modelos de perfil de carga: *modelos hora-do-dia* e *modelos dinâmicos*. Combinar estes dois modelos também é possível.

- **Modelos Hora-do-dia:** Os modelos hora-do-dia definem a carga $z(t)$ em cada tempo discreto t do período de previsão de duração T por uma série temporal $\{z(t), t=1, 2, \dots, T\}$. Em sua forma mais simples, o modelo hora-do-dia armazena T valores de carga baseado no comportamento da carga observado anteriormente.
- **Modelos Dinâmicos:** Os modelos dinâmicos de carga reconhecem o fato que a carga não é apenas uma função da hora do dia, mas também do seu comportamento mais recente, bem como de suas entradas aleatórias e climáticas.

A modelagem com redes neurais que utilizaremos neste trabalho pode ser classificada como um modelo perfil de carga dinâmico.

2.9- Técnicas de Previsão

A classificação dos modelos de previsão de carga para médio, curto e curtíssimo prazo, segundo as técnicas utilizadas, vem ganhando grande ênfase na literatura. Dentro desta ótica, os modelos podem ser classificados como:

1. **Regressão linear múltipla.**
2. **Alisamento exponencial.**
3. **Séries temporais estocásticas.**
4. **Métodos de espaço de estado (ou Filtro de Kalman).**
5. **Sistemas especialistas.**
6. **Redes neurais artificiais**

2.9.1- Regressão linear múltipla

Nesses modelos a carga é expressa em termos de variáveis exógenas que podem ser climáticas (temperatura, luminosidade, umidade do ar, etc.) e não climáticas (número de consumidores, atividade econômica, renda, etc.). A expressão utilizada para representar a carga é dada por:

$$Y(t) = a_0 + a_1x_1(t) + a_2x_2(t) + \dots + a_nx_n(t) + \varepsilon(t) \quad (2.1)$$

onde: $Y(t)$ = carga elétrica no instante t .

$x_i(t)$ = variáveis exógenas.

$\varepsilon(t)$ = variável aleatória com média zero e variância constante.

As variáveis explicativas x_i são identificadas através de uma análise de correlação entre as variáveis e a carga. Quando o número de tais variáveis candidatas é muito grande, uma técnica estatística conhecida como *Stepwise* [32] é utilizada para selecionar, entre todas, as variáveis significativas. A estimação dos coeficientes é feita geralmente por métodos de mínimos quadrados e os testes estatísticos tais como *teste-F* e *t-Student* são usados para avaliar o nível de significância dos coeficientes a serem utilizados pelo modelo. Muitos trabalhos utilizaram esta técnica para previsão no horizonte de médio prazo [12, 17-19, 22, 29, 33, 34]. É importante destacar que o uso de modelos de regressão para previsão de carga exige que se faça previamente a previsão das variáveis exógenas, o que tem limitado o uso dessa técnica.

2.9.2- Alisamento Exponencial

Nestes modelos usa-se um conjunto de funções temporais $f(t)$ para modelar a carga $Y(t)$ no instante t , como:

$$Y(t) = \beta^T(t) \cdot f(t) + \varepsilon(t) \quad (2.2)$$

onde $\beta(t)$ é um vetor de coeficientes que são estimados usando-se métodos de mínimos quadrados ponderados para os N intervalos amostrados mais recentemente, "T" é o operador transposto e $\varepsilon(t)$ é uma variável aleatória. Uma característica desses modelos é que as estimativas dos coeficientes e as previsões podem ser atualizadas recursivamente pelas respectivas equações:

$$\hat{\beta}(N+1) = L^T \cdot \hat{\beta}(N) + F^{-1} \cdot f(0) \cdot [Y(N+1) - \hat{Y}(N)] \quad (2.3)$$

$$\hat{Y}(N+1+l) = f^T(l) \cdot \hat{\beta}(N+1) \quad (2.4)$$

onde F e L são matrizes a serem construídas em função do histórico passado da carga e a função $f(t)$ deve satisfazer a relação $f(t)=L \cdot f(t-1)$. Os trabalhos que utilizaram esse modelo para previsão no horizonte de curto-prazo foram [36, 37]. Gupta *et al* [24] complementaram essa técnica com o envolvimento de variáveis climáticas.

2.9.3- Séries Temporais Estocásticas

Esse é um modelo dos mais populares aplicado à previsão de curto prazo, tendo sido também aplicado com sucesso em previsões de longo e médio prazo como mostram [20, 27, 35]. Nesta técnica a carga é modelada como a saída de um filtro linear cuja entrada é uma série temporal estocástica $a(t)$ com média zero e variância constante desconhecida, usualmente chamado de ruído branco, como mostra a figura.2.3. Os modelos ajustados são conhecidos como modelos ARIMA(p,d,q) (Auto-Regressivo Integrado de Médias Móveis). No caso mais geral, supondo-se que a carga possui componentes sazonais estocásticas, são ajustados modelos SARIMA(p,d,q)(P,D,Q)^S, onde (p,P) determinam as ordens das parcelas auto-regressivas e auto-regressiva sazonal do modelo, (q,Q) determinam as ordens das parcelas de médias móveis e médias móveis sazonal e S é o período.

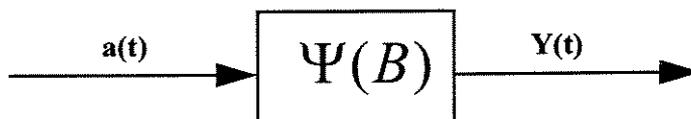


Figura.2.3: Modelo de Séries Temporais para Carga

Aqui $\Psi(B)$ é a função de transferência do filtro, B é o operador retardo tal que $B.Y(t)=Y(t-1)$.

No caso mais geral a relação entre a carga $Y(t)$ e o ruído é dada por [26]:

$$\phi(B). \Phi(B^S). \nabla^d . \nabla_s^D . Y(t) = \theta(B). \Theta(B^S). a(t) \quad (2.5)$$

onde $\phi(B)$ e $\Phi(B^S)$ são polinômios de graus p e P respectivamente, e $\theta(B)$ e $\Theta(B^S)$ são polinômios de graus q e Q respectivamente. ∇ é o operador diferença de ordem definido por $\nabla Y(t) = Y(t) - Y(t-1)$ ou usando a definição do operador retardo tem-se $\nabla Y(t) = (1-B).Y(t)$. Conseqüentemente a diferença de ordem d é definida como $\nabla^d Y(t) = (1-B)^d . Y(t)$ onde d é um número inteiro qualquer. Uma definição similar é

feita para $\nabla_s^D = (1 - B^s)^D$. Tais operadores são necessários quando a série $Y(t)$ é não estacionária, ou seja apresenta tendência ou sazonalidade.

Os trabalhos que aplicaram essa técnica para previsão de curto-prazo são [15, 16, 28, 38-40].

2.9.4- Método de Espaço de Estado (ou Filtro de Kalman)

Este método pode ser visto como uma generalização das duas técnicas apresentadas, permitindo representar modelos com coeficientes variáveis. Nesse método a carga é modelada como uma variável de estado de um sistema linear descrito por dois sistemas de equações, as equações de estado e de observação, são expressas como:

$$X(t+1) = A(t) X(t) + W(t) \quad (2.6)$$

$$Y(t) = H(t) X(t) + V(t) \quad (2.7)$$

onde $X(t)$ é o vetor de estado, $A(t)$ é a matriz de transição de estado $W(t)$ e $V(t)$ são ruídos brancos com matrizes de covariância $Q(t)$ e $R(t)$ respectivamente conhecidas. $Y(t)$ é a carga medida no instante t e $H(t)$ é a matriz que relaciona $X(t)$ com $Y(t)$. A estimativa *a posteriori* do processo é dada por [41]:

$$X(t/t) = X(t/t-1) + K(t) [Y(t) - H(t) X(t/t-1)] \quad (2.8)$$

onde:

$$K(t) = P(t/t-1) H(t)^T [H(t) P(t/t-1) H(t)^T + R(t)]^{-1} \quad (2.9)$$

$$P(t/t) = [1 - K(t) H(t)] P(t/t-1) \quad (2.10)$$

A previsão *a priori* é dada por:

$$X(t+1/t) = A(t) X(t/t) \quad (2.11)$$

$$P(t+1/t) = A(t) P(t/t) Q(t)^T + Q(t) \quad (2.12)$$

onde $K(t)$ é denominado ganho do filtro de Kalman e $P(t/t-1)$ é a matriz de covariância do vetor de erro entre o valor atual e a estimativa do processo $\varepsilon(t/t-1) = X(t) - X(t/t-1)$. É claro que métodos de espaço de estado tornam-se muito atrativos para previsões em tempo real devido à prioridade recursiva do filtro de Kalman. As principais dificuldades de aplicação desse método estão na identificação do sistema, em particular as matrizes de covariância $Q(t)$ e $R(t)$. Alguns trabalhos que utilizaram esse método são [25, 42-44].

2.9.5- Sistemas Especialistas

A avaliação do operador determina a previsão final e explícita que as fórmulas matemáticas são inadequadas para descrever o mecanismo de modelagem. Esta é uma área potencial para a aplicação de sistemas especialistas que emulariam as regras seguidas pelo operador.

Sistemas especialistas são uma nova técnica que tem surgido como resultado do avanço no campo de inteligência artificial (AI) nas últimas décadas. Em resumo, um sistema especialista é um programa computacional (não um algoritmo), o qual tem a flexibilidade de permitir a incorporação de regras heurísticas provenientes da experiência de especialistas (operadores). Isso significa que esse sistema pode analisar, fazer conjecturas e ter sua base de conhecimento expandida a medida que novas regras são elaboradas. Trabalhos que utilizaram esse tipo de técnica foram [45-47].

2.9.6- Redes Neurais Artificiais

Nas últimas duas décadas, esforços consideráveis têm sido feitos para desenvolver aproximações razoáveis para prever cargas horárias com o horizonte de um dia ou uma semana à frente. Estes esforços também são divididos em duas categorias; as aproximações estatísticas e as aproximações por sistemas especialistas.

Existe uma longa história de aplicações de redes neurais à sistemas de potência; a primeira aplicação apareceu em 1975 quando Dillon *et al* [48], apresentou um trabalho

sobre o uso de máquinas auto-organizáveis para previsão de carga na conferência PSCC-5 em Cambridge. Depois disso houve uma longa pausa e somente a partir de 1988 novos trabalhos apareceram. Uma breve discussão sobre estes trabalhos serão vistos no capítulo 4 seção 4.2.

Uma análise comparativa às técnicas apresentadas anteriormente é feita em [49]. No entanto tal análise certamente jamais poderá ser feita de uma forma conclusiva uma vez que uma técnica que possa parecer inadequada em uma determinada situação ou localidade poderá vir a ser a técnica mais adequada em outra circunstância.

Por exemplo, técnicas que utilizam análise de regressão são extremamente adequadas quando se quer avaliar efeitos climáticos na carga. No entanto, em certas regiões do Brasil sabemos que a influência climática é muito pequena no consumo de energia elétrica quer seja por causa da pouca variabilidade de temperatura, quer seja por causa da baixa renda per capita no país, que acarreta um baixo estoque de aparelhos de refrigeração (ar-condicionados) ou de aquecimento.

De um ponto de vista macroscópico, as técnicas de previsão para médio e curto prazo podem ser reagrupadas em duas categorias. A primeira delas considera a carga como sendo composta de duas componentes, uma componente determinística que pode ser expressa em função do tempo ou em termos de variáveis exógenas e uma outra componente de natureza estocástica. À componente determinística é dado o nome de carga padrão e à componente estocástica é dado o nome de carga residual. Essa classe de modelos denomina-se modelos que utilizam carga padrão.

A segunda categoria de modelos são os que consideram a carga como um processo totalmente estocástico. Nessa segunda classe de modelos é possível também incorporar variáveis exógenas através da teoria de processos estocásticos multivariados.

2.10- Considerações Práticas

Até aqui foram discutidos, especialmente, aspectos práticos na formulação e seleção do modelo e algoritmo de previsão. A performance do modelo pode ser continuamente monitorada; entretanto, uma vez estabelecida uma estrutura confiável do modelo, uma deterioração da performance do modelo pode ser corrigida primeiro por um ajuste fino dos parâmetros do modelo. Mudanças na estrutura do modelo não são feitas freqüentemente, mesmo que a escolha do próprio modelo já tenha sido feita.

As exigências computacionais associadas com a previsão de carga a curto e curtíssimo prazos são modestas. Um disco de armazenamento de dados precisa estar disponível para os dados do histórico de carga e condições climáticas usados para inicialização do modelo de previsão e subsequentemente para a atualização.

Na fase de seleção e formulação do modelo, a primeira consideração a ser feita é o objetivo da previsão, isto é, a natureza da previsão, o horizonte de previsão desejado e a intenção do uso da previsão. Um modelo de previsão pode ser exigido para a previsão dos picos diários de carga do sistema, valores da carga do sistema para uma hora específica e energia diária e/ou semanal do sistema. O uso de um modelo que agregue várias técnicas de previsão tende a fornecer um sistema eficaz de verificações e resultados num aumento da confiabilidade da previsão.

Para um modelo em particular, o bom senso é o teste primeiramente aplicado. Mas questões básicas devem ser respondidas, tais como: O modelo faz sentido? Todos os fatores que afetam a carga do sistema estão expostos de maneira explícita ou implícita? Estas questões poderiam receber respostas afirmativas antes de uma análise a mais.

Uma consideração importante na formulação e/ou seleção de modelos de previsão apropriados é a parcimônia do modelo, como o número de variáveis independentes, a facilidade da previsão e a incerteza de cada variável independente e o número de parâmetros ajustáveis.

Em geral, modelos com poucas variáveis independentes e parâmetros ajustáveis são preferíveis. Tais modelos são de fácil inicialização, atualização, modificação e operação.

Uma consideração a mais na formulação/seleção do modelo é a demanda de dados. A demanda de dados associada com vários modelos é fortemente ligada com a natureza do modelo.

Existe um dilema na demanda de dados. Por um lado, é desejável desenvolvê-la como uma relação tão permanente quanto possível entre as variáveis dependentes e independentes. Isto requer, necessariamente, um conjunto de dados do histórico cobrindo um longo período. Isto é evidenciado em sistemas cujo comportamento da carga é muito estável de ano para ano.

Por outro lado, há a necessidade do modelo ser flexível o bastante para considerar quaisquer mudanças no comportamento da carga. Isto impõe a necessidade de um pequeno conjunto de dados cobrindo apenas o mais recente período, tal que qualquer padrão que não ocorra mais devido às mudanças de mercado possa ser excluído do histórico. Esta consideração pode ser especialmente relevante para programas de gerenciamento da demanda, em geral.

Os modelos com intensivo uso de muitos dados tem um impacto negativo nos aspectos de uso e atualização destes dados. Geralmente modelos com menores necessidades de dados são preferíveis. E ainda, precisa-se ter em mente que a base de dados precisa também conter informação sobre dias especiais do ano que tem uma periodicidade anual.

Algumas questões estão claramente envolvidas na formulação de um modelo de carga ou para um modelo particular de uma concessionária. Dado o estado do campo da modelagem e previsão de carga hoje, é, entretanto, razoável tentar desenvolver um

modelo com capacidade de descrever a forma de carga, bem como as condições climáticas e aspectos dinâmicos e aleatórios.

Modelos que descrevem apenas o pico de carga, ou que não modela explicitamente os efeitos climáticos, embora mais simples de desenvolver e atualizar, não oferecem a flexibilidade e a precisão dos modelos mais gerais. No estágio de formulação do modelo, pode-se diminuir a escolha de modelos para aqueles mais adequados às necessidades do usuário baseados no tipo de dados e facilidades computacionais disponíveis.

Pode-se também ser notado que a maioria dos modelos podem ser identificados e operados dentro de necessidades aceitáveis de computação e de dados, tal que este critério é provavelmente não crítico. O último critério será então, a performance do modelo de previsão com dados atuais, onde algumas coisas são difíceis de prever sem experimentação.

Antes do procedimento com a fase de identificação, a carga precisa ser examinada para comportamentos anormais que podem ser causados por eventos tais como greves, *blackouts*, dias de eleição ou programas especiais na TV. Tal “comportamento anormal” precisa ser identificado e retirado da base de dados inicial. Neste estágio, a entrada da experiência em previsão de carga dos operadores é essencial.

No próximo capítulo apresentaremos algumas discussões sobre redes neurais artificiais.

Capítulo 3

REDES NEURAIS ARTIFICIAIS

Este capítulo apresenta um breve histórico sobre redes neurais, os fundamentos teóricos das redes neurais artificiais e suas principais características, além de fornecer o processo de aprendizagem pelo algoritmo de retropropagação de erro e discutir os problemas inerentes a ele.

3.1 - Fundamentos das Redes Neurais Artificiais

As redes neurais artificiais (RNA's) têm este nome porque são inspiradas em modelos biológicos, isto é, elas são compostas por elementos que reproduzem algumas funções elementares do neurônio biológico [50].

A grande motivação do estudo de RNA's é a possibilidade de processamento em paralelo e de forma distribuída (levando a um aumento de velocidade). Outra característica das RNA's é que as mesmas não precisam ser programadas, pois se auto-organizam, aprendendo tarefas associativas e classificatórias através do treinamento repetitivo (supervisionado ou não) usando exemplos ou conjunto de dados. As aplicações das RNA's, em geral, são voltadas para reconhecimento de padrões que computadores convencionais não executam com facilidade [51].

A computação neural é um paradigma computacional alternativo da computação usual (baseada na sequência de instruções programadas), que foi introduzida por Von Newmann e tem sido a base de quase todas as máquinas computacionais até hoje. A computação neural é inspirada pelo conhecimento da neurociência, embora não tente ser biologicamente realista a nível de detalhes [52].

3.2- Histórico

A história desta sorte de idéias em psicologia origina-se com Aristóteles. Como uma base para modelagem computacional nós podemos considerar o artigo de McCulloch e Pitts em 1943 [53], que introduziu o modelo de processamento paralelo, como o mostrado na figura.3.1, como o marco inicial do estudo de redes neurais.

Por volta de 1960 havia muita atividade de pesquisas principalmente com a equipe de Frank Rosenblatt, focalizando o problema de como encontrar pesos apropriados w_{ij} para tarefas computacionais específicas. Eles concentraram o seu trabalho nas rede chamadas *perceptrons*, onde as unidades foram organizadas dentro de camadas com conexões diretamente alimentadas entre uma camada e a próxima. Um exemplo pode ser visto através da figura.3.1. Redes muito similares chamadas *adelines* foram inventadas por volta da mesma época por Widrow e Hoff (1960 e 1962) [54, 55].

Para a mais simples classe de *perceptrons* sem qualquer camada intermediária, Rosenblatt [56] foi capaz de conseguir a convergência de um **algoritmo de treinamento** (um meio de alterar os pesos das sinapses iterativamente tanto quanto fosse necessário para uma tarefa desejada).

Rosenblatt demonstrou um teorema sobre aprendizado do *perceptron* em 1962. O teorema de aprendizado do *perceptron* de Rosenblatt pode ser enunciado como: **“Um *perceptron* pode aprender qualquer coisa que possa representar”**. É importante distinguir aprendizado de representação. Representação refere-se a habilidade de um perceptron (ou outra rede) de simular uma função específica. Aprendizado requer a existência de um procedimento sistemático para ajustar os pesos da rede para produzir aquela função [50].

Muitas pessoas expressaram grande interesse e entusiasmo e esperança de que tais máquinas pudessem ser a base da inteligência artificial.

Foi “um duro golpe” para o teorema de aprendizagem do *perceptron*, forçosamente apontado por Minsky e Papert [57] em 1969 em seu livro “*Perceptrons*”, a seguinte descoberta: o teorema obviamente se aplicava apenas aos problemas cuja estrutura é *capaz de aprender*. Minsky e Papert mostraram que alguns problemas computacionais não poderiam ser resolvidos pelo *perceptron* de uma única camada de Rosenblatt. O simples problema do OU-EXCLUSIVO por exemplo não podia ser resolvido.

Rosenblatt tinha estudado também estruturas com mais camadas de neurônios e acreditava que elas poderiam superar as limitações dos *perceptrons* simples. Entretanto, não havia um algoritmo de treinamento conhecido que poderia determinar os pesos necessários para se aprender uma tarefa. Minsky e Papert duvidaram que este algoritmo pudesse ser encontrado e considerou ser preferível explorar outras abordagens para a inteligência artificial. Com isto a maioria da comunidade científica deixou o paradigma das redes neurais por mais de vinte anos.

Entretanto houveram alguns pesquisadores que continuaram a desenvolver a teoria de redes neurais na década de 70, tais como Anderson, Moser, Willshaw, Marr e Kohonen. Grossberg fez uma compreensiva reformulação no problema geral de aprendizado das redes. Marr desenvolveu teorias de redes do cerebelo, neocortex cerebral e hipocampo, designando funções específicas para cada tipo de neurônio.

Uma outra linha de pesquisa pode ser traçada por Cragg e Temperley. Eles reformularam a rede de McCulloch-Pitts. Entretanto, a influência vinda dos *perceptrons* de Rosenblatt ainda vigorava nesta época pois, baseados em suas teorias vários pesquisadores foram desenvolvendo algoritmos para o ajuste dos pesos das redes neurais multi-camada, culminando com o surgimento do algoritmo de retro-propagação do erro ou *backpropagation* (BP).

Após a descoberta do BP o interesse na área de redes neurais voltou a crescer, desta vez de forma mais madura e consciente. Nos últimos anos, toda teoria de redes neurais tem sido transformada em aplicações, sendo que novas corporações dedicadas à

comercialização desta tecnologia têm aparecido. O crescimento nas atividades de pesquisa tem sido astronômico e novas áreas de atuação das redes neurais têm surgido, tais como: processamento de imagens e voz, reconhecimento de padrões, diagnósticos médicos, sistemas de potência e previsões de séries temporais [7, 58].

3.3- O Neurônio Artificial

O cérebro humano é composto por cerca de 10^{11} **neurônios** de muitos tipos diferentes. A figura.3.1 mostra um esquema genérico de um neurônio simples. Um emaranhado de fibras nervosas chamadas **dendritos** estão conectadas ao **corpo celular** ou **soma**, onde o núcleo celular está localizado. Como uma extensão do corpo celular uma única fibra longa chamada de **axônio**, cujas extensões formam **arcos** ou **arborizam** em várias ramificações. No final destes estão os transmissores das **junções sinápticas**, ou simplesmente **sinapses**, que vão para outros neurônios. Os receptores destas junções em outras células podem ser encontradas nos dendritos e nos próprios corpos celulares. O axônio de um neurônio típico faz muitas milhares de conexões com sinapses de outros neurônios.

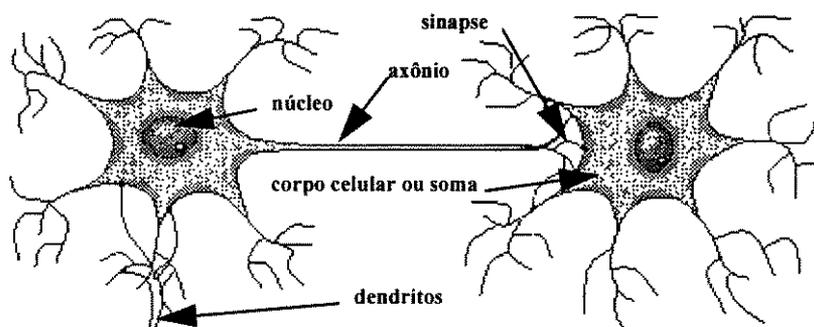


Figura.3.1: Esquema de dois neurônios típicos interligados

A transmissão do sinal de uma célula a outra é feita através de um complexo processo químico no núcleo onde transmissores de substâncias específicas são acumuladas de um lado da junção. O efeito é o aumento ou diminuição do potencial elétrico dentro do corpo celular da célula receptora. Se este potencial alcança um limite, um pulso ou **ação potencial** de intensidade e duração fixa é enviada através do axônio.

Nós então dizemos que a célula foi “disparada”. O pulso propaga através da arborização axonal para as junções sinápticas de outras células. Depois do disparo, a célula tem um tempo de espera chamado **período refratário**. Depois a célula pode novamente ser “disparada”.

O neurônio artificial foi projetado para imitar o princípio de funcionamento de um neurônio biológico como foi descrito acima. Em suma, um conjunto de entradas (V_j) são aplicadas ao neurônio artificial, cada uma representando a saída de outros neurônios ou uma informação externa fornecida à rede como mostra a figura.3.2. Cada entrada é multiplicada por um peso correspondente (w_{ij}), peso da sinapse, gerando entradas ponderadas, de forma análoga à capacidade das *sinapses biológicas*. Então é feito a soma ponderada destes sinais, obtendo-se o valor h_i que será comparado com o valor limite, limiar (l_i), para ativação do neurônio ($V_i = g(h_i)$). Caso este valor exceda o valor do limiar do neurônio, ele se ativará, caso contrário ele ficará inativo.

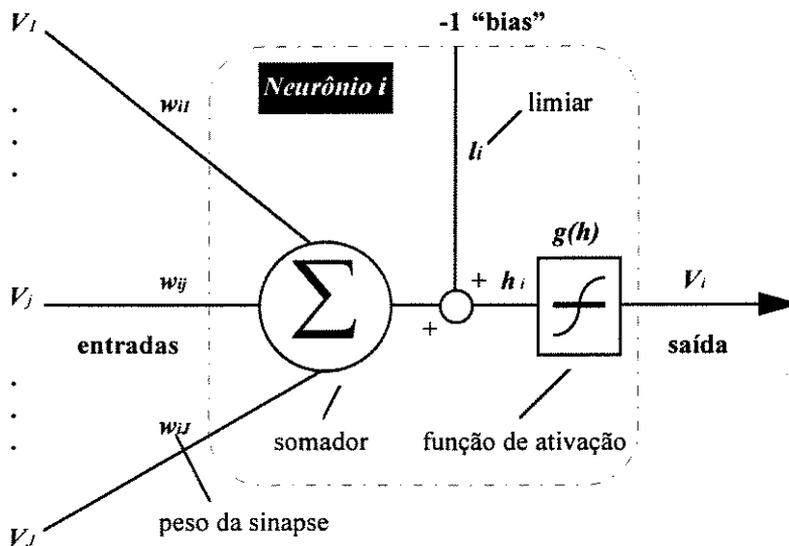


Figura.3.2: Esquema de um neurônio artificial ou *perceptron*.

O peso w_{ij} representa o peso da sinapse conectando o neurônio j ao neurônio i . Ele pode ser positivo ou negativo correspondendo a uma sinapse excitatória ou inibitória respectivamente. O parâmetro l_i é o valor do limiar para o neurônio i ; a soma

ponderada das entradas das sinapses e seus pesos contando com influência do limiar pode ser expressa por:

$$h_i = \sum_{j=1}^J w_{ij} \cdot V_j - l_i \quad (3.1)$$

onde J é o número de neurônios que estão conectados ao neurônio i e V_j é o valor da saída do neurônio j , conectado ao neurônio i .

O valor de h_i passa por uma **função de ativação**, $V_i = g(h_i)$, propagando para os neurônios posteriores o valor de ativação do neurônio i , onde $g(.)$ é a função de ativação.

3.4- Função de Ativação

O número $V_i = g(h_i)$ é chamado de estado ou ativação do neurônio i . Esta função processa o sinal h_i para produzir a saída do neurônio, V_i . Esta função pode ter várias formas: uma simples função linear, uma função limiar (função degrau); ou ainda uma função que simule mais precisamente as características não-lineares do neurônio biológico.

Para $g(.)$ sendo uma função linear temos:

$$g(h_i) = k \cdot h_i, \quad (3.2)$$

onde k é uma constante.

Caso $g(.)$ seja uma função degrau, a saída seria:

$$\begin{cases} g(h_i) = 1, & \text{se } h_i > 0 \\ g(h_i) = 0, & \text{caso contrário} \end{cases} \quad (3.3)$$

No caso de se querer imitar com maior precisão os neurônios biológicos, costuma-se utilizar funções não-lineares, denominadas em inglês funções *squashing*. Elas caracterizam-se por não deixarem a saída, V_i , exceder certos valores limites, os

quais, em geral, são menores que os valores de h_i . A função *squashing* mais usada é a função logística ou sigmóide (figura.3.3.a), a qual é representada matematicamente por:

$$g(x) = 1/(1+e^{-x}) \quad (3.4)$$

Outra função de ativação *squashing* bastante usada é a tangente hiperbólica (figura.3.3.b). Ela possui uma forma similar da sigmóide e é freqüentemente usada por biólogos para o modelo matemático de ativação das células nervosas. Como função de ativação dos neurônios artificiais a função é representada matematicamente por:

$$g(x) = \tanh(x) \quad (3.5)$$

Apesar de ser semelhante à função sigmóide, a tangente hiperbólica possui valores de saída bipolares, característica que é benéfica para alguns tipos de RNA's [15].

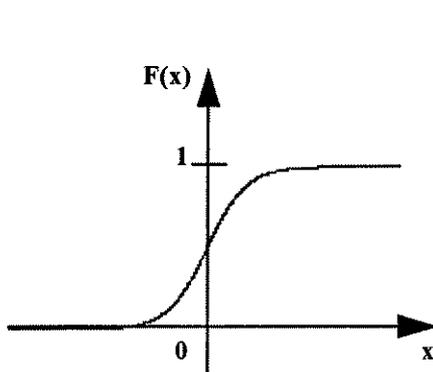


Figura.3.3.a: Função sigmóide

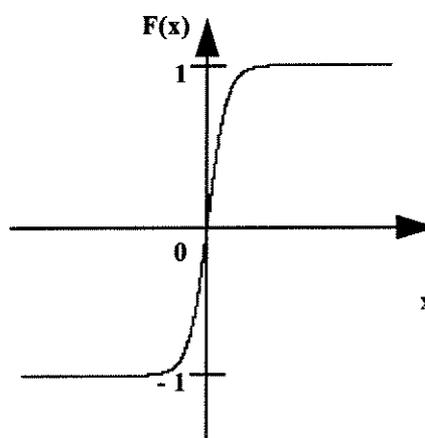


Figura.3.3.b: Função tangente hiperbólica.

É interessante ressaltar que este modelo simples de neurônio artificial ignora diversas características do neurônio natural, tais como a não consideração dos atrasos de tempo que afetam a dinâmica do sistema - as entradas produzem saídas imediatas - e a não inclusão dos efeitos de sincronismo ou de modulação de frequência - característica que alguns pesquisadores acham de fundamental importância. Apesar destas limitações,

as RNA's formadas por simples neurônios artificiais possuem atributos semelhantes aos do sistema biológico, como a capacidade de aprendizado e generalização, podendo-se dizer que a essência do funcionamento do neurônio natural foi absorvida.

3.5- Topologia das Redes Neurais Artificiais

Dentre as várias topologias de RNA's encontradas na literatura, podemos dividi-las basicamente em duas classes:

- Não-recorrentes
- Recorrentes

As RNA's não-recorrentes são aquelas que não possuem realimentação de suas saídas para as suas entradas e por isso são também ditas "sem memória". A estrutura das RNA's ditas não-recorrentes é em camadas, podendo estas serem formadas por uma (camada única) ou mais camadas (multi-camada).

Redes neurais multi-camadas contêm um conjunto de neurônios de entrada, uma camada de saída e uma ou mais camadas escondidas ou intermediárias. Alguns autores [52, 59] não contam a camada de entrada pelo fato desta apenas distribuir padrões. Entretanto neste trabalho a camada de entrada sempre será contada.

A camada com os neurônios que fornecem a saída da rede é chamada de camada de saída. As camadas que não possuem ligações diretas nem com a entrada, nem com a saída são chamadas de camadas intermediárias ou escondidas. Não existem conexões ligando um neurônio de uma camada a outro de uma camada prévia e nem entre os neurônios de uma mesma camada, veja figura.3.4. As RNA's de uma camada são também chamadas de *perceptrons* [52, 59]. As RNA's multi-camadas serão estudadas com mais detalhes nas seções 3.7 e 3.8.

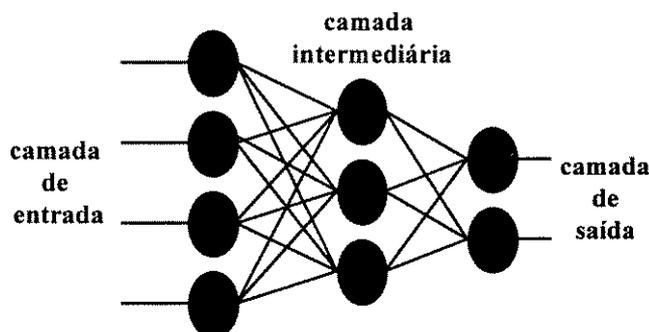


Figura.3.4: Exemplo de uma RNA Não-Recorrente Totalmente Conectada

As RNA's recorrentes são redes mais gerais que contêm realimentação das saídas anteriores. Além disso, suas estruturas não são obrigatoriamente organizadas em camadas e quando são, estas redes podem possuir interligações entre neurônios da mesma camada e entre camadas não consecutivas, gerando interconexões bem mais complexas que as RNA's não-recorrentes, como mostra a figura abaixo:

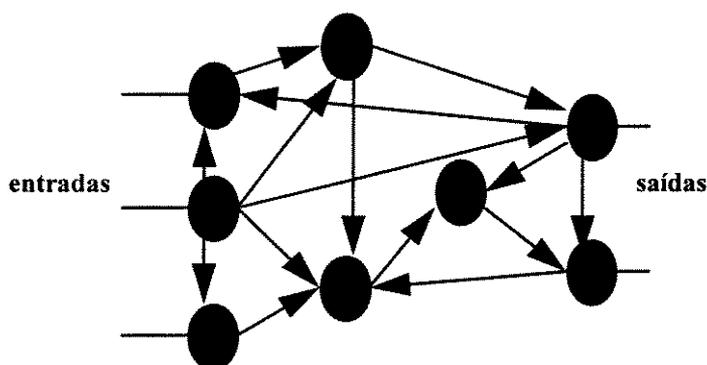


Figura.3.5: Exemplo de uma RNA Recorrente.

As RNA's recorrentes, também chamadas de RNA's com memória por possuírem realimentação, respondem a estímulos dinamicamente, isto é, após aplicar uma nova entrada, a saída é calculada e então realimentada para modificar a entrada. Para as redes estáveis, este processo é repetido várias vezes, produzindo pequenas mudanças nas saídas, até ficarem constantes. Todavia, as RNA's recorrentes nem sempre são estáveis, mesmo com entradas constantes.

O fato de não se conseguir prever quais redes seriam estáveis foi um problema que preocupou os pesquisadores até o início da década de 80 quando Cohen e Grossberg provaram um poderoso teorema para definir quando as RNA's eram estáveis [59]. Este

teorema determina que para as RNA's recorrentes alcançarem um estado estável é necessário que possuam conexões simétricas.

Contribuições importantes também foram dadas por John Hopfield [60], tanto para a teoria como para a prática, sendo que algumas configurações passaram a ser chamadas de redes de Hopfield em sua homenagem.

Hinton *et al* [61] também tiveram seus méritos, introduzindo regras gerais de treinamento para redes, denominadas por eles de máquina de Boltzmann devido ao cálculo das saídas dos neurônios artificiais ser efetuado através de probabilidade com distribuição de Boltzmann.

3.6- Tipos de Treinamento

A habilidade de aprender é a característica humana mais interessante que as redes neurais têm, isto é, a capacidade de capturar. Ao invés de especificar todos os detalhes de uma computação, tem-se a possibilidade de treinar uma rede para fazer esta computação. Isto significa que pode-se tratar problemas onde regras apropriadas são muito difíceis de se conhecer a priori.

O aprendizado é um conceito muito importante na caracterização de RNA's. Aprendizado é definido como uma modificação realizada na matriz de pesos W de uma rede, no sentido de otimizar um mapeamento dos padrões de entrada e saída. Embora em casos simples fosse possível ajustar a matriz W manualmente, de modo a obter um mapeamento único, é desejável que este processo se dê de forma auto-organizada, ou seja, a rede, através de um algoritmo apropriado, auto-ajusta os seus pesos durante a progressão do aprendizado [51].

Os procedimentos de treinamento que levam as RNA's a aprender determinadas tarefas podem ser classificados em duas classes de treinamento:

- Supervisionado
- Não-supervisionado

No treinamento tipo **supervisionado**, uma série de pares de padrões de entrada e saída (conjunto de treinamento) são apresentados simultaneamente à rede, em cada iteração do processo de treinamento. A rede computa então uma função de discrepância ou erro entre o padrão de saída desejado e o obtido naquela iteração, e ajusta os pesos das sinapses em função da magnitude do erro [51].

Um dos algoritmos de aprendizado mais conhecidos é o algoritmo de **retro-propagação** ou *backpropagation* (BP), o qual é utilizado neste trabalho e que será estudado em detalhes na seção 3.8.

No treinamento tipo **não-supervisionado**, os exemplos (ou padrões) são apresentados apenas na camada de entrada da rede, e ela se auto-organiza de modo a conseguir uma separação máxima entre as classes naturalmente existentes nos padrões [2]. O processo de treinamento extrai as propriedades estatísticas do conjunto de treinamento e agrupa os padrões semelhantes em classes [1].

Desta forma, a rede se comporta como um **algoritmo de aglomeração** (*cluster analysis*) do tipo divisivo. Um exemplo deste tipo de rede é o quantizador vetorial de Kohonen (1984) [51].

3.7- Redes Neurais Multicamadas

A capacidade de representação das RNA's com uma única camada foi superada na década de 60, a partir do perfeito entendimento do problema da separabilidade linear [62]. Entretanto, apesar de se ter descoberto como aumentar a capacidade de representação da rede com o acréscimo de novas camadas, faltava ainda uma forma de treinar a rede. Esta forma só foi descoberta em meados da década de 70, com a criação de vários algoritmos, entre eles o de retro-propagação ou *backpropagation*

(BP). O BP se tornou um dos algoritmos mais populares para o treinamento, sendo em parte responsável pelo ressurgimento do interesse na área de RNA's.

O algoritmo básico de aprendizado BP é empregado em uma RNA com topologia do tipo não-recorrente utilizando *perceptrons*¹, geralmente com três camadas (entrada, escondida e saída), heteroassociativo², com conexões de alimentação direta (*feed-forward*) da camada de entrada para a camada escondida e desta para a camada de saída. Uma representação de uma rede de três camadas é mostrada na figura.3.6.

O objetivo básico do treinamento de uma rede neural é realizar um mapeamento $\mathfrak{R}^K \rightarrow \mathfrak{R}^I$, através de uma função não-linear, $f(\cdot)$, relacionando P pares entrada-saída (X^p, Y^p) tal que

$$Y^p = f(X^p), \quad \forall p = 1, \dots, P \quad (3.6)$$

onde $Y^p = [y_1^p, \dots, y_I^p]$ e $X^p = [x_1^p, \dots, x_K^p]$ são os vetores de entrada e saída desejados respectivamente, do conjunto de treinamento da RNA referentes ao padrão p . K e I são os números de neurônios nas camadas de entrada e saída respectivamente.

Esta função poderia ser, por exemplo, a resposta de uma planta não-linear que está sendo simulada. Inúmeras outras aplicações poderiam ser citadas. Para a maioria delas a função $f(\cdot)$ não pode ser realizada exatamente. A RNA realiza uma aproximação

$$\tilde{Y}^p = \tilde{f}(X^p) \approx Y^p \quad (3.7)$$

onde $\tilde{Y}^p = [\tilde{y}_1^p, \dots, \tilde{y}_I^p]$ é o vetor de saída da RNA quando o padrão X^p é apresentado à rede.

¹ Perceptrons - São neurônios artificiais cuja estrutura está mostrada na figura.3.2.

² Heteroassociativo - Em uma rede heteroassociativa, a natureza dos padrões de entrada e saída são diferentes.

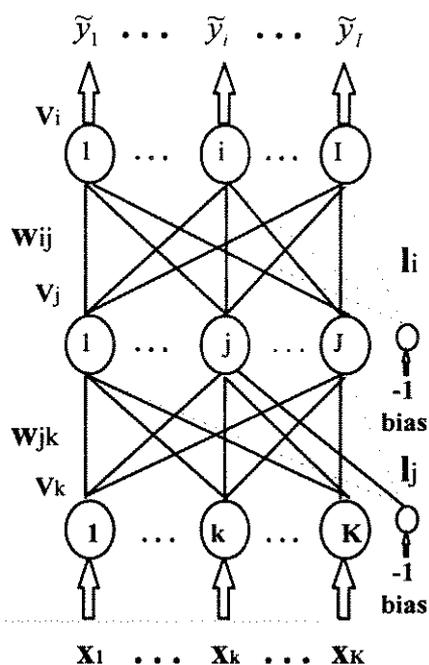


Figura.3.6: Esquema de uma RNA de 3 camadas.

A qualidade da aproximação realizada pode ser medida por uma função de custo qualquer $E(W)$. O problema da aprendizagem é encontrar o valor de W que minimiza $E(\cdot)$ [63].

3.8- Algoritmo de Retropropagação do Erro (BP)

O algoritmo de retro-propagação ou *backpropagation* foi inventado independentemente por vários pesquisadores, Bryson e Ho (1969) [62], Werbos (1974) [64], Parker (1985) [65] e Rumelhart *et al* (1986) [66, 67]. Uma abordagem aproximada foi proposta por Le Cun (1985) [68]. O algoritmo fornece uma forma de alteração dos pesos w_{ij} em qualquer rede diretamente alimentada para aprender a relação entre pares entrada-saída (X^p, Y^p) de um conjunto de treinamento. O algoritmo baseia-se no método do gradiente descendente, o qual tenta minimizar uma função de custo $E(W)$ que é uma medida dos erros da saída.

Consideremos uma RNA multi-camadas diretamente alimentada (RNA-MCDA), como a mostrada na figura 3.6. A convenção e notação são mostradas na própria figura.3.6; a saída de cada unidade são denotadas por $V_i = \tilde{y}_i^p$, para camada de saída, V_j , para a camada escondida e $V_k = x_k^p$, para a camada de entrada. As conexões são, w_{jk} , entre a camada de entrada e a camada escondida e w_{ij} entre a camada escondida e a camada de saída. Note que o índice i sempre refere-se a uma unidade da camada de saída, j para a camada escondida e k , para a camada de entrada.

Designamos diferentes padrões pelo superescrito p , sendo P o número de padrões do conjunto de treinamento. Cada valor x_k^p de entrada pode ser binário (0/1 ou ± 1) ou valores contínuos. K , J e I são o número de neurônios nas camadas de entrada, escondida e saída respectivamente.

Quando um padrão p é apresentado à entrada da rede, a unidade j , na camada escondida, computa a seguinte soma ponderada das unidades da camada de entrada conectadas a ela:

$$h_j^p = \left(\sum_{k=1}^K w_{jk} \cdot V_k^p \right) - l_j \quad (3.8)$$

e processa (função de ativação) na sua saída

$$V_j^p = g(h_j^p). \quad (3.9)$$

Cada unidade (neurônio) de saída i recebe os sinais V_j^p de cada unidade da camada escondida conectada à ela e executa a soma ponderada

$$h_i^p = \left(\sum_{j=1}^J w_{ij} \cdot V_j^p \right) - l_i \quad (3.10)$$

e produz a saída final da rede

$$\tilde{y}_i^p = V_i^p = g(h_i^p). \quad (3.11)$$

onde $g(\cdot)$ é a função de ativação de cada unidade e l são os seus limiares.

Geralmente a função de medida de erro ou função de custo é dada por

$$E(W) = \frac{1}{2} \sum_{p=1}^P \sum_{i=1}^l (\tilde{y}_i^p - y_i^p)^2, \quad (3.12)$$

embora qualquer outra função de custo possa ser considerada.

Podemos observar que a equação 3.12 é contínua e diferenciável para qualquer peso, podendo assim ser utilizado o método do gradiente descendente de modo a minimizar $E(W)$ e encontrar os pesos adequados para a rede.

O método do gradiente descendente sugere mudanças nos pesos w_{ij} por uma quantidade proporcional ao gradiente de $E(\cdot)$ dada por

$$\Delta w_{ij} = -\eta \cdot \frac{\partial E(W)}{\partial w_{ij}} = \eta \cdot \sum_{p=1}^P \delta_i^p \cdot V_j^p \quad (3.13)$$

onde η é conhecido como taxa de aprendizagem [59], e $\delta_i^p = g'(h_i^p) \sum_j w_{ij}^p \cdot \delta_j^{p-1}$.

Em geral, com um número arbitrário de camadas, a regra de correção dos pesos da sinapses entre a camada $c-1$ e c do BP sempre será

$$\Delta w_{q,o}^{c-1,c} = \eta \cdot \sum_{p=1}^P \delta_{o,c}^p \cdot V_{q,c-1}^p \quad (3.14)$$

onde $V_{q,c-1}^p$ é a saída da unidade q da camada $c-1$ e $\delta_{o,c}^p$ é o delta da unidade o da camada c . O significado do delta depende da camada a que ele se refere, sendo em geral uma informação da direção de descida da função de custo no seu espaço de definição.

Sumarizamos abaixo o algoritmo BP passo a passo mostrando as duas seqüências básicas do algoritmo: a seqüência progressiva (*forward*) e a seqüência regressiva (*backward*). Tomando um padrão p no tempo (isto é, correções incrementais).

Consideremos uma rede com M camadas, $m=1,2,\dots,M$ e utilizando V_i^m para a saída da i -ésima unidade da m -ésima camada. V_i^0 é sinônimo da entrada x_i para a i -ésima entrada. Note que os superescritos m 's representam as camadas e não os padrões.

O valor w_{ij}^m representa a conexão de V_j^{m-1} a V_i^m . Logo o algoritmo BP pode ser dividido em passos da seguinte forma:

Seqüência progressiva:

passo.1- Inicialize os pesos com valores aleatórios.

passo.2- Escolha um padrão do conjunto de treinamento e aplique-o à camada de entrada ($m=0$) tal que

$$V_k^0 = x_k^p \quad \forall k \quad (3.15).$$

passo.3- Propague o sinal progressivamente através da rede usando

$$V_i^m = g(h_i^m) = g\left(\sum_j w_{ij}^m \cdot V_j^{m-1}\right) \quad (3.16), \text{ para cada } i \text{ e } m \text{ até a saída}$$

V_i^M ser calculada.

Seqüência regressiva:

passo.4- Compute os deltas para a camada de saída

$$\delta_i^M = g'(h_i^M)[y_i^p - V_i^M] \quad (3.17)$$

Note que δ_i^M é função da saída atual V_i^M e da saída desejada y_i^p para o padrão p considerado.

passo.5- Compute para as camadas precedentes por retro-propagação dos erros os deltas de cada unidade da seguinte forma

$$\delta_i^{m-1} = g'(h_i^{m-1}) \sum_j w_{ji}^m \cdot \delta_j^m \quad (3.18), \text{ para } m=M, M-1, \dots, 2,$$

até um delta ser calculado para cada unidade.

passo.6- Use $\Delta_{ij}^m = \eta \cdot \delta_i^m \cdot V_j^{m-1}$ (3.19), para corrigir todos os pesos de acordo com $w_{ij}^{novo} = w_{ij}^{velho} + \Delta_{ij}^m$.

passo.7- Volte ao *passo.2* e repita tudo para o próximo padrão.

A seqüência de passos acima é repetida até que todos os padrões sejam apresentados à rede e esta compute sua respectiva saída caracterizando assim uma “época”³.

Após cada época computa-se a função de custo comparando-a com o seu valor na época anterior. Se o valor desta diferença for menor que uma certa tolerância pode-se considerar a rede como treinada. Outro critério de parada do algoritmo é limitar o número de épocas para um dado conjunto de treinamento [52].

³ Época - seqüência de apresentação de todos padrões do conjunto de treinamento à rede neural durante o treinamento.

3.8.1- Problemas com o Algoritmo BP

Apesar do grande sucesso das aplicações do algoritmo BP e da sua enorme popularidade, muitos problemas ainda o acompanham. Entre eles, os que mais se destacam são:

Período de treinamento: O treinamento pode demorar um longo período, principalmente para problemas complexos, sendo que não há garantias que depois deste tempo o treinamento tenha sido um sucesso. O termo *momentum* pode ser uma alternativa para diminuir o tempo de treinamento, assegurando um maior grau de estabilidade para a rede. Foi proposto por Rumelhart, Hinton e Willians em 1986 e se mostrou bastante eficiente em alguns casos. O termo *momentum* foi incorporado ao BP utilizado neste trabalho e sua utilização será discutida no capítulo 4. Outras alternativas também foram propostas por Parker em 1987, bem como por Stornetta e Huberman [52, 59].

Mínimos locais: O treinamento é um processo de otimização por gradiente, e sofre dos males inerentes a estes processos. Um dos problemas é que como $E(W)$ é não-linear podem existir mínimos locais [63]. O algoritmo de retro-propagação não garante que o erro mínimo seja alcançado durante o processo de treinamento, mas apenas algum erro local mínimo.

Além disso, se a taxa de aprendizagem for grande podem surgir oscilações do erro global da rede durante o treinamento, que fica aumentando e diminuindo, sem convergir. Pode acontecer, também, que o erro global para o qual a rede converge é inaceitavelmente alto (as classificações resultantes após treinamento serão erradas em um número muito grande de vezes). Em ambos os casos diz-se que a rede é incapaz de aprender [51].

Uma saída para este problema é implementar na rede um processo de *simulated annealing*. Neste processo é incluído na rede neural (ou no algoritmo de treinamento) um ruído randômico que ocasiona “saltos” nos valores das

sinapses que permitem escapar dos mínimos locais, “vales” pouco profundos na superfície de $E(W)$. Ao longo do treinamento a amplitude do ruído vai sendo reduzida e o ponto de operação tende a ficar preso no “vale” mais profundo (o que apresenta menor probabilidade de escape) que corresponde ao mínimo global de $E(W)$. O problema é que o treinamento usando *simulated annealing* é muito mais complexo e lento que os métodos convencionais *steepest descent*, razão pela qual este método não foi utilizado neste trabalho [63].

Um tipo comum de mínimo local é um em que dois ou mais erros compensam os outros. Este mínimo não é muito profundo, apenas um ruído (flutuação randômica) é suficiente para gerá-lo. Uma simples e muito eficiente solução é usar correções incrementais, escolhendo os padrões em ordem aleatória do conjunto de treinamento. Então a média dos padrões é evitada e a ordem aleatória gera o ruído [52].

Paralisia da rede: A forma de $E(W)$ depende evidentemente da função à minimizar. Mas os valores de $|w_{ij}|$ podem levar a valores muito grandes de $|h_i|$. Isto, devido à forma da sigmóide, leva a um valor de $dg(h_i)/dh_i$ (e, conseqüentemente, a um gradiente de $E(W)$) muito pequenos. Se w_{ij} for deslocado para uma região de gradiente pequeno o treinamento praticamente para (fica paralizado) por que como Δw_{ij} é proporcional ao gradiente, Δw_{ij} vai ser muito pequeno. Limitar os valores de $|w_{ij}|$ e $|h_i|$ são soluções normalmente adotadas para sinalizar ou evitar a ocorrência de paralisia no treinamento [69].

Devido às deficiências do algoritmo de retro-propagação, inúmeras variações têm sido propostas nos últimos anos [6, 52, 59]. Todavia, até agora, nenhuma têm solucionado de forma definitiva e confiável os problemas do BP.

Capítulo 4

PREVISÃO DE CARGA A CURTO PRAZO UTILIZANDO RNA's

Este capítulo apresenta o estado da arte de aplicações de redes neurais à previsão de carga a curto-prazo. Logo após, define-se o problema de previsão de carga no horário de ponta tratado neste trabalho. Por fim, propõe-se quatro algoritmos para se resolver o problema acima citado.

4.1 - Aplicação de RNA's à Previsão de Carga a Curto

Devido à sua capacidade em manipular dados imprecisos e ruidosos, as RNA's vêm se tornando uma poderosa ferramenta para previsão de carga, muitas vezes com vantagens sobre vários modelos estatísticos.

Nas últimas duas décadas, esforços consideráveis têm sido feitos para desenvolver aproximações razoáveis para prever cargas horárias com o horizonte de um dia ou uma semana à frente. Recentemente inúmeros trabalhos utilizando a modelagem de RNA's para previsão de carga a curto-prazo foram publicados, entretanto destacamos nesta seção, os mais importantes. Trabalhos considerando horizontes de curtíssimo-prazo são raros, tendo sido encontrado apenas um artigo na literatura, embora este trabalho não seja uma aplicação de RNA's à previsão de carga.

4.2- O Estado da Arte

Park *et al* [69], utilizaram uma *multilayer perceptron network* para previsão, tendo como entrada da RNA dados climáticos (temperatura). Park propôs uma RNA para aprender a relação entre as temperaturas passada, presente e futura e a carga. Para achar a carga prevista, a RNA interpola os dados carga-temperatura num conjunto de treinamento. Para isto o problema da previsão de carga foi subdividido em três sub-

problemas, isto é, foram desenvolvidas três redes neurais especializadas, cada uma para prever o pico de carga, o total de carga do dia e outra para o perfil horário da carga durante o dia, para horizontes de uma a 24 horas à frente. A entrada das redes foi colocado a variável dia da previsão além das variáveis de carga e clima. A média dos erros percentuais absolutos ficou em torno de 1.4 e 2.1% para os horizontes de uma e 24 horas à frente respectivamente.

Lu *et al* [70] compararam a eficácia das RNA's para previsão com horizontes de uma hora à frente com a previsão com horizonte de 24 horas à frente, analisando se o modelo utilizando redes neurais, é ou não, dependente da região de atuação da concessionária de energia elétrica. Para isto, eles utilizaram redes *feedforward multilayer perceptrons* para fazer a previsão uma e 24 horas à frente de dias de semana e finais de semana, utilizando como dados de entrada, carga, temperatura, hora do dia e dia da semana. A média dos erros percentuais absolutos ficou em torno de 3.0 e 2.5% para os horizontes uma e 24 horas à frente respectivamente.

Chen *et al* [71], propõem uma rede neural não totalmente conectada, composta de quatro sub-redes menores para previsão uma hora à frente em patamares de uma hora. Eles mostraram que redes neurais totalmente conectadas com alguns nós na camada escondida são capazes de “capturar” as características da carga, mas a rede não é adaptativa o bastante para responder às mudanças bruscas de temperatura. Um aumento no número de nós e camadas escondidas não mostraram melhora significativa. E além disso o treinamento tornou-se demasiadamente longo. Para a rede não totalmente conectada, o tempo de treinamento é diminuído devido à não conexão de todas as sinapses. Utilizou-se dados de carga, clima, hora do dia e dia da semana como variáveis de entrada. A média dos erros percentuais absolutos ficou em torno de 1.2%.

A desvantagem deste modelo está na particularidade de sua concepção. Cada área de atuação de um sistema de potência possui características próprias que possibilitarão que outros tipos de redes neurais com arquiteturas diferentes sejam projetadas para melhorar a precisão na previsão de carga.

De maneira a conseguir uma rede que se adapte às mudanças no comportamento da carga com o decorrer dos dias, Dillon *et al* [72] proporam um modelo de rede neural adaptativa para previsão de carga a curto-prazo 1 hora à frente. Neste modelo, ele treina a rede neural com os dados de carga dos últimos 20 dias típicos para prever a carga do 21º dia. Depois, a rede é retreinada com a inclusão do 21º dia e a exclusão do 1º dia do conjunto de treinamento e assim sucessivamente. Utilizou-se dados de clima, hora do dia e dia da semana como variáveis de entrada. A média dos erros percentuais absolutos ficou numa faixa de 4.1 a 4.7%.

Para a rede proposta em [72], os dados são “esquecidos” e um novo dado é aprendido. Depois de feita a previsão, o modelo corrige a carga prevista através de um algoritmo de correção.

Para previsão 24 horas à frente, Lee *et al* [73] separaram os dias em duas classes: dias de semana e finais de semana. Através do estudo destas classes, separaram novamente os dias em cinco subclasses com dias com padrões de perfil de carga semelhantes. Eles testaram várias estruturas de RNA's com 2 e 3 camadas, para as diversas subclasses encontradas. Então proporam dois métodos para fazer a previsão 24 horas à frente apenas com variáveis de carga como entrada da rede. O primeiro método é estático onde todos os intervalos horários de carga são previstos simultaneamente, com 24 neurônios na saída da rede. O segundo método é adaptativo e tem como saída da rede apenas a hora que está sendo prevista. A média dos erros percentuais absolutos ficou em torno de 2.0% para previsão de carga horária para os dois métodos.

A desvantagem do segundo método utilizado por Lee *et al*, é que a rede é alimentada na entrada com o valor da carga prevista nas horas anteriores inserindo o ruído do erro da previsão da própria rede para a hora posterior.

Para a previsão horária em um dia, o padrão do perfil da carga do dia a ser previsto precisa ser determinado.

Afim de encontrar um padrão de carga horária adequado para o dia em que a carga horária será prevista, Hsu e Yang [74] desenvolveram uma rede neural baseada nos mapas auto-organizáveis de Kohonen para identificar aqueles dias com padrão de carga similar, agrupando-os em classes. Então, duas redes neurais *feedforward* são treinadas para um dia específico [75]. Elas foram utilizadas para prever o pico e o vale de carga durante o dia, respectivamente, com os dados climáticos e de carga como entrada da rede. A média dos erros percentuais relativos ficou em torno de 1.2%.

O padrão de carga horária para um dia específico é obtido pela média dos padrões de carga de alguns dias anteriores do mesmo tipo, ou seja, pertencentes à mesma classe.

Uma característica especial destes mapas auto-organizáveis é que eles são capazes de identificar um novo tipo de padrão de carga antes que os operadores possam reconhecer o novo padrão do dia.

No desenvolvimento desta categoria de modelos, os fatores mais relevantes foram os padrões de carga horária, os picos e os vales de carga [74].

Na tentativa de eliminar a subjetividade e a heurística da maioria dos modelos de previsão de carga, Erkmen *et al* [76], assim como em [74], desenvolveram uma rede neural híbrida. A primeira rede tem aprendizagem do tipo não-supervisionada e serve para classificar os padrões de carga a partir de um histórico de carga horária. A outra rede, supervisionada, realiza a previsão horária uma hora à frente a partir do padrão já pré-estabelecido anteriormente.

Para fazer esta classificação Erkmen *et al* [76], proporam uma rede de uma só camada altamente conectada que utiliza um algoritmo de aprendizagem modificado daquele utilizado em [74]. Nesta rede as unidades também possuem conexões laterais e a determinação da unidade vencedora é diferente daquele processo utilizado em [74]. Este algoritmo visa diminuir a entropia do sistema e o tempo computacional gasto neste

processo através da expansão de Karhunen-Loeve. O algoritmo detecta padrões de carga inesperados.

A vantagem destes modelos está na determinação de padrões de carga não facilmente detectados pelos operadores do sistema. Outra vantagem destes modelos para detecção de padrões de carga, é que eles podem ser utilizados também em dados de carga diários e semanal, desde que esses dados estejam disponíveis e sejam fornecidos à rede neural não-supervisionada.

Peng *et al* [77] proporam uma estratégia para seleção dos padrões de treinamento da rede utilizando a medida da mínima distância para identificar os padrões de carga dos dias semelhantes. Para a previsão, eles usaram um algoritmo com uma combinação de termos lineares e não-lineares através de uma rede neural apropriada para isto, para fazer a previsão horária com horizonte de 24 horas à frente. A média dos erros percentuais absolutos ficou numa faixa de 1.8 a 2.8%.

Alguns dos modelos de previsão de carga são modelos dinâmicos, tais como os modelos auto-regressivos integrados de médias móveis (ARIMA). A principal característica dos modelos dinâmicos de carga é que eles são efetivos para as previsões de carga de dia típico, mas eles são inapropriados para dias especiais [20, 78].

Karinitionatakis *et al* [79] proporam uma rede neural recorrente de alta ordem que é capaz de detectar rapidamente, possíveis variações abruptas da carga ou importantes mudanças do perfil de carga devido a dias especiais (finais de semana e feriados). Esta rede foi utilizada para previsão 1 hora à frente utilizando dados de carga e clima como variáveis de entrada da rede. A média dos erros percentuais absolutos ficou em torno de 0.70%. Entretanto o sistema usado para os testes desta rede é de porte pequeno.

A vantagem deste modelo está na sua capacidade de acompanhamento do perfil de carga, já que este modelo não necessita da definição do padrão de carga que está

ocorrendo. Logo, a rede neural não precisa ser retreinada periodicamente, pois os pesos da rede são adaptados para os novos perfis de carga que surgirão.

Entretanto, alguns problemas podem surgir para este modelo, como a falta de dados mensuráveis para o vetor de entrada e a consideração de dados de entrada redundantes que não influenciam o comportamento como um todo. Isto pode causar problemas em todo processo de previsão devido a uma complexidade adicional que será exigida.

Enquanto Dillon *et al* [72] proporam uma rede que aprende e esquece os padrões de carga à medida que o tempo passa, Mohammed *et al* [80] proporam um algoritmo que considera a sazonalidade da carga anual e treina diversas redes especializadas em cada estação e em cada dia da semana. O treinamento é feito com os dados do mesmo período do ano anterior e corrigidos de acordo com os novos dados que vão sendo obtidos do sistema, através de uma adaptação diária, semanal e mensal. A média dos erros percentuais absolutos ficou em torno de 2.5%.

Kermanshahi *et al* [81], também treinaram redes especializadas em cada dia da semana para previsão 24 horas à frente em base horária. A rede utilizada por eles usou dados de clima, carga e mês do ano como variáveis de entrada. A rede é treinada com os dados de um ano e faz a previsão em outro, aprendendo a sazonalidade dos dados com período anual.

Alguns autores testaram diversos parâmetros que determinam o desempenho das RNA's em previsão de carga. Ho *et al* [82], testaram o efeito da taxa de aprendizagem, termo momento e número de padrões apresentados em cada iteração sobre a eficiência e precisão do *backpropagation* para previsão uma hora à frente utilizando dados de carga e clima como variáveis de entrada. O uso do termo momento adaptativo mostrou bons resultados. A média dos erros percentuais absolutos ficou numa faixa de 0.9 a 1.8%.

Kiartzis *et al* [83] testaram a eficácia de diferentes tipos de função de ativação na precisão da previsão 24 horas à frente, não havendo encontrando significativas

diferenças no erro da previsão. Entretanto a função sigmóide apresentou os melhores resultados para os casos testados.

Papalexopoulos *et al* [84], compararam o método de regressão linear múltipla e redes neurais artificiais para previsão 24 horas à frente. As RNA's tiveram melhor desempenho que os métodos de regressão.

Satoh *et al* [85] proporam um modelo de previsão combinando RNA's com regressão linear múltipla (MLR). A previsão é feita pela MLR e corrigida por uma RNA, para previsão 24 horas à frente. A média dos erros percentuais absolutos ficou em torno de 2%, incluindo dias especiais.

Para previsão a curtíssimo prazo, abordando a previsão de carga com redes neurais não foi encontrado nenhum trabalho sobre o assunto.

4.3- Previsão do Período de Demanda de Ponta

Como pôde ser observado na seção anterior, a maioria dos trabalhos citados fazem a previsão a curto-prazo com horizonte de uma, 24 e 168 horas à frente em patamares horários.

O problema abordado nesta dissertação é especificamente prever o perfil da carga no período de demanda de ponta com discretização menor do que uma hora, para os horizontes de previsão que serão descritos posteriormente.

Tradicionalmente, no Brasil e em diversas lugares do mundo, como vimos na seção anterior, a previsão de carga a curto prazo é feita em base horária; ou seja, a previsão dá a estimativa da carga média para cada hora do próximo dia. Em função desta carga média, a programação da operação determina uma geração média para cada usina para esta hora.

Entretanto, nos intervalos em que há grande variação na demanda de carga, principalmente nos períodos de entrada e saída do período de ponta, a demanda média pode ser muito diferente do valor instantâneo. Como pode ser observado na figura.4.1, por exemplo no período das 18 às 19 horas. Sendo assim, a geração instantânea necessária nestes intervalos de tempo pode ser muito diferente da geração média programada para o respectivo intervalo de tempo.

Esta diferença pode ser de grande risco, pois o sistema pode não estar preparado para tal geração tanto em termos de equipamentos como também de intercâmbio com outros sistemas. Para se trabalhar com uma curva de carga mais realista, as empresas estão adotando uma discretização menor que uma hora nos intervalos de grande variação da demanda, de modo a se obter uma curva mais fiel ao comportamento da carga.

A previsão de carga com discretização menor que uma hora traz uma dificuldade adicional em relação à previsão com base horária, pois quanto menor o intervalo de discretização, mais importante se torna as componentes de alta frequência da curva de carga. Para alguns sistemas, como o sudeste brasileiro, esta previsão torna-se mais complicada pois o período de demanda de ponta coincide com o final do dia, quando a carga de iluminação começa a crescer. Se no final do dia o tempo estiver nublado, então a carga de iluminação começa antes, modificando assim a entrada do período de ponta.

A figura.4.1 mostra a diferença da curva de carga para discretização de 10 minutos e 1 hora.

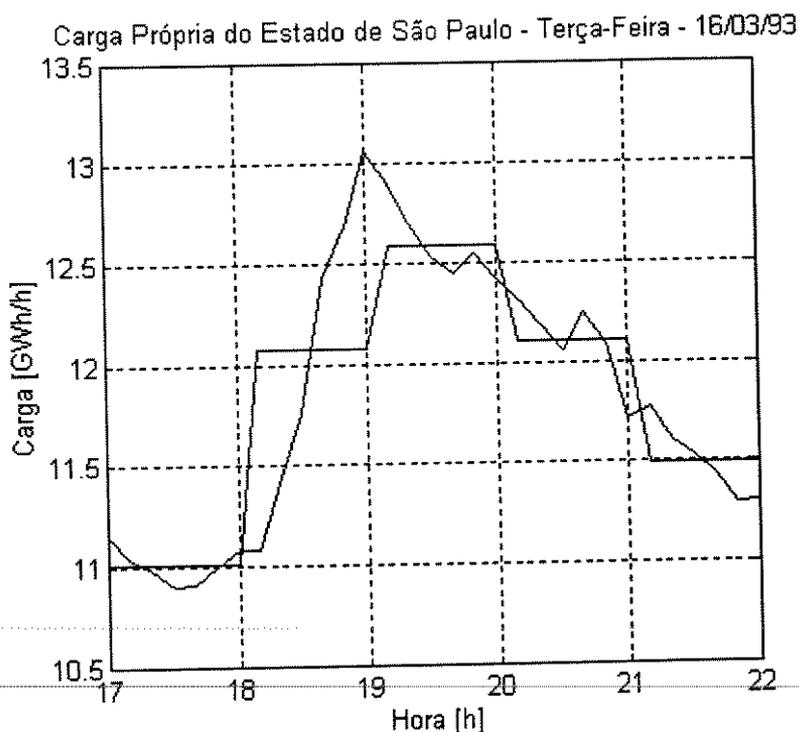


Figura.4.1: Comparação entre a Curva de Carga com Discretização de 10 minutos e 1 hora.

Note que o pico de carga realmente ocorre no intervalo entre as 18:50 e 19 horas e não entre as 19 e às 20 horas como sugere a curva com discretização de 1 hora.

Faremos a previsão do período de demanda de ponta com discretização menor que uma hora, para os seguintes horizontes de previsão:

- **Previsão 10 minutos à frente** (curtíssimo-prazo ou tempo-real) - Às 17:00 horas do dia da previsão inicia-se o treinamento da RNA para prever a carga do intervalo 10 minutos à frente. Logo após encerrado o, é realizada a previsão. O procedimento se repete até que todo o perfil seja previsto.
- **Previsão 1 hora à frente** (curto-prazo) - Às 15:00 horas do dia da previsão inicia-se o treinamento da RNA. Logo após encerrado o treinamento é realizada a previsão.

- **Previsão 24 horas à frente** (curto-prazo) - Às 15:00 horas do dia anterior a previsão inicia-se o treinamento da RNA. Logo após encerrado o treinamento é realizada a previsão.

Na previsão 24 horas à frente, há uma consideração importante: a influência de fatores climáticos sobre a carga, inferida dos valores observados no passado, só terá interesse prático quando o fator climático em questão puder ser previsto com boa precisão para um horizonte equivalente [86].

4.4- Algoritmos Propostos para Previsão da Ponta Via RNA

Uma metodologia de previsão de carga bem sucedida é aquela que consegue extrair dos dados históricos as suas características básicas, as quais supõe-se que permanecerão válidas para o futuro [10].

A maioria dos modelos desenvolvidos para previsão de carga com RNA's abordam a questão da previsão para um horizonte de 24 horas à frente com discretização horária, e pouquíssima atenção tem sido dada ao problema de previsão a curtíssimo prazo com discretização de alguns minutos.

Para a previsão do perfil da ponta com discretização menor que uma hora, propomos resolver o problema da previsão para os horizontes citados acima, através de dois métodos diferentes:

- **Método estático:** Uma RNA é treinada de maneira a fornecer todo o perfil da ponta de uma única vez. As variáveis de entrada devem conter informações significativas sobre todo perfil.
- **Método dinâmico:** Uma RNA é treinada com as informações referente apenas ao intervalo que se está prevendo. Assim, durante cada intervalo do perfil de ponta uma nova RNA será treinada. As variáveis de entrada devem

conter informações significativas apenas sobre o intervalo a ser previsto. Assim, haverão tantos treinamentos com a RNA quanto forem o número de intervalos a serem previstos no período de ponta.

Para cada método citado acima, podemos realizá-los de duas maneiras dependendo do conjunto de treinamento e testes das RNA's. Logo, dois procedimentos são sugeridos para se resolver a previsão pelos dois métodos citados acima.

4.4.1– Algoritmo *Moving Window* (MWS)

Rahman *et al* [87] sugerem que a precisão da previsão requer que todas RNA's sejam desenvolvidas para cada diferente hora do dia e para cada dia da semana. Tal complexidade aumenta o risco de uma falha total na previsão. Com o intuito de avaliar esta especialização das RNA's para o problema de previsão a curto prazo, propomos o algoritmo *MWS*.

Sabendo que a carga do próximo dia tem uma forte correlação com as cargas dos últimos dias, propomos um algoritmo que considera os dados mais recentes do sistema. Neste algoritmo o conjunto de treinamento da RNA para previsão estática ou dinâmica é composto com os dados dos últimos d dias típicos anteriores ao dia a ser previsto. Para se prever o dia $d+1$, o dia d é incorporado ao conjunto de treinamento e o dia mais “antigo” do conjunto de treinamento é eliminado deste.

Note-se que dentro de um horizonte de curto prazo, a adaptabilidade a nível diário permite que as variações da curva horária observadas na transição entre as estações sejam rapidamente incorporadas pelo sistema de previsão à medida em que ocorrem.

O algoritmo *MWS* treina uma RNA apenas para os dados dos últimos dias típicos. Para o próximo dia típico a ser previsto a RNA é retreinada com os dados dos últimos dias típicos que já inclui o que aconteceu no dia anterior e esquece os padrões do dia mais antigo.

Cada RNA treinada por *MWS* é na verdade uma rede especializada para a previsão dos intervalos desejados, portanto com menos variáveis de entrada e o conjunto de treinamento com padrões semelhantes. Portanto, esta rede obtém uma interpolação dos padrões do seu conjunto de treinamento, pois estes são bem semelhantes em praticamente todo período de previsão.

A escolha do número de dias típicos no conjunto de treinamento foi realizado por tentativa-e-erro. Nos testes realizados, variou-se o número de dias típicos do conjunto de treinamento de 2 a 20; na maioria dos casos os menores erros foram obtidos considerando-se os últimos 10 dias típicos.

Um fator relevante tem que ser considerado para as previsões de carga com este algoritmo. O tempo de retreinamento da rede é um fator muito importante, pois para a previsão a curtíssimo-prazo, a rede deve ter o período de treinamento curto para que a previsão seja fornecida ao operador a tempo deste poder tomar alguma decisão. Caso contrário, a informação não servirá para nada.

4.4.2- Algoritmo Ano-Ano (AAS)

Como o padrão de consumo do usuário de energia elétrica é, em geral, semelhante de ano para ano, propomos outro algoritmo que treina uma rede neural que aprende as sazonalidades da carga anual de um sistema de potência. Neste algoritmo, o conjunto de treinamento da RNA é montado com os padrões de carga do período correspondente do ano passado.

Por exemplo, para se fazer a previsão do 1º semestre deste ano por este algoritmo, uma RNA é treinada com os padrões de carga correspondentes a todo o 1º semestre do ano passado. Assim, uma rede é treinada apenas uma vez por período a ser previsto.

O algoritmo *AAS* treina uma rede neural para fazer a previsão de um certo período do ano, baseado no que aconteceu no ano anterior. Portanto, sempre que é feita a previsão do perfil da ponta com este algoritmo, a RNA treinada vai na verdade classificar o padrão de entrada e fornecer um perfil de carga que se assemelha com as classes dos padrões aprendidos no seu conjunto de treinamento. Se este padrão for diferente daqueles aprendidos, a RNA vai generalizar a sua saída. Esta RNA treinada por *AAS* aprende as componentes sazonais e cíclicas da carga que aparecem durante o ano e tenta absorver a relação entre carga-clima.

De um ano para outro, o número de consumidores aumenta com o subsequente aumento do consumo de energia do sistema. De maneira a manter a uniformidade dos dados de treinamento (carga de 1993) em comparação com os dados do conjunto de testes (carga de 1994), a carga de 1993 foi convertida em um valor equivalente a 1994 considerando um aumento de carga da ordem de 4% verificado nos dados do histórico. Este valor pode ser obtido também comparando-se as cargas médias dos últimos dois meses de 1992 e 1993. Assim pudemos considerar a taxa de crescimento anual da carga de um ano para outro para o algoritmo *AAS*.

Duas variações dos algoritmos *MWS* e *AAS* então, serão testados para realizar a previsão do período de demanda de ponta:

1. *Algoritmo MWS-ALL* - para previsão estática 1 e 24 horas à frente.
2. *Algoritmo MWS-PPT* - para previsão dinâmica 1, 24 horas e 10 minutos à frente.
3. *Algoritmo AAS-ALL* - para previsão estática 1 e 24 horas à frente.
4. *Algoritmo AAS-PPT* - para previsão dinâmica 1 e 24 horas à frente.

As figuras 4.2 a 4.5 mostram o fluxograma de cada algoritmo citado acima.

Algoritmo MWS-ALL

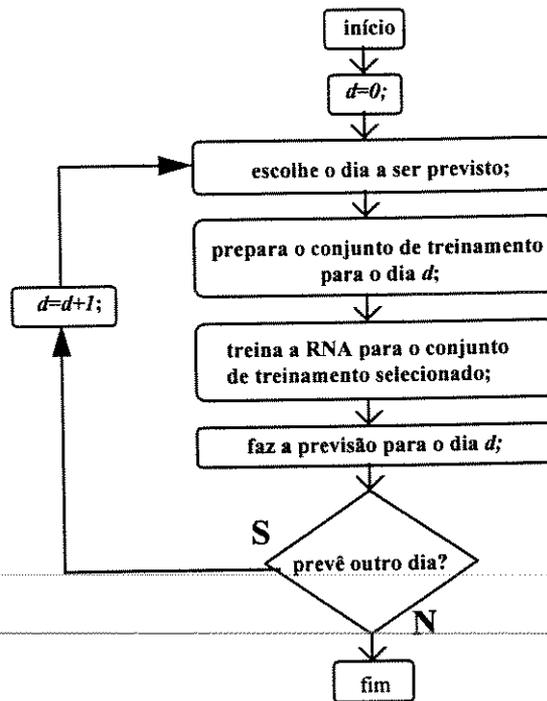


Figura.4.2: Fluxograma do Algoritmo *MWS-ALL*

Algoritmo AAS-ALL

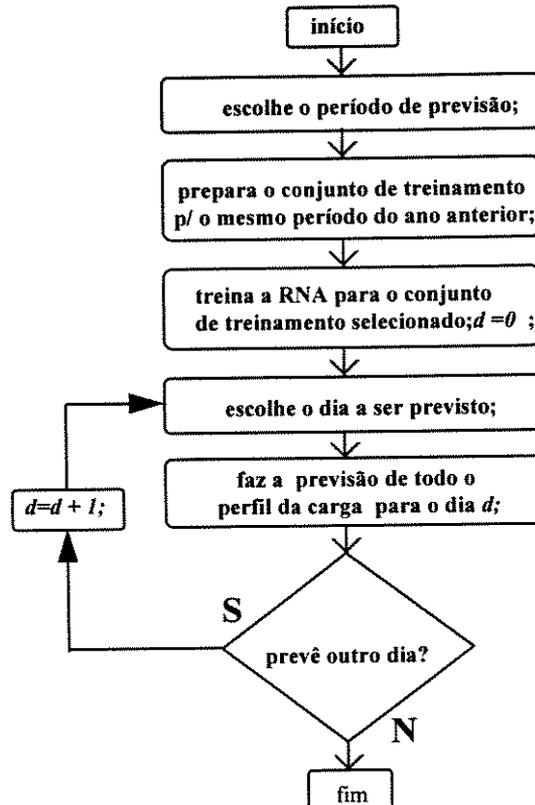


Figura.4.3: Fluxograma do Algoritmo *AAS-ALL*

Algoritmo MWS-PPT

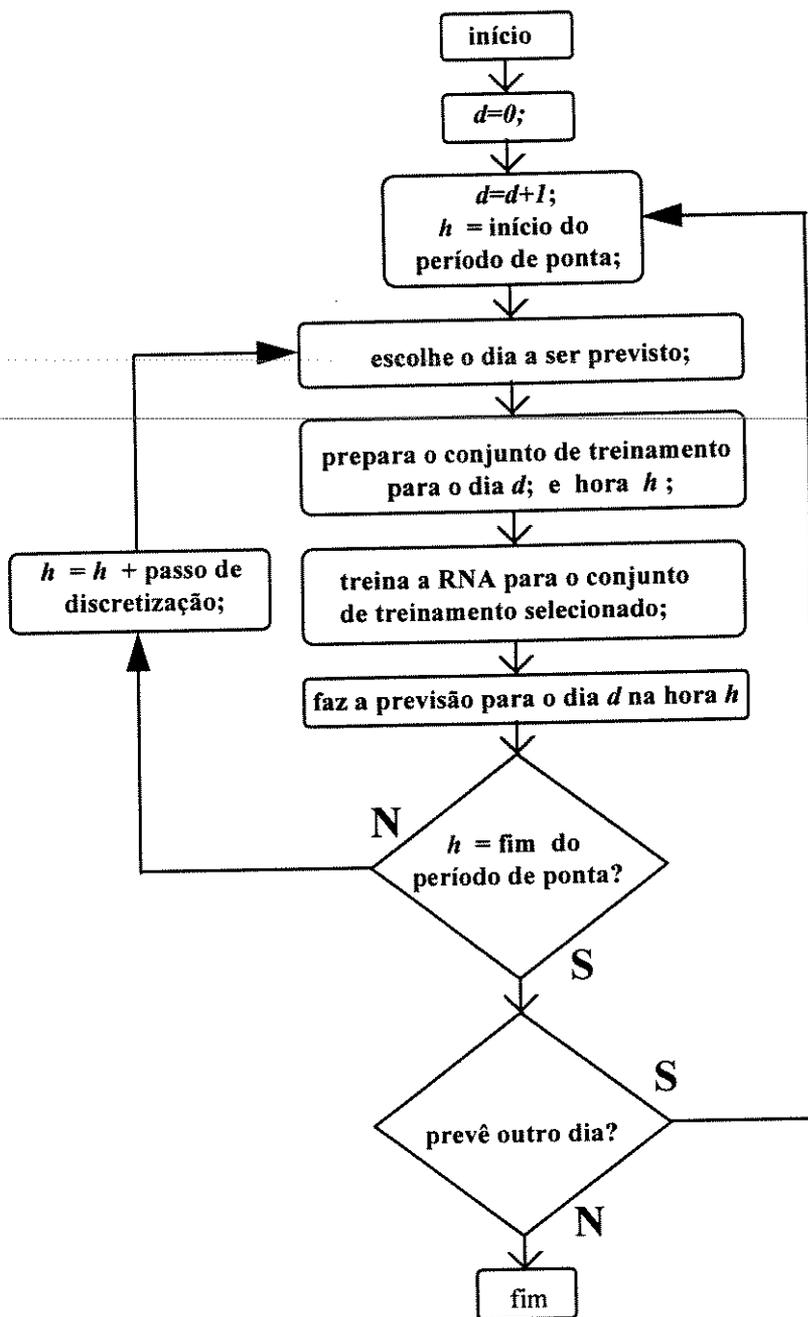


Figura.4.4: Fluxograma do Algoritmo *MWS-PPT*

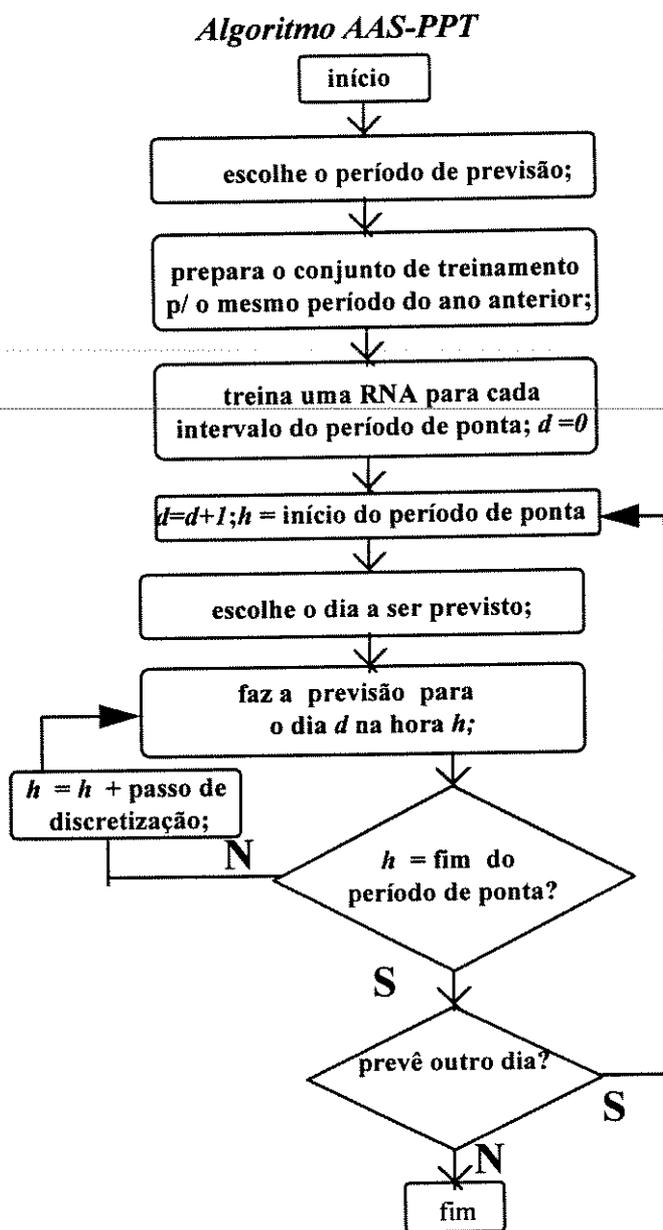


Figura.4.5: Fluxograma do Algoritmo AAS-PPT

4.5- Comentários Adicionais

Geralmente, na literatura, tem-se feito uma RNA generalizada para aprender todos os fatores que determinam o comportamento da carga durante um certo período do ano. Entretanto, para sistemas de potência em países em desenvolvimento como o Brasil, de um ano para outro o comportamento do consumidor de energia elétrica muda a uma taxa maior do que em outros países.

Fatores como a instabilidade da economia, inserção de novas tecnologias de produção industrial, cargas especiais e novos produtos em residências, como TV a cabo, computadores e vídeos cassetes mudam a cultura de utilização de energia elétrica de um ano para outro, dependendo da velocidade de entrada no mercado de tais produtos.

Também na literatura, observamos que muitos autores consideram redes com elevado número de neurônios na camada de entrada e saída. Isto faz com que o treinamento seja lento. Portanto, retrainar a rede torna-se caro computacionalmente. Entretanto, se a rede for menor, o retraining pode ser considerado como uma solução factível para adaptar os dados à nova carga, ao invés de se usar procedimentos e fatores para correção dos dados previstos por uma RNA, para adaptar a previsão às novas exigências do sistema.

Para o estudo de caso, mostrado no capítulo 5, procuramos avaliar o desempenho das RNA's para a solução do problema da previsão do período de ponta com discretização menor que uma hora, de forma a obter uma curva mais realista do ponto de vista do comportamento da carga. Para isto, analisamos o comportamento das RNA's para os horizontes de previsão propostos e a maneira como cada algoritmo realiza a previsão.

Procurou-se neste capítulo, apresentar de forma abreviada os principais procedimentos para previsão de carga a curto prazo utilizando RNA's. Assinale-se ainda, que os trabalhos apresentados aqui foram desenvolvidos via de regra para ambientes de demanda energética diversos do ambiente brasileiro.

Capítulo 5

ESTUDO DE CASO

Este capítulo apresenta a aplicação de RNA's através dos algoritmos propostos no capítulo 4 aos dados do sistema CESP. Algumas discussões são feitas quanto ao tipo de RNA's utilizadas na previsão. Em seguida faz-se uma análise dos dados históricos obtidos e da seleção das variáveis de entrada.

5.1- Introdução

No capítulo anterior, abordamos alguns métodos que foram desenvolvidos para previsão de carga a curto-prazo. Passemos agora a aplicar a metodologia proposta no capítulo anterior à previsão de carga do período de demanda de ponta baseado nos dados do sistema CESP (Companhia Energética do Estado de São Paulo).

A área de responsabilidade da CESP é o Estado de São Paulo, situado na região sudeste brasileira. O sistema CESP, faz parte do sistema interligado S/SE brasileiro e está ligado aos sistemas FURNAS, CPFL, ELETROPAULO, ITAIPU e CEMIG. A CESP tem uma capacidade instalada de 11.860 MW em 19 usinas, sendo a principal delas a usina de Ilha Solteira no rio Paraná com 3.240 MW de potência instalada. O mercado da CESP corresponde a uma carga própria média de 10.000 MW (sendo em torno de 14.000 MW de pico anual - 1993 e 1994), com crescimento de 4 % ao ano nos últimos 4 anos.

5.2- Período de Demanda de Ponta para o Sistema CESP

Utilizaremos os dados referentes ao sistema CESP para resolver o problema de previsão de carga no período de demanda de ponta para os horizontes de previsão citados no capítulo 4, seção 4.3.

O período de demanda de ponta para o sistema CESP começa às 17 horas e se estende até às 22 horas. A CESP discretiza o seu período de ponta em patamares de 10 em 10 minutos. Portanto, este período está sub-dividido em 31 intervalos de tempo.

No caso do sistema sudeste brasileiro, esta previsão torna-se mais complicada pois o período de demanda de ponta começa em torno das 17/18 h, e coincide com o final do dia, quando a carga de iluminação começa a crescer. Se no final do dia o tempo estiver nublado, então a carga de iluminação começa antes, modificando assim a entrada do período de ponta. Isto é bastante complicador, primeiramente por não se dispor de histórico específico sobre as condições climáticas ao final do dia, e segundo, uma boa previsão sobre o nível de luminosidade ao final do dia nem sempre é fácil, principalmente em períodos do ano em que o tempo é muito instável.

Na figura.5.1, mostramos a curva média do período de demanda de ponta da CESP para os anos de 1993 e 1994. Podemos observar que a carga do sistema possui três períodos distintos, os quais destacamos nas figuras 5.1 e 5.2. Podemos descrever estes três períodos como:

- **Período pré-ponta:** Este período vai das 17 até às 17:40 horas aproximadamente; é caracterizado por uma variância relativamente baixa durante o ano e sofre grande influência da luminosidade no final do dia;
- **Período da ponta:** Este período vai das 17:40 até às 20 horas aproximadamente; é caracterizado por uma elevada variância durante o ano e sofre grande influência da luminosidade no final do dia;
- **Período pós-ponta:** Este período vai das 20 às 22 horas; é caracterizado por uma baixa variância durante o ano e sofre pouca ou nenhuma influência da luminosidade no final do dia;

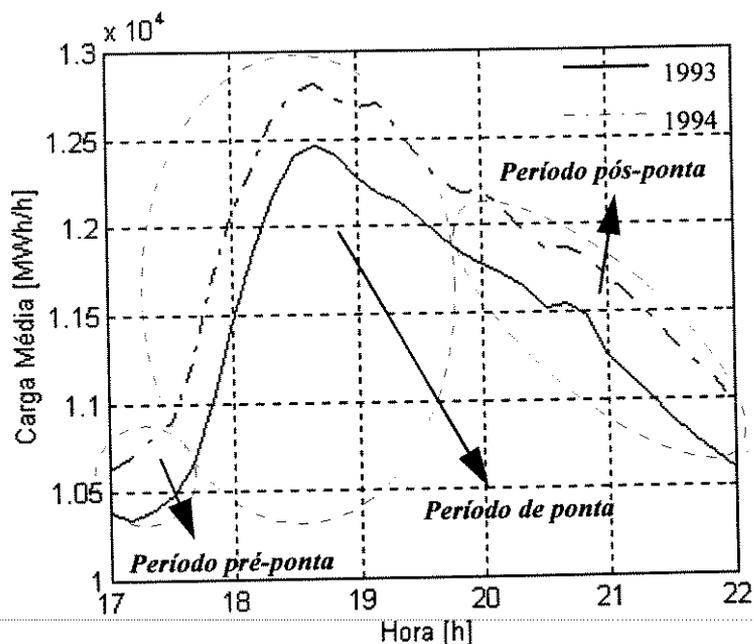


Figura.5.1: Perfil de Carga no Período de Demanda de Ponta da CESP - 1993 e 1994.

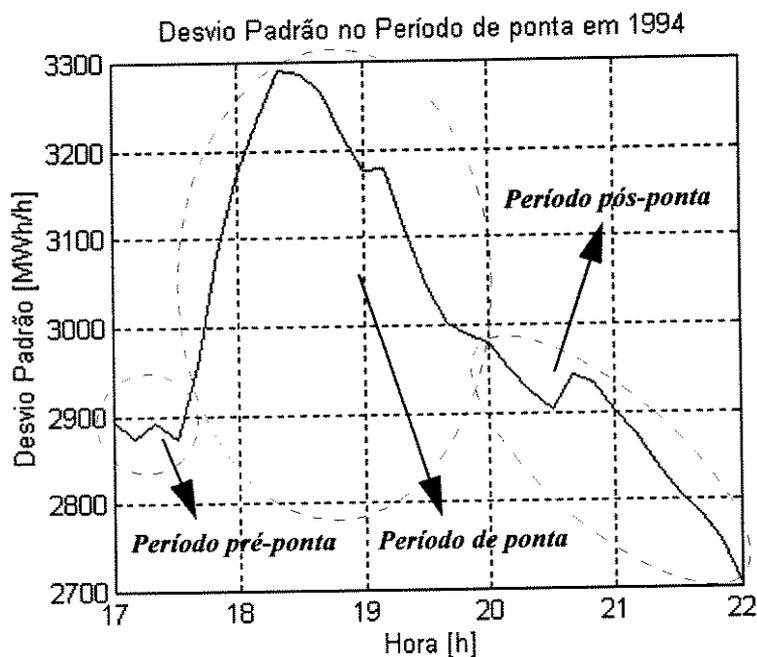


Figura.5.2: Desvio Padrão do Período de Ponta do Sistema CESP de 1994

5.2.1- Consideração de Efeitos Especiais para Previsão na Hora da Ponta

Antes de analisarmos os históricos de carga e de dados climáticos, destacaremos algumas particularidades da curva de carga para o período de ponta.

5.2.1.a- Efeito da Hora do Pôr-do-Sol

Os períodos pré-ponta e de ponta propriamente dito, possuem uma alta variância durante o ano. Isto se deve à grande flutuação da carga neste período influenciada fortemente pela luminosidade no final do dia.

A hora em que ocorre a máxima demanda durante o dia desloca-se ao longo do ano devido basicamente às variações sazonais do horário do crepúsculo. No verão, os dias são mais longos e o sol se põe tarde por volta das 19 horas. O pico de carga do dia acontece logo após o anoitecer devido a entrada da carga de iluminação e o aumento no consumo residencial.

No decorrer do ano com a chegada do inverno, o dia vai anoitecendo cada vez mais cedo. Em julho, por exemplo, o pôr-do-sol ocorre por volta das 17:30 horas. Logo, o pico de carga ocorre agora por volta das 18 horas, havendo assim um deslocamento do pico de carga durante o ano em função da hora do ocaso do sol, como pode ser observado na figura.5.3 que mostra a curva de carga para o período de ponta do sistema CESP para os dias 10/03/94 (verão) e 10/07/94 (inverno).

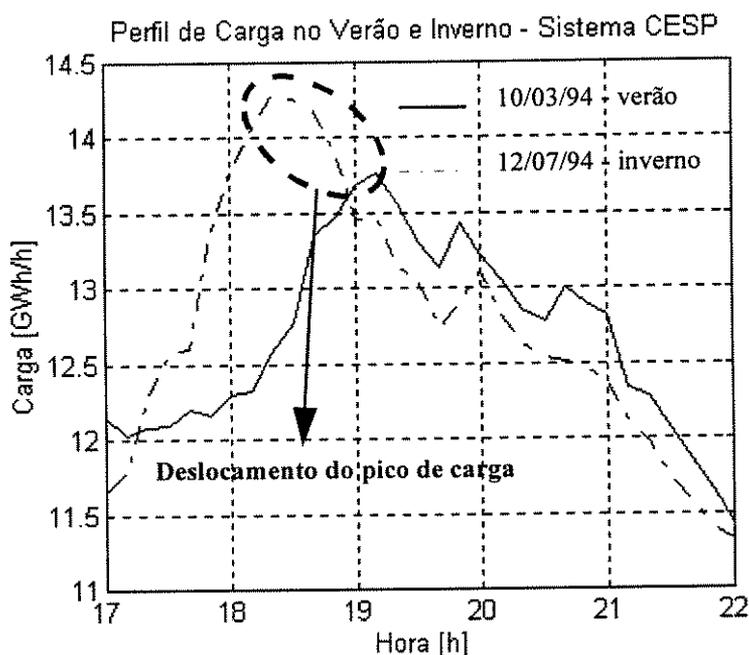


Figura.5.3: Perfil da Ponta no Verão e Inverno de 1994 do Sistema CESP

5.2.1.b- Efeito do Dia da Semana

É causado pelo ciclo trabalho-descanso da população com periodicidade de uma semana.

Para fazer a previsão do perfil de carga em patamares horários, a entrada de carga no início das segundas-feiras é bem diferente da mesma tomada de carga dos demais dias úteis da semana. O comportamento dos dias de semana pós-feriados é bem semelhante aos das segundas-feiras.

Entretanto, para o período de ponta as segundas-feiras são semelhantes às terças, quartas, quintas e sextas-feiras. Os sábados, domingos, feriados e dias especiais formam os outros grupos de dias semelhantes perfazendo um total de quatro classes distintas. No entanto, a preocupação deste trabalho é com os dias chamados típicos¹ e que não sejam finais de semana e nem feriados.

A figura 5.4, mostra os perfis médios no período de ponta dos quatro grandes grupos citados acima.

5.2.1.c- Efeito dos Feriados

A carga de feriados apresenta um comportamento bastante diferente não só nos dias equivalentes na mesma semana e nas semanas anteriores. É desejável manter em arquivo um calendário com os feriados do ano, de maneira a adotar-se automaticamente um procedimento adequado tanto para os feriados como para os dias antecedentes e subseqüentes.

¹ Dia típico - Dia que não seja feriado e que não apresente anormalidades na curva de carga causadas por fatores diversos, como jogos da copa do mundo.

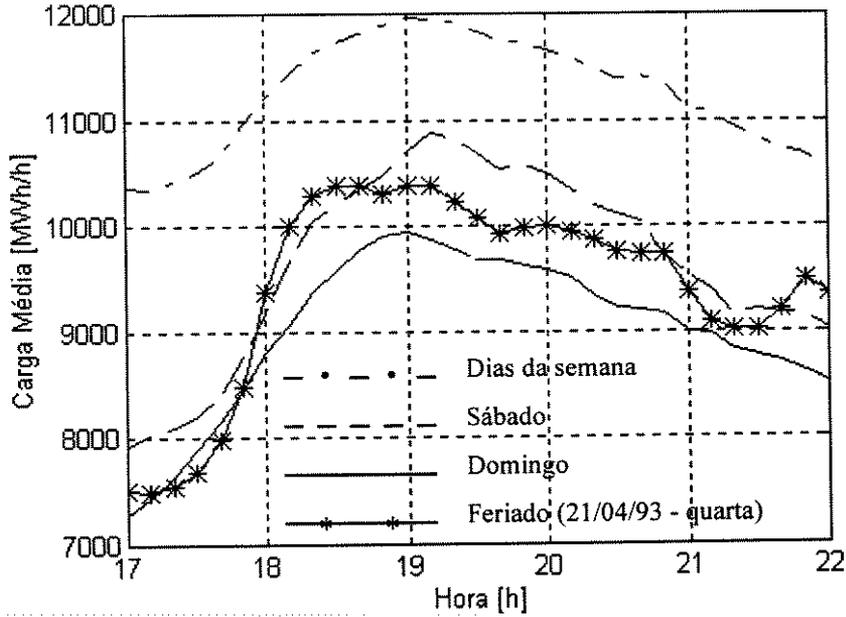


Figura.5.5: Perfil da Ponta das Classes de Padrões da Carga do Sistema CESP

A figura.5.6, mostra a curva de dois feriados de 1993 e 1994. A Sexta-Feira Santa ocorre sempre no mesmo dia da semana, entretanto o dia 21 de abril em cada ano ocorre em um dia da semana diferente: quarta-feira em 1993 e quinta-feira em 1994.

Note que os feriados podem ser divididos em mais sub-classes. O perfil de carga do dia 21 de abril (Tiradentes) é semelhante para todos os anos mas é diferente da Sexta-Feira Santa como podemos ver na figura.5.6.

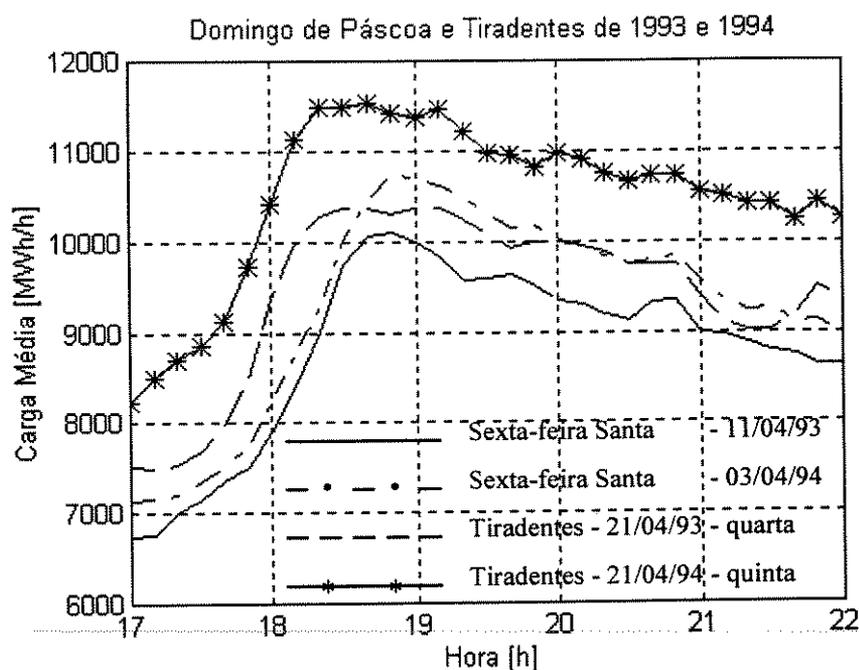


Figura.5.6: Feriados de Tiradentes e Sexta-Feira Santa de 1993 e 1994.

5.2.1.d- Efeito do Horário de Verão

Observou-se uma modificação no nível e na forma da curva de carga nos períodos de ponta, a partir do ano em que passou-se a adotar no Brasil o horário de verão, isto é, o adiantamento de uma hora nos relógios no final da primavera. Um cuidado especial deve ser tomado na correta interpretação dessa influência sobre a carga.

Podemos observar na figura.5.7 o deslocamento da curva de carga ocasionado pelo término do horário de verão em 1993 no sistema CESP.

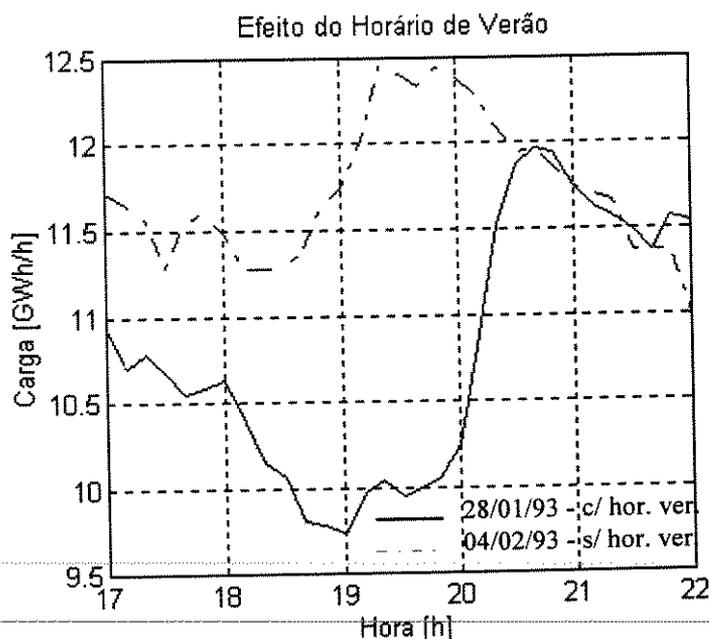


Figura.5.7: Horário de Verão e Seu Efeito Sobre a Curva de Carga

5.2.1.e- Comentários adicionais

O presente trabalho se propõe a discutir a aplicabilidade e eficiência do uso de RNA's sobre a previsão de carga no período de demanda de ponta do sistema CESP.

A carga elétrica no Brasil apresenta algumas características próprias. O padrão de consumo brasileiro, comparativamente às séries de consumo utilizadas na maior parte da literatura existente a respeito do tema, possui patamares per capita significativamente menores. Seu potencial de modulação diária em relação a fatores climáticos, por exemplo, é bastante baixo, apesar de haver alguma influência, principalmente, da temperatura e da cobertura nebulosa sobre a curva de carga na hora da ponta [86].

Outro fator de diferenciação da carga brasileira é o fato do Hemisfério Sul estar submetido a variações climáticas entre estações apreciavelmente inferiores às do Hemisfério Norte. Por isso, a componente de carga sensível à temperatura, no Brasil, é menos significativa do que nas séries analisadas na maior parte dos trabalhos que se referem a previsão de carga a curto-prazo.

Com relação aos erros inerentes a se projetar a carga por um método automático baseado em inferências históricas, pode-se comentar o seguinte:

- Uma das fontes de erro é a relevância das variáveis de entrada, que pode deixar de modelar algumas características da série; quando há pouco ganho em exatidão, pode não valer a pena tornar mais complexo o modelo.

- Erros derivam também das flutuações aleatórias, sem correlação serial, inerentes ao próprio processo físico, ou a imprecisões introduzidas durante o processo de medição da série real.

5.3- Análise dos Dados da CESP

Os dados disponíveis para testar os algoritmos propostos na previsão da carga do sistema CESP, são os seguintes:

5.3.1- Dados de carga

- **Histórico de carga horária.** Refere-se ao perfil de carga própria do Estado de São Paulo durante o dia. Cada valor deste histórico corresponde à carga média entre um intervalo e outro. Logo, a carga de 12500 MWh/h às 15 horas é a carga média entre às 14 e 15 horas. Os dados disponíveis correspondem à curva de carga diária em patamares de uma hora dos anos de 1992, 1993 e 1994. Dados cedidos pela CESP.

- **Histórico de carga no horário de ponta.** Refere-se ao perfil de carga própria do Estado de São Paulo do período de demanda de ponta (das 17 às 22 horas). Cada valor deste histórico corresponde à carga média entre um intervalo e outro da ponta. Logo, a carga de 11650 MWh/h às 18:40 horas é a carga média entre às 18:30 e 18:40 horas. Os dados disponíveis correspondem à curva de carga na ponta em

patamares de dez minutos dos anos de 1992(de 01/07 a 31/12), 1993(de 01/01 a 08/06) e 1994(01/01 a 23/08). Dados cedidos pela CESP.

5.3.2- Dados climáticos

- **Histórico de temperatura.** Refere-se às temperaturas máximas e mínimas diárias, registradas em: São Paulo (capital): 1993 (de 01/01 a 30/09); Campinas: 1992 (o ano todo), 1993 (o ano todo) e 1994 (de 01/01 a 30/11); Ribeirão Preto: 1992 (de 01/01 a 31/04); Pindorama: 1992 (o ano todo) e 1993 (de 01/01 a 31/07). Dados cedidos pelo Instituto Agronômico de Campinas - IAC.

- **Histórico da hora do pôr-do-sol.** Refere-se à hora do pôr-do-sol. Os dados correspondem à hora do pôr do-sol no aeroporto internacional de Guarulhos em São Paulo. Dados cedidos pelo Ministério da Aeronáutica.

Outros dados climáticos foram conseguidos, tais como, chuva nas últimas 24 horas, temperatura média, temperatura de relva, umidade relativa do ar, etc, na cidade de São Paulo, Ribeirão Preto e Pindorama. Entretanto estes dados não estão totalmente disponíveis de forma contínua no tempo e sim, em períodos de alguns meses.

As reduções da luminosidade, devido à intensa cobertura nebulosa, podem afetar significativamente a forma da curva de carga no período de ponta. Seriam, portanto, necessários os dados de cobertura do céu, luminosidade e pluviosidade no final do dia. Contudo estes dados não puderam ser obtidos.

5.3.3- Outros dados

Ainda outros dados são necessários para a aplicação das RNA's nos algoritmos propostos. Os dados referentes ao dia, mês, ano, informações sobre feriados e dias especiais e dia da semana foram obtidos do calendário brasileiro de 1993 e 1994.

Os históricos não estão totalmente disponíveis de forma contínua no tempo e sim, em sessões de alguns meses. Portanto, devido à este problema, o período selecionado para se testar as redes neurais como ferramenta de previsão de carga do período de ponta, foi o primeiro semestre de 1993 e o primeiro semestre de 1994 (de janeiro a julho), cobrindo assim tanto o ápice do verão como do inverno. Estes dois períodos são os únicos que estão com todos os dados disponíveis.

5.4- Considerações no Projeto de RNA's

Sabe-se que a arquitetura das RNA's a nível de entrada e saída, é fortemente determinada pela aplicação. Podemos observar na literatura o uso de RNA's multi-camadas diretamente alimentadas utilizando o BP como algoritmo de aprendizagem.

O pouco entendimento das complexas interrelações entre os parâmetros que influenciam o desempenho da rede, como função de ativação, função de custo, parâmetros de controle do gradiente descendente e do tempo de treinamento e arquitetura da rede, permite um grau de liberdade extra no projeto de uma RNA. Este grau de liberdade na escolha dos parâmetros adequados, torna as aplicações altamente dependentes da experiência do projetista e da experimentação manual para que se obtenham sucesso [1].

5.4.1- Fatores determinantes no desempenho das RNA's

O desempenho das RNA's, em geral, é avaliado em função de três fatores:

- Convergência
- Generalização
- Escalabilidade

A convergência é um fator importantíssimo para o bom desempenho das RNA's, pois diz respeito à capacidade das RNA's de aprender um determinado conjunto de dados definido no procedimento de treinamento. A convergência é o que indica quando o procedimento de treinamento é capaz de aprender a classificação definida nos padrões apresentados, sob que condições da rede e sob quais exigências computacionais. RNA's com topologias fixas que utilizam o algoritmo de retro-propagação provam a convergência mostrando que um tempo limite, quando o tempo de treinamento tende a infinito, o erro minimizado pelo método do gradiente descendente tende a zero [51].

A generalização diz respeito à habilidade das RNA's de reconhecer padrões fora do conjunto de treinamento e é a principal propriedade que deve ser exigida de um sistema neural. Apesar das várias aplicações bem sucedidas de RNA's descritas na literatura [5, 6, 58, 88], as condições para que as boas generalizações sejam encontradas ainda não são totalmente conhecidas.

Por fim, a escalabilidade, que diz respeito a ambos, convergência e generalização, é relativa ao procedimento de ajustar os parâmetros das RNA's para que elas consigam convergir e generalizar bem. Sabe-se que para uma RNA ter um bom desempenho, ela deve ser grande o suficiente para aprender o problema (convergir), mas também pequena o bastante para generalizar bem.

5.4.2- Controle dos fatores determinantes no desempenho das RNA's

O controle dos fatores determinantes de desempenho se dá pela escolha adequada dos parâmetros da rede como: função de ativação, função de custo, parâmetros de controle do gradiente descendente e do tempo de treinamento e arquitetura da rede, os quais, como dito anteriormente, variam de aplicação para aplicação e não possuem regras de escolha muito bem definidas.

5.4.2.a - Padronização

Apesar de não controlar diretamente os fatores determinantes de desempenho das RNA's, a padronização pode, frequentemente, melhorar a performance das RNA's através da remoção de características insignificantes do conjunto de dados, pois informações relevantes das séries são, geralmente, independentes de valores de compensação (*offset*) e de desvios padrão. Estas características podem obscurecer a tarefa das RNA's. Não faremos a padronização das variáveis de carga e clima, para não retirar destes dados informações preciosas dos dados do sistema.

A padronização, entretanto, admite várias formas e exige cuidados, já que é muito fácil remover, inadvertidamente, informações preciosas dos dados. Por outro lado, a padronização correta dos dados pode melhorar consideravelmente o desempenho de uma rede neural [1].

5.4.1.b- Normalização

Antes de se utilizar o conjunto de treinamento, os valores de X^p e Y^p devem ser todos normalizados em uma faixa adequada, por exemplo, convertidos para valores que fiquem entre 0 e 1, inclusive. Uma forma de se realizar esta normalização é aplicar a seguinte fórmula para cada elemento X^p do vetor:

$$x_{k,norm}^p = \frac{x_k^p - x_{min}}{x_{max} - x_{min}} \quad (5.1)$$

onde x_{min} é o valor mínimo encontrado ou determinado em X e x_{max} é o valor máximo encontrado ou determinado em X . A mesma fórmula se aplica ao conjunto de saída Y . Outras funções de normalização e transformação (por exemplo, logaritmo) podem ser utilizadas.

VARIÁVEL	Máximo	Mínimo
Carga	14.500 [MWh/h]	6.000 [MWh/h]
Temperatura máxima	45 [°C]	25 [°C]
Temperatura mínima	20 [°C]	0 [°C]
Hora do pôr-do-sol	19:00 [h]	17:30 [h]
Dia da semana	1	0
Hora da previsão	22:00 [h]	17:00 [h]
Horário de verão	1	0

Tabela.5.1: Valores Máximos e Mínimos das Variáveis de Entrada e Saída.

A normalização é importante, porque se um ou mais neurônios tiverem valores de ativação muito maiores ou menores do que os demais, isto poderá produzir valores de erro desproporcionais, e levar a um desbalanço da rede, conseqüentemente prejudicando a sua convergência [51].

A normalização foi feita de acordo com a equação 5.1 para os valores máximo e mínimo mostrados na tabela.5.1 acima.

5.4.1.c- Funções de ativação

É possível utilizar outras funções de ativação além da sigmóide, tais como logística, sigmóide segmentada, senóides ou cossenóides, cossenóides truncadas, etc. O importante é que seja uma função não-linear. A escolha da função também pode ter repercussões sobre a velocidade de aprendizado, principalmente pelo número de operações em ponto flutuante [51].

No entanto, para previsão de carga, alguns autores [83] já testaram o desempenho de várias destas funções de ativação não havendo ganho de desempenho significativo. Portanto a função sigmóide foi usada em todas as RNA's deste trabalho.

5.4.1.d- Funções de custo

A função de custo ou medida de erro, informa o grau de convergência das RNA's. Através dela, é possível saber o quanto as RNA's estão aprendendo os padrões apresentados durante o treinamento.

Alguns autores [52, 63] utilizam a *SSE* (*Sum Square Error*) como função de custo. A *SSE* é dada pela fórmula:

$$SSE = \frac{1}{2} \sum_{p=1}^P \sum_{i=1}^I (alvo_i^p - saida_i^p)^2, \quad (5.2)$$

onde I é o número de neurônios na camada de saída.

A função da equação 5.2 não é a única escolha possível para a função de custo. Nós podemos trocá-la por qualquer outra função diferenciável que é minimizada quando seus argumentos são iguais, e a derivada corresponde a regra de correção dos pesos das sinapses [52].

Para escolher a melhor função de custo para a previsão da ponta do sistema CESP, testamos três funções de custo citadas na literatura, escolhendo a função *SSE* como função de custo para todas as RNA's utilizadas neste trabalho.

5.4.1.e- Parâmetros de controle do gradiente descendente e do tempo de treinamento

Um importante mecanismo para controlar o desempenho do aprendizado das RNA's é a escolha dos parâmetros de controle do gradiente descendente e do tempo de treinamento. Vários cuidados devem ser tomados para se obter um treinamento mais rápido que evite regiões de mínimos locais, platôs e paralisia.

O tamanho dos pesos iniciais aleatórios do treinamento é importante. Se eles forem muito grandes a sigmóide irá saturar logo no início do treinamento, e o sistema ficará preso num mínimo local (ou num platô muito plano) perto do ponto inicial [52].

O gradiente descendente pode ser muito lento se a taxa de aprendizagem η for muito pequena, e pode oscilar grandemente se η for grande. Há muitas maneiras de tratar este problema, incluindo trocar o gradiente descendente por algoritmos de minimização mais sofisticados. Entretanto, uma maneira muito simples de tratar o problema é adicionar um termo de momento (*momentum*) [89], que é frequentemente eficaz e comumente usado.

A idéia é dar a cada conexão w_{ij} alguma inércia ou momento, tal que ele tende a mudar na direção de descida de $E(\cdot)$, ao invés de oscilar grandemente a cada iteração. Então a taxa de aprendizado eficaz pode ser grande sem oscilações divergentes. Este esquema é implementado por dar uma contribuição da mudança ocorrida no passo anterior da seguinte forma:

$$\Delta w_{ij}(t+1) = -\eta \cdot \frac{\partial E}{\partial w_{ij}} + \alpha \cdot \Delta w_{ij}(t) \quad (5.3)$$

onde o parâmetro momento α precisa estar entre 0 e 1.

Segundo alguns autores “a escolha de η e α é uma arte”, uma vez que η depende da função a aproximar [63].

A abordagem usual é checar se a correção de um peso de uma sinapse particular realmente decresceu a função de custo. É preferível aumentar η por uma constante, mas decrescê-la geometricamente para permitir uma queda rápida quando necessário. Logo η pode ser corrigida da seguinte forma a cada época

$$\Delta \eta = \begin{cases} +a, & \text{se } \Delta E < 0; \\ -b \cdot \eta, & \text{se } \Delta E > 0; \\ 0, & \text{caso contrario.} \end{cases} \quad (5.4)$$

onde ΔE é a diferença da função de custo, e a e b são constantes apropriadas geralmente com os valores de 0.1 para cada uma delas [52].

O termo momento α , foi considerado constante e igual a 0.1 para todas as RNA's utilizadas neste trabalho.

Embora sejam inicializados aleatoriamente no início do aprendizado, os limiares de cada neurônio podem percorrer gamas mais baixas ou mais elevadas, conforme as características da rede neural empregada [51]. Este parâmetro também pode ter influência na capacidade e velocidade de aprendizado.

Os valores dos limiares são em geral constantes e regulam a sensibilidade de um neurônio a variações em suas funções de ativação. Neste trabalho escolhemos corrigir o valor de cada limiar da mesma maneira que corrigimos os pesos das sinapses, considerando o limiar como a contribuição de um neurônio de saída sempre igual a -1 e com conexão com todos os outros neurônios da próxima camada como mostrado na figura 3.6.

Quanto à maneira de se corrigir os pesos das sinapses foi escolhido o método incremental (*pattern-by-pattern*) onde os pesos são corrigidos sempre depois que cada padrão é apresentado à rede.

Deve-se sempre evitar treinamentos prolongados demais (*overtraining*). A rede é treinada com o conjunto de pares de treinamento e a performance obtida, medida com o conjunto de pares de teste. Isto é necessário porque um treinamento longo leva a uma super-especialização da rede (especialmente no caso em que poucos pares entrada-saída são disponíveis para o treinamento) que piora a performance quando os dados de teste são apresentados.

O treinamento deve ser paralisado sempre que o erro do conjunto de testes começar a aumentar mesmo que a função de custo do conjunto de treinamento decresça

[63]. Limitamos assim o número de épocas do treinamento em 100 de acordo com a observação da curva $E(.)$ X *Nr. de Iterações* dos testes realizados.

5.4.1.f- Dimensionamento da RNA

A arquitetura da rede é muito importante para as capacidades de convergência e generalização das RNA's. A escolha de um arquitetura adequada, nem tão pequena e nem tão grande, é a principal responsável pelo sucesso de uma aplicação. Para se obter uma boa habilidade de generalização, deve-se projetar a rede colocando dentro dela o máximo conhecimento possível sobre o problema (isto é, a topologia do espaço de entradas) e limitar o número de conexões adequadamente.

O dimensionamento da rede depende muito do tipo de função a aproximar, e a forma mais usada para realizá-lo ainda é a tentativa-e-erro, embora várias técnicas de se encontrar as arquiteturas apropriadas já foram propostas [52, 90], inclusive o uso de algoritmos genéticos [91]. Neste trabalho escolhemos a melhor topologia para cada algoritmo por tentativa-e-erro. Os resultados são apresentados no capítulo 6. As redes com três camadas conseguiram uma previsão mais precisa em comparação com as redes de duas camadas, apesar da sua menor velocidade de treinamento. Redes com quatro camadas se demonstraram muito lentas no treinamento além de obterem erros médios percentuais absolutos elevados.

Portanto, as redes a serem utilizadas nos algoritmos *MWS* e *AAS* serão multi-camadas diretamente e totalmente conectadas de três camadas, utilizando o algoritmo *backpropagation* do tipo incremental com taxa de aprendizagem adaptativa e termo momento constante. Todos os neurônios possuem função sigmóide como função de ativação. A função de custo utilizada em todos os treinamentos foi a *SSE* (equação 5.2). O número de épocas por treinamento foi limitado em 100. O critério de parada do treinamento foi o número de épocas ou a diferença entre os *SSE*'s de duas iterações consecutivas menor que 10^{-4} .

5.5- Seleção das Variáveis de Entrada.

Um dos mais importantes passos na construção de um modelo de rede neural para previsão de carga é a seleção das variáveis de entrada. Não há uma regra geral que possa ser seguida neste processo. É totalmente dependente do julgamento do projetista e de sua experiência e é processado quase inteiramente por tentativa-e-erro [84].

Deve-se, entretanto, verificar se existem variáveis relevantes para o problema de classificação. Isto pode ser feito através de testes estatísticos de associação e correlação. Assim, podem ser incluídas apenas as variáveis que apresentam um alto grau de associação positiva ou negativa, e retiradas aquelas que não apresentam associação ou correlação significativas [51].

Para os dados do histórico de carga horária foram feitas a decomposição espectral e a função de autocorrelação deste histórico para os anos de 1993 e 1994, mostradas nas figuras.5.8 a 5.11.

Podemos observar nestas figuras que há uma componente de 24 horas e 168 horas com uma correlação muito forte com a hora atual. Ou seja, o valor da carga em uma hora qualquer tem grande influência do valor da carga que ocorreu 24 e 168 horas atrás. Portanto estas variáveis, bem como a carga algumas horas antes, serão consideradas para a previsão.

Com relação às variáveis climáticas, as temperaturas máximas e mínimas foram consideradas. Para a previsão 24 horas à frente, foi montado um histórico de temperaturas previstas com a inserção de um ruído no histórico real de temperatura, de forma a simular a previsão de temperatura, já que este histórico de temperaturas previstas não pôde ser obtido.

Vários problemas foram enfrentados na consideração das variáveis, temperatura máxima e mínima ocorridas no dia. Entre eles, podemos destacar, além da falta de continuidade dos dados, o problema de saber de qual localidade a temperatura será

considerada. Poderia se fazer uma média das temperaturas da cidade de São Paulo, Campinas, Pindorama e Ribeirão Preto. Entretanto, seriam necessários muitos neurônios (8), para considerar esta variável de uma forma completa. Como os dados de carga são referentes a todo o Estado de São Paulo, resolveu-se considerar a temperatura referente às regiões com maior participação no montante da demanda de energia.

O histórico de temperatura da cidade de São Paulo está muito deficiente e como praticamente a curva de temperatura máxima e mínima de São Paulo e Campinas são semelhantes, foram selecionados apenas os históricos de temperatura máxima e mínima da cidade de Campinas como informação de temperatura para as redes treinadas.

.....
A figura.5.12 mostra as temperaturas máximas e mínimas do 1º semestre de 1993 da cidade de Campinas.

5.6- Testes Realizados

Foram realizados diversos testes para se avaliar o desempenho dos algoritmos *AAS* e *MWS* na previsão de carga (estática e dinâmica) do período de demanda de ponta para os horizontes de 10 minutos, uma e 24 horas à frente com os dados dos históricos de carga do sistema CESP e dados climáticos.

A arquitetura e as variáveis de entrada da rede foram escolhidas em sua grande maioria por tentativa-e-erro de maneira a identificar as variáveis que realmente são relevantes para a previsão para que as RNA's treinadas pelos algoritmos, tenham o menor número de neurônios possível e sejam rápidas e robustas.

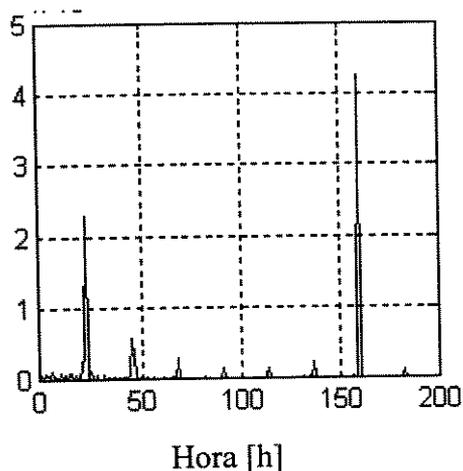


Figura.5.8: Decomposição Espectral da Carga Horária de 1993.

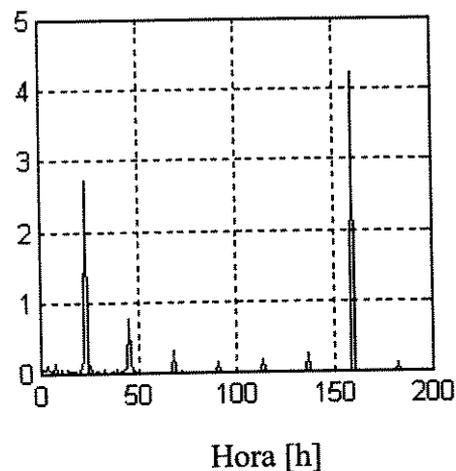


Figura.5.10: Decomposição Espectral da Carga Horária de 1994.

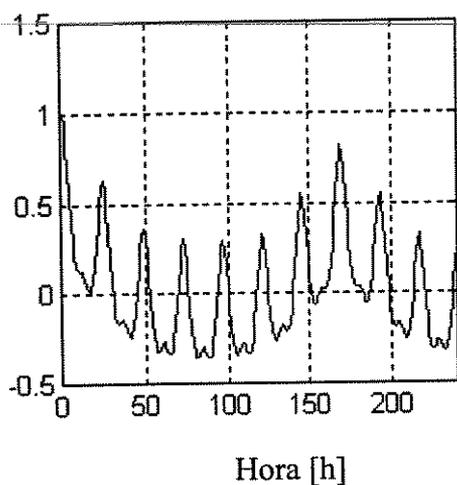


Figura.5.9: Função de Auto-Correlação da Carga Horária de 1993.

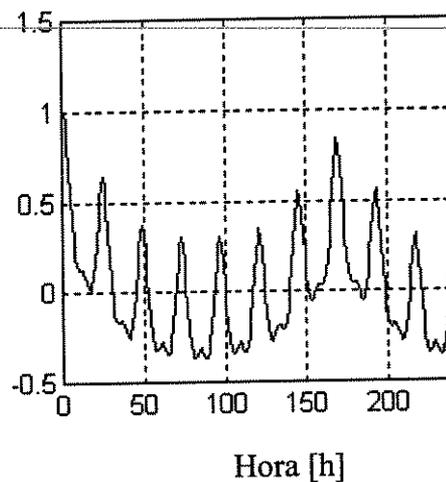


Figura.5.11: Função de Auto-Correlação da Carga Horária de 1994.

Os testes foram feitos de modo a encontrar as melhores arquiteturas para os algoritmos propostos do ponto de vista da precisão da previsão e do desempenho das RNA's, para os horizontes de previsão considerados. Os resultados e sua análise são mostrados no capítulo 6.

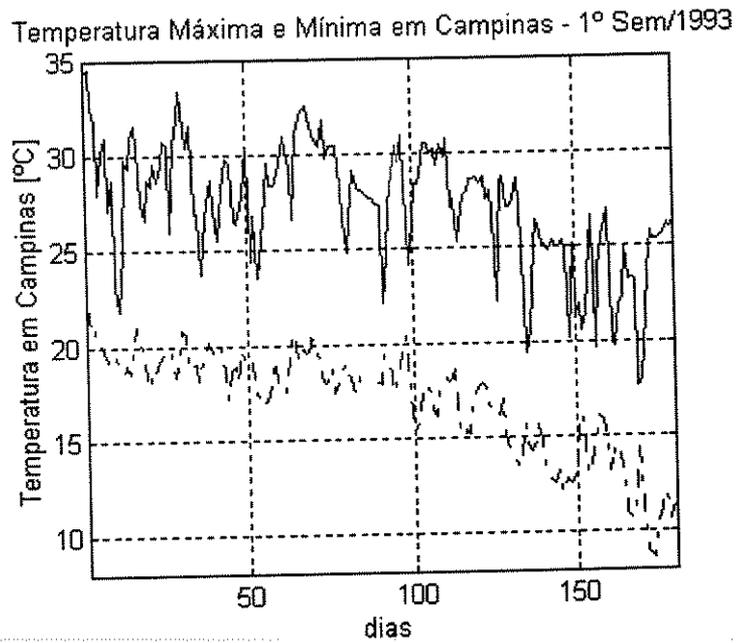


Figura.5.12: Temperatura Máxima e Mínima do 1º Semestre de 1993 em Campinas.

Note que os testes com o algoritmo *AAS* só foram possíveis para o 1º semestre de 1994 por insuficiência de dados. Para o algoritmo *MWS* os testes foram feitos para o 1º semestre de 1993 e 1994.

Capítulo 6

ANÁLISE DOS RESULTADOS

Este capítulo apresenta os resultados obtidos da aplicação de RNA's através dos algoritmos propostos no capítulo 4, para o sistema CESP. Apresenta-se os métodos de avaliação dos resultados. Em seguida analisa-se os resultados do ponto de vista do desempenho das RNA's e da exatidão da previsão para os horizontes de curto e curtíssimo prazos.

6.1- Introdução

Durante o estudo de caso foram feitos vários testes com RNA's de três camadas utilizando BP como algoritmo de aprendizagem. Como a modelagem de RNA's é fortemente dependente da aplicação, propomos resolver o problema de previsão de carga no período de demanda de ponta com os algoritmos *MWS* e *AAS* para os horizontes de curto e curtíssimo prazos.

Nestes testes foram manipulados vários parâmetros das RNA's em busca de resultados cada vez melhores. Estas variações foram feitas de maneira sistemática para que fosse possível analisar se as mudanças tinham sido vantajosas e quais os fatores responsáveis por esta melhora.

Para validação estatística das respostas das RNA's foram rodados cada caso 100 vezes. Os valores mostrados nas tabelas, referem-se aos piores valores para estes 100 testes.

6.1.1- Métodos de Avaliação dos Resultados

Com o objetivo de avaliar os resultados das RNA's de forma consistente, foram realizadas medidas diferentes de erro, onde cada medida observou os resultados por um prisma diferente. As medidas utilizadas na avaliação dos resultados foram os seguintes: comparação gráfica, diagramas de dispersão ou *scattergrams*, erro médio

percentual (*MPE - Mean Percent Error*), erro médio percentual absoluto (*MAPE - Mean Absolute Percent Error*), desvio padrão do erro percentual absoluto (*SDAPE - Standard Deviation Absolute Percent Error*), valor máximo do erro absoluto (*MAXAPE - Maximun Absolute Percent Error*), número de previsões com erro maior que 800 MW (*NPM800*).

A seguir, faz-se uma rápida descrição de cada uma dessas medidas ressaltando-se como elas devem ser utilizadas, quais suas vantagens e seus problemas.

6.1.1.a- Comparação gráfica

É a mais simples das avaliações. A comparação gráfica apenas plota, em um mesmo gráfico, os resultados alvo e os previstos pelas RNA's para uma análise visual. A figura.6.1 mostra a comparação gráfica da previsão 24 horas à frente estática do dia 20/04/94 através do algoritmo *MWS-ALL*.

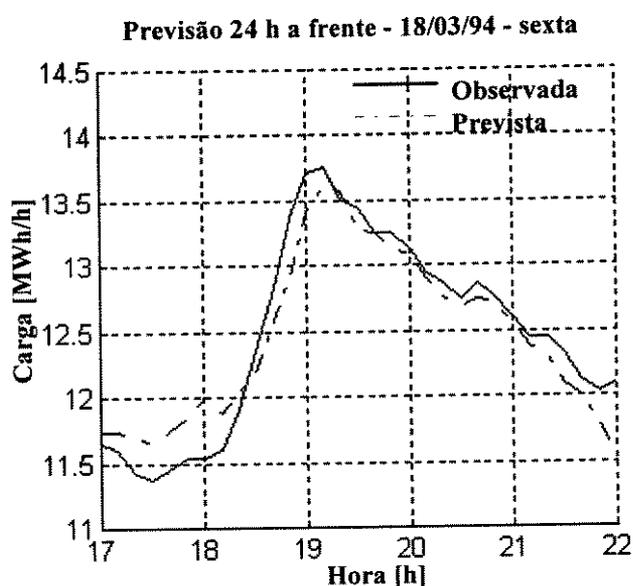
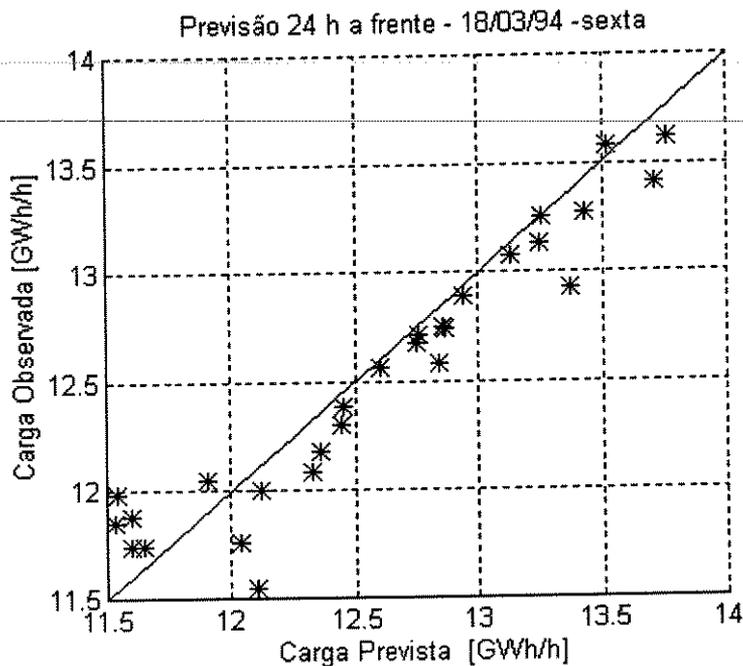


Figura.6.1: Comparação Gráfica Entre os Resultados Previstos e Observados Para o Algoritmo *MWS-ALL*.

6.1.1.b- Diagramas de Dispersão ou *Scattergrams*

Nesta métrica plota-se os dados como pontos no gráfico com o intuito de mostrar o quão próximo a previsão está do valor alvo. No eixo x pode-se plotar o valor previsto pelas RNA's e no eixo y o valor observado. Idealmente, os pontos devem formar uma linha de 45° que passa pela origem. A razão é óbvia, pois quando se possui um alvo x_i deseja-se que a previsão seja mais precisa possível, o que significa $y_i = x_i$, logo ter-se-á pares cartesianos (x_i, y_i) onde $y_i = x_i$. A figura.6.2 mostra o "scattergrama" da previsão 24 horas à frente estática do dia 20/04/94 através do algoritmo *MWS-ALL*.



$$MPE = \frac{1}{NF} \cdot \sum_{i=1}^{NF} \left(\frac{x_i - f_i}{x_i} \right) \quad (6.1)$$

onde: NF - Número de previsões.

x_i - Carga observada.

f_i - Carga prevista.

A vantagem do MPE é observar se a previsão foi mais pessimista que otimista de acordo com o sinal de MPE . Se a previsão não for tendenciosa, MPE deverá ser nulo aproximadamente.

6.1.1.d- Mean Absolute Percent Error (MAPE)

O $MAPE$ é muito utilizado para medir a média dos desvios relativos entre o valor observado e o valor previsto, motivo pelo qual ele é muito utilizado na literatura como parâmetro de precisão da previsão. O $MAPE$, nos dá uma noção do grau de precisão média da previsão. O $MAPE$ é calculado por:

$$MAPE = \frac{1}{NF} \sum_{i=1}^{NF} APE_i \quad (6.2)$$

onde :

$$APE_i = \left| \frac{x_i - f_i}{x_i} \right| \cdot 100\% \quad (6.3)$$

6.1.1.e- Standard Deviation Absolute Percent Error (SDAPE)

O $SDAPE$ é o desvio padrão do erros percentuais absolutos (APE_i) da previsão em relação ao valor observado. O $SDAPE$ mede o grau de dispersão destes erros percentuais absolutos. O $SDAPE$ pode ser calculado da seguinte maneira:

$$SDAPE = \sqrt{\frac{\sum_{i=1}^{NF} (APE_i - MAPE)^2}{NF - 1}} \quad (6.4)$$

6.1.1.f- *Maximun Absolute Percent Error (MAXAPE)*

O maior valor do MAPE é dado por MAXAPE e corresponde ao maior erro cometido durante todo o conjunto de previsão.

Para um sistema com um grande mercado de energia elétrica, como o caso do sistema CESP, torna-se importante saber qual o erro em *MW* que determina para o operador, se a previsão é boa ou não. Portanto, deve-se considerar este limite de modo a garantir a segurança e confiabilidade da previsão sem colocar em risco a operação do sistema. Então, um dos objetivos do modelo de previsão é garantir uma margem de segurança do erro da previsão. Para a CESP este valor foi considerado de 800 *MWh/h* (8% da carga média da concessionária.) como limite de segurança.

Portanto, previsões com erros superiores a 800 *MWh/h* serão consideradas péssimas previsões. Chamaremos o número de previsões acima deste limite (800 *MWh/h*) de *NPM800*.

6.2- Resultados Obtidos

Dentre as melhores arquiteturas encontradas para cada horizonte de previsão e algoritmo de previsão, destacamos na tabela.6.1 as que obtiveram as melhores performances com relação ao desempenho das RNA's e a exatidão da previsão. Os resultados são melhor detalhados no apêndice A.

O número de previsões variou, de um algoritmo para outro. Para o algoritmo *MWS* foram feitas 2077 previsões em 1993 e 1767 em 1994. Para o algoritmo *AAS* o número de previsões foi de 2108 em 1994.

Analisando os resultados apresentados na tabela.6.1, é possível constatar alguns fatos importantes os quais foram ordenados e são discutidos a seguir.

◆ Quanto ao método de previsão

Podemos observar que o método dinâmico de previsão da ponta obteve sempre o melhor desempenho geral em comparação com o método estático. Analisando a comparação gráfica de cada dia previsto com este método, percebemos um melhor acompanhamento por parte da previsão em relação ao valor observado na série de carga, o que ocasionou um menor *SDAPE* para estes casos.

Algoritmo	Horizonte	Ano	Arquitetura	MPE	MAPE	SDAPE	MAXAPE	0 —4%
				%	%	%	%	
MWS-ALL estático	1 h a f.	1993	3 X 2 X 31	1,29	2,27	1,93	10,20	83,09
		1994	3 X 2 X 31	0,96	2,23	2,13	11,82	82,71
	24 h a f.	1993	4 X 2 X 31	1,38	2,26	1,96	9,74	83,24
		1994	4 X 2 X 31	0,87	2,17	2,13	12,35	83,62
MWS-PPT dinâmico	1 h a f.	1993	6 X 3 X 1	0,72	2,01	1,64	7,65	87,81
		1994	6 X 3 X 1	0,44	1,94	1,86	9,79	86,36
	24 h a f.	1993	6 X 3 X 1	0,71	2,01	1,63	8,72	88,35
		1994	6 X 3 X 1	0,43	1,93	1,86	10,12	86,35
	10 m a f.	1993	5 X 3 X 1	0,75	1,91	1,56	7,61	88,82
		1994	5 X 3 X 1	0,48	1,84	1,77	9,67	88,09
AAS-ALL estático	1 h a f.	1994	10 X 5 X 31	1,56	2,78	1,97	11,68	74,62
	24 h a f.	1994	9 X 5 X 31	1,87	2,97	1,92	12,65	71,19
AAS-PPT dinâmico	1 h a f.	1994	11 X 6 X 1	0,95	2,56	1,74	12,88	81,10
	24 h a f.	1994	11 X 6 X 1	1,07	2,70	2,23	13,25	76,04

Tabela.6.1: Resumo dos Melhores Resultados e RNA's para a Previsão da Ponta do 1º Semestre de 1993 e 1994.

◆ Quanto ao algoritmo de previsão

O algoritmo *MWS* obteve uma performance na previsão sempre superior ao algoritmo *AAS*. As RNA's treinadas neste algoritmo são sempre menores, mais

ⁱ arquitetura (l X m X n) significa: rede feedforward com 3 camadas, tendo l neurônios na camada de entrada, m neurônios na camada intermediária e n neurônios na camada de saída.

especializadas e são treinadas mais rapidamente, sem a necessidade de uma grande massa de dados em seu conjunto de treinamento.

A nível diário, a adaptabilidade do conjunto de treinamento das RNA's do algoritmo *MWS* aos novos padrões de carga, devido à sazonalidade das variáveis de entrada e saída durante o período de treinamento, foi satisfatória, pois o horizonte de 24 horas é muito menor do que o período das componentes sazonais da carga. Logo, o padrão do novo dia típico não se diferencia tanto com o decorrer do ano se este sempre for perto do mais recente dia típico encontrado no conjunto de treinamento da RNA.

Em relação ao tempo de treinamento e requisitos de memória, o algoritmo *MWS* mostrou-se mais eficiente que o algoritmo *AAS*.

◆ Quanto a variável horário de verão

Quando acaba o horário de verão, o próximo dia a ser previsto terá seu período de ponta deslocado em aproximadamente uma hora. O deslocamento do pico de carga ocasionado pelo início/término do horário de verão causa os seguintes problemas:

- *para o algoritmo MWS:*

Para o *MWS-ALL*, pudemos observar que o fim do horário de verão provoca erros consideráveis na previsão do período de ponta devido ao deslocamento do pico de carga, para cerca de uma hora antes. Observando o funcionamento do algoritmo *MWS-ALL*, verificamos que a previsão do primeiro dia pós horário de verão deveria realmente trazer tamanhos erros, pois no seu conjunto de treinamento não apareceria nenhum padrão que pudesse conter esta informação.

Entretanto, a previsão dos demais dias também ocorreu sem que a RNA conseguisse corrigir o pico de carga. O comportamento da carga pós horário de verão só foi assimilada totalmente após 9 últimos dias típicos, já fora do horário de verão, fossem incorporados no conjunto de treinamento que possuía um número máximo de 10

padrões típicos. Ou seja, é necessário que o conjunto de treinamento (de 10 padrões) seja composto majoritariamente com os novos padrões.

Isto ocorreu para as RNA's com e sem a variável horário de verão em 1993 e 1994, para o algoritmo *MWS-ALL* nos horizontes de uma e 24 horas à frente e para o *MWS-PPT* com horizonte de 10 minutos a frente (previsão a curtíssimo prazo).

Com a inclusão desta variável à entrada da rede, pudemos observar uma perturbação grande nas previsões pós horário de verão gerando elevada instabilidade na rede depois do fim do horário de verão para o *MWS-ALL* uma e 24 horas à frente.

Para o *MWS-PPT* com horizontes de uma e 24 horas à frente a inclusão da variável horário de verão desloca o pico de carga mas este deslocamento é instável, o que inviabiliza o seu uso para este período.

O uso de uma regra para correção do pico de carga da previsão ou de outra rede neural para fazer esta correção, seria uma possibilidade de se resolver este problema. Após passados dez dias típicos pós o término do horário de verão, a rede com e sem a variável horário de verão volta a prever com mais precisão.

- *para o algoritmo AAS:*

As RNA's do algoritmo *AAS* fazem uma previsão ruim para os dias entre 31/01 e 10/02 de 1994. Isto acontece devido ao fato do horário de verão não terminar no mesmo dia todos os anos, como mostra a figura.6.3. Esta flutuação faz com que a rede espere que o horário de verão sempre irá terminar na mesma data do ano que foi utilizado para o treinamento da rede. Assim, para o nosso caso, a RNA "esperava" o término do horário de verão de 1994 no dia 31/01 e não no dia 10/02 como ocorreu. Portanto, de 31/01/94 a 10/02/94 a rede estava generalizando padrões os quais não foram aprendidos do conjunto de treinamento e nesta generalização não houve o deslocamento do pico de carga como o esperado.

Não foi possível contornar este problema com os dados históricos disponíveis. Portanto, pode haver ainda um conjunto de variáveis de entrada que faça com que a RNA aprenda a identificar com maior facilidade o fim (ou início do horário de verão).

Uma solução viável seria corrigir os dados previstos, deslocando o período de ponta, para os períodos do ano onde não há coincidência do período de ponta estar em vigor ou não.

Entretanto a inclusão desta variável na entrada da rede não trouxe melhoras na previsão para este algoritmo.

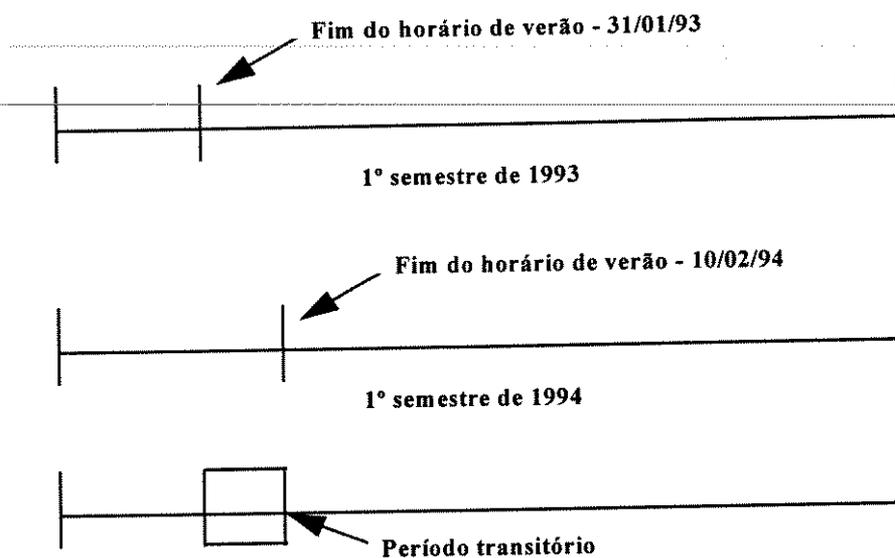


Figura.6.3: Período Transitório do Horário de Verão de 1993 e 1994.

◆ Quanto a variável dia da semana

No intuito de reduzir o número de neurônios na camada de entrada da rede, foi avaliada a possibilidade de não considerar a variável dia da semana nesta. Observou-se que a exclusão desta variável melhora sensivelmente a exatidão da previsão (*MAPE* e *SDAPE*), além de tornar mais rápido o treinamento das RNA's para o algoritmo *MWS*. Concluímos que esta variável pode ser excluída do conjunto de variáveis de entrada da rede tanto para o algoritmo *MWS* como para o algoritmo *AAS* ao contrário do que é observado na maioria dos trabalhos apresentados na literatura que consideram esta variável.

No caso do sistema CESP o perfil da carga no horário de demanda de ponta é semelhante de segunda a sexta, como pode ser visto na figura 5.4. Logo, para os dias úteis é necessário apenas uma RNA.

◆ Quanto aos dados climáticos

A inclusão das variáveis temperatura mínima e da hora-do-por-do-sol não trouxeram melhoras significativas para os algoritmos *MWS*.

Entretanto, a variável temperatura máxima tornou-se importante para as previsões com o *MWS-PPT* 10 minutos, 1 e 24 horas à frente.

• Quanto ao horizonte de previsão

Na seleção das variáveis de entrada, além da consideração das variáveis dia da semana e horário de verão, tentou-se conseguir o melhor conjunto considerando as variáveis de carga e clima, que fossem realmente as mais relevantes sempre tendo em mente obter a menor rede possível em termos de conexões, sem comprometer a exatidão da previsão.

Para os três horizontes de previsão, o algoritmo *MWS-PPT* obteve os melhores resultados. As tabelas 6.2 e 6.3 mostram os resultados obtidos do ponto de vista da previsão e as tabelas 6.4 e 6.5 mostram o desempenho das RNAs do ponto de vista computacional. O período testado foi o 1º semestre de 1993 e 1994.

HORIZONTE	ARQUITETURA	MPE	MAPE	SDAPE	MAXPRE	NPM800	0 ----4%
1 h	6 X 3 X 1	0,72	2,01	1,64	7,65	24	87,81
24 h	6 X 3 X 1	0,71	2,01	1,63	8,72	24	88,35
10 min	5 X 3 X 1	0,75	1,91	1,56	7,61	20	88,82

Tabela.6.2: Performance do algoritmo *MWS-PPT* - 1993.

HORIZONTE	ARQUITETURA	MPE	MAPE	SDAPE	MAXPRE	NPM800	0 ----4%
1 h	6 X 3 X 1	0,44	1,94	1,86	9,79	51	86,36
24 h	6 X 3 X 1	0,43	1,93	1,86	10,12	46	86,35
10 min	5 X 3 X 1	0,48	1,84	1,77	9,67	41	88,09

Tabela.6.3: Performance do algoritmo *MWS-PPT* - 1994.

Horizonte	Arquit.	Nr. de Treinamentos por perfil de ponta	Tempo máximo por treinamento	Tempo máximo de treinamento por perfil de ponta
1 h	6 X 3 X 1	31	24.72 s	260.23 s
24 h	6 X 3 X 1	31	16.52 s	131.20 s
10 min	5 X 3 X 1	31	16,32 s	62,98 s

Tabela.6.4: Tempos de Treinamento - 1993.

Horizonte	Arquit.	Nr. de Treinamentos por perfil de ponta	Tempo máximo por treinamento	Tempo máximo de treinamento por perfil de ponta
1 h	6 X 3 X 1	31	22.70 s	280.87 s
24 h	6 X 3 X 1	31	18.20 s	125.67 s
10 min	5 X 3 X 1	31	16.27 s	58.95 s

Tabela.6.5: Tempos de Treinamento - 1994.

Como pudemos observar das tabelas 6.2 e 6.3, mesmo com um horizonte de previsão menor, não houve uma melhora significativa no *MAPE* da previsão. Analisando desta forma, a previsão com horizonte de 24 horas à frente apresentou os melhores resultados.

Observamos das tabelas 6.4 e 6.5 que o tempo de treinamento para o *MWS-PPT* 10 minutos à frente permite que a previsão seja realizada com uma precedência razoável. Entretanto, o seu *MAPE* desencoraja a sua utilização pois é desejado um *MAPE* em torno de 1% para esta previsão. Estes tempos correspondem ao treinamento mais longo de um conjunto de 100 testes realizados para cada horizonte de previsão.

◆ Quanto ao tipo de treinamento

O algoritmo de aprendizagem *backpropagation* com correção dos pesos da sinapses do tipo incremental com taxa de aprendizagem adaptativa e termo momento constante, proporcionou o treinamento mais rápido com a melhor eficiência do ponto de vista da convergência.

Observando-se o comportamento do *SSE* dos treinamentos realizados, chegamos ao limite de 100 épocas por treinamento para as RNA's dos quatro algoritmos testados. A taxa de aprendizagem inicial sempre foi 0.9. O termo momento constante foi de 0.1.

◆ Quanto a arquitetura das RNA's

Para os quatro algoritmos testados a rede de três camadas (inclusive a entrada; uma camada escondida) *feedforward* obteve os melhores resultados. A rede de quatro camadas se demonstrou muito lenta em seu treinamento e com valores *MAXAPE* muito altos. Quando convergia, não generalizava bem os padrões. Já as redes com duas camadas (sem camada escondida) apesar de terem um treinamento bem rápido, tem um *MAPE* cerca de 1% maior em geral em comparação com as redes de três camadas.

Observou-se que o número de neurônios na camada escondida, tanto para o algoritmo *MWS* como para *AAS*, é em geral dado por:

$$J = \begin{cases} \frac{L+1}{2}, & \text{se } L \text{ impar} \\ \frac{L}{2}, & \text{se } L \text{ par} \end{cases} \quad (6.5)$$

onde J é o número de neurônios na camada escondida e L é o número de neurônios na entrada da rede.

6.3- Comparação com o Modelo da CESP

A sistemática adotada atualmente na CESP para previsão de carga do período de ponta consiste apenas num modelo ARMA e um ARIMA para previsão horária 24 horas à frente. Portanto, não poderemos fazer comparações com este modelo, já que os dois resolvem problemas diferentes do ponto de vista da previsão de carga.

6.4- Considerações Finais

Para efeito de comparação, Gross e Galiana [10] argumentam que um *MAPE* na faixa de 2 a 3% pode ser considerado como uma boa previsão para intervalos de discretização de uma hora. Como a variância dos dados com discretização de 10 minutos é maior do que os dados em base horária, os *MAPES* conseguidos com as RNA's utilizando os algoritmos propostos situados na faixa acima, sugerem que a previsão para os horizontes de curto e curtíssimo prazos pode ser considerada boa, comparando com metodologias semelhantes em base horária.

A consideração da influência climática sobre a carga é um tema relevante, mas carece ainda, no Brasil, de informações climáticas confiáveis e abrangentes o suficiente para que se possa levar a termo uma análise exaustiva.

A falta de dados sobre a luminosidade no final da tarde e de dados sobre a cobertura do céu e pluviosidade certamente aumentou o erro da previsão para os algoritmos de previsão testados. Assim, pesquisas complementares podem comprovar esta hipótese, já que a luminosidade é um fator importante na determinação da carga no final da tarde.

Os programas utilizados nesse trabalho, para identificação e ajuste das RNA's, implementação dos algoritmos de previsão, identificação da autocorrelação da carga e decomposição espectral, visualização e análise dos resultados, determinação das estatísticas e medidas de previsão para cada previsão testada, foram implementados em MATLAB® versão 4.0 para estações de trabalho. Foram utilizados nas simulações estações de trabalho SUNsparc IPX. Os tempos correspondem aos tempos de cpu destas máquinas e estão apresentados no apêndice A.

Analisou-se ao longo deste capítulo a possibilidade de se desenvolver algoritmos eficientes, rápidos e baratos para previsão do período de demanda de ponta com discretização menor que uma hora utilizando RNA's, levando em consideração inclusive as variáveis climáticas mais significativas no nosso ambiente. Verificou-se que

métodos de previsão de fácil desenvolvimento e de baixos requisitos computacionais são factíveis, inclusive com retreinamento da rede, para a carga da CESP, levando a erros aceitáveis sem requisitar intervenção do operador.

A previsão com o algoritmo *MWS* através do método dinâmico para os três horizontes propostos obteve os melhores resultados tanto a nível da precisão da previsão como da performance computacional do algoritmo.

Capítulo 7

CONCLUSÃO

Este capítulo final apresenta as conclusões gerais e mostra as contribuições deste trabalho, bem como sugere possíveis trabalhos futuros.

7.1 - Conclusões Gerais

O objetivo principal deste trabalho foi investigar a solução do problema de previsão de carga do período de demanda de ponta a curto e curtíssimo prazos, com discretização menor que uma hora, através da utilização de Redes Neurais Artificiais, em particular para o sistema CESP.

A principal motivação para o uso de RNA na previsão de carga a curto prazo foi encontrar melhores resultados que os apresentados pelos procedimentos estatísticos convencionais e devido à capacidade das RNA's de extrair as características da carga através da inclusão de variáveis climáticas.

A modelagem das RNA's mostrou-se eficaz para a obtenção de bons resultados na aplicação em previsão de carga a curto prazo. A falta de um maior conhecimento das complexas interrelações entre os diversos parâmetros que influenciam o desempenho da rede, faz com que o sucesso de uma aplicação seja muito dependente da experimentação prática e da experiência prática do projetista. Este, por sua vez, além do conhecimento em engenharia de RNA's, requisito fundamental para se executar ajustes finos adequados, deve possuir um bom conhecimento do domínio da aplicação que garanta uma escolha conveniente tanto das variáveis mais significativas a serem apresentadas às RNA's, bem como do pré-processamento apropriado aos conjuntos de dados.

Convém ressaltar que o sucesso das aplicações está fortemente associado ao pré-processamento dos dados de entrada; um pré-tratamento equivocado ou mesmo não realizado quando necessário, pode ser determinante no fracasso de uma aplicação.

Procurou-se analisar os diversos fatores que influenciam o comportamento da curva de carga na ponta, bem como os modelos propostos na literatura para previsão de carga a curto prazo em base horária via RNA's. Propôs-se então, quatro algoritmos para resolver o problema da previsão do período de ponta em patamares de 10 em 10 minutos.

O estudo de caso sobre o sistema CESP possibilitou testar a eficiência do uso de RNA's à previsão da ponta. A capacidade das RNA's em prever carga no horário de ponta foi comprovada através das previsões para os horizontes de previsão de uma e 24 horas (curto prazo) e 10 minutos à frente (curtíssimo prazo).

Dos quatro algoritmos que foram avaliados para previsão dinâmica e estática do período de demanda de ponta, o método dinâmico realizado através do algoritmo *MWS* mostrou ser mais eficiente do ponto de vista computacional e da exatidão da previsão.

Observou-se que o uso da variável dia da semana é dispensável para este algoritmo diminuindo assim o número de neurônios na camada de entrada da rede.

O algoritmo *MWS* pelo método dinâmico mostrou-se mais capaz de se adaptar às novas exigências e mudanças no padrão de consumo de energia elétrica no horário de ponta, pois as RNA's deste algoritmo necessitam de aprender menos características da variedade de padrões (pois são mais especializadas) que estão no conjunto de treinamento destas redes.

O algoritmo *MWS* mostrou-se robusto e as RNA's tem pequeno tempo de treinamento, mostrando ser versátil para este tipo de aplicação, tanto para a previsão a curto prazo como a previsão de curtíssimo prazo. Entretanto o erro médio percentual absoluto da previsão a curtíssimo prazo é considerado ainda elevado.

Do ponto de vista de modelagem, o uso de RNA's para previsão de carga no período de ponta mostrou-se altamente promissor para este tipo de aplicação.

A falta de dados climáticos mais consistentes sobre as condições de temperatura, cobertura do céu, pluviosidade e luminosidade no final da tarde impediu a avaliação da correlação destas variáveis com a carga no período da entrada do pico de carga.

Com os dados disponíveis, só foi constatado que a temperatura máxima do dia tem uma pequena influência sobre a carga.

7.2- Trabalhos Futuros

Em termos de trabalhos futuros, o uso de redes especializadas em cada intervalo de previsão requer um algoritmo que seja capaz de classificar as entradas de acordo com o intervalo a ser previsto. Isto sugere que um classificador neural pode ser anexado a este algoritmo de forma a atingir um maior índice de automatização na escolha dos dados que farão parte do conjunto de treinamento da rede.

O projeto deste classificador pode ser feito através dos mapas auto-organizáveis de Kohonen. Logo, o sistema de previsão seria composto de um classificador, que classificaria a que grupo ou classe o padrão de carga pertence, e outro previsor onde o algoritmo *MWS* teria um papel fundamental.

A análise da influência das condições de luminosidade e temperatura no final do dia pode diminuir os erros observados no início do período de ponta. As variáveis de entrada como, intensidade luminosa no final do dia, cobertura do céu e pluviosidade seriam consideradas na entrada da rede.

Há ainda um número de projetos e adaptações que podem ser feitas no algoritmo *MWS-PPT* para servir como ferramenta de apoio ao operador para tomada de decisões.

Podemos destacar vários outros trabalhos, como:

O desenvolvimento de um sistema, com interface gráfica interativa que integre num só ambiente todas as fases necessárias à previsão, tais como: pré-processamento dos dados, treinamento e avaliação das previsões do ponto de vista estatístico e visual. Esta ferramenta serviria para apoio ao operador para previsão de carga a curto e curtíssimo prazo.

Dando continuidade ao desenvolvimento do estudo de RNA's aplicadas a previsão de carga do período de demanda de ponta, um número de estudos correlatos, não abordados nesta dissertação, podem ser desenvolvidos:

- Um estudo mais minucioso da série a ser prevista, com um pré-tratamento mais rigoroso da série de carga, buscando-se eliminar todos os sinais expúrios e não significativos da mesma, a fim de fornecer às RNA's valores mais representativos.

- Avaliar outros algoritmos de aprendizagem existentes, em particular aqueles que não apresentam os problemas intrínsecos do método gradiente descendente no que diz respeito ao tempo de treinamento e mínimos locais.

- Utilizar sistemas híbridos com RNA's e uma outra técnica computacional inteligente, como lógica *fuzzy*, algoritmos genéticos e sistemas especialistas. Outra alternativa é utilizar RNA's juntamente com métodos estatísticos, onde as RNA's podem funcionar como classificador do procedimento estatístico mais adequado para a previsão. Métodos estatísticos também podem ser usados para filtrar os dados a serem apresentados às RNA's.

Referências Bibliográficas

- [1] **Abelém, A. J. G.**, "*Redes Neurais Artificiais na Previsão de Séries Temporais*". Tese de Mestrado, DEE/PUC-RIO. Set/1994. 76p.
- [2] **Martins, W., Lima, W. S.**, "*The Use of Artificial Neural Networks in Electrical Engineering*". Electro'95: XI Congresso Chileno de Engenharia Elétrica, Punta Arenas. Nov./95. Anais do Congresso, pp. E-018 a E-023.
- [3] **Sobajic, D. J.**, editor, "*Neural Network Computing for the Electric Power Industry*". Lawrence Erlbaum Associates, London, England. 1993.
-
- [4] **Dillon, T. S.**, "*Artificial neural network applications to power systems and their relationship to symbolic methods*". Electrical Power & Energy Systems pp. 66-72. Vol.13 no.2 abr/1991.
- [5] **Chakraborty, K. et al.** "*Forecasting the behavior of multivariate time series using neural networks*". Neural Networks, EUA, no.5, pp.961-970. 1992.
- [6] **Azoff, E. M.**, "*Reducing error in neural network time series forecasting*". Neural & Computing & Applications, Londres, no.1, pp. 240-247. 1993.
- [7] **Landi, L., Barucci, E.**, "*Artificial neural network for treasury bills rate forecasting*" In: Neural Network in Capital Markets, 1. Londres, 18-19, nov/1993. Anais, Londres, London Business School, 1993.
- [8] **Morettin, P. A., Toloí C. M. de C.**, "*Previsão de Séries Temporais*". Editora ATUAL. São Paulo, 1985. 438p.
- [9] **Mendes, L. E. P.** "*Análise Recursiva de Modelos Lineares de Regressão e Séries Temporais: O método Bayesiano de Previsão.*" Tese de Mestrado, COPPE/UFRJ. Rio de Janeiro - RJ. Dez/78. 131p.

- [10] Gross G., Galiana, F. D., "Short-Term Load Forecasting" , Proceedings of the IEEE, Vol. 75, n. 12, dezembro de 1987.
- [11] Chen, S., Yu, D. C., Moghaddamjo, A. R., "Weather sensitive short-term load forecasting using nanfully connected artificial neural network". IEEE - Transactions on Power Systems. Vol. 07, no. 13, aug/92. pp. 1098-1105.
- [12] Thompson, R. P. "Weather Sensitive Electric Demand and Energy Analysis on a Large Geographically Diverse Power System Application to Short-Term Hourly Electric Demand Forecasting", IEEE - Transaction on Power System, vol. PAS-95, nr.1, pp.385-393. 1976.
- [13] SGPECA/GTPC/ELETROBRÁS, "Modelo Avançado de Previsão de Carga para a Área de Operação dos Sistemas Interligados - MAP", Termo de Referência. Fev/94. 6p.....
-
- [14] GCOI/ELETROBRÁS, "I Jornada Técnica do GTPC ". Subcomitê de Estudos Energéticos - Rio de Janeiro; 24, 25 e 26 de março de 1992.
- [15] Vemuri, S., Hoveida, B., Mohebbi, S., "Short-Term Load Forecasting Based on Weather Load Models", Proceedings IFAC Power Systems. Pp. 315-320.
- [16] Hagan, M. T. And Behr, S. M., "The Time Serie Approach to Short-Term Load Forecasting". IEEE - Transactions on Power Systems, vol.PWRS-2, no.3. pp. 785-791.
- [17] Heinemann, G. T., Nordman, D. A., Plant, E. C. " The Relationship Between Summer Weather and Summer Loads - A Regression Analysis", IEEE Transactions on Power Apparatus Systems, vol. PAS-85 nr.11, pp. 1144-1154. 1966.
- [18] Stanton, K. N. "Medium-Range, Weekly and Seasonal Peak Demand Forecasting by Probalibility Methods", IEEE - Transaction on Power Systems, vol. PAS-90, nr.3, pp. 1183-1189. 1971.
- [19] Rao, S. C. e Singh, S. , "Statistical Correlation Method to Predict Peak Demand" , IE(I) Journal-EL, vol.61, pp. 189-192. 1980.

- [20] Gupta, P. C., "A Stochastic Approach to Peak Power Demand Forecasting in Electric Utility Systems", IEEE - Transactions on Power Apparatus Systems. Vol. PAS-90, no.2, pp. 824-831, abr/1971.
- [21] Peng, T. M., Hubele, N. F., Karady, G. G., "An adaptive neural network approach to one-week ahead load forecasting.". IEEE - Transactions on Power Systems, vol.87, no.3, ago/93, pp. 1.195-1.203.
- [22] Davey, J., Saacks, J. J., Cunningham, G. W., Priest, K. W., "Practical Application of Weather Sensitive Load Forecasting to System Planning", IEEE Transaction on Power Apparatus Systems, vol. PAS-92, nr.3, pp. 971-977. 1972.
-
- [23] Brown, R. G., "Smoothing, Forecasting and Prediction of Discrete Time Series" Prentice Hall Inc. 1965.
- [24] Gupta, P. C., Yamada, K., "Adaptive Short-Term Forecasting of Hourly Loads Using Weather Information", IEEE Transactions on Power Apparatus Systems. vol. PAS-91, nr. 9/10, pp. 2085-2094. 1972.
- [25] Toyoda J., Chen, M. S., "An Application of State Estimation to Short-Term Load Forecasting, Part I: Forecasting Modeling/ Part II: Implementation", IEEE - Transaction on Power Apparatus Systems, vol.PAS-89, no.7. pp. 1678-1687.
- [26] Box, G. E. P., Jenkins, G. M., "Time Series Analysis Forecasting and Control". Holden Day. 553 p. 1970.
- [27] Vemuri, S., Huang, W. L., Nelson, D. J., "On-line Algorithms for Forecasting Hourly Loads of An Electric Utility". IEEE - Trans. on Power Apparatus Systems, vol. PAS-100, no.8. pp. 3775-3784.
- [28] Hagan, M. T., Klein, R., "On-line Maximum Likelihood Estimation for Load Forecasting". IEEE - Trans. on Systems, Man and Cybernetics, vol. SMC-8, no. 9. Pp. 711-715.

- [29] Corpening, S. L., Reppen, N. D., Ringlee, R. J., "Experience with Weather Sensitive Load Models for Short and Long-Term Forecasting", IEEE Transaction Power Apparatus Systems, vol. PAS-92, nr.6. pp. 1966-1972. 1973.
- [30] Goh, T. N., Ong, H. L., Lee, Y. O., "A new approach to statistical forecasting of daily peak power demand". Elec. Power Syst. Res., vol.10, no.2, pp. 145-148, mar/1986.
- [31] Mathewman, P. D., Nicholson, H., "Techniques for load prediction in the electricity-supply". Proc. Inst. Elec. Eng., vol. PAS-85, pp. 1144-1154, nov/1966.
- [32] Drapper, N. R., Smith, H., "Applied Regression Analysis", John Wiley & Sons Inc. 1981.
-
- [33] Cargil, T. F., Meyer, R. A. "Estimating the Demand for Electricity by Time of Day", Applied Economics, nr.3, pp. 233-246. 1971.
- [34] Asbury, C. E., "Weather Load Model for Electric Demand and Energy Forecasting", IEEE Transactions on Power Apparatus Systems, vol. PAS-94, nr.4, pp. 1111-1116. 1975.
- [35] Filho, M. G. A., Carvalho, M. F. H., Barone, J. C., "Modelo para Previsão do Consumo de Energia Elétrica". IX SNPTEE, Grupo VII (GPL), pp. 1-7.
- [36] Lijesen, D. P., Rosing, J., "Adaptive Forecasting of Hourly Loads Based on Load Measurements and Weather Information", IEEE Transactions on Power Apparatus Systems. vol. PAS-90. pp. 1757-1767. 1971.
- [37] Christiaanse, W. R., "Short-Term Load Forecasting Using General Exponential Smoothing", IEEE Transaction on Power Apparatus Systems. vol. PAS-90, nr.2, pp. 900-910. 1971.
- [38] Galiana, F. D., Handschin, E., Flechter, A., "Identification of Stochastic Electric Load Models for Physical Data". IEEE - Trans. on Automation and Control, vol.AC-19, no.6. pp. 887-893.

- [39] Vemuri, S., Huang, W. L., Nelson, D. J., "On-Line Algorithms for Forecasting Hourly Loads of An Electric Utility". IEEE - Transaction on Power Apparatus Systems, vol.PAS-100, no.8. pp. 3775-3784.
- [40] Abu-El-Magd, M. A., Sinha, N. K., "Short-term Load Demand Modelling and Forecasting: A Review", IEEE - Transactions on System, Man and Cybernetics, vol.SMC-12, no.3, pp. 370-382.
- [41] Caines, P. E., "Linear Stochastic Systems", John Wiley & Sons.
- [42] Sharmma, K. L. S., Mahalanabis, A. K., "An Application of Nonlinear Adaptive Estimation Theory in Short-Term Load Prediction". IEEE - PES Summer Meeting & Energy Research, no.C-74-333-1.
-
- [43] Abu-El-Nagd, M. A., Sinha, N. K., "Two New Algorithms for On-Line Modelling and Forecasting of the Load Demand of Multinode Power System". IEEE - Transactions on Power Apparatus Systems, vol.PAS-100, no.7, pp. 3246-3253.
- [44] Campo, R., Ruiz, P., "Adaptive Weather Sensitive Short-term Load Forecasting", IEEE - Transactions on Power Systems, vol.PWRS-2, no.3. pp. 592-600.
- [45] Dehdashti, A. S., Tudor, J. R., Smith, M. C., "Forecasting of Hourly Load by Pattern Recognition: A Deterministic Approach". IEEE - Transactions on Power Apparatus Systems, vol.PAS-101, no.9. pp. 3290-3294.
- [46] Jabbour, K., Riveros, J. F. V., Landesbergen, D., Meyer, W., "Alfa: Automated Load Forecasting Assistant". IEEE - Transaction on Power Systems, vol.3, no.3. pp. 908-914.
- [47] Rahman, S., Bhatnagar, R., "An Expert System Based Algorithm for Load Forecasting". IEEE - Transactions on Power Systems, vol.3, no.2. pp. 392-399.
- [48] Dillon, T.S., Morsztyn, K., Phua, K., " Short term load forecasting using adaptative pattern recognition and self organizing techniques". Fifth Power Systems Computation Conference PSCC Proceedings, Cambridge,U.K.,1-5 Set/75.

- [49] Moghram I., Rahman, S., "Analysis and Evaluation of Five Short-Term Load Forecasting Techniques". Trabalho 89 WM 171-0 PWRS apresentado no IEEE/PES Winter Meeting.
- [50] Noguti, M. Y., "Relatório de Atividades - Projeto: Redes Neurais Para a Previsão de Vazões". DENISIS/FEE/UNICAMP. Jan/1994. 61p.
- [51] Sabbatini, R. M. E., "Aprendizado em Redes Neurais". Núcleo de Informática Biomédica - UNICAMP, jul/1992.
- [52] Hertz, J., Krogh, A., Palmer, R. G., "Introduction To The Theory Of Neural Computation". Addison-Wesley, Santa Fé, EUA. 327p. 1991.
-
- [53] McCulloch, W. S., Pitts, W., "A logical Calculus of Ideas Immanent in Nervous Activity". Bulletin of Mathematical Biophysics 5, pp. 115-133. Reprinted in Anderson and Rosenfeld [1988]. 1943.
- [54] Widrow, B., and Hoff, M. E., "Adaptive switching circuits". In 1960 IRE WESCON Convention Record, part 4, pp. 96-104. New York: IRE. 1960.
- [55] Widrow, B., "Generalization and Information Storage in Networks of Adaline Neurons". In Self-Organizing Systems 1962, Chicago, eds. M. C. Yovits, G. T. Jacobi, and G. D. Goldstein, pp. 435-461. Washington: Spartan. 1962.
- [56] Rosenblatt, F., "Principles of Neurodynamics". New York: Spartan. 1962.
- [57] Minsky, M. L., Papert, S. A., "Perceptrons". Cambridge: MIT Press. 1969.
- [58] Poddig, T., "Short-Term Forecasting of the USD/DM-exchange rate". In: Neural Network in Capital Markets, 1. Londres, nov/1993. Anais do congresso. 1993.
- [59] Wasserman, P. D., "Neural Computing: Theory and Practice". New York: Van Nostrand Reinhold, 1989.

[60] **Hopfield, J. J.**, "*Neurons with grade responses have collective computational properties like those of two-states neurons*". In: National Academy of Science. EUA, 1981. Anais do congresso, pp. 3088-3092. 1984.

[61] **Hinton, G. E., Sejnowski, T. J.**, "*Learning and relearning in Boltzmann machines*". In: Parallel Distributed Processing. Anais do congresso, vol.1, cap.7. 1986.

[62] **Bryson, A. E., Ho, Y.-C.**, "*Applied Optimal Control*". New York: Blaisdell. 1969.

[63] **Calôba, L. P.**, "*Introdução à Computação Neuronal*". 9º Congresso Brasileiro de Automática, UFES - Vitória-ES, Brasil. Anais do congresso, pp. 25-38.

[64] **Werbos, P. J.**, "*Beyond regression: new tools for prediction and analysis in the behavioral science*". Havard University, 1974. PhD Thesis.

[65] **Parker, D. B.**, "*Learning Logic*". Technical report TR-47. Center of Computational Research in Economics and Management Science, MIT-Cambridge, MA. 1985.

[66] **Rumelhart, D. E., Hinton, G. E., Willians, R. J.**, "*Learning representations by back-propagation errors*". Nature, no.323, pp. 533-536. 1986.

[67] **Rumelhart, D. E., Hinton, G. E., Willians, R. J.**, "*Learning internal representations by back-propagation*". In: Parallel Distributed Processing, vol.1, chap.8.1986.

[68] **Le Cun, Y.** "*Une procédure d'apprentissage pour réseau à seuil assymétrique*". In: Cognitiva 85: a la frontiere de l'intelligence artificielle desw sciences de la connaissance des neuroscience, Paris 1985. Anais do congresso, pp. 599-604. 1985.

[69] **Park, D. C., El-Sharkawi, M. A., Marks II, R. J.**, "*Electric load forecasting using an artificial neural network*". IEEE Transactions on Power Systems, vol.6, no.2, mai/1991. pp. 442-449.

[70] **Lu, C. N., Wu, H. T., Vemuri, S.**, "*Neural Network Based Short-Term Load Forecasting*". IEEE Transactions on Power Systems, vol.8, no.1, pp. 336-342. Fev/1993.

- [71] **Chen, S. T., Yu, D. C., Moghaddamio, A. R.**, "*Weather sensitive short-term load forecasting using nonfully connected artificial neural network*". IEEE Transactions on Power Systems, vol.7, no.3, pp. 1098-1105. Ago/1992.
- [72] **Dillon, T. S., Sestito, S., Leung, S.**, "*Short-term load forecasting using an adaptive neural network*". Electrical Power and Energy Systems, vol.13, no.4, pp. 186-192. Ago/1991.
- [73] **Lee, K. Y., Cha, Y. T., Park, J. H.**, "*Short-Term Load Forecasting Using An Artificial Neural Network*". IEEE Transactions on Power Systems, vol.7, no.1, pp. 124-132. Fev/1992.
- [74] **Hsu, Y. Y., Yang, C. C.**, "*Design of neural networks for short-term load forecasting. Part I: Self-organising feature maps for day type identification*". IEE Proceedings-C, vol.138, no.5, pp. 407-413. Set/1991.
- [75] **Hsu, Y. Y., Yang, C. C.**, "*Design of neural networks for short-term load forecasting. Part II: Multilayer feedforward networks for peak load and valley load forecasting*". IEE Proceedings-C, vol.138, no.5, set/1991.
- [76] **Erkmen, I., Özsökmen,** "*A hybrid neural network for short-term load forecasting*". Proceedings of IEEE/NTUA Athens Power Tech Conference: "Planning, Operation and Control of Today's Electric Power Systems", Athens, Greece, set. 5-8/1993. Anais do congresso, vol.II, pp. 811-815. Set/1993.
- [77] **Peng, T. M., Hubele, N. F., Karady, G. G.**, "*Advancement In The Application Of Neural Networks For Short-Term Load Forecasting*". IEEE Transactions on Power Systems, vol.7, no.1, pp. 250-257. Fev/1992.
- [78] **Engle, R. F., Mustafa, C., Rice, J.**, "*Modeling peak electricity demand*". Journal of Forecasting. Vol.11, pp. 241-251. 1992.
- [79] **Kariniotakis, G. N., Kosmatopoulous, Christodoulou, M. A., Strakakis, G. S.**, "*Load forecasting using dynamic high-order neural network*". Proceedings of IEEE/NTUA Athens Power Tech Conference: "Planning, Operation and Control of Today's Electric Power Systems", Athens, Greece, set. 5-8/1993. Anais do congresso, vol.II, pp. 801-805. Set/1993.

- [80] Mohammed, O., Park, D., Merchant, R., Dinh, T., Tong, C., Azeem, A., Farah J., Drake, C., *"Practical Experiences with an Adaptive Neural Network Short-Term Load Forecasting System"*, IEEE - Transactions on Power Systems, vol. 10, no.1, pp. 254-265. Fev/1995.
- [81] Kermanshahi, B. S., Poskar, C. H., Swift, G., McLaren, P., Pedrycz, W., *"Load Forecasting Under Extreme Climatic Conditions"*. Proceedings of IEEE/NTUA Athens Power Tech Conference: "Planning, Operation and Control of Today's Electric Power Systems", Athens, Greece, set. 5-8/1993. Vol.II, pp. 806-810. Set/1993.
- [82] Ho, K.-L., Hsu, Y.-Y., Yang, C.-C., *"Short-Term Load Forecasting Using a Multilayer Neural Network With An Adaptive Learning Algorithm"*. Transaction on Power Systems, vol.7, no.1, fev/1992.
- [83] Kiartzis, S. J., Bakirtz, A. G., Petridis, V., *"Neural Networks Application To Short-Term Load Forecasting"*. Anais do congresso desconhecido. pp. 339-344.
- [84] Papalexopoulos, A., D., Hao, S., Peng, T.-M., *"Application of Neural Network Technology To Short-Term System Load Forecasting"*. Proceedings of IEEE/NTUA Athens Power Tech Conference: "Planning, Operation and Control of Today's Electric Power Systems", Athens, Greece, set. 5-8/1993. Vol.II, pp. 796-800. Set/1993.
- [85] Satoh, R., Tanaka, E., Hasegawa, J., *"Daily Load Forecasting Using A Neural Network Combined With Regression Analysis"*. Anais do congresso desconhecido. pp. 345-352.
- [86] Lobo, R. M. P. De O., *"Modelos para previsão de carga a curto prazo aplicados ao sistema elétrico do Paraná"*. Tese de mestrado - UNICAMP. 98p. Fev/1991.
- [87] Rahman, S., Hazim, O., *"A Generalized Knowledge-Based Short-Term Load Forecasting Technique"*. IEEE Transactions on Power Systems, vol.8, no.2, mai/1993. pp. 508-514.
- [88] Tang, Z., *et al*, *"Time series forecasting using neural network vs. Box Box & Jenkins methodology"*. Artificial Neural Network: Forecasting time series. EUA, IEEE Press pp. 20-27, 1994.

[89] **Plaut, D., Nowlan, A., Hinton, G.,** *"Experiments on Learning by Back Propagation"*. Technical Report CMU-CS-86-126, Department of Computer Science, Carnegie Mellon University, Pittsburgh, PA. 1986.

[90] **Refenes, A. N., Vithlani, S.,** *"Constructive learning by specialization"*. In: ICANN-91, Helsinki, julho de 1991. Anais do congresso. Helsinki, 1991.

[91] **Miller, G. F., Todd, P. M., Hedge, S. U.,** *"Desining neural networks using genetic algorithms"*. In: Third International Conference on Genetic Algorithms. Anais do congresso, pp. 177-179. San Mateo: Morgan Kaufmann 1989.

Apêndice A

ESTATÍSTICAS E DESEMPENHO DAS MELHORES RNA's PARA CADA ALGORITMO PROPOSTO

A.1- Algoritmo *MWS-ALL*

A.1.1- Horizonte 1 hora à frente

- Variáveis de entrada: Carga (histórico horário) às 14 e 15 horas e hora do pôr-do-sol.
- Método: estático
- Topologia: 3 X 2 X 31
- Período: 13/01 a 31/05 (exceto o período transitório do horário de verão)

Ano	Nr. Prev.	MPE [%]	MAPE [%]	SDAPE [%]	MAXAPE [%]	NPM800	Nr.Tr. por Dia	TMaxPT ¹ [s]	TMaxPP ² [s]
1993	2077	1,09	2,38	1,92	10,42	78	1	76,63	76,33
1994	1767	0,75	2,16	2,00	11,82	78	1	76,40	76,40

Tabela.A.1: Desempenho das RNA's para o algoritmo *MWS-ALL* 1 hora à frente

MAPE [%]		Frequência		Percentual	
		1993	1994		
- ∞	----- - 12	0,00	0,00		0,00
- 12	----- - 10	0,00	0,00		0,00
- 10	----- - 8	0,00	0,00		0,00
- 8	----- - 6	1,02	0,74		0,74
- 6	----- - 4	2,05	3,89		3,89
- 4	----- - 2	5,66	8,77		8,77
- 2	----- 0	24,63	28,29		28,29
0	----- 2	31,16	30,19		30,19
2	----- 4	21,40	14,64		14,64
4	----- 6	8,73	7,03		7,03
6	----- 8	3,38	3,31		3,31
8	----- 10	1,89	2,23		2,23
10	----- 12	0,08	0,74		0,74
12	----- ∞	0,00	0,17		0,17

Tabela.A.2: Frequência Percentual do *MAPE* para o algoritmo *MWS-ALL* 1 hora à frente

¹ TMaxPT - Tempo máximo (pior caso) de treinamento para as RNA's deste algoritmo.

² TMaxPP - Tempo máximo de treinamento para previsão do perfil no horário de ponta para um dia.

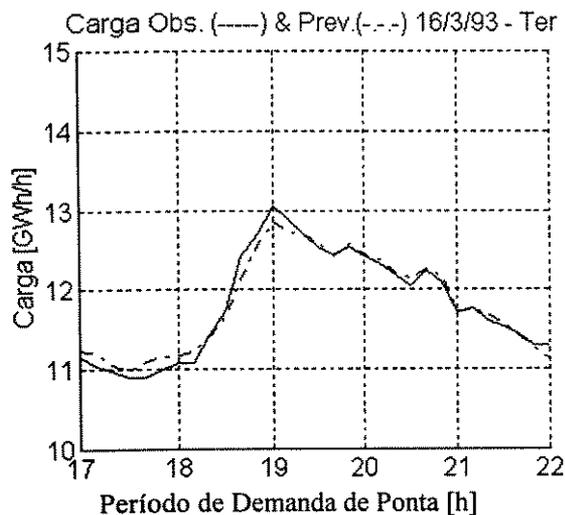


Figura.A.1: Perfil com a melhor previsão
1 hora à frente, 1993.
Algoritmo *MWS-ALL*

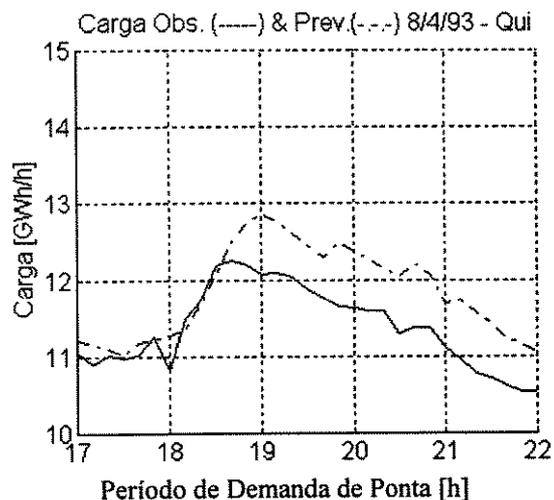


Figura.A.2: Perfil com a pior previsão
1 hora à frente, 1993.
Algoritmo *MWS-ALL*

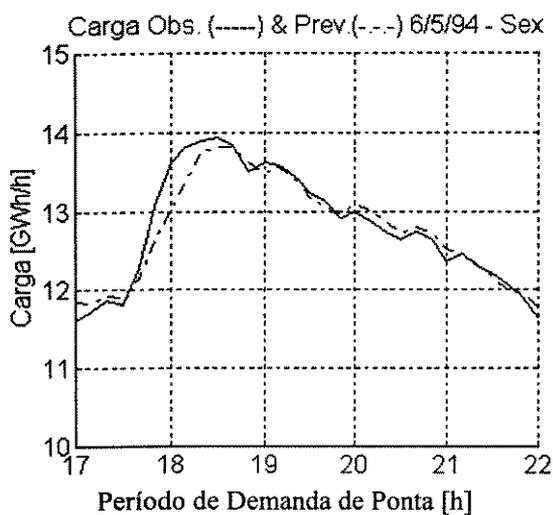


Figura.A.3: Perfil com a melhor previsão
1 hora à frente, 1994.
Algoritmo *MWS-ALL*

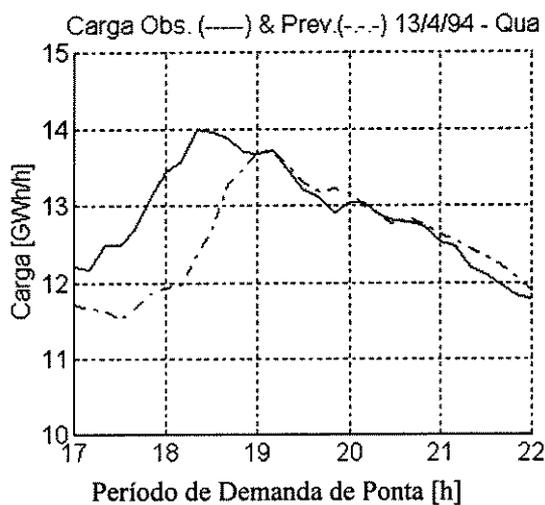


Figura.A.4: Perfil com a pior previsão
1 hora à frente, 1994.
Algoritmo *MWS-ALL*

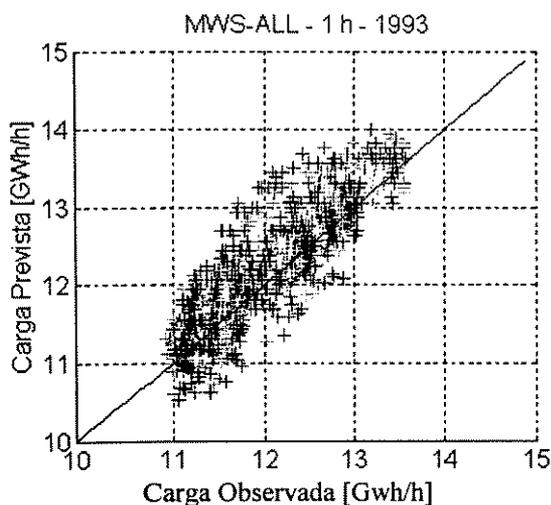


Figura.A.5: Diagrama de dispersão da previsão 1 hora à frente, 1993
Algoritmo *MWS-ALL*

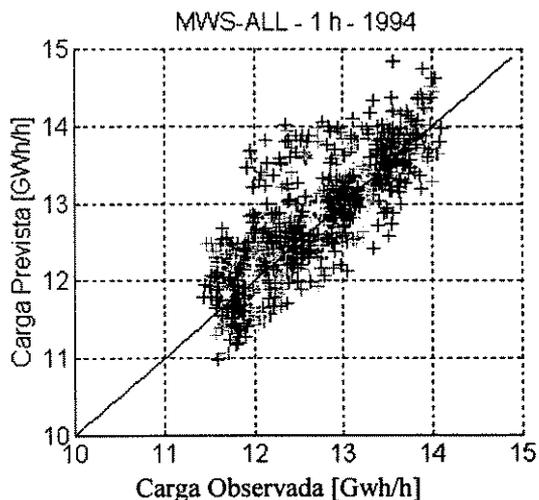


Figura.A.6: Diagrama de dispersão da previsão 1 hora à frente, 1994
Algoritmo *MWS-ALL*

A.1.2- Horizonte 24 horas à frente

- **Variáveis de entrada:** Carga (histórico horário) às 14 e 15 horas, temperatura prevista máxima e hora do pôr-do-sol.
- **Método:** estático
- **Topologia:** 4 X 2 X 31
- **Período:** 13/01 a 31/05 (exceto o período transitório do horário de verão)

Ano	Nr. Prev.	MPE [%]	MAPE [%]	SDAPE [%]	MAXAPE [%]	NPM800	Nr.Tr. por Dia	TMaxPT [s]	TMaxPP [s]
1993	2077	1,03	2,39	1,95	10,45	79	1	177,63	177,63
1994	1767	0,71	2,16	2,03	12,35	85	1	172,78	172,78

Tabela.A.3: Desempenho das RNA's para o algoritmo *MWS-ALL* 24 horas à frente

MAPE [%]		Frequência	
		1993	1994
- ∞	----- - 12	0,00	0,00
- 12	----- - 10	0,00	0,00
- 10	----- - 8	0,00	0,00
- 8	----- - 6	0,63	0,66
- 6	----- - 4	1,65	3,14
- 4	----- - 2	4,88	6,70
- 2	----- 0	22,90	27,05
0	----- 2	33,36	34,08
2	----- 4	22,50	15,38
4	----- 6	8,50	6,70
6	----- 8	3,70	3,06
8	----- 10	1,89	2,23
10	----- 12	0,00	0,91
12	----- ∞	0,00	0,08

Tabela.A.4: Frequência Percentual do MAPE para o algoritmo *MWS-ALL* 24 horas à frente

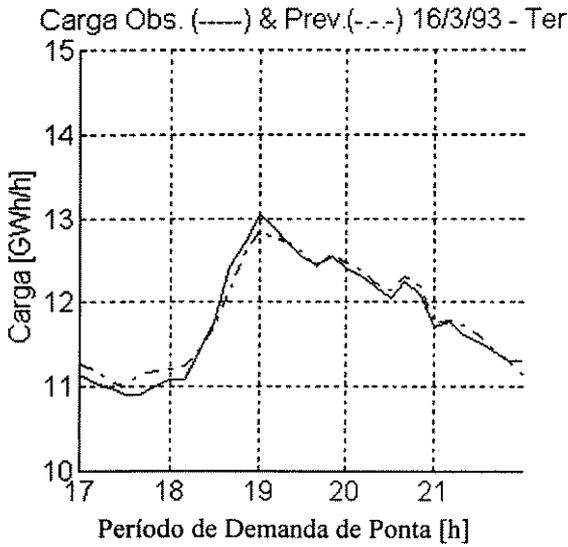


Figura.A.7: Perfil com a melhor previsão 24 horas à frente, 1993. Algoritmo *MWS-ALL*

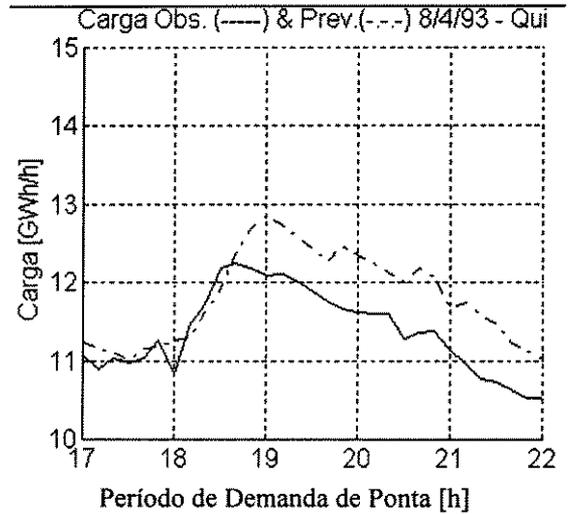


Figura.A.8: Perfil com a pior previsão 24 horas à frente, 1993. Algoritmo *MWS-ALL*

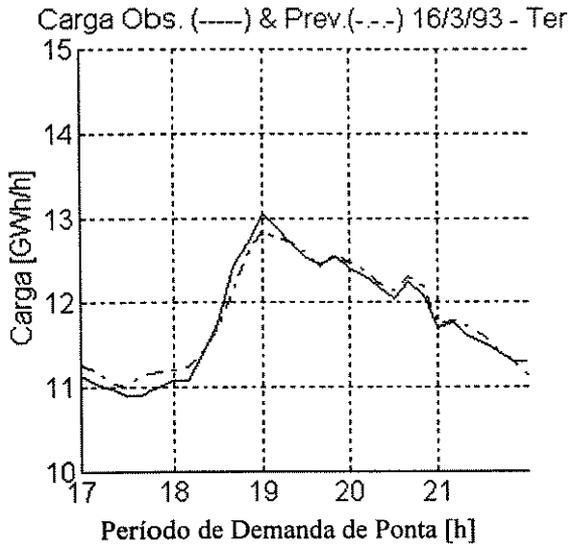


Figura.A.9: Perfil com a melhor previsão
24 horas à frente, 1994.
Algoritmo *MWS-ALL*

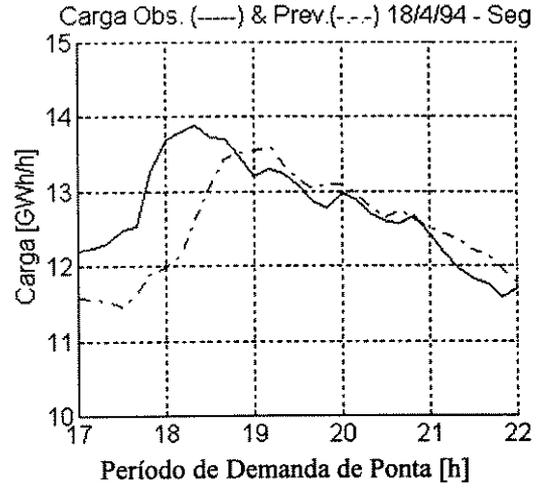


Figura.A.10: Perfil com a pior previsão
24 horas à frente, 1994.
Algoritmo *MWS-ALL*

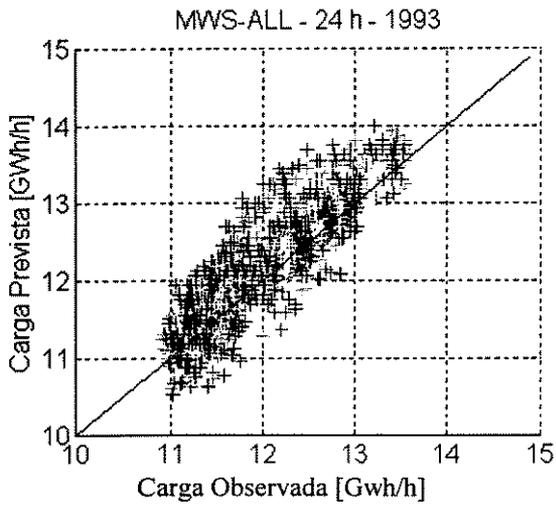


Figura.A.11: Diagrama de dispersão da
previsão 24 horas à frente, 1993
Algoritmo *MWS-ALL*

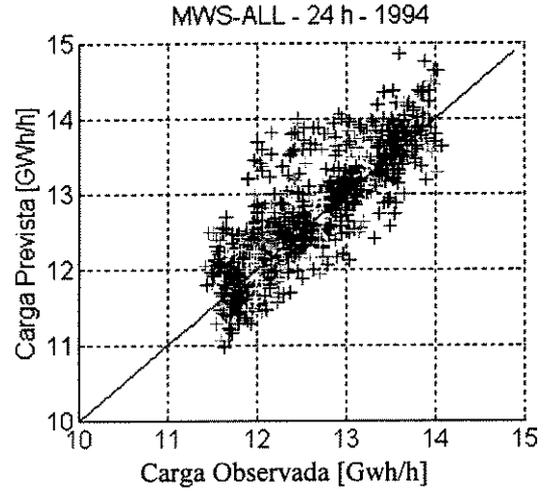


Figura.A.12: Diagrama de dispersão da
previsão 24 horas à frente, 1994
Algoritmo *MWS-ALL*

A.2- Algoritmo *MWS-PPT*

A.2.1- Horizonte 1 hora à frente

- Variáveis de entrada: Carga (histórico 10 minutos) às $x, \dots, x-3$ intervalos do dia anterior; carga às x horas uma semana atrás e temperatura máxima.
- Método: dinâmico
- Topologia: 6 X 3 X 1
- Período: 13/01 a 31/05 (exceto o período transitório do horário de verão)

Ano	Nr. Prev.	MPE [%]	MAPE [%]	SDAPE [%]	MAXAPE [%]	NPM800	Nr. Tr. por Dia	TMaxPT [s]	TMaxPP [s]
1993	2077	0,52	2,16	1,75	9,52	42	31	24,72	260,23
1994	1767	0,30	1,99	1,83	9,79	55	31	22,70	280,87

Tabela.A.5: Desempenho das RNA's para o algoritmo *MWS-PPT* 1 hora à frente

MAPE [%]	Frequência		Percentual	
	1993	1994	1993	1994
$-\infty$ ----- - 12	0,00	0,00	0,00	0,00
- 12 ----- - 10	0,00	0,00	0,00	0,00
- 10 ----- - 8	0,00	0,17	0,00	0,17
- 8 ----- - 6	1,02	0,99	1,02	0,99
- 6 ----- - 4	2,12	4,05	2,12	4,05
- 4 ----- - 2	8,03	8,02	8,03	8,02
- 2 ----- 0	27,85	29,45	27,85	29,45
0 ----- 2	32,42	35,24	32,42	35,24
2 ----- 4	19,51	13,65	19,51	13,65
4 ----- 6	6,61	5,13	6,61	5,13
6 ----- 8	2,44	2,23	2,44	2,23
8 ----- 10	0,00	1,08	0,00	1,08
10 ----- 12	0,00	0,00	0,00	0,00
12 ----- ∞	0,00	0,00	0,00	0,00

Tabela.A.6: Frequência Percentual do *MAPE* para o algoritmo *MWS-PPT* 1 hora à frente

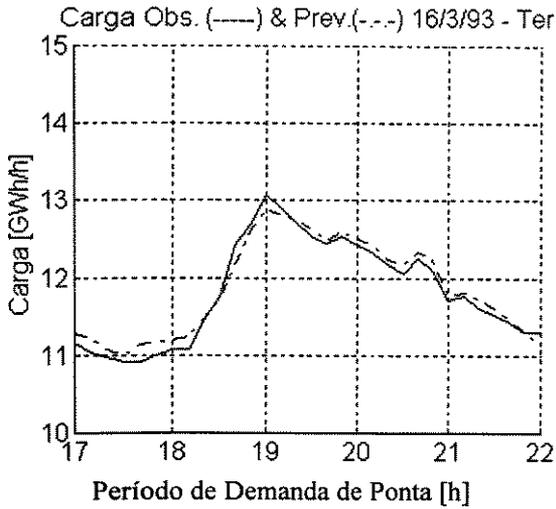


Figura.A.13: Perfil com a
melhor previsão
1 hora à frente, 1993.
Algoritmo *MWS-PPT*

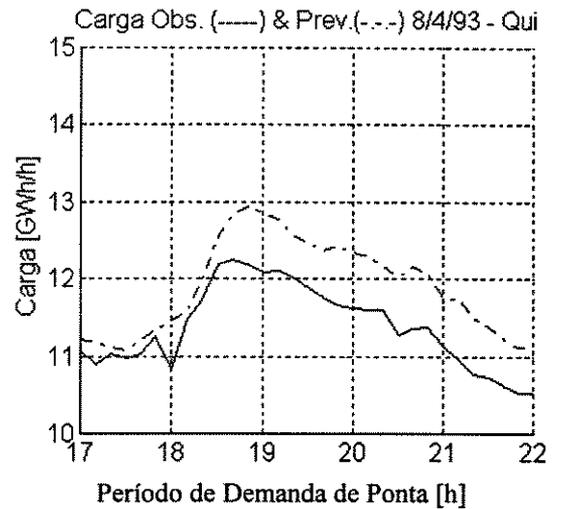


Figura.A.14: Perfil com a pior previsão
1 hora à frente, 1993.
Algoritmo *MWS-PPT*

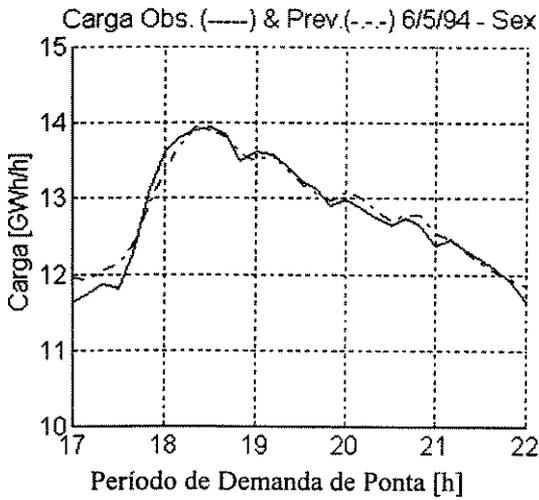


Figura.A.15: Perfil com a
melhor previsão
1 hora à frente, 1994.
Algoritmo *MWS-PPT*

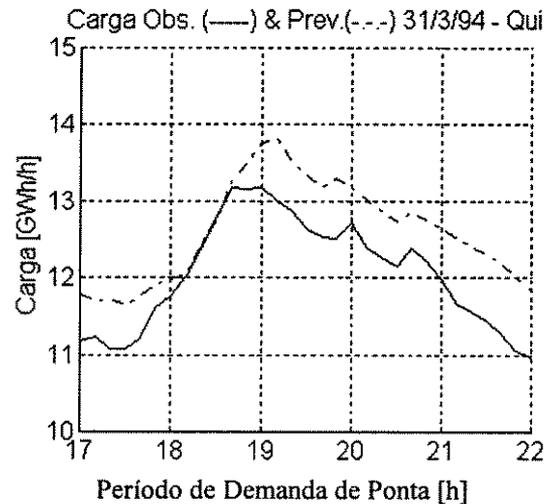


Figura.A.16: Perfil com a pior previsão
1 hora à frente, 1994.
Algoritmo *MWS-PPT*

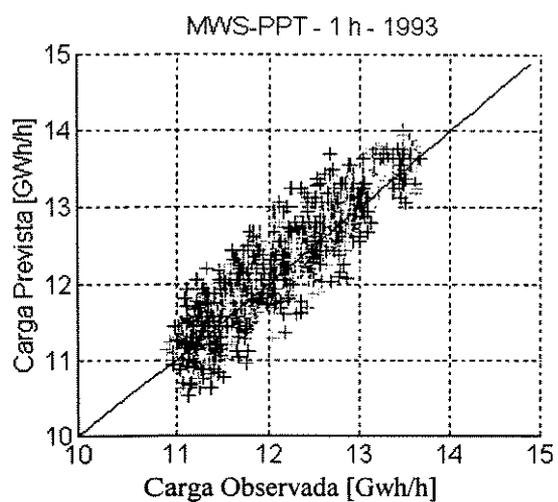


Figura.A.17: Diagrama de dispersão da previsão 1 hora à frente, 1993
Algoritmo *MWS-PPT*

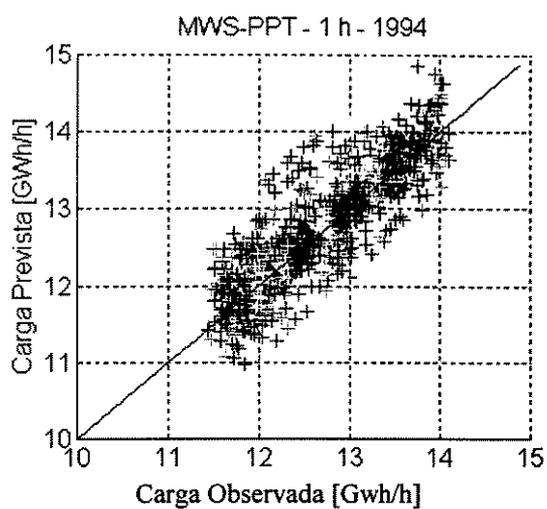


Figura.A.18: Diagrama de dispersão da previsão 1 hora à frente, 1994
Algoritmo *MWS-PPT*

A.2.2- Horizonte 24 horas à frente

- Variáveis de entrada: Carga (histórico 10 minutos) às $x, \dots, x-3$ intervalos do dia anterior à previsão; carga às x horas uma semana atrás e temperatura prevista máxima.
- Método: dinâmico
- Topologia: 6 X 3 X 1
- Período: 13/01 a 31/05 (exceto o período transitório do horário de verão)

Ano	Nr. Prev.	MPE [%]	MAPE [%]	SDAPE [%]	MAXAPE [%]	NPM800	Nr.Tr. por Dia	TMaxPT [s]	TMaxPP [s]
1993	2077	0,55	2,17	1,74	9,75	39	31	16,52	131,20
1994	1767	0,29	1,97	1,82	10,12	51	31	18,20	75,67

Tabela.A.7: Desempenho das RNA's para o algoritmo *MWS-PPT* 24 horas à frente

MAPE [%]		Frequência		Percentual	
		1993		1994	
- ∞	----- - 12	0,00		0,00	
- 12	----- - 10	0,00		0,00	
- 10	----- - 8	0,00		0,17	
- 8	----- - 6	0,94		0,99	
- 6	----- - 4	2,20		4,14	
- 4	----- - 2	7,87		8,11	
- 2	----- 0	27,93		29,61	
0	----- 2	32,57		35,07	
2	----- 4	19,98		13,56	
4	----- 6	6,14		5,21	
6	----- 8	2,28		1,99	
8	----- 10	0,08		1,08	
10	----- 12	0,00		0,08	
12	----- ∞	0,00		0,00	

Tabela.A.8: Frequência Percentual do *MAPE* para o algoritmo *MWS-PPT* 24 horas à frente

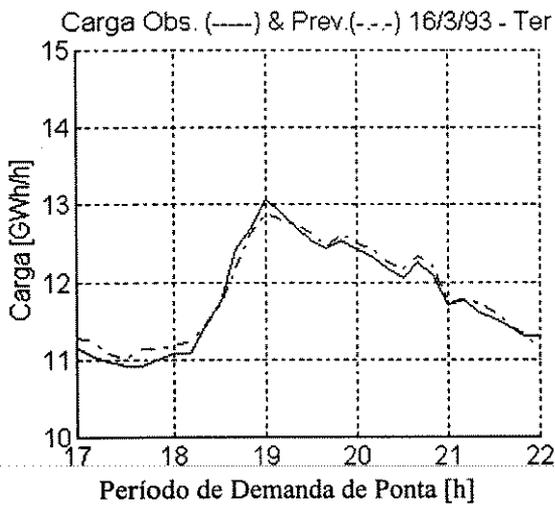


Figura.A.19: Perfil com a melhor previsão 24 horas à frente, 1993. Algoritmo *MWS-PPT*

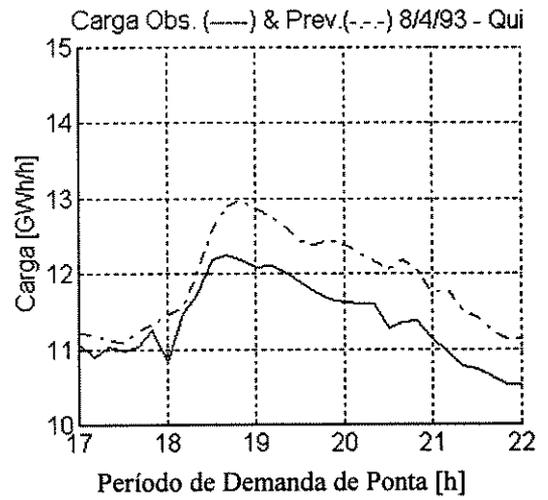


Figura.A.20: Perfil com a pior previsão 24 horas à frente, 1993. Algoritmo *MWS-PPT*

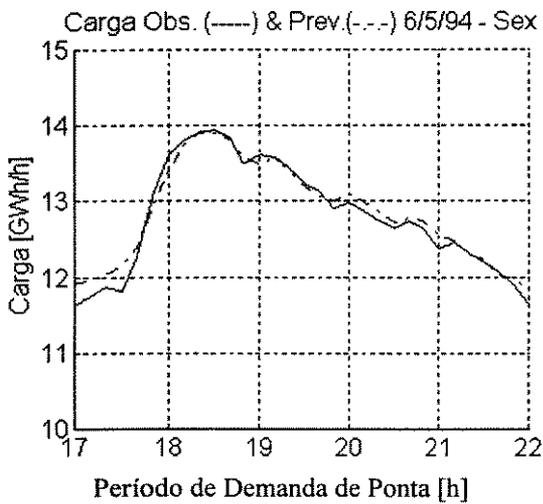


Figura.A.21: Perfil com a melhor previsão 24 horas à frente, 1994. Algoritmo *MWS-PPT*

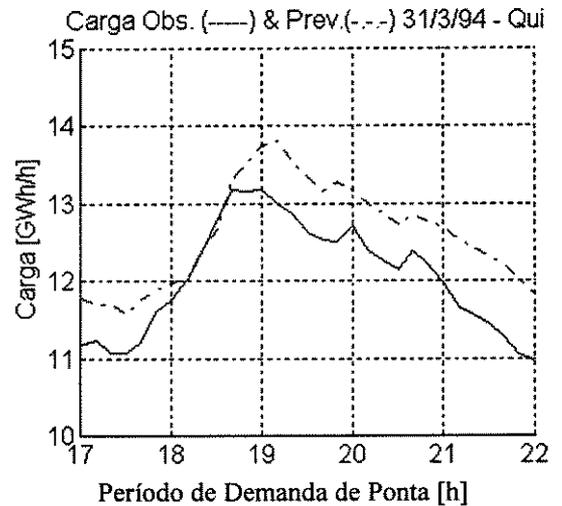


Figura.A.22: Perfil com a pior previsão 24 horas à frente, 1994. Algoritmo *MWS-PPT*

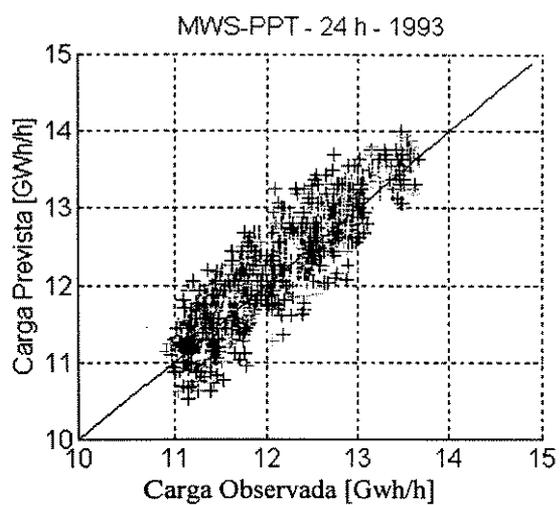


Figura.A.23: Diagrama de dispersão da previsão 24 horas à frente, 1993
Algoritmo *MWS-PPT*

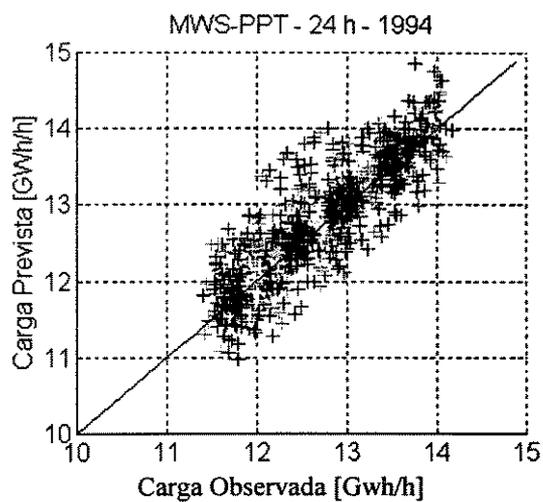


Figura.A.24: Diagrama de dispersão da previsão 24 horas à frente, 1994
Algoritmo *MWS-PPT*

A.2.3- Horizonte 10 minutos à frente

- Variáveis de entrada: Carga (histórico 10 minutos) às $x-1, \dots, x-3$ intervalos; carga às x horas do dia anterior e temperatura máxima.
- Método: dinâmico
- Topologia: 5 X 3 X 1
- Período: 13/01 a 31/05 (exceto o período transitório do horário de verão)

Ano	Nr. Prev.	MPE [%]	MAPE [%]	SDAPE [%]	MAXAPE [%]	NPM800	Nr.Tr. por Dia	TMaxPT [s]	TMaxPP [s]
1993	2077	0,63	2,01	1,71	10,10	42	31	16,32	62,98
1994	1767	0,42	1,91	1,79	11,35	51	31	16,27	38,95

Tabela.A.9: Desempenho das RNA's para o algoritmo *MWS-PPT* 10 minutos à frente

MAPE [%]		Frequência		Percentual	
		1993		1994	
$-\infty$	----- - 12	0,00		0,00	
- 12	----- - 10	0,00		0,00	
- 10	----- - 8	0,00		0,00	
- 8	----- - 6	0,47		0,74	
- 6	----- - 4	2,28		3,14	
- 4	----- - 2	6,92		8,02	
- 2	----- 0	28,64		30,52	
0	----- 2	33,20		36,23	
2	----- 4	20,06		13,32	
4	----- 6	6,53		4,96	
6	----- 8	1,89		2,07	
8	----- 10	0,00		0,99	
10	----- 12	0,00		0,00	
12	----- ∞	0,00		0,00	

Tabela.A.10: Frequência Percentual do *MAPE* para o algoritmo *MWS-PPT* 10 minutos à frente.

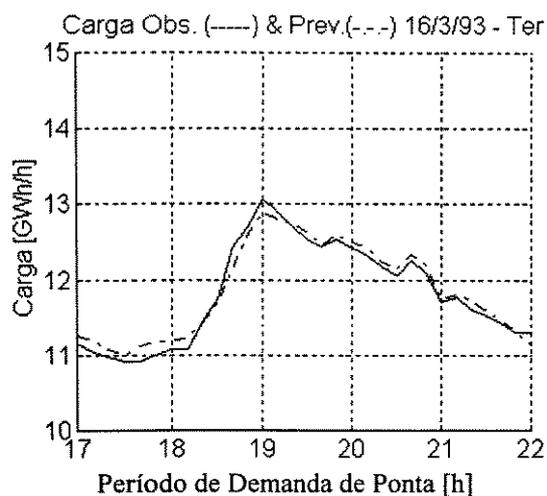


Figura.A.25: Perfil com a melhor previsão 10 minutos à frente, 1993. Algoritmo *MWS-PPT*

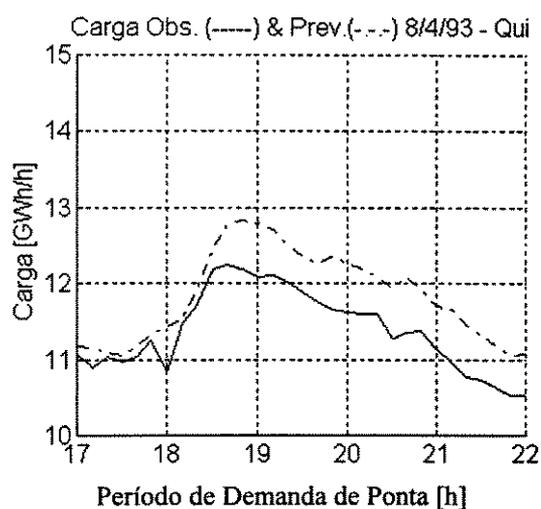


Figura.A.26: Perfil com a pior previsão 10 minutos à frente, 1993. Algoritmo *MWS-PPT*

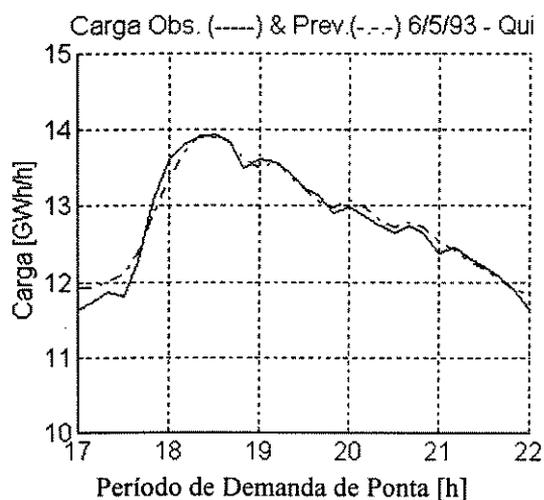


Figura.A.27: Perfil com a melhor previsão 10 minutos à frente, 1994. Algoritmo *MWS-PPT*

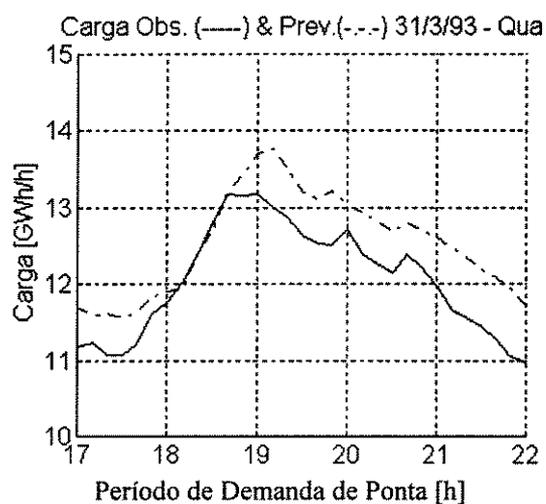


Figura.A.28: Perfil com a pior previsão 10 minutos à frente, 1994. Algoritmo *MWS-PPT*

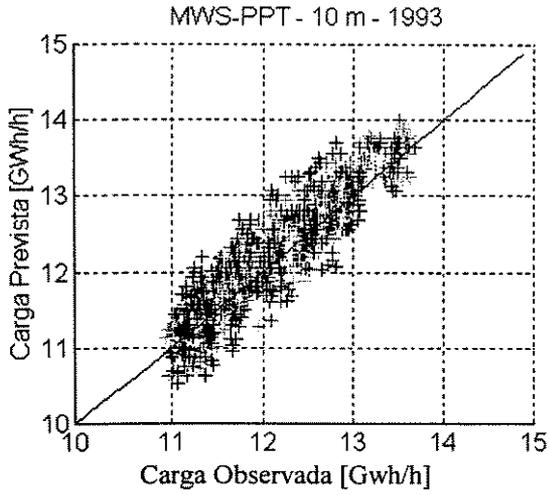


Figura.A.29: Diagrama de dispersão da previsão 10 minutos à frente, 1993
Algoritmo *MWS-PPT*

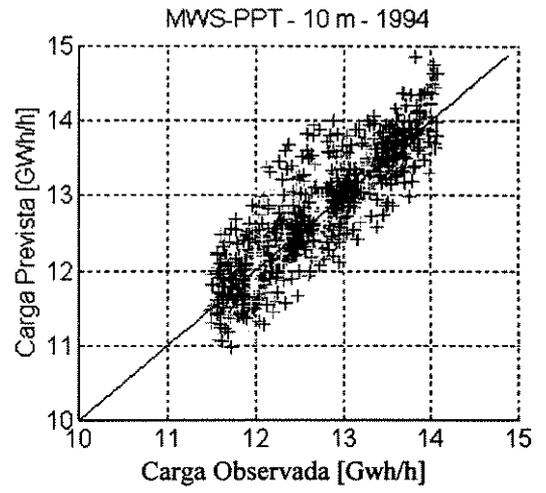


Figura.A.30: Diagrama de dispersão da previsão 10 minutos à frente, 1994
Algoritmo *MWS-PPT*

A.3- Algoritmo *AAS-ALL*

A.3.1- Horizonte 1 hora à frente

- Variáveis de entrada: Carga (histórico horário) às 14 e 15 horas no mesmo dia da previsão, um dia atrás e uma semana atrás; temperatura máxima e mínima; hora do pôr-do-sol e horário de verão.
- Método: estático
- Topologia: 10 X 5 X 31
- Período: 13/01 a 31/05 (exceto o período transitório do horário de verão)

Ano	Nr. Prev.	MPE [%]	MAPE [%]	SDAPE [%]	MAXAPE [%]	NPM800	Nr.Tr. por Ano	Tempo Trein.[h]
1994	2077	1,56	2,78	1,97	11,68	90	1	1:26:41

Tabela.A.11: Desempenho das RNA's para o algoritmo *AAS-ALL* 1 hora à frente

MAPE [%]		Freqência Percentual
		1994
- ∞	----- - 12	0,00
- 12	----- - 10	0,22
- 10	----- - 8	0,43
- 8	----- - 6	1,08
- 6	----- - 4	3,44
- 4	----- - 2	6,02
- 2	----- 0	15,70
0	----- 2	24,30
2	----- 4	28,60
4	----- 6	15,05
6	----- 8	4,52
8	----- 10	0,65
10	----- 12	0,00
12	----- ∞	0,00

Tabela.A.12: Freqüência Percentual do *MAPE* para o algoritmo *AAS-ALL* 1 hora à frente

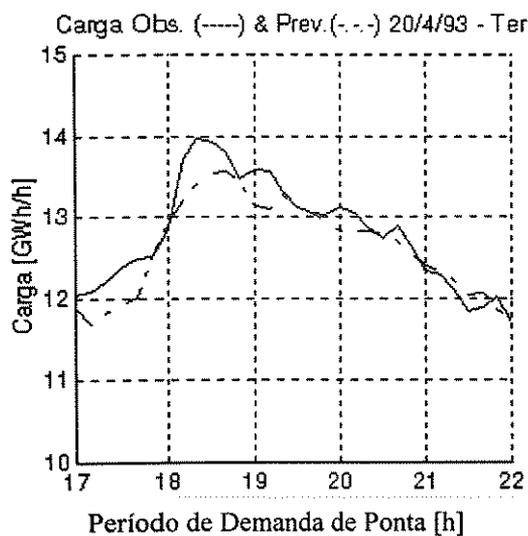


Figura.A.31: Perfil com a melhor previsão 1 hora à frente, 1994. Algoritmo *AAS-ALL*

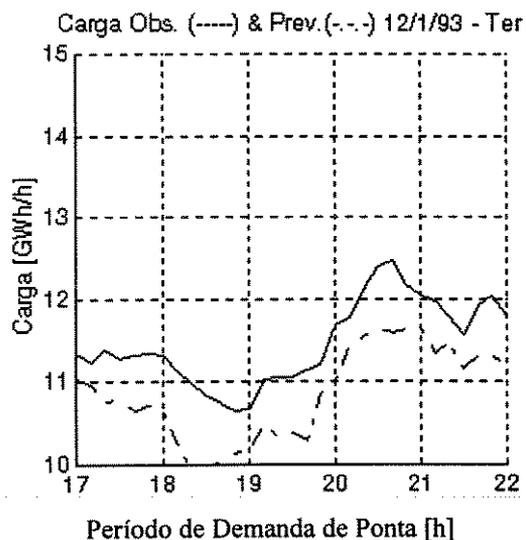


Figura.A.32: Perfil com a pior previsão 1 hora à frente, 1994. Algoritmo *AAS-ALL*

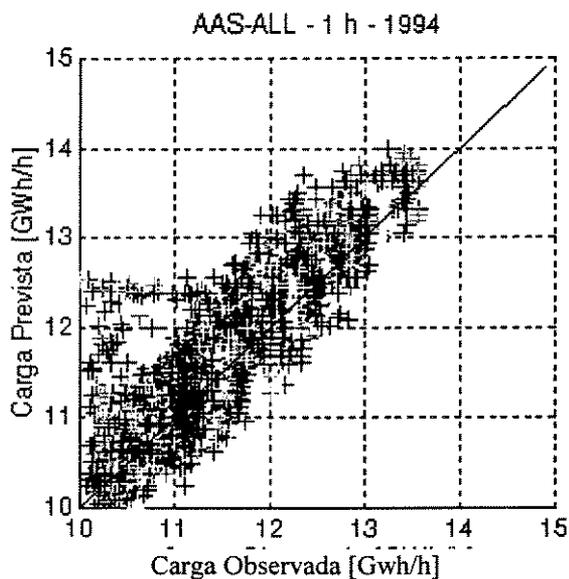


Figura.A.33: Diagrama de dispersão da previsão 1 hora à frente, 1994 Algoritmo *AAS-ALL*

A.3.2- Horizonte 24 horas à frente

- Variáveis de entrada: Carga (histórico horário) às 14 e 15 horas do dia anterior, um dia atrás do dia da previsão e uma semana atrás; temperatura prevista máxima e mínima; hora do pôr-do-sol e horário de verão.
- Método: estático
- Topologia: 9 X 5 X 31
- Período: 13/01 a 31/05 (exceto o período transitório do horário de verão)

Ano	Nr. Prev.	MPE [%]	MAPE [%]	SDAPE [%]	MAXAPE [%]	NPM800	Nr.Tr. por Ano	Tempo Trein. [h]
1994	2077	1,87	2,97	1,92	12,65	132	1	1:26:38

Tabela.A.13: Desempenho das RNA's para o algoritmo *AAS-ALL* 24 horas à frente

MAPE [%]		Freqüência Percentual
		1994
- ∞	----- - 12	0.00
- 12	----- - 10	0.00
- 10	----- - 8	0.43
- 8	----- - 6	0.86
- 6	----- - 4	3.66
- 4	----- - 2	5.81
- 2	----- 0	11.18
0	----- 2	25.38
2	----- 4	28.82
4	----- 6	17.63
6	----- 8	5.59
8	----- 10	0.43
10	----- 12	0.22
12	----- ∞	0.00

Tabela.A.14: Freqüência Percentual do *MAPE* para o algoritmo *AAS-ALL* 24 horas à frente

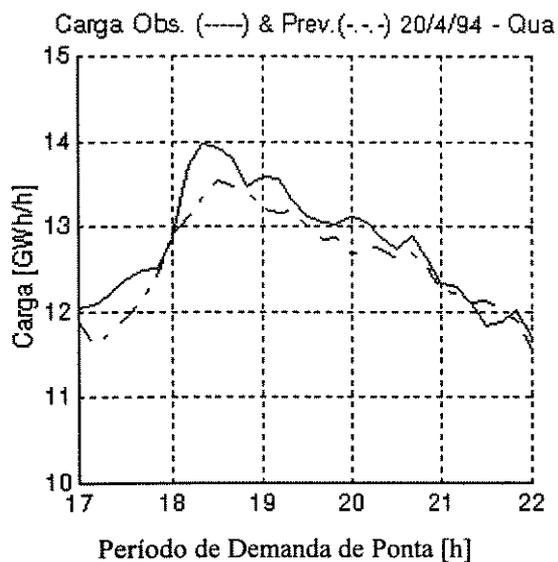


Figura.A.34: Perfil com a melhor previsão 24 horas à frente, 1994. Algoritmo *AAS-ALL*

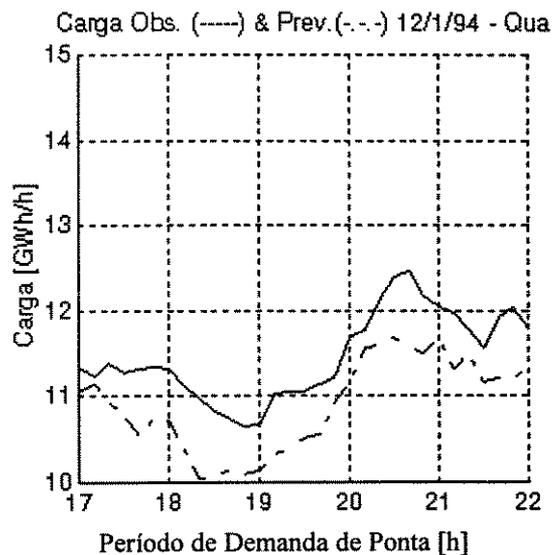


Figura.A.35: Perfil com a pior previsão 24 horas à frente, 1994. Algoritmo *AAS-ALL*

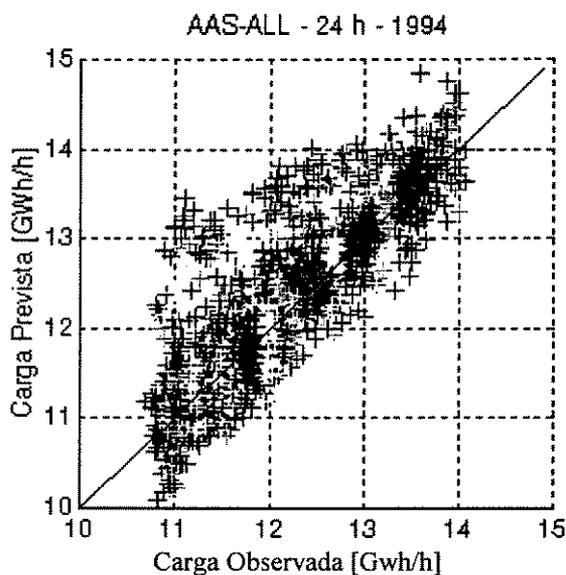


Figura.A.36: Diagrama de dispersão da previsão 24 horas à frente, 1994 Algoritmo *AAS-PPT*

A.4- Algoritmo AAS-PPT

A.4.1- Horizonte 1 hora à frente

- Variáveis de entrada: Carga (histórico 10 minutos) às $x, \dots, x-3$ intervalos atrás do dia anterior; carga na hora x uma semana atrás; temperatura máxima e mínima; hora do pôr-do-sol; horário de verão; hora e minuto.
- Método: dinâmico
- Topologia: 11 X 6 X 1
- Período: 13/01 a 31/05 (exceto o período transitório do horário de verão)

Ano	Nr. Prev.	MPE [%]	MAPE [%]	SDAPE [%]	MAXAPE [%]	NPM800	Nr.Tr. por Ano	Tempo Trein. [h]
1994	2077	0,95	2,56	1,74	12,88	70	1	2:08:21

Tabela.A.15: Desempenho das RNA's para o algoritmo AAS-ALL 1 hora à frente

MAPE [%]		Freqência Percentual
		1994
- ∞	----- - 12	0.00
- 12	----- - 10	0.00
- 10	----- - 8	0.23
- 8	----- - 6	4.38
- 6	----- - 4	5.07
- 4	----- - 2	8.06
- 2	----- 0	16.36
0	----- 2	27.42
2	----- 4	29.26
4	----- 6	8.76
6	----- 8	0.46
8	----- 10	0.00
10	----- 12	0.00
12	----- ∞	0.00

Tabela.A.16: Freqência Percentual do MAPE para o algoritmo AAS-PPT 1 hora à frente

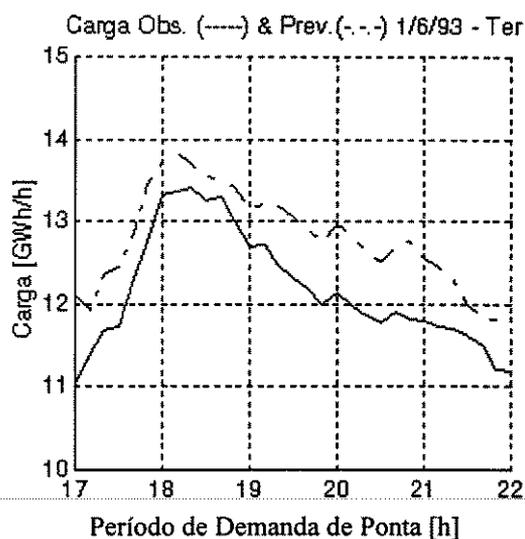
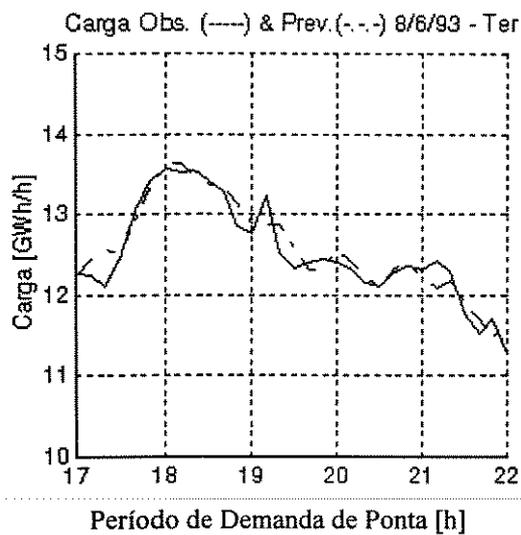


Figura.A.37: Perfil com a melhor previsão 1 hora à frente, 1994. Algoritmo *AAS-PPT*

Figura.A.38: Perfil com a pior previsão 1 hora à frente, 1994. Algoritmo *AAS-PPT*

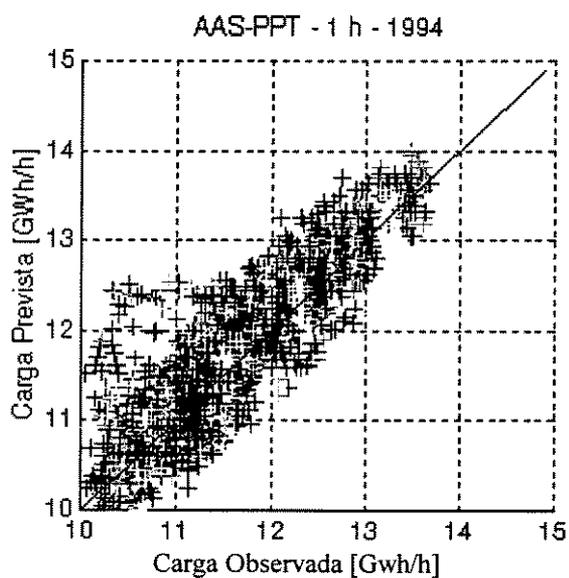


Figura.A.39: Diagrama de dispersão da previsão 1 hora à frente, 1994 Algoritmo *AAS-PPT*

A.4.2- Horizonte 24 horas à frente

- Variáveis de entrada: Carga (histórico 10 minutos) às $x, \dots, x-3$ intervalos atrás do dia anterior da previsão; carga na hora x uma semana atrás; temperatura prevista máxima e mínima; hora do pôr-do-sol; horário de verão; hora e minuto.
- Método: dinâmico
- Topologia: 11 X 6 X 1
- Período: 13/01 a 31/05 (exceto o período transitório do horário de verão)

Ano	Nr. Prev.	MPE [%]	MAPE [%]	SDAPE [%]	MAXAPE [%]	NPM800	Nr.Tr. por Ano	Tempo Trein.[h]
1994	2077	1,07	2,70	2,23	13,25	150	1	1:26:06

Tabela.A.17: Desempenho das RNA's para o algoritmo *AAS-PPT* 24 horas à frente

MAPE [%]		Freqüência Percentual
		1994
- ∞	----- - 12	0.00
- 12	----- - 10	0.00
- 10	----- - 8	1.38
- 8	----- - 6	2.76
- 6	----- - 4	2.76
- 4	----- - 2	9.91
- 2	----- 0	23.73
0	----- 2	26.73
2	----- 4	15.67
4	----- 6	8.99
6	----- 8	7.14
8	----- 10	0.92
10	----- 12	0.00
12	----- ∞	0.00

Tabela.A.18: Freqüência Percentual do *MAPE* para o algoritmo *AAS-PPT* 24 horas à frente

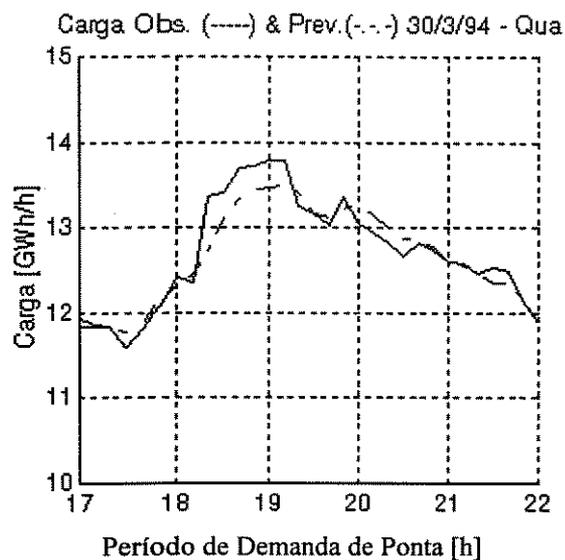


Figura.A.40: Perfil com a melhor previsão 24 horas à frente, 1994. Algoritmo *AAS-PPT*

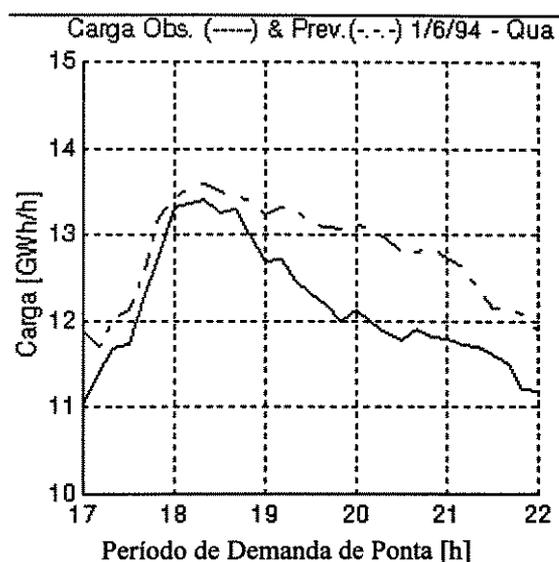


Figura.A.41: Perfil com a pior previsão 24 horas à frente, 1994. Algoritmo *AAS-PPT*

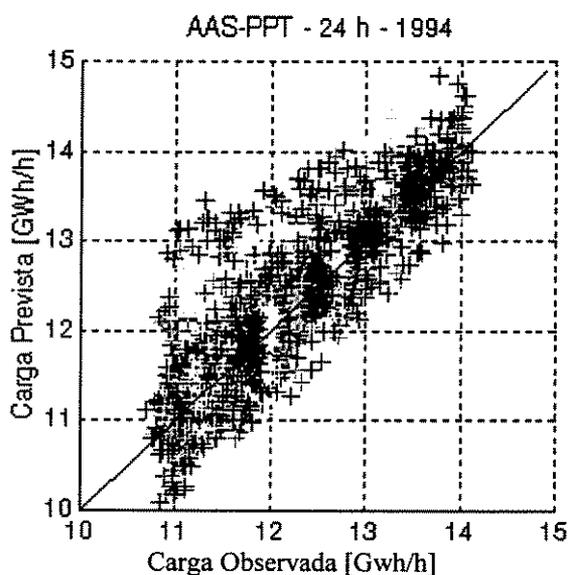


Figura.A.42: Diagrama de dispersão da previsão 24 horas à frente, 1994 Algoritmo *AAS-PPT*