

# Previsão de Fornecimento de Energia Elétrica : Um Estudo de Caso do Ponto de Vista do Barramento

Daniela Dias Rodrigues

14 de julho de 2002

Universidade Federal de Minas Gerais  
Centro de Pesquisa e Desenvolvimento em Engenharia Elétrica  
Programa de Pós-Graduação em Engenharia Elétrica

# **Previsão de Fornecimento de Energia Elétrica : Um Estudo de Caso do Ponto de Vista do Barramento**

Daniela Dias Rodrigues

Dissertação submetida à banca examinadora designada pelo Colegiado do Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal de Minas Gerais como parte dos requisitos necessários à obtenção do grau de Mestre em Engenharia Elétrica.

Orientador: Luis Antônio Aguirre

9 de Abril de 2002

## Resumo

Informações de consumo de energia elétrica minutos ou mesmo horas à frente, são de grande importância para as empresas envolvidas no setor elétrico (desde concessionárias ao operador nacional do sistema, ONS). Com esses dados, pode-se operar de forma otimizada e com melhor qualidade, a energia gerada, transmitida e distribuída aos vários pontos do sistema. Isso, em parte explica o recente interesse pelo tema de previsão de consumo de energia elétrica.

Neste trabalho, ao invés de usar diretamente os dados de energia medidos — quem têm uma forte componente sazonal — será usada uma série denominada “série de resíduos”, onde a componente sazonal é muito menos evidente.

A série de resíduos foi obtida subtraindo da série original a série de ciclos médios. Essa série de ciclos médios foi calculada a partir da classificação automática de padrões via redes de Kohonen.

A contribuição deste estudo está na análise de uma das parcelas da série original, a série de resíduos. Para realizar a previsão de consumo a curto prazo, utilizou-se de forma inovadora, a série de resíduos, e não a usual série original como encontrado na literatura. O objetivo principal não é obter a melhor previsão, mas verificar quanto da série de resíduos é previsível.

Com o objetivo de encontrar resposta para esta questão, foi feito um estudo preliminar de análise de determinismo nesta nova série. Através desta análise, não foi possível identificar a presença de determinismo em todas as janelas da série. Ou seja, os resultados mostram que existem trechos do sinal em que se torna muito difícil fazer uma previsão melhor que o ciclo médio.

Os dados utilizados nesta dissertação, foram de uma série de carga de uma concessionária de Minas Gerais, a Companhia Força e Luz Cataguazes Leopoldina. Os dados foram obtidos no período de 01 de Novembro de 2000 a 31 de Dezembro de 2000, totalizando 61 dias (1464hs). Os testes de previsão foram realizados com o uso do Matlab, através das técnicas de redes neurais artificiais. Essa ferramenta mostrou-se bastante flexível e adequada para o uso em previsores de carga.

Os valores finais de erros de previsão (Erro percentual absoluto médio, MAPE) ficaram na faixa de  $3\% < MAPE < 5\%$  para as janelas onde algum

determinismo foi encontrado. A comparação destes resultados com os publicados na literatura deve ser feita de forma muito cautelosa tendo em vista que os dados utilizados neste trabalho têm algumas características peculiares, como por exemplo o fato da série ser de potência aparente o que sabidamente dificulta a previsão.

## Abstract

Information of energy demand from minutes or hours ahead is very important to the electrical industrial sector (from utility companies to the national operator of the system, ONS). Having this information, they can manage the energy generated, transmitted and distributed to many points of the grid with better quality and in a optimized way. This partially explains why load forecasting has attracted much attention recently.

In this work, instead of using the original load data directly — which has a very strong seasonal component — residuals time series, will be used. The latter data have a less evident seasonal component.

The residual were obtained subtracting from the original load data the average daily cycles. These averages load data were calculated by automatic standard classification using Kohonen network technique.

This research focused mainly on the residuals data, that is, one of the parts of the original load data. The novelty of this work consists on using the residual and not the usual original data as seen in the literature. The main objective isn't to get the better forecasting, but to check how predictable the residual are.

In order to find the answer to this question, a preliminary analysis of determinism on the residuals was done. Through this analysis, it was not possible to identify in some parts of the data, the presence of determinism. In other words, the results show that some windows (parts) of the data are very difficult to get the forecasting result better than the average daily cycle.

The data used in this dissertation, were obtained from one of the utility companies of Minas Gerais State, the Companhia Força e Luz Cataguazes Leopoldina. The data were collected from November 1<sup>st</sup> to December 31<sup>rd</sup> of 2000, a total number of 61 days (1464 hours). The load forecasting was done using Matlab, with artificial neural networks. This technique was found very flexible and adequate to use in load forecasting research.

The forecasts Mean absolute percent error was  $3\% < MAPE < 5\%$  for parts that have some determinism. The comparison of these results with the literature should be done very carefully because the data used in this work have some special features, such as the time series corresponds to apparent power (rather than active power), which is known to be more difficult to predict.

## Agradecimentos

A Deus, pela força que me deste durante todo esse trabalho e em toda minha vida.

Aos meus pais pela paciência e compreensão de minhas faltas em momentos importantes da nossa família.

Aos meus irmãos Rogério e Cíntia, bem como meus amigos, pela confiança e companherismo.

Ao Cris pelo amor, atenção, carinho e paciência, estando sempre ao meu lado.

Um agradecimento especial ao Silvio pela contribuição feita para a realização deste trabalho, em especial no capítulo 5.

À Companhia Força e Luz Cataguases Leopoldina, pelo fornecimento dos dados.

E por fim um agradecimento mais que especial ao Prof. Luis Aguirre pela atenção, coordenação e excelente orientação mesmo eu estando longe.

A todos que caminharam comigo nessa etapa maravilhosa da minha vida.

## Nomenclatura

### Lista de Símbolos

$a_j$	Vetor de entrada do neurônio;
$a_i$	Vetor de saída do neurônio;
$b$	Vetor de <i>bias</i> ;
$d$	Maior atraso em tempo morto;
$D$	Consumo de energia diário;
$e(k)$	Erro no instante $k$ , pode ou não ser branco;
$F$	Qualquer função vetorial não linear;
$g$	Função de ativação;
$I$	Dados de corrente (A);
$k$	Índice de tempo para sistemas discretos;
$L$	Valor de baixa demanda de energia diário;
$N$	Número de épocas para treinamento da rede;
$n_y$	Máximo atraso entre os regressores de saída;
$n_u$	Máximo atraso entre os regressores de entrada;
$n_e$	Máximo atraso entre os regressores de ruído;

$q$	Número de saídas;
$q^{-1}$	Operador de atraso, $y(k)q^{-1} = y(k - 1)$ ;
$r_{pp}(k)$	Função de autocorrelação de $S(k)$ ;
$S$	Série de carga (kVA);
$\tilde{S}(k)$	Série de resíduos;
$\bar{S}(k)$	Série de ciclos médios;
$\bar{S}_1(k)$	Média de domingos;
$\bar{S}_2(k)$	Média dos dias úteis;
$t$	Índice de tempo para sistemas dinâmicos;
$u(k)$	Sinal de entrada no instante $k$ ;
$V$	Dados de tensão (kV);
$y(k)$	Sinal de saída no instante $k$ ;
$W_{j,i}$	Matriz de pesos;
$\eta(t)$	Taxa de aprendizagem;
$\epsilon(n)$	Erro quadrático na saída da rede no instante $n$ ;
$\theta_i$	$i$ -ésimo parâmetro a ser estimado;
$\sigma$	Desvio padrão;
$\psi(k - 1)$	Vetor de regressores que contém observações até o instante $k - 1$ ;



## Lista de abreviações

ARX	Modelo auto-regressivo com entradas exógenas ( <i>autoregressive model with exogenous inputs</i> );
ARMAX	Modelo auto-regressivo, de média móvel com entradas exógenas ( <i>autoregressive moving average model with exogenous inputs</i> );
CAG	Unidade no ONS responsável pelo controle automático da geração;
CFLCL	Companhia Força e Luz Cataguazes Leopoldina;
DE	Unidade responsável pelo despacho econômico;
EE	Energia elétrica;
ERR	Taxa de redução de erro ( <i>error reduction ratio</i> );
IA	Inteligência Artificial;
MAPE	Erro percentual absoluto médio;
NARX	Modelo não-linear auto-regressivo, com variáveis exógenas ( <i>nonlinear autoregressive model with exogenous inputs</i> );
NARMA	Modelo não-linear auto-regressivo com média móvel ( <i>Nonlinear autoregressive moving average model</i> );
NARMAX	Modelo não-linear auto-regressivo, de média móvel com entradas exógenas ( <i>Nonlinear autoregressive moving average model with exogenous inputs</i> );
ONS	Operador nacional do sistema elétrico;

PCCP	Previsão de consumo a curto prazo;
RMSE	Raiz do erro quadrático médio ( <i>Root Mean Squared Error</i> );
SGE	Unidade responsável pelo sistema de gerenciamento de energia;

# Sumário

<b>1</b>	<b>Introdução</b>	<b>7</b>
<b>2</b>	<b>Previsão de Consumo de Energia Elétrica</b>	<b>11</b>
2.1	Visão geral do problema . . . . .	11
2.2	Importância da previsão de consumo . . . . .	14
2.3	Dificuldades na previsão de consumo . . . . .	17
<b>3</b>	<b>Representações Matemáticas</b>	<b>21</b>
3.1	Introdução . . . . .	21
3.2	Modelos de Box & Jenkins . . . . .	21
3.3	Modelos NARMA e NARMAX . . . . .	23
3.4	Outras representações . . . . .	24
3.5	Redes Neurais Artificiais - RNA . . . . .	25
3.5.1	Conceitos gerais . . . . .	25
3.5.2	Breve histórico . . . . .	26
3.5.3	Aplicações de RNAs . . . . .	27
3.5.4	Neurônios biológicos $\times$ neurônios artificiais . . . . .	28
3.5.5	Tipo de redes neurais artificiais . . . . .	30

<i>SUMÁRIO</i>	2
3.5.6 Mapas de características de Kohonen . . . . .	31
3.5.7 Redes MLP . . . . .	35
<b>4 Análise e Classificação de Padrões</b>	<b>41</b>
4.1 Introdução . . . . .	41
4.2 Análise dos dados . . . . .	42
4.3 Classificação via redes de Kohonen . . . . .	46
4.3.1 Resultado da classificação . . . . .	46
4.3.2 Redução da sazonalidade . . . . .	49
4.4 Análise dos Resultados . . . . .	51
<b>5 Análise de Determinismo</b>	<b>54</b>
5.1 Introdução . . . . .	54
5.2 Teste de determinismo . . . . .	55
5.2.1 O algoritmo . . . . .	56
5.3 Resultados dos testes . . . . .	58
5.4 Análise dos resultados . . . . .	60
<b>6 Previsão de Curto Prazo</b>	<b>68</b>
6.1 Introdução . . . . .	68
6.2 Definição da série de carga . . . . .	69
6.3 Seleção dos dados para treinamento e validação . . . . .	70
6.4 Escolha das variáveis de entrada das RNA's . . . . .	70
6.5 Definição da estrutura das RNA's . . . . .	70
6.6 Treinamento . . . . .	72
6.7 Testes e simulações . . . . .	73

<i>SUMÁRIO</i>	3
6.7.1 Testes usando a janela 1 . . . . .	73
6.7.2 Testes usando a janela 5 . . . . .	74
6.7.3 Testes usando a janela 7 . . . . .	76
6.7.4 Testes usando a janela 9 . . . . .	77
6.8 Validação . . . . .	79
<b>7 Considerações Finais</b>	<b>84</b>
7.1 Discussão . . . . .	84
7.2 Conclusões . . . . .	85
7.3 Sugestão para trabalhos futuros . . . . .	85

# Lista de Figuras

2.1	Curvas diárias de sábados . . . . .	18
2.2	Não estacionariedade da série de carga $S(k)$ . . . . .	19
2.3	Autocorrelação de $S(k)$ . . . . .	20
3.1	Neurônio Biológico . . . . .	28
3.2	Representação matemática de um Neurônio Artificial . . . . .	29
3.3	Funções de ativação . . . . .	30
3.4	Exemplos de arquiteturas de RNAs . . . . .	31
3.5	Rede de Kohonen . . . . .	32
3.6	Função chapéu mexicano . . . . .	33
3.7	Estrutura de rede MLP . . . . .	37
3.8	Fases do <i>backpropagation</i> . . . . .	38
3.9	Superfície de erro: <i>backpropagation</i> . . . . .	40
4.1	Série de dados: $S(k)$ . . . . .	43
4.2	Parâmetros de classificação . . . . .	45
4.3	Padrões desejados . . . . .	47
4.4	Resultado da classificação de Kohonen . . . . .	48
4.5	Ciclos médios diários . . . . .	49

4.6	Gráfico das séries: $S(k)$ , $\bar{S}(k)$ e $\tilde{S}(k)$ . . . . .	50
4.7	Espectro de potência das séries de carga . . . . .	51
4.8	Autocorrelação de $S(k)$ e $\tilde{S}(k)$ . . . . .	52
5.1	Teste de determinismo para a janela 1 . . . . .	59
5.2	Teste de determinismo para a janela 2 (a) . . . . .	60
5.3	Teste de determinismo para a janela 2 (b) . . . . .	61
5.4	Teste de determinismo para a janela 8 . . . . .	62
5.5	Teste de determinismo para a janela 5 (a) . . . . .	63
5.6	Teste de determinismo para a janela 5 (b) . . . . .	64
5.7	Teste de determinismo para a janela 7 (a) . . . . .	65
5.8	Teste de determinismo para a janela 7 (b) . . . . .	66
5.9	Teste de determinismo para a janela 9 . . . . .	67
6.1	Estrutura básica definida para as MLP's . . . . .	71
6.2	Arquitetura neural da janela 1 . . . . .	74
6.3	Resultados dos testes de previsão para a janela 1 . . . . .	75
6.4	Arquitetura neural da janela 5 . . . . .	76
6.5	Resultados dos testes de previsão para a janela 5 . . . . .	77
6.6	Arquitetura neural da janela 7 . . . . .	78
6.7	Resultados dos testes de previsão para a janela 7 . . . . .	79
6.8	Resultados dos testes de previsão para a janela 9 . . . . .	80
6.9	Resultado da previsão para as janelas 1, 5, 7 e 9 . . . . .	82

# Lista de Tabelas

5.1	Resultados obtidos nos testes de determinismo . . . . .	63
6.1	Características construtivas de rede para a janela 1 . . . . .	74
6.2	Características construtivas de rede para a janela 5 . . . . .	75
6.3	Características construtivas de rede para a janela 7 . . . . .	76
6.4	Características construtivas de rede para a janela 9 . . . . .	78
6.5	Índices de avaliação MAPE para as janelas 1, 5, 7 e 9 . . . . .	81
6.6	Índices de avaliação RMSE para as janelas 1, 5, 7 e 9 . . . . .	81



# Capítulo 1

## Introdução

Devido ao crescimento e à importância econômica dos sistemas elétricos de potência nos países do mundo todo, estudos de previsão de consumo de energia têm sido um dos tópicos de grande interesse de pesquisa nas áreas de engenharia (Liu et al. (1996); Silva et al. (1997); Lamedica et al. (1996); Mori and Kobayashi (1996); Oliveira et al. (2000)). Para as concessionárias de energia elétrica as informações de consumo desde os horizontes de curto prazo (poucos minutos, horas ou até dias a frente) até os de longo prazo (por exemplo, 20 anos), são de grande importância, pois permitem uma otimização da produção, conseqüente redução de custos de geração, melhor operação conferindo mais segurança e confiabilidade ao funcionamento do sistema energético.

Diante desse quadro, destaca-se o crescimento de estudos para a previsão de consumo a curto prazo (Peng et al. (1993); Khotanzad et al. (1997, 1998); Drezga and Rahman (1998); Silva and Moulin (2000); Hippert et al. (2001)). Isto ocorreu em sua maioria, em função da desregulamentação do setor energético e a conseqüente criação do mercado competitivo no setor. Vários países viveram recentemente esta mudança e com isso a energia elétrica passou a ser uma mercadoria a ser vendida e comprada no mercado energético. Assim sendo, uma mercadoria para ser vendida ou comprada precisa ter no mínimo preço. Para compor este preço existe a necessidade de se saber o quanto será vendido, ou seja, o quanto vai ser necessário produzir para atender a demanda (o consumo). Com tudo isso então, as informações sobre o consumo de energia elétrica minutos, horas a frente tornam-se fundamentais para as concessionárias e para o ONS (Operador nacional do sistema), de for-

ma tal que eles poderão gerenciar o sistema de uma maneira mais otimizada, planejada, qualificada e conseqüentemente com o melhor e mais competitivo preço de seu produto (ou mercadoria) (Makansi (2001)).

No entanto, realizar previsão de consumo de energia é um tarefa árdua, visto que a série de carga é complexa e apresenta comportamentos aleatórios devido à influência de vários fatores (temperatura, estação do ano, etc) (Chow and Leung (1996)). Além disto, as características da série variam em função da época do ano, região do país, potência total dependendo do barramento, entre outros fatores. Um outra dificuldade em realizar a previsão de carga é o banco de dados utilizado para este teste. Os dados medidos muitas vezes não estão completos e muitos possuem erros de medição. Outro fator também importante é a diversidade de consumidores, fazendo com que o comportamento da carga seja diferentemente para cada tipo (exemplo: consumidores residenciais, rurais, industriais e comerciais). Com esta complexidade, torna-se mais difícil encontrar uma técnica capaz de representar a série de forma precisa.

Algumas técnicas têm sido utilizadas para este problema, dentre elas podemos destacar os modelos de séries temporais (Box & Jenkins, 1970), os modelos de regressão, tais como NARMA e NARMAX (Cassini (1999); Aguirre and Aguirre (2000); Gomes et al. (2000)). Outro conjunto de técnicas que também merece destaque são os métodos em lógica difusa (*fuzzy logic*) (Caminhas et al. (1997); Mori and Kobayashi (1996)).

Neste estudo optou-se em utilizar as técnicas baseadas em redes neurais artificiais (RNA) devido a familiaridade do autor com estas técnicas e as mesmas terem sido bastante utilizadas para problemas do setor energético, tais como planejamento, análise de segurança, previsões de vazão, previsão de consumo, entre outros (Sodré and Mota (2000); Teive et al. (2000); Valença and Ludermir (1998b); Reis and Silva (2000)). Essa ferramenta mostrou-se atrativa devido à sua característica marcante na facilidade de adaptação (aprendizado) aos dados.

No entanto, após um intenso estudo literário que abrange o tema de previsão de carga, constatou-se que os autores desses, sempre utilizaram dados de medição real (ou seja, a medição da potência na saída do transformador). A esses dados chamou-se neste estudo de séries originais. Esta série é basicamente composta por duas partes:

1. sazonalidade;

2. flutuações que acontecem superpostas a sazonalidade;

A primeira parte corresponde a presença da sazonalidade, ou seja, a série contém dados que ocorrem sempre diariamente, semanalmente, ou anualmente (exemplo: horário de pico as 21hs). Se por um lado a sazonalidade tem carácter bastante previsível, a segunda parcela que corresponde as flutuações que são superpostas a este ciclo tem um carácter muito mais estocástico.

Esse trabalho está fortemente preocupado com a segunda parte da série original, e portanto uma das análises a serem feitas será uma avaliação desta segunda parcela com relação a presença de determinismo. A esta segunda parcela da série original, chamou-se de série de resíduos.

A análise de determinismo tem sido aplicada em outras áreas do conhecimento (Gomes (2001)). A importância deste problema é evidente tendo em mente que as RNAs são supostas a aprender o comportamento dos dados de entrada, e assim conseguir representar os valores futuros adequadamente na sua saída. Se os dados de entrada não possuem dinâmica determinística, não se pode esperar que a rede consiga obter uma boa previsão de dados. Entretanto, se esses são determinísticos, espera-se que a rede consiga representá-los de forma melhor.

O objetivo deste trabalho, é realizar a previsão de consumo de energia a curto prazo (PCCP) utilizando dados de medição de uma concessionária de Minas Gerais. Para tal foram utilizadas redes neurais artificiais do tipo MLP (*Multlayer Perceptron*). Foi feito um estudo preliminar nos dados de entrada da rede que, no contexto deste trabalho, é uma série de resíduos. Este estudo permitiu uma avaliação da presença de determinismo em tal série. Os resíduos são gerados a partir de uma análise de classificação de padrões, utilizando a ferramenta de redes de Kohonen. A série de resíduos é obtida sem assumir a presença da dinâmica do comportamento sazonal destes dados. Para a análise de determinismo, foram utilizados testes de hipóteses, conhecidos como *surrogate data analysis* (Theiler et al. (1992); Gomes (2001)).

Esta dissertação é dividida como se segue. O Capítulo 2 apresenta os conceitos de previsão de carga, bem como a importância do problema e a motivação deste trabalho. O Capítulo 3 apresenta as técnicas clássicas e modernas utilizadas para a solução de PCCP, destacando o uso e aplicação de redes neurais. O Capítulo 4 descreve a análise de classificação de padrões realizada através do uso de redes de Kohonen. Esta classificação permitiu

a construção de dados ‘residuais’ (série de carga sem sazonalidade) para a análise de determinismo no Capítulo 5. O Capítulo 5 descreve como foi feito o estudo de determinismo utilizando modelagem NARMA e análise de *surrogates*. O Capítulo 6 apresenta os procedimentos realizados via redes neurais artificiais para a previsão de consumo, bem como seus resultados e análise. Apesar do procedimento seguido no Capítulo 6 ser específico a esta dissertação, acredita-se que uma combinação importante esteja nos Capítulos 4 e 5, uma vez que os procedimentos descritos ali não são convencionais em estudos de previsão de consumo de energia elétrica. Finalmente, o Capítulo 7 apresenta as conclusões e sugestões para trabalhos futuros.

## Capítulo 2

# Previsão de Consumo de Energia Elétrica

### 2.1 Visão geral do problema

O crescimento da capacidade instalada, a necessidade de melhores padrões de confiabilidade, além da otimização dos sistemas elétricos, foram alguns dos fatores que impulsionaram a implantação de redes interligadas de energia. Devido à complexidade dos sistemas, constituídos por diversas unidades geradoras e pontos de consumo, a operação dos mesmos é uma tarefa difícil e que exige um elevado esforço técnico. Dentre os fatores que representam as maiores dificuldades na solução ótima do problema de operação dos sistemas elétricos, pode-se citar a variação do consumo de energia elétrica (EE). Desta forma, a previsão do comportamento da carga, ou seja de consumo de energia elétrica, foi e continua sendo um fator crítico para o equacionamento deste problema (Overbye and Weber (2001)).

As informações sobre o consumo de energia, minutos, horas e anos à frente, para cargas em certos pontos (nodais) e a carga total do sistema elétrico, são cada vez mais indispensáveis e tornam-se cada vez mais úteis para as concessionárias. A título de exemplo, estas informações têm um grande valor em sistemas hidrotérmicos pois permitem com um certo grau de tempo disponível suprir as unidades geradoras térmicas para períodos de maior consumo. Este tipo de preparação melhora a qualidade do serviço e diminui a dependência de compra de energia elétrica externa (via linhas de

## CAPÍTULO 2. PREVISÃO DE CONSUMO DE ENERGIA ELÉTRICA<sup>12</sup>

transmissão com energia de outras operadoras). Este novo cenário do setor elétrico o qual apresenta as novas regulamentações do setor e à competição prevista para este processo também são exemplos do quão importante serão estas informações de previsão de carga (Makansi (2001)).

Para as empresas de energia elétrica, estas informações são fundamentais para o despacho de carga e para a otimização do sistema. Assim, a previsão de consumo a curto prazo (PCCP) permite, em princípio, otimizar a produção de energia, reduzir custos de geração, melhorar a coordenação da geração e a programação das trocas de energia dentro da empresa e/ou entre companhias, promovendo um aumento de confiabilidade e economia na operação do sistema (Teive et al. (2000); Sodré and Mota (2000); Oliveira et al. (2000)).

O acompanhamento da operação e o planejamento da expansão dos sistemas elétricos é de vital importância para garantir a oferta de energia com valores acessíveis e qualificados à sociedade. Teive e colegas (2000) destacam a importância da realização do planejamento energético. Este, por sua vez, é definido como um processo de tomada de decisões antecipadas, o que evidentemente requer um tempo para a sua implementação na prática. Para este caso, a previsão é no horizonte de longo prazo, o qual a projeção é baseada em cenários. Estes por sua vez são refeitos constantemente e oferecem uma visão macro dos valores de demanda futura. O planejamento bem executado e baseado em dados seguros e confiáveis torna-se um fator importante para a otimização do sistema elétrico do país. No entanto, o horizonte de longo prazo não será o foco do presente trabalho.

Dentre as formas de classificação de previsão de consumo, podem-se descrever quatro tipos (Reis (1999)):

1. Previsão a Curtíssimo Prazo - o horizonte de estudo é de poucos minutos a uma hora;
2. Previsão a Curto Prazo - nesta classificação a previsão é horária (de 1 hora até 1 semana);
3. Previsão a Médio Prazo - são previsões para o horizonte de um a vários meses;
4. Previsão a Longo Prazo - previsões realizadas em períodos superiores a um (1) ano.

## CAPÍTULO 2. PREVISÃO DE CONSUMO DE ENERGIA ELÉTRICA<sup>13</sup>

Deve-se levar em consideração que esta classificação de horizontes de previsão não é universal e nem única, mas foi definida de forma a obter uma melhor compreensão de sua aplicabilidade, e será doravante adotada ao longo deste trabalho.

Os critérios utilizados para a classificação das previsões são variados, e para tal consideram-se alguns fatores que podem influenciar no comportamento da carga. Dentre tais fatores, podem-se citar:

- influências climáticas tais como temperatura, velocidade do vento, umidade relativa do ar e intensidade luminosa;
- hora do dia, em um dia específico da semana;
- tipo de tarifação;
- estação do ano;
- eventos especiais (festas de fim de ano, grandes eventos esportivos e assim por diante);
- intervenções governamentais (por exemplo, forma de avaliação das tarifas, desregulamentação ou regulamentação do setor elétrico);
- fatores de carácter aleatório.

Considera-se que uma previsão a médio e longo prazo deve levar em consideração aspectos econômicos e decisões políticas (Reis (1999)). Para cada horizonte de classificação das previsões de carga, tem-se ações diferenciadas, de forma a se manter o gerenciamento e o controle do consumo de energia.

Na escala de tempo de segundos, quando ocorrem variações de carga que parecem aleatórias e de pequena magnitude, o objetivo é manter a estabilidade do sistema. Para tal o Controle Automático de Geração (CAG) utiliza informações para monitorar em tempo real a geração do sistema.

Na escala de minutos, quando as variações de carga mais intensas são sentidas, aciona-se o Despacho Econômico (DE) para garantir o atendimento da carga de forma econômica considerando a geração disponível.

Na escala de horas e dias, ocorrem variações de carga ainda maiores. Neste horizonte de tempo ações como entrada e/ou saída de unidades geradoras,

intercâmbios de demanda entre unidades e/ou companhias, bem como as programações de geração hidráulica, coordenação hidrotérmica, pré-despacho, dentre outras, são realizadas.

Finalmente, na escala de tempo de semanas, ações do tipo programação energética e manutenção dos equipamentos são realizadas de forma a garantir o atendimento da carga de maneira mais segura e econômica.

As previsões a curtíssimo e a curto prazo são utilizadas pelo Sistema de Gerenciamento de Energia (SGE) para manter uma estabilidade operacional das estações de energia, das suas unidades geradoras e para planejar as transações do mercado energético. Elas também são necessárias para os estudos de segurança, análise de contingência e gerenciamento de carga (conhecido como *despacho de carga*). Além disto, a previsão a curtíssimo prazo é requerida para o controle de frequência e as funções de despacho econômico. Há os reguladores automáticos de velocidade que, em conjunto com o CAG, ajustam a frequência, constantemente (Charytoniuk and Chen (2000); Reis and Silva (2000)).

As previsões a médio e longo prazo objetivam subsidiar decisões do quanto será necessário investir no sistema elétrico e informar dados para o planejamento da operação do mesmo. Nestas condições de previsão, o planejamento a médio e longo prazo torna-se um fator fundamental para a tomada de decisões futuras sobre o setor energético.

O foco do presente estudo são as Previsões de Consumo a Curto Prazo (PCCP) para os Sistemas de Energia Elétrica. A grandeza a ser prevista será a demanda <sup>1</sup> do sistema num determinado nodo. O estudo da previsão em outros horizontes merecem atenção particular e não fazem parte do escopo da presente dissertação, apesar de que algumas das ferramentas e procedimentos usados e decritos aqui possam ser úteis em tais problemas.

## 2.2 Importância da previsão de consumo

Problemas como racionamento, falta de energia, cortes de atendimento, e mais recentemente, mudança das tarifas de energia elétrica (devido a rees-

---

<sup>1</sup>Para este estudo, o conceito de demanda será definido como sendo o valor de potência elétrica em kVA ofertado pela subestação e não o valor de demanda conceitualmente utilizado pelas concessionárias como sendo o valor de energia consumido em um determinado tempo, normalmente 15 minutos.



truturação do setor energético) são alguns dos fatores que interferem na atual operação dos sistemas energéticos nacionais (Overbye and Weber (2001); Makansi (2001); Ilic et al. (2001)).

A operação real de um sistema de energia e a utilização econômica de suas fontes requer um bom planejamento acompanhado de uma adequada previsão de consumo num vasto período de tempo (desde poucos minutos até vários dias). Desta forma, pode-se notar que a previsão confiável da demanda de longo e médio prazos permitirá a adoção de ações de planejamento que resultarão em investimentos que devem ser implementados dentro de um horizonte de tempo factível. Previsões mal feitas ou equivocadas podem resultar em ações errôneas e de elevado custo. As ações de curto prazo são necessárias para as seguintes aplicações:

- programação da geração básica (principalmente térmicas);
- avaliar a segurança do sistema em qualquer instante;
- fornecer dados para as funções de programação e transporte de potência;
- fornecer informação adequada para o despachante.

Assim a PCCP deve fornecer dados para as funções de programação de forma a permitir a realização do transporte e/ou troca de energia entre as unidades e/ou companhias de forma mais econômica, consistente com requisitos de confiabilidade, restrições operacionais e limitações físicas, ambientais e de equipamentos. Desta forma, a PCCP pode atuar nos seguintes segmentos energéticos:

- Sistemas hidráulicos: a PCCP é necessária para a realização da função de despacho de águas (ou seja, a programação hidráulica). Também fundamental a este problema é a previsão do nível de reservatórios em função da precipitação pluviométrica;
- Sistemas térmicos: a PCCP é necessária para que o pré-despacho funcione com estratégias de horas de custo baixo para a entrada e saída de unidades de modo a sempre atender a carga prevista;
- Sistemas hidrotérmicos de energia elétrica: a PCCP é importante para a coordenação hidrotérmica realizar sua programação de operação entre várias unidades geradoras de forma a ter custos operacionais reduzidos.

Outro aspecto de grande importância da PCCP que merece destaque é na avaliação preditiva da segurança do sistema de potência. A PCCP permite uma análise *off line*, a qual estuda as condições futuras às quais o sistema torna-se vulnerável. Assim, o despachante, diante de uma informação prevista poderia em princípio, tomar decisões preventivas no sistema (por exemplo: deslocamento de geração, remanejamento de consumo, dentre outras).

Várias técnicas têm sido utilizadas para obtenção de previsões de carga. Dentre as técnicas utilizadas para previsão, podem-se citar os modelos de séries temporais que objetivam identificar padrões de comportamento nos dados disponíveis e extrapolá-los (Box & Jenkins, 1970). Tem-se também os modelos de regressão, que visam identificar funções de comportamento entre as variáveis dependentes (de resposta) e as independentes (preditoras), e utilizá-las para prever as variáveis dependentes através do uso dos valores futuros das independentes. Dentre estes modelos de regressão podemos citar aqueles que utilizam técnicas não lineares, tais como NARMAX (*nonlinear autoregressive moving average model with exogenous variables*) utilizados em (Gomes et al. (2000); Aguirre and Aguirre (1999, 2000)).

As técnicas baseadas em inteligência artificial (*IA*) são facilmente encontradas na modelagem e previsão de variáveis climáticas, hidrológicas (Valença and Ludermir (1998b,a)), energéticas (Charytoniuk and Chen (2000); Oliveira et al. (2000); Sodré and Mota (2000); Teive et al. (2000); Reis and Silva (2000)), bolsas de valores, séries econômicas, dentre outros.

Redes Neurais Artificiais apresentam-se como uma tecnologia atrativa pois sua característica marcante é a facilidade para se adaptar aos dados. Em (Valença and Ludermir (1998b)) destaca-se a utilização desta técnica no uso das previsões de vazões, visto que esta informação é de plena importância para a capacidade de geração futura do sistema. Em (Reis and Silva (2000); Reis (1999)) os autores utilizam RNA para realizar o pré-processamento de séries temporais para previsão de consumo a curto prazo. Já em (Teive et al. (2000)), esta técnica é utilizada como solução do problema de síntese de redes elétricas, na qual a rede determina o plano de expansão mínimo de custo (ou seja, a rede busca uma solução econômica e tecnicamente adequada para o sistema), buscando atender todas as cargas do sistema. Em geral, uma RNA possui várias características, tais como o aprendizado através de exemplos, flexibilidade, capacidade de generalização, tolerância a falhas e rápida implementação.

Os métodos baseados em lógica difusa (*fuzzy logic*) também merecem

destaque. A lógica difusa apresenta-se como uma ferramenta a qual tem se mostrado mais adequada para tratar determinadas incertezas sobre a informação do que a teoria das probabilidades. Assim, esta ferramenta trabalha com uma estrutura de alto nível, através das regras *if-then* (Braga et al. (1998); Caminhas et al. (1997)).

Em (Sodré and Mota (2000)), os autores utilizam técnicas baseadas em *neuro-fuzzy* (ANFIS) para estudos de Estabilidade de Tensão. O sistema do tipo *neuro-fuzzy* é uma combinação de redes neurais com a lógica difusa. Esta combinação procura integrar os benefícios de ambos os sistemas. Outras aplicabilidades desta técnica também podem ser encontradas em controle de processos, detecção de falhas (Caminhas et al. (1997)), sistemas de automação, entre outras.

Uma visão mais detalhada das técnicas utilizadas para a previsão de carga será dada no próximo capítulo.

## 2.3 Dificuldades na previsão de consumo

A variação de demanda (ou atendimento de carga) é um processo que não tem um modelo matemático prontamente definido que possa ser utilizado para realizar previsões de cargas futuras. Existem diversas variáveis e comportamentos inerentes a este processo, que tornam o problema da previsão de consumo bastante difícil.

As séries de carga do sistema de potência possuem um comportamento com alto índice de aleatoriedade, devido à mesma possuir muitos graus de liberdade, ou seja, é um sistema que sofre influências de muitas variáveis, tais como: fatores meteorológicos (temperatura, umidade do ar, por exemplo), temporais (horário de maior ou menor demanda), aleatórias e determinísticas. Esta aleatoriedade pode ser observada na Figura 2.1, em que foram traçados os gráficos de consumo de energia elétrica em 9 (nove) dias de Sábado dos meses de Novembro e Dezembro de 2000 de uma das subestações da Cia. Força e Luz Cataguases Leopoldina (MG), a subestação MCS (Mercês). O tempo de amostragem desses dados é de 1 hora, para um período de 24 horas. Pela análise desta figura verifica-se que não existem duas curvas iguais, apesar de todas corresponderem ao dia de Sábado. Se este estudo fosse ampliado para todos os dias da semana, aumentando o espaço amostral, seria verificado que as curvas não serão coincidentes, comprovando assim um certo grau de

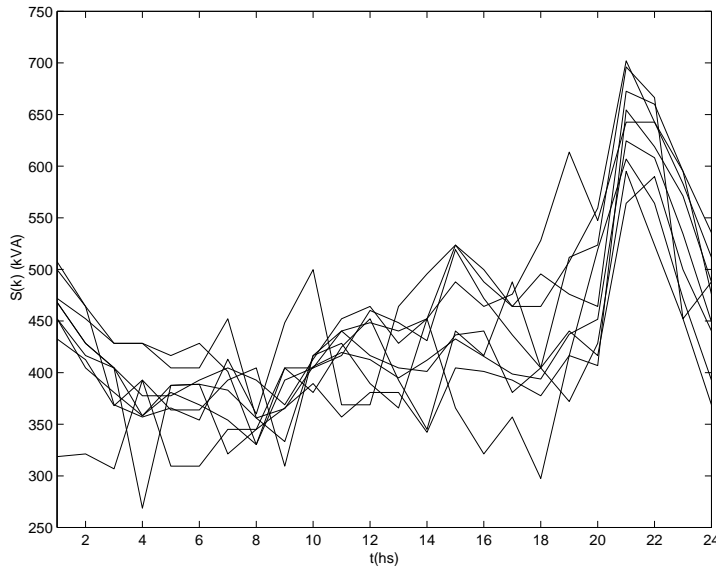


Figura 2.1: Curvas diárias de sábados.

aleatoriedade do sistema. Esta aleatoriedade, na realidade, reflete o grande número de graus de liberdade no sistema estudado. No entanto, não se pode afirmar que o sinal seja 100% aleatório.

Fatores do tipo calendários tais como as estações do ano, dias de semana e as considerações dos horários diários, (por exemplo: horários de pico de energia) (Jen-Lun and Terrence (1993)), influenciam o comportamento da demanda. Desta forma, a série de carga possui ciclos de tempos diversos, ou seja, ocorrem vários fenômenos em escala de tempo diferentes.

Pode-se dizer também que as séries de carga são não estacionárias. O conceito de estacionariedade está, de certa forma, ligado ao conceito de invariância temporal. Ou seja, ser invariante no tempo não quer dizer que o sistema é estático, mas sim que o mesmo tem uma dinâmica fixa que descreve a evolução temporal (Haykin and VanVeen (1999)). Desta forma, uma das maneiras de se avaliar se uma série temporal é ou não estacionária é utilizando a ferramenta de autocorrelação (Papoulis (1991)).

A autocorrelação é uma medida de dependência temporal entre o valor corrente e os valores passados da série original. Assim, a função de autocorrelação de uma série temporal estacionária não depende da janela de dados

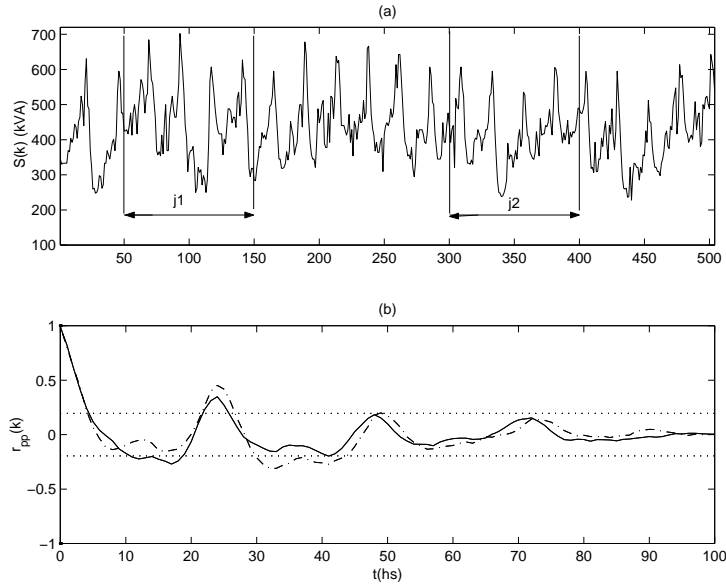


Figura 2.2: Não estacionariedade da série de carga  $S(k)$ : (a) Série de carga  $S(k)$  com as janelas  $j_1$  e  $j_2$ , (b) Autocorrelação de  $S(k)$  para as duas janelas.

usada no cálculo, desprezando-se o efeito de bordas. Na Figura 2.2 (a), tem-se duas janelas de dados com o mesmo número de observações. Considerando  $r_{p1}(k)$  e  $r_{p2}(k)$  as funções de autocorrelação das janelas  $j_1$  e  $j_2$ , respectivamente, pode-se dizer que  $S(k)$  é estacionária, se:

$$r_{p1}(k) = r_{p2}(k), \quad \forall k. \quad (2.1)$$

A Figura 2.2(b) mostra o autocorrelograma da série de carga  $S(k)$  para as duas janelas mostradas na Figura 2.2(a). Esta figura sugere que a série não é estacionária, pois a mesma não satisfaz a condição de estacionariedade (2.1). Técnicas como a diferenciação são utilizadas para melhorar a estacionariedade da série (Reis (1999); Reis and Silva (2000)).

A Figura 2.3 apresenta a série de carga  $S(k)$  e sua autocorrelação para uma janela de 504 horas (equivalente a 3 semanas). Observa-se que a série contém componentes periódicas (a autocorrelação apresenta picos a cada 24 e 168hs), comprovando assim a sazonalidade diária e semanal da série de cargas.

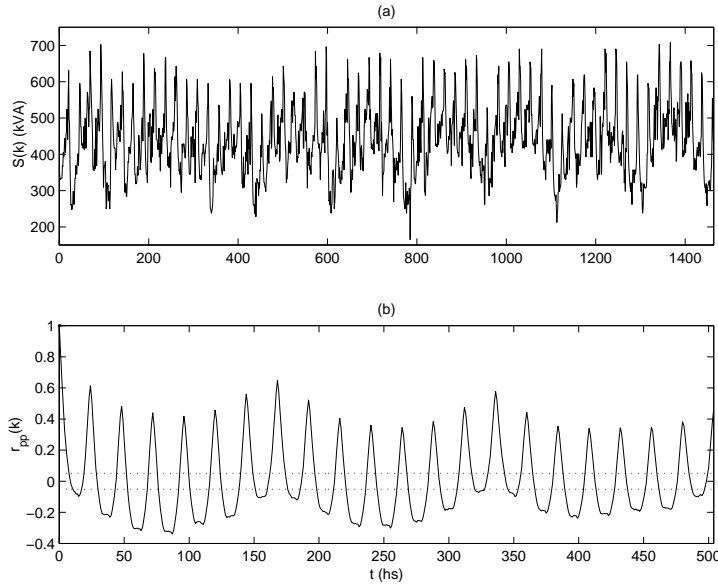


Figura 2.3: (a) Série de carga  $S(k)$  (serieP.txt), (b) Função de autocorrelação da série  $r_{pp}(k)$  para uma janela de 3 semanas ( $k=504$ hs). As distâncias entre os picos coincidem com o período de 24hs, como era de se esperar.

Todas estas características e comportamentos devem ser considerados ao se realizar uma previsão de consumo no sistema elétrico a partir de um registro histórico, ou seja, de uma série temporal. Entretanto torna-se uma tarefa difícil considerar todas estas variáveis ao mesmo tempo. Para fins de análise de PCCP, muitas variáveis são desprezíveis tais como as sazonalidades anuais e as variáveis climáticas.

O capítulo seguinte aborda as ferramentas de Redes Neurais Artificiais do tipo MLP (*Multilayer Perceptron*), as quais serão utilizadas, no contexto deste trabalho, para realizar a PCCP.

## Capítulo 3

# Representações Matemáticas

### 3.1 Introdução

Na interação de um indivíduo e um determinado sistema é necessário conhecer as variáveis envolvidas e como as mesmas são relacionadas, ou seja, é necessária a obtenção de um modelo para representar o sistema em questão (Ljung (1987)). A modelagem matemática é a área do conhecimento que estuda maneiras de desenvolver e implementar modelos matemáticos de sistemas reais. Em Aguirre (2000), o autor descreve técnicas de modelagem utilizadas em vários sistemas.

Neste capítulo serão abordadas algumas representações matemáticas utilizadas na literatura para realizar previsão de consumo. Dentre os modelos existentes destacamos os modelos de Box & Jenkins e os modelos que utilizam técnicas baseadas em inteligência artificial tais como redes neurais e lógica difusa (*NeoFuzzy*).

### 3.2 Modelos de Box & Jenkins

Considere o seguinte modelo geral:

$$A(q)y(k) = \frac{B(q)}{F(q)}u(k) + \frac{C(q)}{D(q)}\nu(k), \quad (3.1)$$

sendo  $q^{-1}$  o operador de atraso, de forma que  $y(k)q^{-1} = y(k-1)$ ,  $\nu(k)$  ruído branco,  $A(q)$ ,  $B(q)$ ,  $C(q)$ ,  $D(q)$  e  $F(q)$  os polinômios definidos a seguir:

$$\begin{aligned} A(q) &= 1 + a_1q^{-1} + \dots + a_{n_y}q^{-n_y}; \\ B(q) &= b_1q^{-1} + \dots + a_{n_u}q^{-n_u}; \\ C(q) &= 1 + c_1q^{-1} + \dots + c_{n_e}q^{-n_e}; \\ D(q) &= 1 + d_1q^{-1} + \dots + d_{n_d}q^{-n_d}; \\ F(q) &= 1 + f_1q^{-1} + \dots + f_{n_f}q^{-n_f}. \end{aligned} \quad (3.2)$$

O modelo de Box & Jenkins pode ser obtido a partir do modelo geral 3.1 tomando-se  $A(q) = 1$  e os demais polinômios arbitrários, resultando em:

$$y(k) = \frac{B(q)}{F(q)}u(k) + \frac{C(q)}{D(q)}\nu(k), \quad (3.3)$$

Neste modelo, as funções de transferência do sistema  $B(q)/F(q)$  e do ruído  $C(q)/D(q)$  não têm parâmetros comuns, ou seja, são independentemente parametrizadas. Desta forma, pode-se dizer que este método faz uma avaliação do perfil de comportamento dos dados existentes, e a partir destes, gera um modelo a ser usado, sendo necessário para tal o conhecimento e experiência do analista. As seguintes etapas descrevem como este método se aplica à série temporal para realizar a previsão da série:

- *determina* a estrutura dos modelos a partir da série;
- *estima* os parâmetros do modelo;
- *realiza testes estatísticos* para verificação do modelo;
- *faz previsão* de valores futuros da série.

Dentre vários estudos que utilizam a metodologia de Box & Jenkins, pode-se citar os modelos de regressão, com destaque para aqueles que utilizam técnicas não lineares, tais como NARMA (*nonlinear autoregressive with moving average*) e NARMAX polinomial (Leontaritis and Billings (1985a,b)). A seção seguinte faz uma melhor descrição destes modelos.



### 3.3 Modelos NARMA e NARMAX

Para sistemas lineares discretos utilizam-se modelos de equações de diferenças lineares, que relacionam variáveis de entrada e saída do sistema, pode-se citar os modelos ARX (*autoregressive model with exogenous inputs*), ARMAX (*autoregressive moving average model with exogenous inputs*), dentre outros. Analogamente, pode-se citar os modelos utilizados para descrever sistemas não lineares que são os modelos NARX (*nonlinear autoregressive model with exogenous inputs*) e NARMAX (*nonlinear autoregressive moving average model with exogenous inputs*).

Os modelos NARMAX discretos no tempo explicam o valor da saída  $y(k)$  em função dos valores prévios dos sinais de saída, entrada e ruído branco. Assim sendo, pode-se escrever este modelo da seguinte forma:

$$y(k) = F[y(k-1), \dots, y(k-n_y), u(k-1), \dots, u(k-n_u), e(k-1), \dots, e(k-n_e)] + e(k), \quad (3.4)$$

sendo  $n_y$ ,  $n_u$  e  $n_e$  os maiores atrasos em  $y$ ,  $u$  e  $e$ , respectivamente,  $m$  é o número de saídas,  $r$  é o número de entradas,  $e(k)$  é o ruído branco e  $F$  é qualquer função vetorial não linear. Um caso particular do modelo NARMAX é quando não se tem o modelo de ruído, sendo assim esse modelo passa a chamar-se de NARX e pode ser descrito da seguinte forma:

$$y(k) = F[y(k-1), \dots, y(k-n_y), u(k-d), \dots, u(k-n_u)], \quad (3.5)$$

sendo  $d$  o maior atraso em tempo morto.

Os modelos polinomiais NARMA são modelos não lineares que utilizam o valor corrente da série temporal  $y(k)$  em função dos seus valores nos períodos anteriores, dos valores correntes e anteriores de um ruído. Tais modelos podem ser descritos da seguinte forma:

$$y(k) = \theta_0 + \sum_{i_1=1}^{M_r} \theta_{i_1} x_{i_1}(k) + \sum_{i_2=1}^{M_r} \sum_{i_1=1}^{M_r} \theta_{i_1 i_2} x_{i_1}(k) x_{i_2}(k) + \dots + \sum_{i_1=1}^{M_r} \dots \sum_{i_i=i_{i-1}}^{M_r} \theta_{i_1 \dots i_i} x_{i_1}(k) \dots x_{i_i}(k) + e(k). \quad (3.6)$$

Em que  $x_1(k) = y(k-1)$ ,  $x_2(k) = y(k-2)$ , ...,  $x_{n_y+1}(k) = e(k-1)$ , ..., e  $x_n(k) = e(k-n_e)$ ,  $M_r = n_y + n_e$ ,  $y(k)$  são os dados,  $e(k)$  é o ruído e  $n_y$ ,  $n_e$  são

respectivamente os maiores atrasos de cada uma destas duas variáveis. As constantes  $\theta_i$  são os parâmetros que devem ser estimados a partir dos dados. Análises de determinismo (Aguirre and Aguirre (2000); Gomes et al. (2000)) e previsão de preços da carne bovina são alguns exemplos de aplicações deste tipo de modelo (Aguirre (2000)).

Neste trabalho, utilizou-se os modelos NARMA para realizar a análise de determinismo descrita no capítulo 5. Tais modelos são da seguinte forma:

$$\hat{y}(k) = \hat{\psi}^T(k-1)\hat{\theta}_y. \quad (3.7)$$

Para este modelo,  $\hat{\psi}(k-1)$  é o vetor de regressores composto de variáveis tomadas até o instante  $(k-1)$ , de forma que a predição atual de  $\hat{y}(k)$  é normalmente referida como predição livre, ou *free run prediction*. Uma descrição mais detalhada desta modelagem matemática pode ser encontrado em Gomes (2001) e Lima et al. (2002).

O número de possíveis candidatos nos modelos polinomiais descritos pela equação 3.6, normalmente, é muito grande, mesmo para valores moderados de  $M_r$  e  $l$ . Sendo assim, se todos os regressores forem incluídos no modelo, ter-se-á um modelo muito grande para se trabalhar. Além disto, a estimação direta dos parâmetros envolvendo todos os termos candidatos, geralmente resulta em mal condicionamento da matriz regressora, (Aguirre (2000)). Desta forma deve-se utilizar um método para determinar quais termos, dentre os candidatos, devem ser incluídos no modelo. Dentre os métodos existentes pode-se citar o uso da *Taxa de Redução de Erro* (ERR) (Chen et al. (1989)).

Maiores detalhes destes modelos não lineares e sua aplicabilidade podem ser encontrados em (Billings (1980); Leontaritis and Billings (1985a,b)) e (Aguirre and Aguirre (2000); Cassini (1999)).

### 3.4 Outras representações

As técnicas de Redes Neurais Artificiais (RNA) foram escolhidas nesta dissertação como ferramenta fundamental para a realização da previsão de consumo de energia elétrica. Estas técnicas serão descritas mais detalhadamente na seção 3.5. Entretanto, não se pode deixar de destacar o uso das técnicas baseadas na lógica difusa (NeoFuzzy), a qual também é tema de vários artigos

e estudos recentes na área de inteligência artificial (Caminhas et al. (1997); Mori and Kobayashi (1996)).

A técnica da lógica difusa iniciou-se em 1965, quando Lofti A. Zadeh (Zaded (1965)) introduziu o conceito formal de conjunto nebuloso. Os conjuntos nebulosos são a base dos sistemas nebulosos, diferindo em um certo grau, dos conjuntos clássicos, em que os limites são valores determinísticos, ou seja, um determinado valor por exemplo, pertence ou não a um determinado conjunto, não havendo um meio termo. Os conjuntos nebulosos têm como principal característica, o fato de não apresentarem limites determinísticos, ou seja, existe uma transição suave em relação à pertinência ou não de um elemento dentro de um conjunto nebuloso.

Estudos de análise de segurança da estabilidade de tensão, bem como pesquisas para detecção e diagnóstico de falhas utilizando lógica difusa, e na maioria das vezes, combinando estas com as redes neurais, são exemplos recentes da utilização desta técnica nos últimos anos (Sodré and Mota (2000); Caminhas et al. (2001)).

## 3.5 Redes Neurais Artificiais - RNA

### 3.5.1 Conceitos gerais

Para realizar a previsão de consumo de energia elétrica, foi escolhida a ferramenta de redes neurais artificiais (RNAs). Esta escolha se deve ao fato das RNAs serem um tipo de ferramenta com grande aplicabilidade e de grande interesse de estudos, conforme descrito na seção 3.5.3.

As RNAs representam uma forma de computação não algorítmica, baseada em processamento paralelo e distribuído, que em sua essência, procuram imitar a estrutura do cérebro humano (Braga et al. (1998)).

Estas redes são formadas por diversos nodos (neurônios), que são conectados entre si através de ligações que são normalmente unidirecionais, podendo formar diversas camadas, dependendo da complexidade do problema a ser resolvido. As conexões entre estes elementos são dotadas de pesos, que nada mais são que um tipo de ponderação que tentam refletir a importância destas conexões em relação à ativação de um determinado nodo.

Para a utilização das RNAs na solução de problemas complexos, como

por exemplo a previsão de consumo (existem  $n$  fatores que influenciam neste problema), faz-se necessário, primeiramente, uma fase de aprendizado da rede. Nesta fase ajustam-se os pesos de cada conexão da rede, basicamente através de exemplos de entradas/saídas, de forma que a rede responda aos estímulos do ambiente externo. Além da capacidade de aprendizado, outra característica das RNAs é a sua capacidade de generalização, que é a habilidade da mesma em apresentar soluções coerentes, a partir de novos dados de entrada, diferentes dos utilizados durante a fase de aprendizado.

De uma forma geral, as principais características das RNAs são:

- aprendizado através de exemplos;
- adaptabilidade;
- capacidade de generalização;
- tolerância a falhas;
- rápida implementação.

### 3.5.2 Breve histórico

A teoria das redes neurais artificiais teve origem na década de 40, quando o psiquiatra e neuroanatomista Warren McCulloch e o matemático Walter Pitts publicaram em 1943 o trabalho “A Logical Calculus of the Ideas Immament in Nervous Activity” (Mc Culloch and Pitts (1943)). Em 1949, Donald Hebb (Hebb (1949)) mostrou que através de variações dos pesos de entradas dos nodos era possível dotar as redes neurais da capacidade de aprendizado.

Um dos trabalhos mais importantes no desenvolvimento das RNAs foi apresentado por Frank Rosenblatt em 1958 (Rosenblatt (1958)), onde o autor criou a estrutura denominada *Perceptron*, composta por unidades treináveis. De maneira similar, em 1960 Widrow e Hoff (Widrow and Hoff (1960)) sugeriram a Regra Delta de aprendizado de RNAs, que tem como base o método do gradiente para a minimização do erro na saída de um neurônio com resposta linear (Braga et al. (1998)).

Porém, em 1969, Minsky e Papert (Minsky and Papert (1969)) criticaram severamente algumas tarefas que o Perceptron não era capaz de executar, uma vez que este só é capaz de resolver problemas linearmente separáveis. Isto

levou a um certo descrédito em relação às RNAs, tendo havido uma drástica redução de trabalhos nesta área na década de 70.

Foi quando John Hopfield, em 1982, publicou um artigo (Hopfield (1982)) analisando as redes simétricas usando mecânica estatística e analogias da física, marcando o renascimento das redes neurais. Apesar de ter sido este artigo responsável em parte pela retomada das pesquisas nesta área, foi a descrição do algoritmo de treinamento *Backpropagation*, em 1986, (Rumelhart et al. (1986)) que mostrou que a visão de Minsky e Paper era bastante pessimista, uma vez que as RNAs multicamadas eram capazes de resolver problemas mais complexos: problemas não linearmente separáveis (Braga et al. (1998)).

Assim, a partir da segunda metade da década de 80, houve uma sensível retomada dos trabalhos relativos às RNAs, motivados principalmente pelo grande avanço da microeletrônica, através do desenvolvimento de sistemas digitais com grande capacidade de processamento.

Autores como Fukushima (1975) e Kohonen (1974, 1982), entre vários outros, também contribuíram intensamente para o estabelecimento das redes neurais. Para um histórico mais abrangente ver (Haykin (1999)).

### 3.5.3 Aplicações de RNAs

Com o grande interesse pelo uso das técnicas de RNAs observado nos últimos anos, diversas aplicações foram desenvolvidas para as mais variadas finalidades. Dentre estas aplicações pode-se citar o uso da RNA para identificação de sistemas dinâmicos, previsão, controle de processos, classificação, entre outros (Khotanzad et al. (1997, 1998); Caminhas et al. (1997); Chowdhury and Aravena (1998)).

Pham e Xing (1995) escrevem exemplos de aplicações de RNAs para o uso nas áreas de identificação, previsão e controle, destacando-se o capítulo 5, onde é descrito o uso desta técnica para realizar previsão de bolsa de valores. Outros exemplos do uso de RNA para realizar previsão, em especial na área de sistema elétrico, podem ser encontrados em (Oliveira et al. (2000); Reis and Silva (2000); Reis (1999); Sodr e and Mota (2000); Teive et al. (2000); Valena and Ludermir (1998a,b); Charytoniuk and Chen (2000); Jen-Lun and Terrence (1993)).

Nas aplicações para processos de classificação e identificação de padrões, o uso de RNA pode ser encontrado em (Howell (1993); Chowdhury and Aravena (1998)), em que estes estudos utilizam em particular as Redes de Kohonen para classificar os padrões desejados.

Desta forma verifica-se que as técnicas RNA têm sido aplicadas em várias áreas de pesquisa e aplicação direta, aumentando assim a versatilidade e a confiabilidade de seus resultados.

### 3.5.4 Neurônios biológicos × neurônios artificiais

#### Neurônios biológicos

O neurônio biológico é formado basicamente pelo corpo da célula, pelos dendritos (ramificações de entrada) e pelo axônio (terminal de saída). Os dendritos têm a função de receber as informações (impulsos nervosos), provenientes de outros neurônios e conduzi-las até o corpo celular. Após o processamento dos sinais recebidos pelos neurônios, novos impulsos são gerados, passando para outros neurônios através dos axônios. O ponto de junção entre o axônio de um neurônio A e o dendrito de um outro neurônio B é chamado de *sinapse*. As sinapses funcionam como válvulas, e são capazes de controlar a transmissão de impulsos entre os nodos das redes neurais. A Figura 3.1 apresenta uma ilustração de um neurônio biológico.

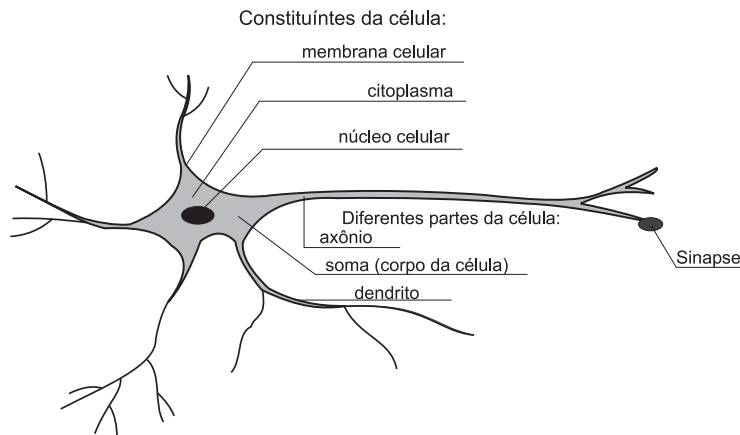


Figura 3.1: Neurônio Biológico.

Em relação ao cérebro humano, tem-se em torno de  $10^{11}$  neurônios, cada um podendo estar conectado a milhares de outros neurônios, o que facilita o estabelecimento de processamentos paralelos de informações e a formação de redes muito complexas. Assim, as RNAs tentam reproduzir as funções das redes biológicas, buscando implementar seu comportamento básico e sua dinâmica (Braga et al. (1998)).

### Neurônios Artificiais

O neurônio artificial é um modelo com  $n$  terminais de entrada (representando os dendritos) e  $m$  terminais de saída (representando os axônios). Para simular o comportamento das sinapses, que ocorre basicamente quando a soma dos impulsos que chegam aos neurônios através dos diversos dendritos ultrapassa um certo limiar, os terminais de entrada do neurônio artificial têm pesos a eles associados. Assim, estes pesos podem ser negativos ou positivos (sinapses inibitórias ou excitatórias). A Figura 3.2 apresenta um neurônio artificial típico.

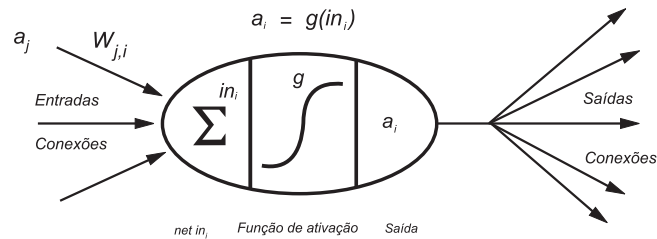


Figura 3.2: Representação matemática de um Neurônio Artificial (Reis (1999)).

O corpo do neurônio é representado por um mecanismo que faz a soma dos valores das entradas ( $a_j$ ), ponderados por seus pesos ( $W_{j,i}$ ), que, por meio de uma função de ativação ( $g$ ), determina se a saída ( $a_i$ ) deste neurônio artificial dispara ou não. Matematicamente, pode-se descrever esta computação como (Reis (1999)).

$$in_i = \sum_j W_{j,i} \cdot a_j = W_i \bullet a_i \quad (3.8)$$

$$a_i \leftarrow g(in_i) = g\left(\sum_j W_{j,i} \cdot a_j\right) \quad (3.9)$$

### Funções de ativação

Diferentes funções matemáticas podem ser obtidas para a função de ativação ( $g$ ). Três possíveis funções de ativação utilizadas em RNA são do tipo Passo, Degrau e Sigmoidal, a saber:

$$\text{Passo}(x) = \begin{cases} 1, & \text{se } x \geq t \\ 0, & \text{se } x < t \end{cases}$$

$$\text{Degrau}(x) = \begin{cases} +1, & \text{se } x \geq 0 \\ -1, & \text{se } x < 0 \end{cases}$$

$$\text{Sigmoidal}(x) = \frac{1}{1 + e^{-x}}.$$

A Figura 3.3 ilustra estes três tipos de funções de ativação.

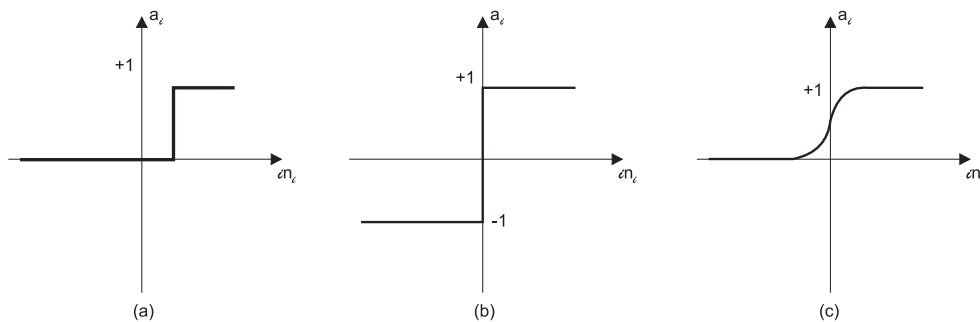


Figura 3.3: Três tipos de funções de ativação: (a) função passo, (b) função degrau e (c) função sigmoidal.

### 3.5.5 Tipo de redes neurais artificiais

Os tipos de redes neurais artificiais são definidos de acordo com a arquitetura da rede, que por sua vez restringe-se ao tipo de problema a ser tratado por ela.



As arquiteturas das RNAs podem ser definidas de acordo com o número de camadas (camada única ou múltiplas camadas), tipo de conexões (*feedforward* ou *feedback*) e ainda pela sua conectividade, ou seja, se são fracamente ou completamente conectadas. A Figura 3.4 mostra alguns tipos de arquiteturas das RNAs.

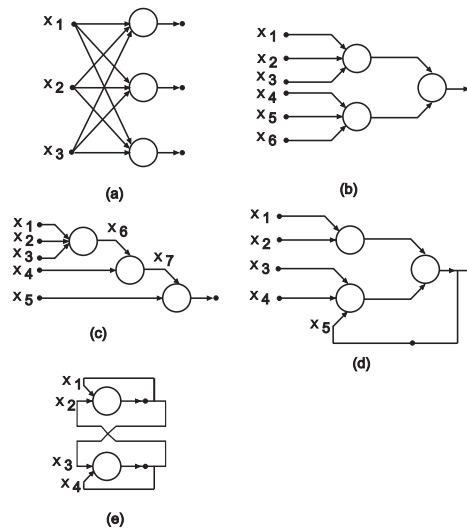


Figura 3.4: Exemplos de arquiteturas de RNAs.

Na próxima seção tem-se detalhadas as estruturas mapas de características de Kohonen e a estrutura do tipo MLP (Perceptron de Múltiplas Camadas). Estes dois tipos de RNA foram escolhidos para realizar a classificação e previsão da série de carga  $S(k)$ , respectivamente.

### 3.5.6 Mapas de características de Kohonen

As redes de Kohonen, ou mapa de características de Kohonen, são amplamente utilizadas para classificar padrões (Braga et al. (1998)). Estas redes são também conhecidas como redes mapas auto-organizáveis, SOM (do inglês, *Self-Organising Maps*) e foram desenvolvidas por Teuvo Kohonen na década de 80 (Kohonen (1989)).

Aplicações desta rede podem ser encontradas em uma variedade de problemas incluindo detecção de falhas (Caminhas et al. (1997)), mapa de características (Howell (1993)), classificação de falhas em sistemas de energia

(Chowdhury and Aravena (1998)), entre outros.

### Estrutura

A estrutura das redes de Kohonen é do tipo auto-organizável, ou seja, consiste em um arranjo de neurônios o qual induz relações de vizinhança (similaridade) entre os neurônios (Fritzke (1991)). Nesta rede, os nodos de saída se orientam geralmente em uma grade ou arranjo bidimensional, podendo algumas vezes ser unidimensional. A grade bidimensional tem a forma de uma superfície plana, onde os nodos de saída estão organizados em linhas e colunas. Esta organização depende de sua aplicabilidade. A Figura 3.5 ilustra uma rede de Kohonen típica.

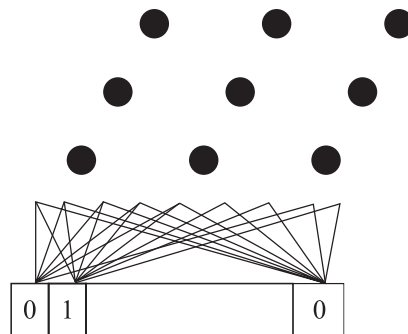


Figura 3.5: Rede de Kohonen típica para uma entrada binária (Braga et al. (1998)).

Cada nodo da rede recebe todas as entradas e funciona como um discriminador ou extrator de características. Para tal, esta rede trabalha com o chamado aprendizado por competição.

A fim de ilustrar como é feita esta relação de vizinhança e, conseqüentemente, a formação das classes ou padrões distintos, a Figura 3.6 mostra uma função conhecida em RNA como ‘chapéu mexicano’. Segundo esta função, cada neurônio influencia o estado de ativação de seus neurônios vizinhos de três formas possíveis:

- excitatória: se os vizinhos estão próximos a ele (exemplo: dentro da área de raio  $R_1$ );
- inibitória: se os vizinhos estão fora da área anterior, mas dentro de uma segunda área (exemplo: dentro da área de raio  $R_2$ , mas fora da área de raio  $R_1$ , sendo  $R_2 > R_1$ );
- levemente excitatória: se os vizinhos estão fora das áreas anteriores, mas dentro de uma terceira área (exemplo: dentro da área de raio  $R_3$ , mas fora das áreas de raios  $R_1$  e  $R_2$ , sendo  $R_3 > R_2 > R_1$ );

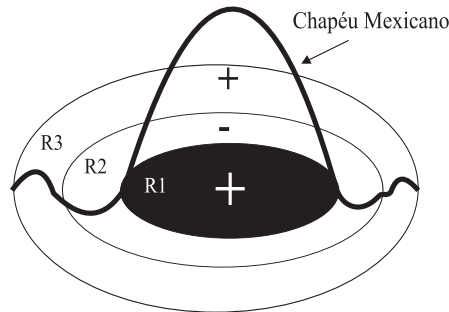


Figura 3.6: Tipos de estímulos produzidos de acordo com a função chapéu mexicano.

A rede funciona então da seguinte forma: quando um padrão de entrada  $\mathbf{p}$  é apresentado, a rede procura a unidade mais parecida com  $\mathbf{p}$ . Assim sendo, existe uma competição entre as unidades de saída para decidir qual delas será a vencedora e, conseqüentemente, terá sua saída ativada e seus pesos atualizados no treinamento. Durante o seu treinamento, a rede aumenta a semelhança do nodo escolhido e de seus vizinhos ao padrão  $\mathbf{p}$ . Desta forma, a rede constrói um mapa topológico em que os nodos que estão topologicamente próximos respondem de forma semelhante a padrões de entrada semelhantes. Com isto, a rede de Kohonen consegue um resultado semelhante àquele obtido pela utilização do ‘chapéu mexicano’.

A equação (3.10) mostra o estado de ativação de um nodo, sendo este a distância entre seu peso e o vetor de entrada. A função de ativação da rede de Kohonen é baseada na medida da distância Euclidiana.

$$g_j = \sum_{i=1}^n | a_j - W_{j,i} |, \quad (3.10)$$

sendo:  $W_{j,i} \in \mathbb{R}^2$  é o peso da conexão entre o elemento de entrada  $a_j$  e o nodo  $j$ .

### Treinamento

O treinamento da rede de Kohonen é do tipo competitivo e não supervisionado (tem-se apenas o vetor de entradas) e pode ser descrito como a seguir (Braga et al. (1998)):

1. inicializar pesos ( $W_{j,i}$ ) de forma aleatória, com zeros ou com valores previamente classificados (neste estudo, inicializou-se com dois padrões distintos: os valores carga leve ( $L(kVA)$ ) e consumo total ( $D(kVA)$ ) respectivamente das classes de domingos e dias úteis;
2. inicializar parâmetros: taxa de aprendizagem ( $\eta(t)$ ) e número de épocas para treinamento ( $N$ );
3. escolher os dados para treinamento ( $a_j$ ) (no caso presente serão 7 dias, equivalente a 168hs);
4. para  $m=1$  até  $N$  fazer
  - (a) para cada 24hs (1 dia) identifique os parâmetros de classificação  $a_j$  (no caso, serão os valores de carga leve ( $L(kVA)$ ) e consumo total ( $D(kVA)$ ) em 1 dia);
    - i. definir o nodo vencedor calculando a distância mínima entre  $a_j$  e os pesos  $W_{j,i}$ , utilizando a equação (3.10);
    - ii. para os nodos com menor distância entre  $a_j$  e  $W_{j,i}(t)$ , atualizar os pesos deste nodo e de seus vizinhos utilizando a seguinte equação:
 
$$W_{j,i}(t+1) = \begin{cases} W_{j,i}(t) + \eta(t)(a_j - W_{j,i}), & \text{se } i \in \Lambda(t) \\ W_{j,i}(t), & \text{caso contrario.} \end{cases} \quad (3.11)$$
    - iii. Se o número de épocas for múltiplo de  $N$ , então reduzir a taxa de aprendizado  $\eta(t)$  e a área de vizinhança.
5. até o mapa de características não mudar.

Para a equação (3.11), tem-se:  $W_{j,i}$  é o peso da conexão entre o elemento de entrada  $a_j$  e o nodo  $j$ ,  $\eta(t)$  é a taxa de aprendizado,  $\Lambda(t)$  é a vizinhança do nodo vencedor e  $t$  o intervalo de tempo.

Cada vez que um novo padrão de treinamento é apresentado à rede, os nodos competem entre si para ver qual gera a maior saída (aquele cuja distância entre o vetor peso e o parâmetro de classificação  $a_j$  é menor). Definido o nodo vencedor, inicia-se o processo de atualização dos pesos. Neste processo, apenas o nodo vencedor e seus vizinhos dentro de um certo raio ou vizinhança atualizam seus pesos. Durante o treinamento, a taxa de aprendizado e o raio da área de vizinhança são continuamente decrementados. A equação (3.11) mostra como os pesos são atualizados.

No capítulo seguinte, será descrito como a rede de Kohonen foi utilizada para classificar padrões semanais da série de carga da subestação de MCS (MG), e como esta classificação é utilizada para a realização da previsão de consumo deste sistema.

### 3.5.7 Redes MLP

As redes MLP (do inglês *Multilayer Perceptron*) são oriundas das redes do tipo *Perceptron* e *Adaline* (Rosenblatt (1958); Widrow and Hoff (1960)). Estas redes foram desenvolvidas com o intuito de resolver problemas mais complexos, ou seja, problemas não linearmente separáveis.

Segundo Cybenko, (1989), uma rede com uma camada intermediária permite a implementação de qualquer função contínua. Já a utilização de duas camadas intermediárias permite a aproximação de qualquer função (Cybenko (1988)). Porém, dependendo da distribuição de dados, a rede pode convergir durante o treinamento para um mínimo local, ocasionando a não implementação de uma determinada função.

Assim sendo, estas redes apresentam camadas intermediárias que servem como uma forma de simplificar o problema para a camada de saída, sendo que o número de camadas intermediárias está relacionado com a complexidade do problema.

O treinamento destas redes é uma tarefa relativamente complexa, já que só se sabe com certeza a magnitude do erro cometido pela camada de saída. O erro cometido pelas camadas escondidas é apenas uma estimativa, que tem

como base o erro da camada de saída.

Para treinar as redes multicamadas, foi proposto o algoritmo chamado *backpropagation* (Rumelhart et al. (1986)), tendo como base o gradiente descendente. Este algoritmo necessita que a função de ativação seja contínua, diferenciável e, preferencialmente, não decrescente. A função de ativação deve informar os erros cometidos pela rede para as camadas anteriores com a maior precisão possível. A solução encontrada foi o uso de funções de ativação do tipo sigmoidal (Braga et al. (1998)).

Nesta seção, serão descritas as redes do tipo MLP, sendo este tipo de RNA utilizado para a realização da previsão de consumo, foco do presente estudo.

A fim de ilustrar a grande variedade de aplicações das redes MLPs, pode-se citar alguns exemplos:

- previsão de ações na bolsa e controle de robôs (Pham and Xing (1995)).
- previsão de energia elétrica (Reis and Silva (2000); Reis (1999); Hippert et al. (2001); Liu et al. (1996)).
- reconhecimento de caracteres (Denker (1996)).
- segurança em transações com cartões de crédito (Reategui and Campbell (1994)).
- diagnóstico médico (Burke et al. (1995)).

A seguir será feita a descrição da arquitetura da MLP bem como o algoritmo utilizado para treinamento desta rede.

### Arquitetura das redes tipo MLPs

As arquiteturas das redes MLP são tipicamente organizadas em camadas. Cada camada tem uma matriz de pesos  $W_{j,i}$ , um vetor de *bias*  $b$  e um vetor de saída  $y$ . A Figura 3.7 apresenta a uma rede MLP típica. Trata-se de uma rede *forward* com uma camada de entrada, uma camada escondida e uma camada de saída. Para este tipo de arquitetura, as saídas de uma camada são as utilizadas como entradas das camadas seguintes.

A camada de entrada da rede recebe os dados de entrada, sendo que a camada escondida realiza a função de simplificar o trabalho para a camada de saída, que por sua vez produz a saída da rede (apresenta os resultados).

Um dos principais aspectos relacionados ao projeto de redes MLP diz respeito à função de ativação utilizada, que deve ser não-linear e diferenciável. Isto porque, uma rede com mais de uma camada cujos nodos utilizam funções de ativação lineares é equivalente a uma rede de uma única camada. A necessidade de que a função de ativação seja derivável, se deve à utilização do método do gradiente para a determinação dos ajustes dos pesos da rede (Braga et al. (1998)).

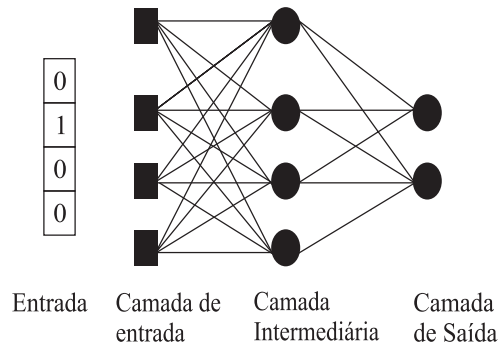


Figura 3.7: Estrutura típica de Rede MLP para uma entrada binária.

### Funcionalidade das MLPs

Todo o “conhecimento” adquirido pela RNA está armazenado nos pesos relativos às conexões (nodos) entre as suas camadas (Rich and Knight (1994)). Conseqüentemente, o processo de aprendizagem consiste em apresentar os dados que representam o problema a ser tratado pela rede e ajustar os pesos de suas conexões de maneira iterativa (via algoritmo de treinamento), com o objetivo de capturar o conhecimento contido nos dados.

Deve-se salientar que a escolha da topologia da rede, ou seja, o número de camadas escondidas, o número de nodos em cada camada é uma tarefa árdua, podendo ser encontrados na literatura vários estudos com a sugestão de métodos para a identificação mais apropriada da topologia da rede (Hippert et al. (2001); Liu et al. (1996); Braga et al. (1998)).

### Treinamento das MLP's: Backpropagation

O método de treinamento *backpropagation* foi, sem dúvida, o grande impulsionador para a retomada dos estudos relativos à RNAs (Rumelhart et al. (1986)). Ele é baseado na regra delta proposta por Widrow and Hoff (1960), sendo também chamada de regra delta generalizada.

O *backpropagation* é um algoritmo supervisionado, que utiliza a entrada e saída desejadas, para se obter a saída da rede através de um mecanismo de correção de erros (método do gradiente), o que corresponde aos ajustes dos pesos da rede. A idéia básica é que o erro de cada camada escondida seja determinado realimentando-se os erros a partir da camada de saída.

Mais especificamente, o treinamento de uma rede MLP com o *backpropagation* ocorre em duas fases, sendo que, em cada fase, a rede é percorrida num sentido. Estas fases são chamadas de fase *forward* e fase *backward*.

A fase *forward* é utilizada para definir a saída da rede para um dado padrão de entrada. A fase *backward* utiliza a saída desejada e a saída fornecida pela rede na fase *forward* para atualizar os pesos de suas conexões. A figura 3.8 ilustra estas etapas.

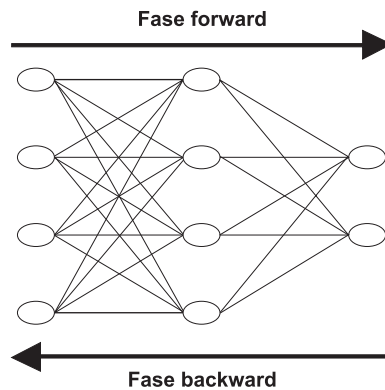


Figura 3.8: Fases do *backpropagation*.

O objetivo deste treinamento é a minimização do erro quadrático cometido pela rede. Assim, se considerarmos a soma dos erros quadráticos na saída da rede no instante  $n$ , tem-se:

$$\epsilon(n) = \frac{1}{2} \sum_{i=1}^q e_i^2(n), \quad (3.12)$$



sendo  $\epsilon$  a soma dos erros quadráticos,  $n$  o número da iteração,  $q$  número de saídas e  $e_i$  o erro da saída  $i$ .

A variação dos pesos da rede é função do gradiente descendente do erro em relação ao peso, como pode ser observado pela equação (3.13).

$$\Delta W_{j,i} \propto -\frac{\partial \epsilon}{\partial W_{j,i}}, \quad (3.13)$$

sendo  $W_{j,i}$  o peso entre o nodo  $j$  e o nodo  $i$ .

Com base nas equações (3.12) e (3.13), obtêm-se as equações para ajuste dos pesos das conexões da rede MLP (Braga et al. (1998)), conforme descrito a seguir:

$$W_{j,i}(t+1) = W_{j,i} + \eta \delta_i(t) a_j(t), \quad (3.14)$$

sendo que, para o nodo de saída tem-se:

$$\delta_i = (d_j - y_j) f' \left( \sum_{j=1}^n a_j W_{j,i} \right). \quad (3.15)$$

Para o nodo intermediário pode-se dizer que:

$$\delta_i = f' \left( \sum_{j=1}^n a_j W_{j,i} \right) \sum_l \partial_l W_{j,l}. \quad (3.16)$$

Pensando-se em termos de superfície de erro, cada combinação de pesos e limiares corresponde a um ponto sobre a superfície de erro. A Figura 3.9 ilustra graficamente uma superfície de erro do *backpropagation*. A solução ótima será o ponto mais baixo. Na verdade, procura-se a minimização do erro cometido pela rede através do ajuste dos pesos e limiares da mesma, de forma a se caminhar para os pontos mais baixos da superfície de erro, utilizando para isto o método do gradiente, conforme mencionado anteriormente. Porém, dependendo das condições iniciais, pode-se atingir mínimos locais <sup>1</sup>, o que é indesejável.

### Dificuldades de treinamento do backpropagation

Vários problemas poderão ocorrer durante esta fase de treinamento da MLP, sendo que os principais podem estar relacionados com os seguintes itens:

<sup>1</sup>Mínimos locais são pontos na superfície de erro que apresentam uma solução estável, embora não seja a saída correta (Braga et al. (1998)).

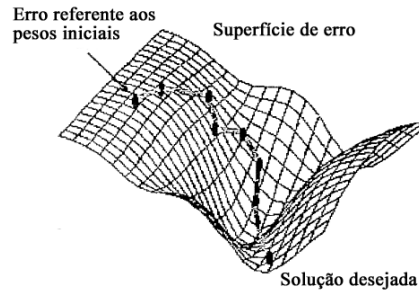


Figura 3.9: Exemplo de uma superfície de erro obtida a partir de uma MLP treinada pelo *backpropagation* (Braga et al. (1998)).

- seleção dos parâmetros de treinamento;
- critério de parada de treinamento;
- frequência de ajustes dos pesos;
- sobreparametrização (*overfitting*).<sup>2</sup>

Para a solução destes problemas, existem critérios e técnicas que ajudam a melhorar o desempenho do algoritmo, como por exemplo o uso de técnicas de *pruning*, para evitar o problema de *overfitting*, ou usar conhecimento a priori quando este estiver disponível (Amaral (2001); Reed (1993); Henrique et al. (2000)).

Como visto, o *backpropagation* apresenta uma série de deficiências, o que não lhe tira os méritos em relação ao alavancamento do estudo de RNAs a partir de sua aplicação, sendo ainda, um dos algoritmos mais utilizados para o treinamento de redes MLP.

O capítulo 6 descreve como as redes MLP foram utilizadas neste estudo para a realização da previsão de consumo de energia. Esta previsão foi realizada após uma preparação dos dados originais com o uso das técnicas de redes de Kohonen descrita no capítulo seguinte.

---

<sup>2</sup>Problema relacionado à piora da taxa de aprendizagem da rede, ou seja, ocorre uma redução da capacidade de generalização da rede.

## Capítulo 4

# Análise e Classificação de Padrões

### 4.1 Introdução

A série de carga (ou demanda de energia) do sistema elétrico sofre os efeitos de vários fatores que influenciam o seu comportamento. Dentre estes fatores podem-se destacar aqueles que são de origem temporal, tais como as estações do ano, periodicidade diária-semanal, variações sazonais, considerações dos horários diários (por exemplo: horários de pico de energia), dentre outros (Jen-Lun and Terrence (1993)).

Assim sendo, este capítulo tem como finalidade fazer uma análise do comportamento semanal dos dados de carga de uma das subestações da Cia. Força e Luz Cataguases-Leopoldina (MG), a subestação MERCES. Esta análise será feita com o objetivo de se obter uma série temporal residual da qual serão eliminados os ciclos diários.

Optou-se em fazer a redução da sazonalidade diária da série original  $S(k)$  devido ao fato que os dados finais obtidos nesta etapa serão utilizados para realizar a previsão de carga no horizonte de 1 hora até 12 horas. Assim, o procedimento para realizar esta previsão não será feito utilizando-se os dados originais da série de carga, mas sim será realizada uma previsão a partir dos dados de resíduo desta nova série. O estudo desta série de resíduos será feito utilizando ferramentas de *surrogate* (Theiler et al. (1992)), e será verificado

o grau de determinismo ou aleatoriedade desta parte do sinal. Esta análise de *surrogate* e previsão será objeto de outros capítulos e, portanto, não serão apresentados maiores detalhes deste processo.

O método do cálculo das médias semanais foi escolhido para se obter a redução do efeito de sazonalidade. Este método consiste no cálculo da média de dias afins da semana e a subtração desta média dos dias semanais correspondentes. Para tal, é necessário identificar os dias afins da semana como padrões de comportamento (ou seja, classificar as segundas-feiras, terças-feiras, e por aí em diante), calcula-se a média destes dias e por fim faz a subtração destes dados médios com os respectivos dias semanais. Esta classificação de padrões foi realizada através do uso das redes de Kohonen, conforme descrito a seguir.

## 4.2 Análise dos dados

Os dados disponíveis para estudo são dados de medição de demanda de energia coletados no período de 1 de Novembro a 31 de Dezembro de 2000. Foram observados os valores de tensão (kV) e corrente (A), com tempo de amostragem de 1 hora. A Figura 4.1, mostra esta série de dados para uma demanda de potência  $S(k) = V(k) \times I(k)$  (kVA).

Segundo Reis, (1999), o pré-processamento <sup>1</sup> dos dados de séries temporais é importante para que se tenha certeza do adequado fornecimento de padrões de entrada da rede para a realização da previsão de carga. Entretanto, este não é o foco de estudo deste trabalho, e, portanto, os padrões de entrada não sofrerão nenhuma análise mais profunda de pré-processamento. Contudo, foi realizada uma análise dos dados, de forma a reduzir sua sazonalidade.

Para redução da sazonalidade da série  $S(k)$  mostrada na Figura 4.1, utilizou-se o método do cálculo das médias diárias para uma semana. O valor desta nova série resíduo, chamada de  $\tilde{S}(k)$ , foi calculado conforme a

---

<sup>1</sup>Compreende-se como pré-processamento dos dados como sendo um tratamento dado a estes antes de se fazer a previsão, como por exemplo: elimina-se as medições erradas, normaliza-se e/ou pode-se utilizar técnicas de diferenciação nos dados, etc.

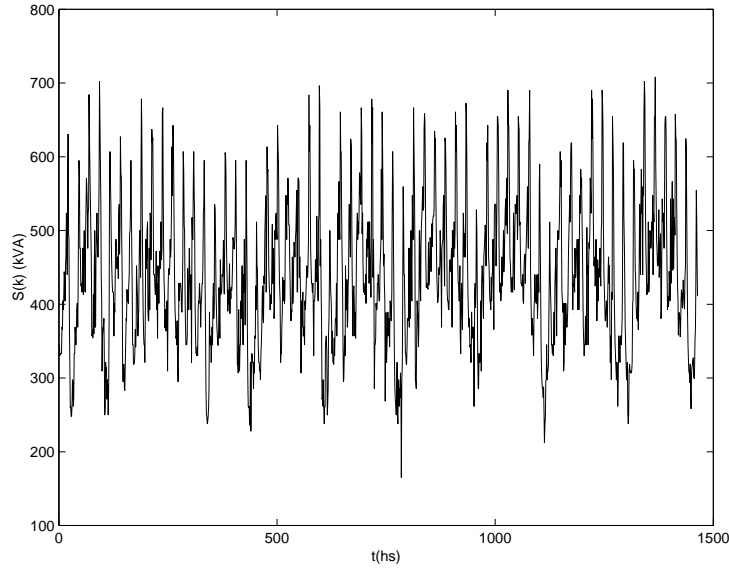


Figura 4.1: Série de Carga  $S(k)$  (kVA)  $\times$  tempo  $t$  (horas).

equação (4.1) a seguir:

$$\begin{aligned}\tilde{S}(k) &= S(k) - \bar{S}_j(k), \\ j &= 1, \dots, c \\ k &= 1, \dots, 24\end{aligned}\quad (4.1)$$

sendo:  $S(k)$  um trecho de 24hs da série original,  $\bar{S}_j(k)$  é um trecho de 24 horas (um dia) obtido como a média de cada dia afim da semana, hora a hora, da série  $S(k)$  e  $c$  é o número de classes em que os dias da semana podem ser classificados. Este valor ainda precisa ser encontrado. A equação (4.2) mostra como é feito o cálculo de  $\bar{S}_j(k)$ .

Supondo que  $S_j^i(k)$ ,  $k = 1 \dots 24$ ;  $j = 1 \dots c$ , seja o  $i$ -ésimo trecho de 24 horas da classe  $j$ , então

$$\begin{aligned}\bar{S}_j(k) &= \frac{\sum_{i=1}^I S_j^i(k)}{I}, \\ k &= 1, \dots, 24\end{aligned}\quad (4.2)$$

em que  $I$  é o número total de trechos de 24 horas na classe  $j$ .

Desta forma, para cada 24 valores, uma janela média,  $\bar{S}_j(k)$ , será subtraída da série original  $S(k)$ , conforme descrito pela equação (4.1). Um dos

problemas a ser investigado neste capítulo é a classificação automática da série em janelas médias  $\bar{S}_j(k)$ .

Com o objetivo de classificar estes dias semanais, foi analisado o comportamento da série diária durante os períodos de 1 semana. Sabe-se que a série de carga possui variações devido à sazonalidade, ou seja, pode-se identificar padrões de comportamento parecidos, os quais a série temporal parece obedecer durante os mesmos dias da semana (normalmente o comportamento de toda as segundas-feiras são muito parecidas, os domingos, sábados, etc.). Estes comportamentos são resultantes de eventos periódicos que ocorrem semanalmente<sup>2</sup>, como por exemplo o comportamento dos dias úteis (de segunda à sexta-feira).

Inicialmente, houve a tentativa de identificar os 7 padrões (os 7 dias da semana) na série de carga da Figura 4.1. Ou seja, tentou-se verificar se o algoritmo de classificação conseguiria distinguir as segundas-feiras das terças-feiras, e assim por diante. Para tal, tornou-se necessário identificar os parâmetros que melhor classificam a série para cada um destes padrões.

Foi realizada uma análise de agrupamento utilizando alguns parâmetros<sup>3</sup> para identificar aquele ou aqueles que melhor classificam a série. Dentre os parâmetros estudados, os escolhidos foram: o consumo de energia diário ( $D$ ) e o valor de baixa demanda de energia em 1 dia ( $L$ ). A Figura 4.2, mostra estes parâmetros para a série mostrada na Figura 4.1, enquanto a equação (4.3) descreve como estes valores foram calculados.

$$\begin{aligned}
 D(j) &= \frac{\sum_{i=1}^{24} P(i + (j - 1) \times 24)}{24}, \\
 L(j) &= \min \sum_{i=1}^{24} P(i + (j - 1) \times 24), \\
 &\quad j = 1, \dots, n \\
 &\quad i = 1, \dots, 24,
 \end{aligned} \tag{4.3}$$

sendo  $S(k)$  o vetor dos dados da série de carga e  $n$  o número de dias da semana desta série.

Os dados mostrados na Figura 4.2 serão utilizados para realizar a classificação dos dias da semana via redes de Kohonen.

<sup>2</sup>Estes eventos periódicos, além de semanais, podem ser anuais, diários, horários, etc.

<sup>3</sup>Os parâmetros para uma análise prévia foram: valor de pico de energia, horário de pico de energia, consumo em 24 horas, valor de baixa demanda e horário de baixa demanda.

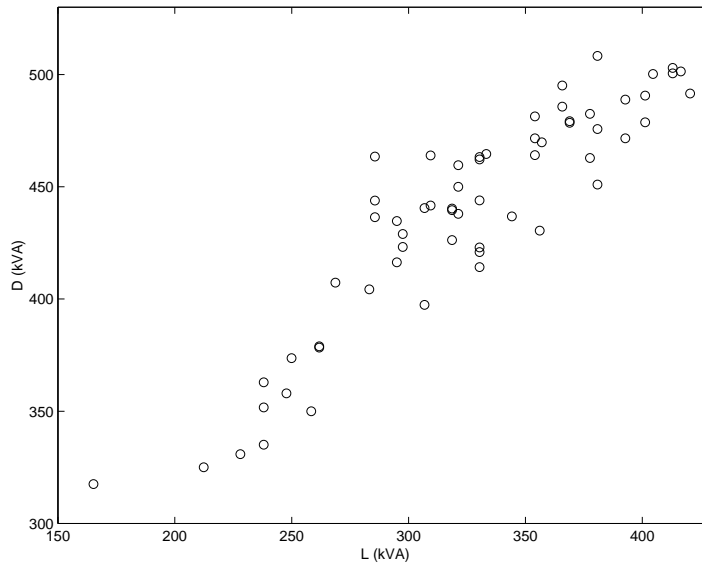


Figura 4.2: Valores de consumo,  $D$  (kVA), em função de carga leve,  $L$  (kVA), da série de carga  $S(k)$ .

Após uma análise prévia dos dados, verificou-se que os dias da semana (de segunda a sexta-feira) possuem um comportamento bastante similar entre si, indentificando assim como um padrão distinto, o qual foi designado neste estudo como *padrão dos dias úteis*. Outro padrão também foi facilmente identificado, o *padrão dos domingos*, sendo agrupado com os dias de feriados, devido também ao fato destes apresentarem comportamentos semelhantes entre si. Assim, conforme mostrado na seção 4.3, os padrões de segunda à sexta, e os padrões de domingos e feriados serão inclusos nos padrões de *dias úteis* e *domingos*, respectivamente. Os dias de sábados, foram identificados, a princípio, como sendo um outro padrão distinto, pois seu comportamento parece estar entre os dois padrões pré-definidos anteriormente.

Para realizar a classificação destes padrões foi utilizada a rede de Kohonen descrita na seção 3.5.6. A seção seguinte descreve como foi realizado este procedimento de classificação de padrões automático.

### 4.3 Classificação via redes de Kohonen

Na seção 3.5.6, foram descritas as redes de Kohonen (Braga et al. (1998)), ou mapa de características de Kohonen, os quais são amplamente utilizados para classificar padrões. Assim sendo, optou-se por utilizar esta técnica para classificar de forma automática os dias da semana.

Conforme descrito anteriormente, os padrões a serem encontrados na série de carga da Figura 4.1 foram agrupados, a princípio, em três tipos: dias úteis (de segunda a sexta-feira), sábados e domingos (incluído neste último padrão, os feriados). Para encontrar os padrões desejados, implementou-se um algoritmo utilizando a técnica de treinamento da rede de Kohonen conforme descrito na subseção 3.5.6.

Os dados de treinamento escolhidos foram os dados de  $D(\text{kVA})$  e  $L(\text{kVA})$  das 5 últimas semanas da série  $S(k)$  (Figura 4.1) e os dados de validação como sendo os dados de  $D(\text{kVA})$  e  $L(\text{kVA})$  do restante da série (aproximadamente 3 semanas). Estes parâmetros foram calculados conforme descrito na equação (4.3) e apresentados na Figura 4.2.

Inicialmente, foi escolhido um peso de entrada para cada padrão de classificação, ou seja, peso de dias úteis ( $W_u$ ), peso de sábados ( $W_s$ ) e peso de domingos ( $W_d$ ). Estes pesos não foram iniciados aleatoriamente, mas sim calculados com base nos dados de classificação ideal mostrado na Figura 4.3. Esta Figura mostra a saída desejada da série. Os dias de feriados (os quais neste caso são dias úteis) podem ser facilmente identificados nesta figura com um comportamento bastante semelhante aos dias de domingos, o que comprova a idéia de se agrupar este padrão ao padrão de domingos. Vale lembrar que o treinamento de Kohonen é do tipo não-supervisionado, e portanto a classificação mostrada na Figura 4.3 não é fornecida à rede, mas será utilizada apenas como forma ilustrativa (gráfica) de verificar se a rede está classificando corretamente estes padrões.

#### 4.3.1 Resultado da classificação

O resultado que a rede apresentou para a classificação desejada de três padrões (dias úteis, sábados e domingos (este inclui os feriados)) é o apresentado na Figura 4.4. Verifica-se que a rede não conseguiu distinguir o padrão dos sábados dos demais dias úteis. Obteve-se a classificação de apenas dois pa-



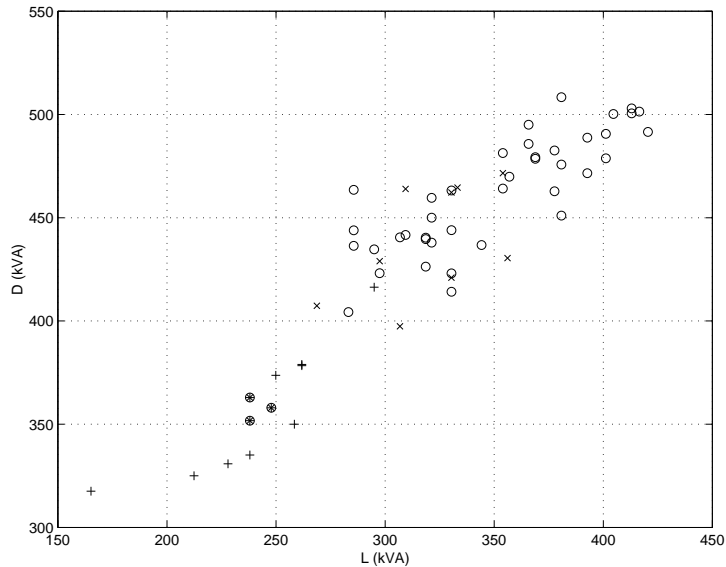


Figura 4.3: Saída desejada: (o) dias úteis, (x) sábados, (+) domingos e (\*) feriados.

drões distintos: dias úteis e sábados por um lado e domingos e feriados, por outro. Além disso, conforme análise prévia, a rede comprovou o comportamento dos dias de feriados identificados como padrões de domingo. Outras análises foram feitas utilizando outros parâmetros de classificação (horário de pico de energia, horário de baixa demanda e inclinação da curva de carga) e mesmo assim a rede não conseguiu classificar claramente os padrões de sábado e feriados. Desta forma, o estudo de classificação foi alterado para identificar apenas dois padrões distintos de comportamento: os dias úteis e os domingos.

A fim de se obter um melhor resultado (menor erro de classificação vide Figura 4.3) de classificação destes dois padrões (dias úteis e domingo), foi acrescentado mais um vetor de peso, ou seja, os pesos de entrada para cada padrão foram definidos como sendo:  $W_{u1}$  e  $W_{u2}$  para o padrão de dias úteis, e  $W_{d1}$  e  $W_{d2}$  para o padrão de domingos. A base de cálculo destes novos pesos foi a mesma realizada anteriormente, ou seja, visualizando os padrões pré-definidos na Figura 4.3.

Após estas alterações, a rede de Kohonen conseguiu identificar de forma precisa os padrões desejados (dias úteis e domingo), apresentando erros de 2%

para a classificação do padrão dias úteis e erro de 8,33% para a classificação do padrão de domingo. A Figura 4.4 mostra como foi definida esta classificação utilizando uma taxa de aprendizado ( $\eta = 0,4$ ) e número de épocas ( $N = 7$ ).

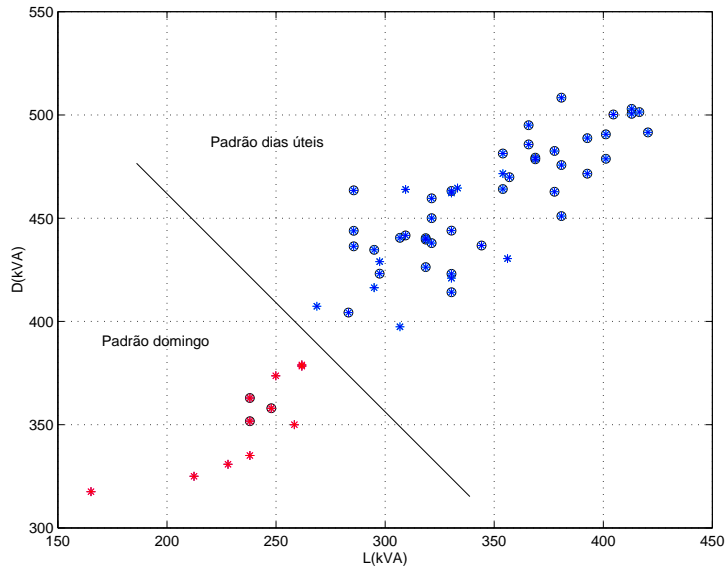


Figura 4.4: (o) dias úteis, ( $\times$ ) sábados, (+) domingos e (\*) feriados. Classificação via Redes de Kohonen: o padrão de dias úteis e o padrão de domingos, separados por uma reta para melhor visualização.

Verifica-se que este resultado (classificação de dois padrões) pode estar ligado ao banco de dados que é tipicamente de uma localidade com pequena população e que tem uma economia não industrializada (urbana com características rurais). Assim, era se esperar encontrar dois grupos: dias com trabalho (padrão de dias úteis) e dias sem trabalho (padrão de domingos). Desta forma em trabalhos posteriores pode-se estudar a influência das características de desenvolvimento sócio-econômico para classificação dos dias semanais.

A próxima seção descreve como os resultados da classificação foram utilizados para o cálculo da nova série  $\tilde{S}(k)$  sem o comportamento sazonal diário-semanal.

### 4.3.2 Redução da sazonalidade

Após realizada a classificação dos padrões de domingo e dias úteis na série de carga  $S(k)$ , esses dados foram utilizados para calcular a média de cada respectivo padrão ( $\bar{S}_1(k)$ , média de domingos e  $\bar{S}_2(k)$ , média dos dias úteis). A Figura 4.5 mostra estes ciclos médios obtidos. O cálculo para se obter  $\tilde{S}(k)$  foi realizado utilizando-se a equação (4.1) com estas respectivas médias de acordo com a classificação diária realizada pela rede de Kohonen.

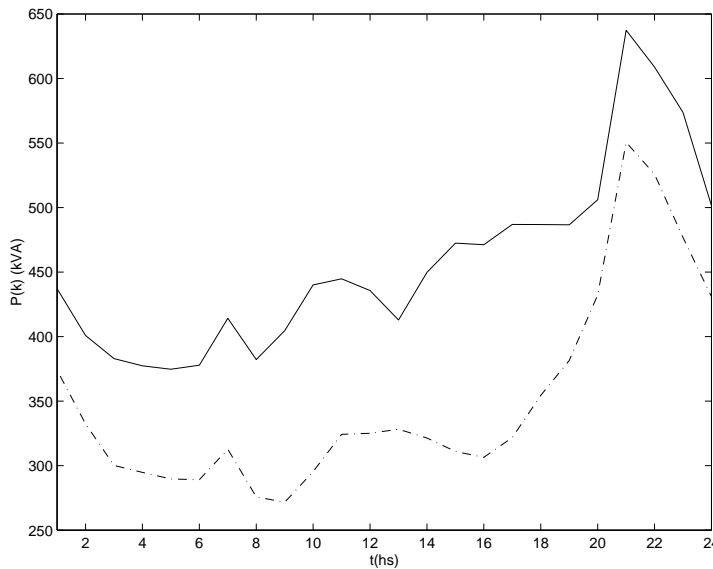


Figura 4.5: Ciclos médios diários de domingos (---)  $\bar{S}_1(k)$  e (—) dias úteis  $\bar{S}_2(k)$ .

A Figura 4.6 mostra de forma gráfica como foi obtida a série  $\tilde{S}(k)$  através da equação 4.1. Observa-se que ao retirar os ciclos médios  $\bar{S}(k)$  (Figura 4.6(b)) da série original  $S(k)$  (Figura 4.6(a)), obtêm-se uma nova série  $\tilde{S}(k)$  (Figura 4.6(c)) com bem menos sazonalidade quando comparado a Figura 4.6(a), série original  $S(k)$ . Para uma melhor compreensão deste resultado, pode-se verificar através do espectro de potência para estas duas séries, Figura 4.7. A componente de frequência principal em 1 dia ( $1/24 = 0,0417$ ) desapareceu no gráfico (b) com a nova série  $\tilde{S}(k)$ , confirmando a análise feita acima. Desta forma, optou-se em fazer uma rede para realizar a previsão de carga a partir de  $\tilde{S}(k)$ .

Foi feita uma análise de autocorrelação desta nova série  $\tilde{S}(k)$ , bem como

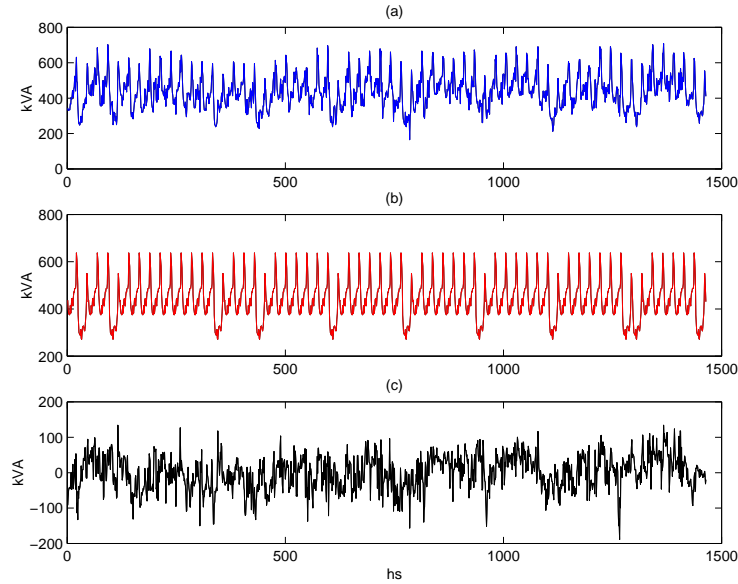


Figura 4.6: (a) Série original  $S(k)$ , (b) Série de ciclos médios  $\tilde{S}(k)$ , (c) Nova série  $\tilde{\tilde{S}}(k)$ .

a autocorrelação da série  $S(k)$ , afim de se obter mais informações de seus respectivos comportamentos.

A função de correlação entre dois sinais (ou o mesmo sinal, no caso de autocorrelação) é a medida da dependência temporal entre eles. Assim, pode-se descrever a função de autocorrelação ( $r(k)$ ) de um sinal  $u(k)$  como sendo:

$$r_{uu}(k) = \lim_{2M \rightarrow \infty} \frac{1}{2N + 1} \sum_{i=-N}^N u(i)u(i + k), \quad (4.4)$$

sendo  $k$  um número inteiro e, na prática,  $N$  deve ter um valor elevado (Papoulis (1991)).

A Figura 4.8(a) mostra a autocorrelação  $r(k)$  das duas séries  $S(k)$  e  $\tilde{S}(k)$  para uma janela de 168hs (equivalente a uma semana). Verifica-se que para a série  $S(k)$ , o sinal de autocorrelação é maior que o sinal de autocorrelação de  $\tilde{S}(k)$  até um período de aproximadamente 8hs. Além disto, o primeiro tem um decaimento mais forte em relação ao segundo, que por sua vez mantém-se mais constante por um maior período dentro deste intervalo de comparação. Deve-se destacar também, a forte presença da sazonalidade diária (picos de

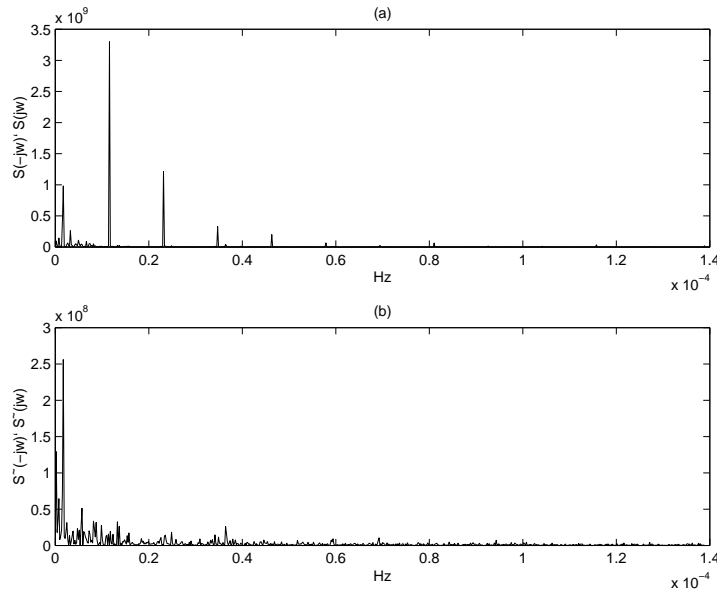


Figura 4.7: (a) Espectro de potência da Série original  $S(k)$ , (b) Espectro de potência da Nova série  $\tilde{S}(k)$ , com um tempo de amostragem  $f=1$  amostra/3600s = 0,00027Hz.

24hs) em  $S(k)$ .

A Figura 4.8(b) mostra os mesmos sinais de autocorrelação plotados em (a), mas agora numa janela de 24hs (1 dia). Neste intervalo de comparação observa-se que o sinal de autocorrelação de  $S(k)$  apresenta-se maior em relação ao sinal de  $\tilde{S}(k)$  até o período de aproximadamente 8hs e após 19hs. A sazonalidade diária de  $S(k)$  poderia justificar o aumento da autocorrelação deste sinal em relação a  $\tilde{S}(k)$ . Entretanto, em relação às primeiras horas do dia seria precipitado chegar-se a maiores conclusões neste momento.

#### 4.4 Análise dos Resultados

Esta seção objetivou analisar e classificar padrões de comportamento semanal dos dados coletados experimentalmente na subestação de MCS. Esta avaliação consistiu em uma análise de eliminação de sazonalidade deste sinal de forma a obter-se uma série de carga a qual será utilizada para realizar a

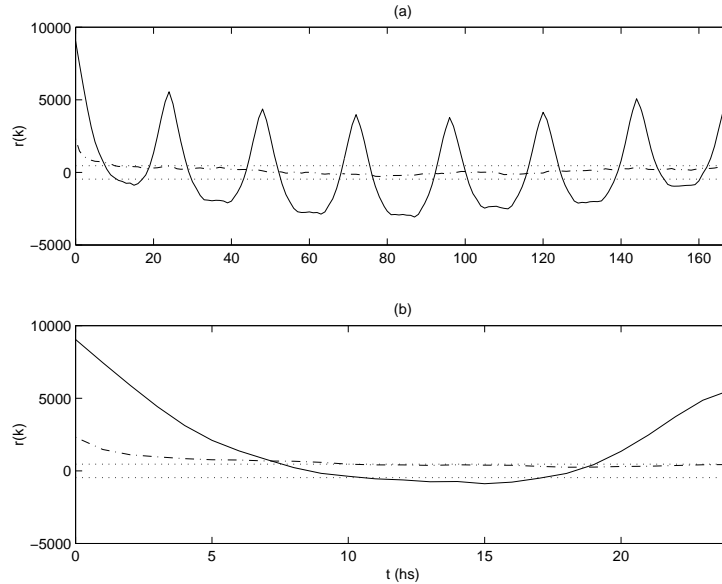


Figura 4.8: (a) Autocorrelação de  $S(k)$  (-) e  $\tilde{S}(k)$  (-.-) para uma janela de 1 semana (168hs), (b) Autocorrelação de  $S(k)$  (-) e  $\tilde{S}(k)$  (-.-) para uma janela de 1 dia (24hs)

previsão de consumo a curto prazo (horizonte de 1 hora até 1 semana) em um capítulo posterior.

Para obtenção destes novos dados, foi necessário realizar um estudo de classificação de padrões, o qual foi realizado via redes de Kohonen. Esta rede conseguiu identificar 2 padrões de comportamento distintos na série de carga ( $S(k)$ ): domingos (que inclui os feriados) e dias úteis (que inclui os sábados), utilizando-se para tal dois parâmetros de identificação: o valor de consumo em um dia (vetor  $D$ ) e o valor de baixa demanda de energia ( $L$ ). Realmente, pode-se verificar que o comportamento de consumo de energia para os dias de domingo apresenta diferenças em sua demanda de baixo atendimento (vetor  $L$ ), bem como seu consumo ao longo do dia (vetor  $D$ ) em relação aos dias úteis. A título de exemplo, o valor médio de baixo consumo nos dias de domingo é aproximadamente igual a 230kVA, enquanto para os dias úteis este valor é 350kVA. Assim, utilizando estes parâmetros de classificação, conseguiu-se identificar estes dois padrões de comportamento (dias úteis e domingo) da série com erros de 2% e 8,33%, respectivamente. Estes erros podem ser explicados em função do caráter fortemente estocástico da série e

também pelo fato dos dados disponíveis para este estudo não serem muitos.

A partir destes padrões de classificação foi realizado o cálculo da nova série  $\tilde{S}(k)$ . Através da análise do espectro de potência e da autocorrelação desta nova série, pôde-se constatar que seu comportamento parece ser menos sazonal, ou seja, o que se obteve foram dados com as flutuações da série original.

Desta forma, com os resultados obtidos neste trabalho, será avaliado em uma próxima etapa como estes novos dados se comportam em uma análise de determinismo e finalmente será realizada a previsão de carga a curto prazo a partir destes dados.

## Capítulo 5

# Análise de Determinismo

### 5.1 Introdução

Os dados de uma série de carga são normalmente muito complexos, devido a forte influência de fatores externos em seu comportamento. Desta forma não se pode partir da premissa de que a série temporal é determinística<sup>1</sup>. Assim sendo, considerou-se um fator muito importante quando se fala em previsão, a caracterização do sistema (no caso, da série) em questão, como possuindo traços determinísticos em sua dinâmica. Para esta verificação utilizou-se o teste de determinismo baseado em modelagem NARMA e análise de *surrogates* (Gomes et al. (2000)).

Esse procedimento consiste basicamente na geração das seqüências artificiais (*surrogates*) a partir de  $\tilde{S}(k)$ , sendo estas, por construção, aleatórias. Essas seqüências por sua vez, serão utilizadas para testar a hipótese nula. A hipótese nula é que os dados são aleatórios (Engbert (2000)).

Neste trabalho, o desenvolvimento desse teste foi realizado com o uso de um algoritmo de identificação desenvolvido pelo grupo de Modelagem, Análise e Controle de Sistemas Não Lineares, MACSIN, da Universidade Federal de Minas Gerais (UFMG) (Gomes (2001); Lima et al. (2002)). Estudos e aplicações utilizando análises de determinismo e *surrogates* podem ser encontrados em (Gomes et al. (2000); Liangyue et al. (1995); Engbert (2000); Gomes (2001)).

---

<sup>1</sup>Determinístico significa que a série não possui um comportamento aleatório.



## 5.2 Teste de determinismo

O programa utilizado para esse teste, realiza o mesmo a partir de dados originais, que no caso são os dados da série  $\tilde{S}(k)$ , que foi obtida no capítulo anterior via redes de Kohonen.

A idéia principal que norteia essa técnica é a seguinte: a partir da série temporal ( $\tilde{S}(k)$ ) é gerada uma seqüência aleatória artificial, denominada *surrogate*. Essa nova seqüência retém algumas propriedades dos dados originais, como por exemplo, a função de densidade de potência espectral. A *surrogate* é, por construção, aleatória e representa a “hipótese nula”, que pode ser rejeitada se os dados originais possuírem alguma característica determinística.

Neste trabalho 16 *surrogates* com fases embaralhadas (*phase randomized surrogate sets*) e 16 *surrogates* com ajuste de amplitude (*amplitude-adjusted surrogate sets*) são utilizadas, não se descartando, no entanto, a possibilidade de utilização de outros tipos de seqüências bem como a criação de uma especial para a série em questão.

Durante o desenvolvimento deste teste, foram utilizados modelos NARMA para realizar a predição das séries originais e das respectivas *surrogates*. O modelo usado foi o descrito pela equação (3.7), descrita no capítulo 3.

Para se caracterizar o aspecto determinístico da série temporal, um índice é usado após a modelagem não linear e predição livre, tanto da série original quanto das *surrogates*. Se os resultados obtidos para os dados reais (neste caso,  $\tilde{S}(k)$ ), através dos modelos NARMA, forem significativamente diferentes daqueles obtidos para as *surrogates*, que são seqüências aleatórias, é sinal de que há uma característica determinística nos dados reais. Por outro lado, se a predição sobre os dados reais assemelha-se àquela sobre as *surrogates*, é uma indicação que os dados originais não possuem uma assinatura determinística.

A fim de se comparar o desempenho dos modelos sobre os dados reais com o sobre as *surrogates*, o índice RMSE (*root mean square error*) foi utilizado:

$$RMSE = \frac{\sqrt{\sum_{k=1}^L (y(k) - \hat{y}(k))^2}}{\sqrt{\sum_{k=1}^L (y(k) - \bar{y}(k))^2}}, \quad (5.1)$$

em que  $\hat{y}(k)$  é o valor estimado de  $y(k)$ ,  $\bar{y}(k)$  é o valor médio do sinal medido sobre uma janela de dados estimados de comprimento  $L$ . Esse índice compara o desempenho da predição do modelo obtido com o preditor trivial, que é o

valor médio. Valores de  $RMSE$  menores que a unidade indicam um melhor desempenho quando comparados com o preditor trivial, e sugerem que a série em estudo é, em alguma medida, previsível. A partir desta comparação, pode-se rejeitar, ou não, a hipótese nula.

A seguir será descrito como esse procedimento é realizado.

### 5.2.1 O algoritmo

O algoritmo utilizado para o teste de determinismo consiste basicamente em três estágios, a saber:

1. geração das *surrogates* a partir de  $\tilde{S}(k)$ ;
2. modelagem NARMA;
3. predição e teste de hipóteses.

Neste algoritmo, a série  $\tilde{S}(k)$  foi dividida em 9 janelas, e cada uma delas possui 224 pontos, ou seja, 224hs. Assim sendo, o primeiro passo consiste na construção das *surrogates* de cada uma destas janelas.

O segundo passo consiste na modelagem NARMA. Nesta etapa o algoritmo subdivide cada uma das 9 janelas em duas partes, chamadas:

- janela de modelagem;
- janela de predição.

Sendo a janela de modelagem escolhida como os dados representativos de uma semana, ou seja, equivale a 168 pontos (7 dias  $\times$  24hs) e a janela de predição, por sua vez, possui 56 pontos restantes para completar a janela com 224 pontos.

Uma vez divididas as janelas, obtêm-se famílias de modelos NARMA para cada uma delas, ou seja, o algoritmo consegue gerar e identificar o melhor modelo NARMA utilizando cada uma das janelas de modelagem, 168 pontos. Os 56 pontos restantes (janela de predição) foram utilizados para o cálculo do índice RMSE, que será o parâmetro indicador ou não da presença

de determinismo nos dados analisados (Gomes (2001); Lima et al. (2002)). Para maiores detalhes desta técnica de modelagem utilizando esse algoritmo NARMA, pode-se consultar (Aguirre (2000)).

Após a modelagem e previsão, o algoritmo calcula o índice RMSE para os dados de  $\tilde{S}(k)$  e para as seqüências de *surrogates*, além de calcular também o RMSE médio e o desvio padrão para cada uma das seqüências artificiais.

O resultado final apresenta uma figura para cada janela, a qual pode-se comparar o desempenho do índice RMSE dos dados originais com o RMSE médio das *surrogates*, e assim concluir a existência ou não da característica determinística na janela em teste.

Para classificar cada uma das 9 janelas, utilizou-se os seguintes paradigmas (Gomes et al. (2000)):

1. Paradigma 1:

A hipótese nula de dinâmica estocástica é rejeitada se o desempenho do modelo em relação à predição da série  $\tilde{S}(k)$  for superior ao desempenho médio em relação à predição das *surrogates* que possuem modelos estáveis. Nesta classificação a hipótese nula é rejeitada e, portanto, a janela pode ser considerada com características determinísticas.

2. Paradigma 2:

Nesta classificação, o RMSE das *surrogates* e seu valor médio apresentam-se próximos da unidade. Neste caso, a hipótese nula não pode ser rejeitada e não é possível julgar a característica determinística para estes dados.

3. Paradigma 3:

Valores do índice *RMSE* oscilam entre valores maiores ou menores que a unidade, indicando uma presença de não-estacionariedade. Neste caso, não se pode obter conclusões relevantes sobre determinismo para um estes dados.

A seguir, tem-se os resultados obtidos utilizando os procedimentos listados nesta seção.

### 5.3 Resultados dos testes

As nove janelas de 224 pontos juntamente com as 16 *amplitude adjusted surrogates* e as 16 *phase randomized surrogates*, num total de 33 séries para cada janela, foram submetidas ao teste de determinismo conforme explicado na seção anterior. Alguns resultados estão apresentados nas Figuras 5.1 à 5.9, sendo o índice RMSE calculado para janelas deslizantes de 6hs e 12hs. É importante citar o fato de que os limites de confiança estabelecidos são limites de mais e menos um desvio padrão, que foi calculado tomando-se o desvio padrão médio ponto a ponto entre todas as *surrogates*.

Na Figura 5.1, o paradigma 1 pode ser observado para a janela 1 <sup>2</sup>. As predições sobre os dados originais são estatisticamente diferentes das predições sobre as *surrogates* até a janela de 7hs, o que sugere a rejeição da hipótese nula nos dados analisados. À medida que a janela utilizada para o cálculo do índice *RMSE* é deslizada ao longo de toda série nota-se que a qualidade da predição feita pelo modelo se deteriora e tende a se igualar à qualidade daquela feita sobre as *surrogates*.

As Figuras 5.2 e 5.3 visualizam respectivamente, as previsões realizadas para a janela 2 <sup>3</sup>, para uma janela deslizante de 6hs e 12hs. Observa-se que, nestas Figuras, o desempenho é oscilatório (melhor ou pior que 1 para diversos valores de RMSE), indicando um possível comportamento não-estacionário. Pode-se dizer que esta janela encontra-se entre os paradigmas 2 e 3, mas por conservadorismo optou-se por classificá-la como paradigma 3.

A Figura 5.4 evidencia a verificação do paradigma 2 para a janela 8 <sup>4</sup>. As predições sobre os dados originais não são qualitativamente diferentes daquelas feitas sobre as *surrogates*, conforme pode ser visto no gráfico (c) (RMSE). Com este resultado não se pode rejeitar a hipótese nula, e conseqüentemente não se pode definir a existência ou não de determinismo nos dados.

Por outro lado, a Figura 5.5 (janela 5 <sup>5</sup>), mostra um caso onde os valores do RMSE, estão plotados para uma janela deslizante de 6hs. Observa-se que somente nas primeiras previsões verifica-se erros grandes, sendo posteriormente as previsões enquadradas nos limites de confiança  $\pm\sigma$ . Desta forma, esta janela pode ser identificada como sendo do tipo paradigma 2.

---

<sup>2</sup>Janela 1: medições no período de 1h a 224hs.

<sup>3</sup>Janela 2: medições no período de 140hs a 363hs.

<sup>4</sup>Janela 8: medições no período de 980hs a 1203hs.

<sup>5</sup>Janela 5: medições no período de 560hs a 783hs.

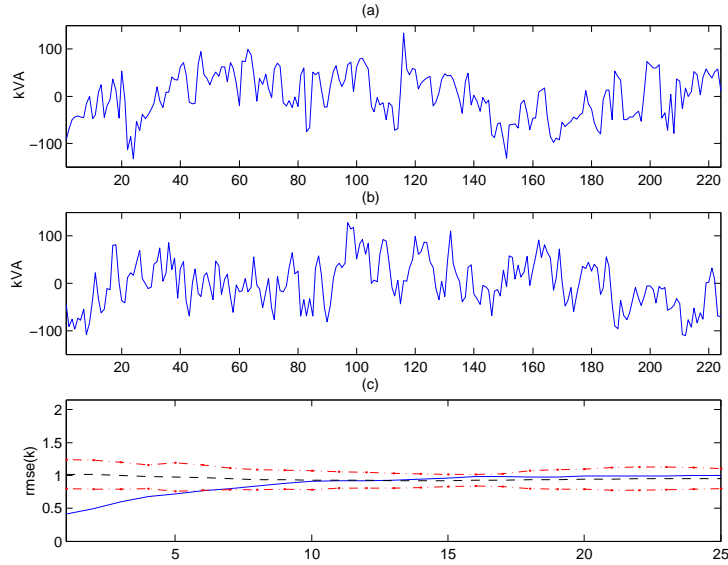


Figura 5.1: resultados do teste de determinismo para a janela 1, considerando uma janela deslizante de 12hs. (a) Série  $\tilde{S}(k)$ , (b) *surrogates* de  $\tilde{S}(k)$ , (c) (-) índice RMSE, (- -) RMSE médio, (-.-) limites de confiança  $\pm\sigma$ .

Visualizando esta mesma janela, mas agora com o índice RMSE em janelas de 12hs, na Figura 5.6. Verifica-se que ao aumentar a janela deslizante de 6hs para 12hs, filtra-se os erros cometidos nas primeiras previsões, fazendo com que se confirme a presença do paradigma 2 nesta janela. Ou seja, estes dados, podem apresentar características que permitem uma não rejeição à hipótese nula.

A Figura 5.7, mostra o desempenho do modelo para a janela 7<sup>6</sup>, para uma janela de deslizamento de 6hs. Verifica-se a forte tendência à rejeição de hipótese nula nestes dados até o horizonte de aproximadamente 6hs, classificando-os como sendo paradigma 1. No entanto, esta apresenta um comportamento diferente para a janela deslizante de 12hs. A Figura 5.8 mostra este resultado. Observa-se que agora o comportamento destes dados não mais permite uma rejeição à hipótese nula, sendo esta classificada no paradigma 2.

A janela 9<sup>7</sup>, conforme pode ser visto na Figura 5.9, pode ser classificada

<sup>6</sup>Janela 7: medições no período de 840hs a 1063hs.

<sup>7</sup>Janela 9: medições no período de 1120hs a 1343hs.

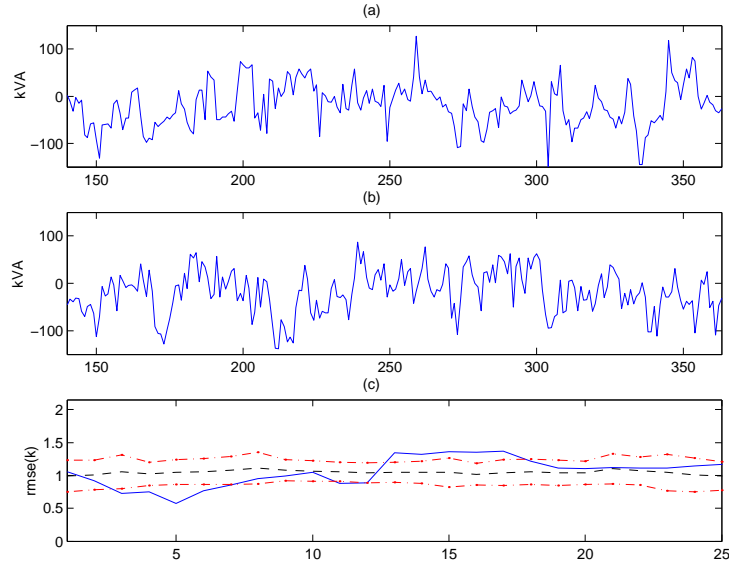


Figura 5.2: resultados do teste de determinismo para a janela 2, considerando uma janela deslizante de 6hs. (a) Série  $\tilde{S}(k)$ , (b) *surrogates* de  $\tilde{S}(k)$ , (c) (-) índice RMSE, (- -) RMSE médio, (-.-) limites de confiança  $\pm\sigma$ .

como paradigma 1, visto que os resultados apresentados, considerando uma janela deslizante de 12hs, mostram uma possível rejeição à hipótese nula.

Como uma síntese dos resultados obtidos nos testes de determinismo, a tabela 5.1 apresenta a classificação de cada uma das nove janelas, nas quais a série  $\tilde{S}(k)$  foi dividida.

## 5.4 Análise dos resultados

Neste capítulo, o objetivo é identificar se a série obtida no capítulo 4 ( $\tilde{S}(k)$ ), possui comportamento típico de sistemas com características determinísticas. Caso esta característica seja identificada, deve-se esperar que a partir desta série será possível a realização da previsão de consumo de energia com resultados satisfatórios. Caso este comportamento determinístico não seja evidente, a previsão da série de resíduos via redes neurais provavelmente não obterá resultados relevantes.

Desta forma, para realizar esta avaliação, utilizou-se um algoritmo de

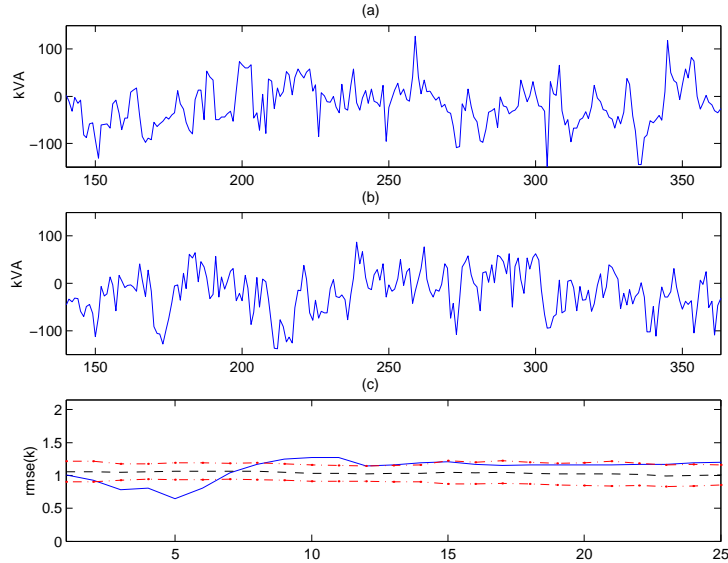


Figura 5.3: resultados do teste de determinismo para a janela 2, considerando uma janela deslizante de 12hs. (a) Série  $\tilde{S}(k)$ , (b) *surrogates* de  $\tilde{S}(k)$ , (c) (-) índice RMSE, (- -) RMSE médio, (-.-) limites de confiança  $\pm\sigma$ .

identificação NARMA. Este algoritmo possibilitou esta avaliação a partir de uma comparação entre as janelas de *surrogates* da série e a série por si só ( $\tilde{S}(k)$ ). Para isto, foi utilizado o índice RMSE. Através deste índice, pôde-se classificar as 9 janelas construídas pelo algoritmo, conforme os paradigmas descritos na seção 5.2.1 e os resultados mostrados na Tabela 5.1.

Assim sendo, com os resultados obtidos nesta tabela, pode-se concluir que a série  $\tilde{S}(k)$  possui característica determinística para algumas janelas, como por exemplo a janela 1 (até provavelmente 3 a 5hs segundo a Figura 5.1) e que este comportamento não é para todas as janelas da série. Ou seja, existem partes (janelas) as quais não se pode rejeitar a hipótese nula, e portanto, não é possível classificá-la, como por exemplo a janela 8 (ver Figura 5.4). Além disso janelas como a Figura 5.2 possuem um comportamento não-estacionário, e portanto não se pode concluir nada a respeito de determinismo nestes dados.

Portanto, para fins de previsão de consumo de energia utilizando a série  $\tilde{S}(k)$  via técnicas de redes neurais, deve-se esperar que os resultados apresentados pelas redes deverão ser compatíveis com a identificação de determinis-

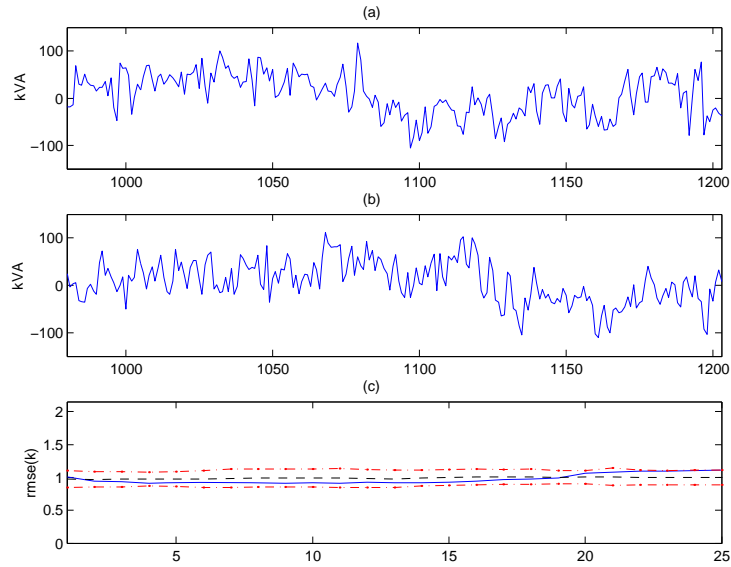


Figura 5.4: resultados de determinismo para a janela 8, considerando uma janela deslizante de 12hs. (a) Série  $\tilde{S}(k)$ , (b) *surrogates* de  $\tilde{S}(k)$ , (c) (-) índice RMSE, (- -) RMSE médio, (-.-) limites de confiança  $\pm\sigma$ .

mo nestas janelas.

A seguir, o capítulo 6 descreve a implementação das RNA's para fins de previsão de consumo de energia elétrica. Vários testes foram realizados e seus resultados serão apresentados e comentados.



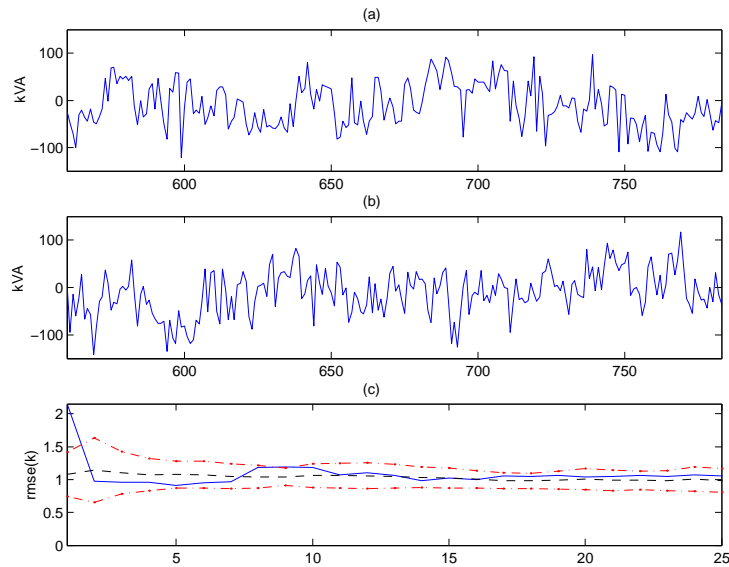


Figura 5.5: resultados de determinismo para a janela 5, considerando uma janela deslizante de 6hs. (a) Série  $\tilde{S}(k)$ , (b) *surrogates* de  $\tilde{S}(k)$ , (c) (-) índice RMSE, (- -) RMSE médio, (-.-) limites de confiança  $\pm\sigma$ .

Tabela 5.1: Resumo dos resultados obtidos nos testes de determinismo para as 9 janelas da série  $\tilde{S}(k)$ .

Janela n°	Paradigma 1		Paradigma 2		Paradigma 3	
	6hs	12hs	6hs	12hs	6hs	12hs
1	X	X				
2					X	X
3			X	X		
4			X	X		
5			X	X		
6	X	X				
7	X			X		
8			X	X		
9	X	X				

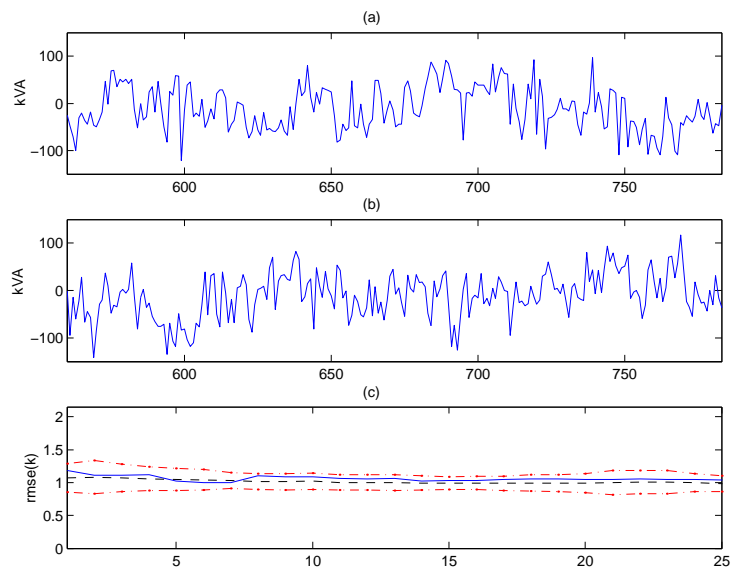


Figura 5.6: resultados de determinismo para a janela 5, considerando uma janela deslizante de 12hs. (a) Série  $\tilde{S}(k)$ , (b) *surrogates* de  $\tilde{S}(k)$ , (c) (-) índice RMSE, (- -) RMSE médio, (-.-) limites de confiança  $\pm\sigma$ .

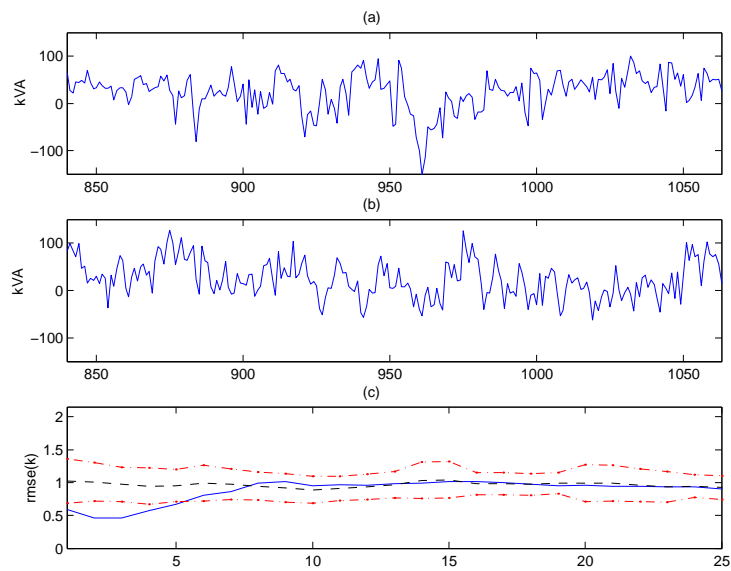


Figura 5.7: resultados de determinismo para a janela 7, considerando uma janela deslizante de 6hs. (a) Série  $\tilde{S}(k)$ , (b) *surrogates* de  $\tilde{S}(k)$ , (c) (-) índice RMSE, (- -) RMSE médio, (-.-) limites de confiança  $\pm\sigma$ .

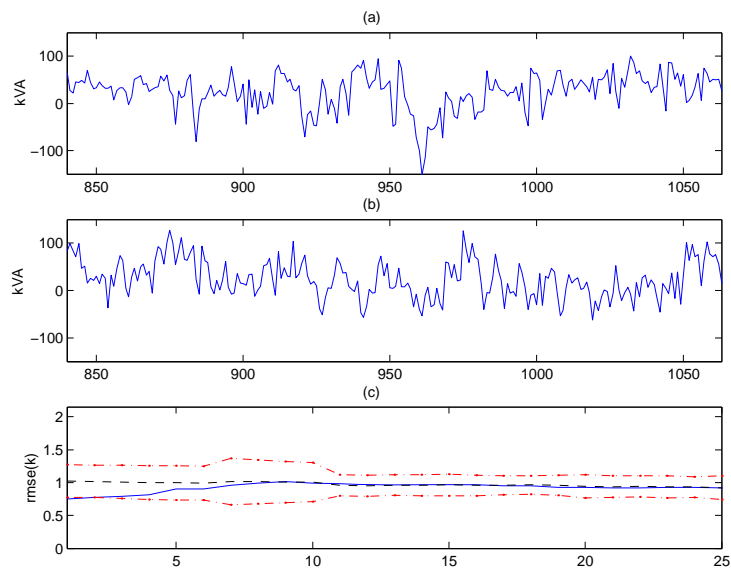


Figura 5.8: resultados de determinismo para a janela 7, considerando uma janela deslizante de 12hs. (a) Série  $\tilde{S}(k)$ , (b) *surrogates* de  $\tilde{S}(k)$ , (c) (-) índice RMSE, (- -) RMSE médio, (-.-) limites de confiança  $\pm\sigma$ .

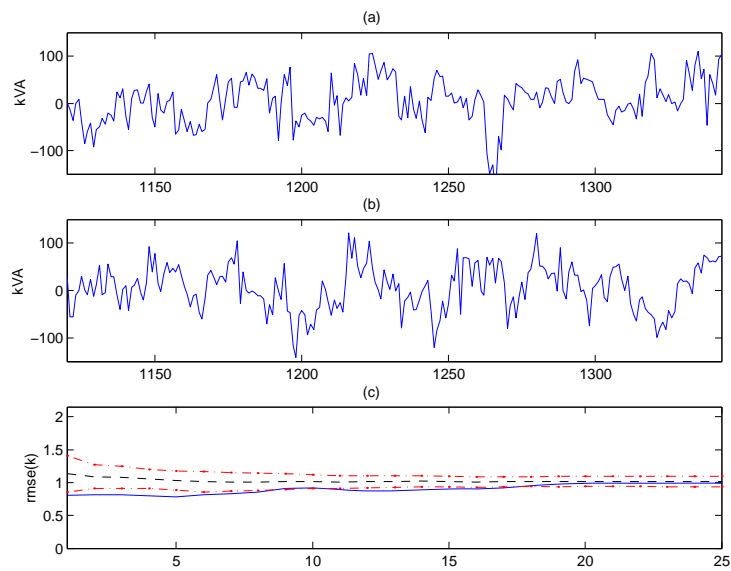


Figura 5.9: resultados de determinismo para a janela 9, considerando uma janela deslizante de 12hs. (a) Série  $\tilde{S}(k)$ , (b) *surrogates* de  $\tilde{S}(k)$ , (c) (-) índice RMSE, (- -) RMSE médio, (-.-) limites de confiança  $\pm\sigma$ .

## Capítulo 6

# Previsão de Curto Prazo

### 6.1 Introdução

Neste capítulo serão apresentadas e detalhadas as etapas de desenvolvimento do previsor implementado para este estudo. Como será visto, redes neurais artificiais do tipo MLP são uma importante ferramenta em tais previsores. Vários testes foram simulados em ambiente MATLAB e seus resultados serão apresentados e analisados. A implementação realizada foi dividida nos seguintes passos, que na seqüência serão descritos nas seções 6.2 a 6.8, respectivamente.

1. definição da série de carga;
2. seleção dos dados para treinamento e validação;
3. identificação das variáveis de entrada das RNAs;
4. definição da estrutura das RNAs;
5. treinamento;
6. testes e simulações;
7. validação.

## 6.2 Definição da série de carga

Os dados disponíveis para este estudo foram coletados a partir de um sistema de medição capaz de obter leituras horárias correspondente às medições de potência  $S(k) = V(k) \times I(k)$  (kVA), totalizando 1.464 medições.

Durante a aquisição de dados ocorreram algumas falhas de medição (não incomuns com sistemas reais), em alguns horários ou até mesmo durante todo um dia. Para completar a série, foi feito o cálculo da média dos valores de  $S(k)$ , para  $k = 1 \cdots 24$ , correspondente a cada dia da semana (por exemplo, fez-se a média das medições de todas as segundas-feiras às 16hs, às 17hs, 18hs e assim por diante). Desta forma completou-se a série  $S(k)$  com valores médios para os dias e horários sem medição.

Após as análises de classificação (capítulo 4), construiu-se uma nova série  $\tilde{S}(k)$ , chamada de série de resíduos, a qual possui menor sazonalidade que a primeira. Com os estudos de determinismo desta nova série, no capítulo 5, constatou-se que a mesma também possui carácter determinístico em algumas janelas, o que sugere que as redes obterão melhores resultados de previsão para estas janelas do que para as demais não-determinísticas.

Assim sendo, a partir da série  $\tilde{S}(k)$ , subdividida em 9 janelas conforme descrito no capítulo 5, realizou-se os testes de treinamento e validação via redes MLP, para cada uma destas janelas. É importante ressaltar que, nenhum procedimento de agrupamento de dados (tipo: dias úteis da semana, domingos, etc) foi utilizado para o treinamento da MLP. Desta forma, das 9 janelas geradas foram escolhidas as janelas 1, 5, 7 e 9 para a realização da previsão.

A escolha destas janelas foi feita com base na análise do capítulo 5 e os resultados apresentados na Tabela 5.1. Diante desta tabela, optou-se por escolher duas janelas com características determinísticas (janelas 1 e 9), uma janela com característica de não rejeição a hipótese nula, mas com tendência ao comportamento não estacionário (janela 5), e por fim optou-se pela janela 7, por apresentar uma não rejeição à hipótese nula, mas tendendo a uma característica determinística. Assim sendo, com estas janelas pode-se observar comportamentos diferentes, com escalas de tempo diferentes, e espera-se que a rede possa retornar informações relevantes sobre estes diferentes comportamentos.

Após os resultados obtidos para estas janelas, torna-se necessário com-

parar os resultados da previsão de resíduo,  $\tilde{S}(k)$ , obtidos pela rede, com os valores da série original  $S(k)$ . Para tal, utilizou-se a equação 4.1 de forma a se obter agora os valores previstos de  $S(k)$ . Portanto, a cada período de 24hs soma-se os valores das médias respectivas a cada período ( $\bar{S}_1(k)$  e  $\bar{S}_2(k)$ , respectivamente, as médias de domingos e as médias dos dias úteis, identificadas no capítulo 4). Desta forma, foi possível obter os valores previstos correspondentes na escala da série original  $S(k)$ .

### 6.3 Seleção dos dados para treinamento e validação

O banco de dados é composto por 1464hs (equivalente a 61 dias). Este banco foi dividido em 9 janelas de 224hs, as quais foram escolhidas para testes de previsão 4 destas. Para cada janela, optou-se por escolher os dados referentes a uma semana (168hs, ou seja 168 dados de medição), para fins de treinamento, e, para fins de validação, as restantes horas para completar uma janela de 224hs, ou seja 56hs.

### 6.4 Escolha das variáveis de entrada das RNA's

As RNA's obtidas realizam previsões de dados futuros a partir de dados passados. Assim sendo, deve-se escolher qual ou quais atrasos (*lags*) serão utilizados no processo de aprendizagem da RNA.

Alguns autores utilizam vários valores de atrasos para treinar a rede (Lamedica et al. (1996); Drezga and Rahman (1999)). No entanto, para este estudo tal valor foi definido como variando de 5 até 12hs.

### 6.5 Definição da estrutura das RNA's

Conforme descrito na seção 3.5.7, uma das maiores dificuldades em se definir a estrutura de uma rede é o dimensionamento mais adequado de sua topologia. Após definida a topologia inicial, são realizados vários refinamentos a fim de se obter uma topologia de rede que melhor represente o conjunto de dados de treinamento, mas que também consiga uma aproximação com boa capacidade de generalização.



Desta forma a estrutura da MLP ficou definida, inicialmente, conforme mostra a Figura 6.1.

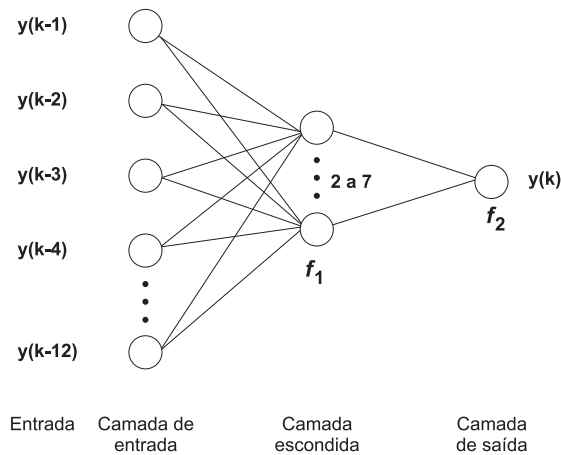


Figura 6.1: Estrutura básica definida para as MLP's

Existem na literatura vários procedimentos para se obter a melhor, ou a mais adequada topologia de rede para o caso de previsão de carga (Hippert et al. (2001)). Entretanto, como não existe um consenso entre eles, e os dados utilizados neste estudo são residuais (sem ciclos de sazonalidade), o que requer uma nova análise, optou-se por empregar a seguinte metodologia:

- [passo 1]: criar uma estrutura com, no mínimo, dois neurônios na camada oculta;
- [passo 2]: realizar os testes de treinamento e validação armazenando os erros correspondentes a cada teste;
- [passo 3]: aumentar o número de neurônios na camada oculta em 1 unidade, mantendo todos os demais parâmetros inalterados (ex: regressor, épocas, etc.);
- [passo 4]: voltar ao primeiro passo até que o número de neurônios na camada oculta seja igual a 7;
- [passo 5]: alterar a função de ativação da saída  $f_2$  para singular ou hiperbólica e voltar ao passo 1;

- [passo 6]: Avaliar os resultados e escolher um novo parâmetro a ser alterado até que se encontre um valor aceitável para o erro RMSE <sup>1</sup> de validação.

De posse dos resultados obtidos no passo 6, analisa-se os resultados (i.e., os erros de previsão). A partir deste ponto, escolhe-se um valor fixo para o número de neurônios da camada oculta e qual função de ativação para  $f_2$  obteve melhor resultado. Diante disto, escolhe-se o próximo parâmetro a ser alterado substituindo os passos 3, 4 e 5 pelos passos descritos pelos novos parâmetros. Estes parâmetros podem ser:

- atraso: aumenta-se o número de atrasos em uma unidade, iniciando em 5 até 12, mantendo os demais parâmetros fixos;
- épocas: aumenta-se o número de épocas em 100 unidades, iniciando em 100 até 600, mantendo os demais parâmetros fixos.

Após estes passos, a estrutura que apresentar melhor desempenho define a melhor estrutura de rede para aquela janela. Este desempenho foi avaliado utilizando alguns índices. Esse assunto será tratado na seção 6.8.

## 6.6 Treinamento

O algoritmo utilizado para realizar o treinamento das MLPs foi o *backpropagation*, sendo que foram utilizadas as funções *trainlm.m* e *simuff.m* do MATLAB para tal. Desta forma, apenas alguns parâmetros foram dados como entrada ao algoritmo, estes parâmetros foram:

- atraso: tempo de atraso para iniciar a rede. Conhecido como *lags*. A rede foi treinada para os atrasos variando de 5 a 12, em todas as janelas;
- nodos: número de nodos da camada escondida. Este número variou de 2 a 7;

---

<sup>1</sup>Este índice permite avaliar o desempenho da rede. A seção 6.8, descreve maiores detalhes sobre este índice.

- função de ativação:  $f_1$  e  $f_2$ . Sendo  $f_1$  a função de ativação da camada escondida, e  $f_2$  da camada de saída. Foram escolhidas respectivamente para  $f_1$  e  $f_2$ , as funções tangente hiperbólica e linear;
- épocas: quantas vezes a rede será treinada. Esses valores variaram de 100 a 600, de cem em cem;
- janela: escolha da janela a ser testada (variando de 1 a 9).

Os pesos utilizados para o treinamento do *backpropagation* foram iniciados aleatoriamente, sendo armazenados em um banco de dados para a reconstrução da rede.

## 6.7 Testes e simulações

Conforme descrito no início deste capítulo, foram escolhidas 4 janelas da série  $\tilde{S}(k)$  para serem testadas. Várias topologias de redes via MLP foram testadas, sendo que os dados de entrada, a arquitetura das redes, entre outros parâmetros, variaram significativamente. A seguir cada subseção descreverá os testes realizados e seus respectivos resultados.

### 6.7.1 Testes usando a janela 1

A janela 1 possui os dados de medição no período de 1h a 224hs, ou seja, contém as medições de 01 de novembro até 9 de dezembro de 2000 da série  $\tilde{S}(k)$ .

Conforme pôde ser visto no capítulo 5, esta janela apresenta características determinísticas. Espera-se, portanto, ser possível encontrar uma rede capaz de representar valores futuros para série de resíduos.

Após alguns testes variando os dados de entrada conforme descrito nas seções 6.5 e 6.6 (regressor, nodos, número de épocas, etc.), constatou-se que a melhor configuração de rede, apresentando resultados aceitáveis, é a arquitetura mostrada na Figura 6.2. Esta topologia possui as características da tabela 6.1:

O resultado deste teste está mostrado na Figura 6.3 (a) e (b).

Tabela 6.1: Características construtivas de rede para a janela 1.

atrasos	nodos	$f_1$	$f_2$	épocas
12	2	tang. hiperbólica	tang. hiperbólica	500

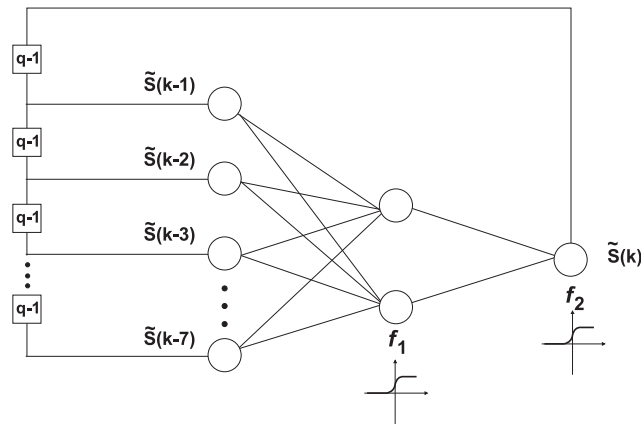


Figura 6.2: Arquitetura neural da janela 1.

Pode-se verificar através do gráfico 6.3 (a), que a rede perde sua capacidade de previsão após um período de aproximadamente 10hs, ou seja após este horizonte a rede não consegue mais acompanhar a série de resíduos. Tal resultado parece confirmar a classificação mostrada na tabela 5.1.

### 6.7.2 Testes usando a janela 5

A janela 5 possui os dados referentes à medição no período de 560hs a 783hs, ou seja, as medições do dia 24 de novembro até o dia 03 de dezembro de 2000 da série  $\tilde{S}(k)$ .

Devido à sua tendência a característica não-estacionária, observada no capítulo anterior, foi uma tarefa árdua encontrar uma arquitetura mais adequada para esta janela. Assim sendo, após vários testes, optou-se por uma configuração parecida com a da janela 1, variando apenas o número de neurônios na camada escondida de 2 para 3 unidades. Sua arquitetura é a mostrada na Figura 6.4 e suas características construtivas de rede estão na tabela 6.2.

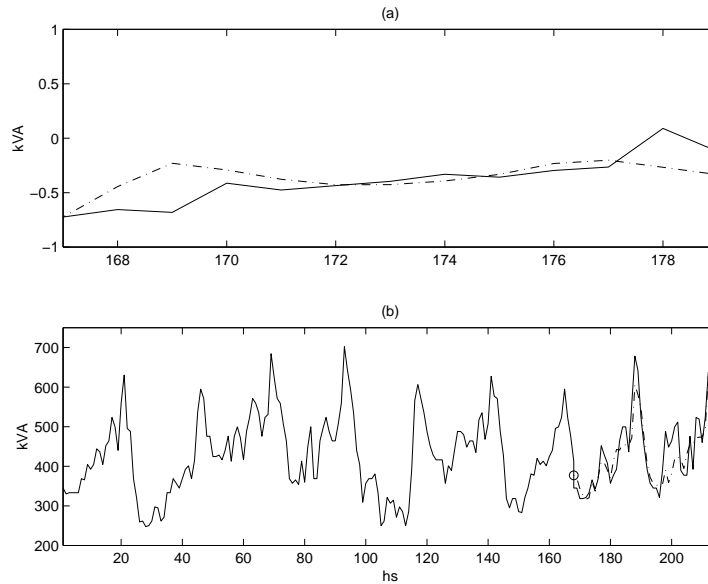


Figura 6.3: (a) (-.-) resultados dos testes de previsão para a janela 1, (-) a partir dos dados da série  $\tilde{S}(k)$ ; (b) (-.-) previsão de resíduos mais ciclo médio ( $\tilde{S}(k) + \bar{S}(k)$ ), (-) comparados com os dados originais da série  $S(k)$ , (o) iniciando no ponto marcado.

Tabela 6.2: Características construtivas de rede para a janela 5.

atrasos	nodos	$f_1$	$f_2$	épocas
12	3	tang. hiperbólica	tang. hiperbólica	500

O resultado do teste usando esta janela está mostrado na Figura 6.5 (a). Sendo que o resultado final, ou seja, a previsão comparada com os dados reais,  $S(k)$ , está mostrado na mesma Figura 6.5 no gráfico (b).

Concluindo esta subseção, através do gráfico 6.5 (a) pode-se visualmente verificar que a rede não consegue mapear adequadamente os dados futuros para esta janela no curto prazo. Entretanto, aparentemente, leva-se a pensar que a rede aprendeu esta “tendência de médio prazo” apesar de não ser capaz de aprender o comportamento determinístico de curto prazo, pois o mesmo não deve existir.

Com os resultados apresentados no gráfico 6.5 (b), verifica-se que houve

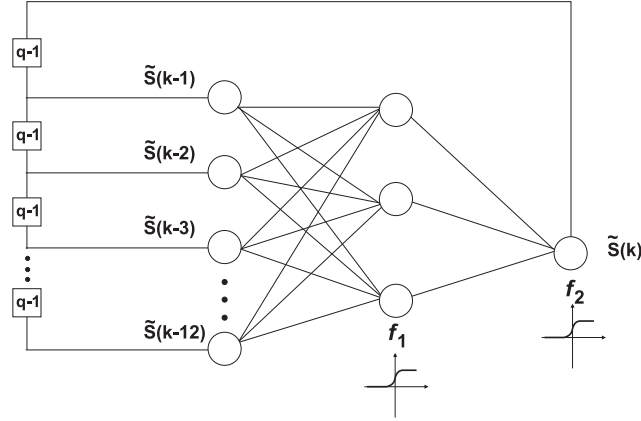


Figura 6.4: Arquitetura neural da janela 5.

um deslocamento da previsão em relação a série original. Isto se deve ao fato do comportamento da média ( $\bar{S}(k)$ ) que, somada ao sinal de resíduo, influencia nos resultados finais. Uma análise mais detalhada sobre estes resultados será apresentada na seção 6.8.

### 6.7.3 Testes usando a janela 7

A janela 7 possui as medições referentes ao período de 840hs a 1063hs, ou seja, contém os dados de medição de 06 de dezembro até 14 dezembro de 2000 da série  $\tilde{S}(k)$ .

Vários testes foram realizados, sendo que dentre estes testes, foi escolhida a configuração da rede mostrada da Figura 6.6, a qual possui as seguintes características:

Tabela 6.3: Características construtivas de rede para a janela 7.

atrasos	nodos	$f_1$	$f_2$	épocas
7	2	tang. hiperbólica	tang. hiperbólica	500

Os resultados para esta configuração estão mostrados nos gráficos (a) e (b) da Figuras 6.7.

Observa-se que os resultados obtidos para esta janela, vide gráfico 6.7 (a), foram considerados bons diante de um horizonte relativamente curto,

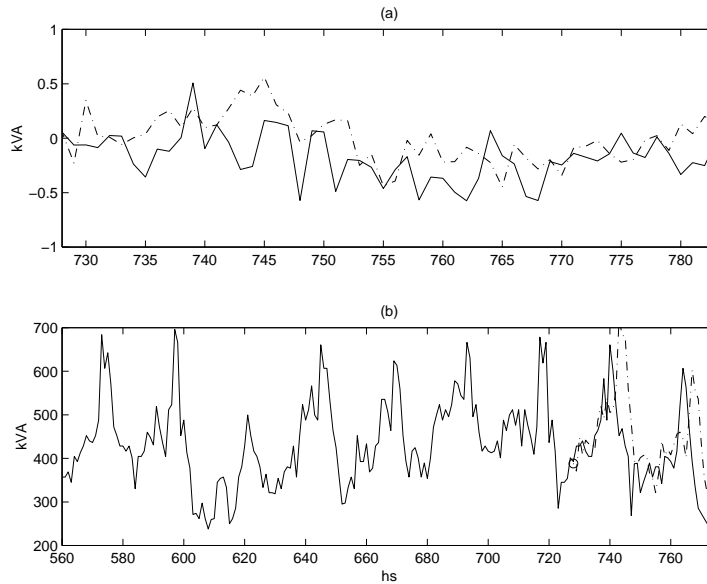


Figura 6.5: (a) (-.-) resultados dos testes de previsão para a janela 5, (-) a partir dos dados da série  $\tilde{S}(k)$ ; (b) (-.-) previsão de resíduos mais ciclo médio ( $\tilde{S}(k) + \bar{S}(k)$ ), (-) comparados com os dados originais da série  $S(k)$ , (o) iniciando no ponto marcado.

aproximadamente 5hs. A partir desse instante, a rede perde sua capacidade de generalização e não consegue mais prever os dados futuros. Entretanto o resultado final, visto pelo gráfico (b) desta mesma figura, mostra que, apesar de não ser possível rejeitar a hipótese nula para estes dados (vide capítulo 5), a previsão realizada pela MLP apresenta bons resultados até o horizonte de aproximadamente 5hs, caracterizando assim uma possível presença de determinismo para este horizonte de previsão.

#### 6.7.4 Testes usando a janela 9

Finalmente, chega-se à última janela escolhida para os testes de previsão. A janela 9 possui dados referentes ao período de 1120hs a 1343hs, ou seja, medições no período de 17 de dezembro até dia 26 de dezembro de 2000, da série  $\tilde{S}(k)$ . Neste conjunto de dados, é importante citar a existência de um feriado nacional: 25 de dezembro.

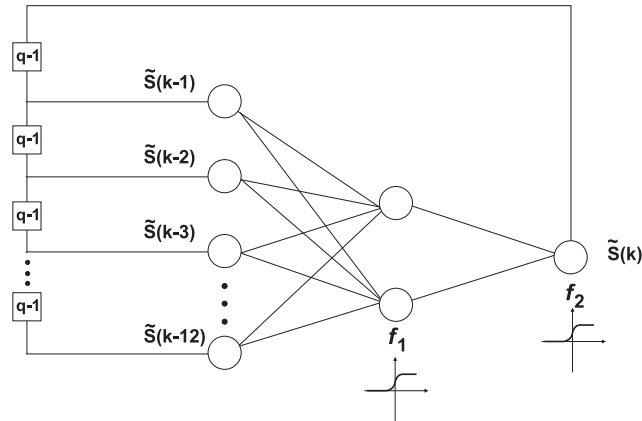


Figura 6.6: Arquitetura neural da janela 7.

Para esta janela, conseguiu-se encontrar uma rede com as características listadas na tabela 6.4, sendo sua arquitetura neural muito parecida com a da janela 1 (vide Figura 6.2), com alteração apenas para o trazo de 12hs para 5hs e a função de ativação da saída de tangente hiperbólica para função linear.

Tabela 6.4: Características construtivas de rede para a janela 9.

atrasos	nodos	$f_1$	$f_2$	épocas
5	2	tang. hiperbólica	função linear	500

A janela 9 apresentou resultados bastante semelhantes à janela 7, conforme pode ser visto na Figura 6.8.

O resultado mostrado no gráfico 6.8 (a), mostra um comportamento parecido com a janela 7 até o horizonte de aproximadamente 5hs. Já o gráfico 6.8 (b), permite a visualização dos resultados finais desta previsão.

Uma análise mais detalhada de todos os resultados das 4 janelas mostradas nas Figuras 6.3 a 6.8, será feita na seção seguinte com o uso dos índices de desempenho dos previsores expostos para cada uma destas janelas.



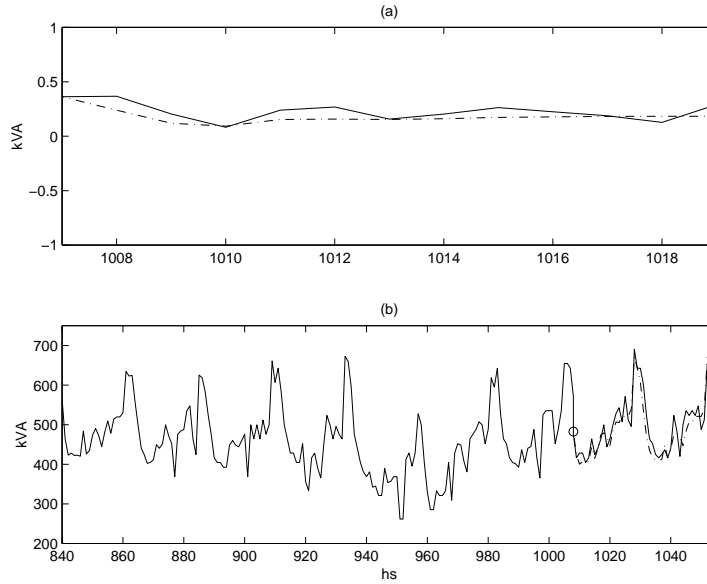


Figura 6.7: (a) (-.-) resultados dos testes de previsão para a janela 7, (-) a partir dos dados da série  $\tilde{S}(k)$ ; (b) (-.-) previsão de resíduos mais ciclo médio ( $\tilde{S}(k) + \bar{S}(k)$ ), (-) comparados com os dados originais da série  $S(k)$ , (o) iniciando no ponto marcado.

## 6.8 Validação

Para avaliar o desempenho das redes para cada teste descrito anteriormente, foram utilizados dois índices de avaliação de previsores: MAPE (erro percentual absoluto médio) e o RMSE (a raiz quadrada do erro médio quadrático). O primeiro deles é usualmente utilizado pelas concessionárias de energia elétrica para fins de avaliação do desempenho de previsores de carga, no entanto ambos são bastante usados em vários estudos de previsão (Drezga and Rahman (1998, 1999); Liu et al. (1996); Reis (1999)). As equações (6.1) e (6.2), descrevem como estes índices são obtidos, sendo o segundo índice, o mesmo descrito pela equação (5.1). Este índice avalia os resultados da previsão para a série de resíduos.

$$MAPE = \frac{1}{L} \sum_{k=1}^L \frac{|y(k) - \hat{y}(k)|}{y(k)} \times 100\%, \quad (6.1)$$

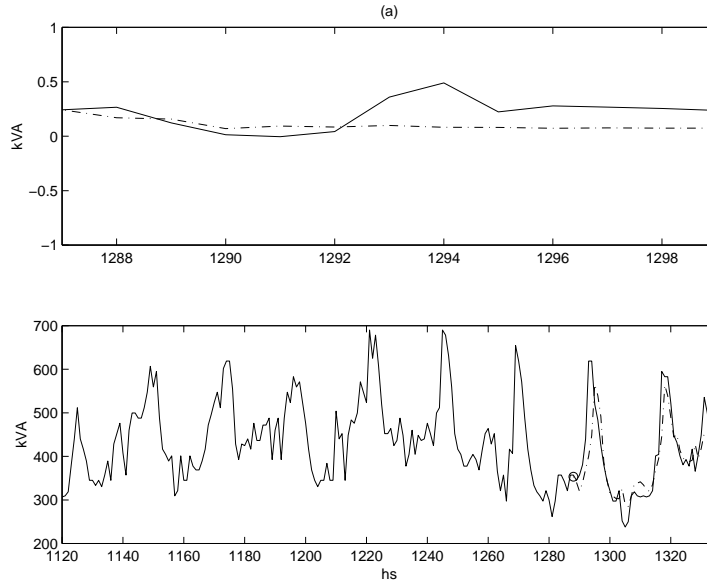


Figura 6.8: (a) (-.-) resultados dos testes de previsão para a janela 9, (-) a partir dos dados da série  $\tilde{S}(k)$ ; (b) (-.-) previsão de resíduos mais ciclo médio ( $\tilde{S}(k) + \bar{S}(k)$ ), (-) comparados com os dados originais da série  $S(k)$ , (o) iniciando no ponto marcado.

$$RMSE = \frac{\sqrt{\sum_{k=1}^L (y(k) - \hat{y}(k))^2}}{\sqrt{\sum_{k=1}^l (\tilde{y}(k) - \bar{y}(k))^2}}, \quad (6.2)$$

sendo que  $\hat{y}(k)$  é o valor previsto do valor real  $y(k)$ ,  $l$  o comprimento do vetor de dados previstos e  $\tilde{y}(k)$  o resíduo de  $y(k)$ .

Desta forma, após a submissão do teste de previsão via redes neurais para cada janela conforme descrito no capítulo anterior, estes índices foram utilizados para avaliar o desempenho da rede para cada uma das janelas. Os cálculos dos mesmos foram realizados para os horizontes de  $l = 5$ hs,  $l = 8$ hs,  $l = 10$ hs e  $l = 12$ hs, sendo estes apresentados nas Tabelas 6.5 e 6.6. Os índices  $MAPE_{rm}$  e  $MAPE_m$ , representam respectivamente os erros de previsão para os resultados obtidos pela rede com os valores de previsão de resíduos mais ciclo médio ( $\tilde{S}(k) + \bar{S}(k)$ ) e os erros de previsão utilizando apenas os ciclos médios.

Tabela 6.5: Resultados dos índices de avaliação MAPE para as janelas 1, 5, 7 e 9.

<i>janela</i>	<i>1</i>		<i>5</i>	
1	$MAPE_{rm} (\%)$	$MAPE_m (\%)$	$MAPE_{rm} (\%)$	$MAPE_m (\%)$
5	5,62	19,62	6,47	10,53
8	4,33	17,05	6,81	9,64
10	4,45	14,88	6,61	9,97
12	4,55	13,28	7,72	11,18
<i>janela</i>	<i>7</i>		<i>9</i>	
1	$MAPE_{rm} (\%)$	$MAPE_m (\%)$	$MAPE_{rm} (\%)$	$MAPE_m (\%)$
5	2,94	8,36	17,21	18,65
8	2,94	8,99	19,01	21,68
10	2,62	8,52	20,81	19,34
12	3,02	8,17	19,03	21,52

Tabela 6.6: Resultados dos índices de avaliação RMSE para as janelas 1, 5, 7 e 9.

<i>RMSE</i>				
<i>l</i>	<i>janela 1</i>	<i>janela 5</i>	<i>janela 7</i>	<i>janela 9</i>
5	0,4455	1,1406	0,5438	1,1982
8	0,3588	1,5071	0,5101	1,7998
10	0,4376	1,4181	0,4582	1,8776
12	0,4401	1,5323	0,5107	1,8564

Avaliando estes resultados, observa-se através da Tabela 6.5, os baixos valores do  $MAPE_{rm}$  em relação ao  $MAPE_m$  em quase todas as janelas, com exceção dos resultados para a janela 9. O valor de erro máximo considerado aceitável para este índice, neste estudo, foi de 5%.

Fazendo uma análise dos resultados das quatro janelas apresentados na Tabela 6.5 e visualizando os gráficos da Figura 6.9, pode-se concluir que os melhores resultados apresentados pela rede foram aqueles utilizando os dados da janela 7 e janela 1, sendo para a janela 7 o melhor entre eles. Ou seja, considerando um horizonte de 8hs, a rede obteve valores de erro de previsão em torno de 2,94% para a janela 7 e 4,45% para a janela 1. Já as janelas 5 e 9, apresentaram, para este horizonte de previsão, valores na ordem de 6,81%

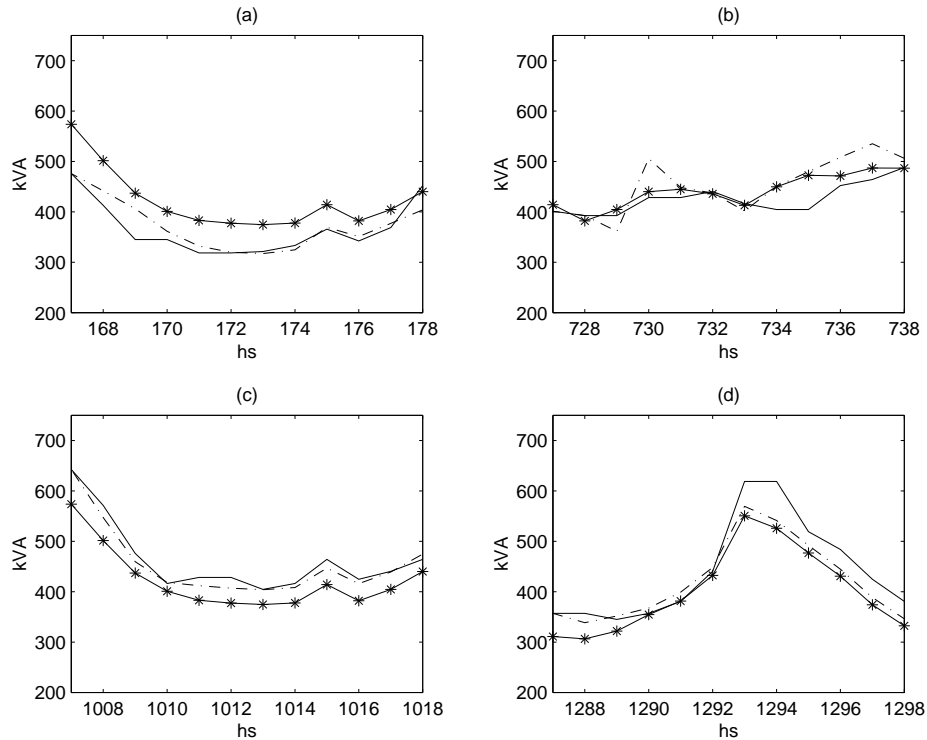


Figura 6.9: Resultado da previsão para as janelas 1 (a), 5 (b), 7 (c) e 9 (d), sendo: (-) Série original  $\tilde{S}(k)$ , (-.-) previsão de resíduos mais ciclo médio  $(\tilde{S}(k) + \bar{S}(k))$  e (-\*-) ciclo médio.

e 21,67%, respectivamente.

Pelos gráficos 6.9(a) e (c), respectivamente das janelas 1 e 7, verifica-se que o resultado apresentado pela rede é mais vantajoso que se utilizar os valores médios, pois a rede consegue “corrigir” os desvios entre a série original e o ciclo médio. Entretanto, os gráficos 6.9(b) e (d), respectivamente das janelas 5 e 9, mostram que o ciclo médio obtém resultados semelhantes ou melhores que a rede nos horizontes avaliados.

A Tabela 6.6 mostra os resultados de previsão de resíduo avaliados em relação aos valores de RMSE. Para valores de RMSE menores que a unidade, os dados contidos na respectiva janela são considerados previsíveis, ou seja, leva-se a pensar que a rede tem a capacidade de obter resultados representativos de previsão para estes dados.

Assim sendo, verifica-se que os índices de RMSE para as janelas 1 e 7 são menores que a unidade, sendo o índice para a janela 7 o menor entre eles. Já as janelas 5 e 9 apresentam valores maiores que 1. Desta forma, comparando estes índices com os resultados apresentados pela rede através da Figura 6.9, pode-se confirmar os respectivos comportamentos esperados da rede. Ou seja, para as janelas 1 e 7 que apresentaram índices de RMSE menores que 1, a rede obteve melhores resultados de previsão. Enquanto que para a janela 5, a qual obteve um comportamento não-estacionário, os resultados apresentados pela rede foram piores.

Um observação deve ser citada com relação à janela 9. No capítulo 5, os dados contidos nesta janela se apresentaram como sendo do tipo determinístico, e no entanto, a rede não conseguiu obter resultados melhores que o ciclo médio (vide gráfico 6.9 (d)). Deve-se lembrar que estes dados contêm um dia de feriado (25 de dezembro) e que este resultado pode ter sido influenciado por este dia não-comum. Entretanto, casos como este devem ser avaliados em estudos futuros.

## Capítulo 7

# Considerações Finais

### 7.1 Discussão

Neste estudo a previsão de consumo de energia foi obtida com o uso de técnicas de redes neurais. Este procedimento pode ser encontrado em vários trabalhos da literatura, sendo que, entre elas existem vários tipos de procedimentos e técnicas de execução para obter o resultado com melhor precisão. Para realizar análises de previsão, a maioria dos autores, para não dizer em sua totalidade até o momento, fez-se uso de séries, chamadas neste trabalho, de séries originais (séries com medições reais de energia). Ou seja, conjecturase que esta contém informações de comportamento determinístico, devido à influência de sua sazonalidade. No entanto, sabe-se que a série de carga é bastante complexa, visto que a mesma sofre influência de vários fatores em seu comportamento.

Desta forma, diante do exposto acima, optou-se por fazer uma avaliação inovadora em uma série de dados de energia elétrica suspeitando este comportamento determinístico. Assim sendo, realizaram-se testes de previsão a partir de dados residuais, ou seja, foi gerada uma série não-sazonal, para avaliar o comportamento determinístico deste sinal. Este comportamento deve ser tal que permita fornecer informações à rede neural, para que esta consiga obter resultados representativos em sua previsão.

Os resultados mostram que a série de resíduos utilizada apresenta-se com melhor dados de previsão que se utilizasse a média destes valores. Ou seja,

pode-se esperar que as previsões de consumo de energia podem ser obtidas não apenas com o uso das séries originais, mas também através do uso das chamadas séries de resíduos.

## 7.2 Conclusões

Esse trabalho veio a contribuir de forma inovadora, como se obter a previsão de consumo de energia elétrica a partir de dados sem o comportamento sazonal (com as séries de resíduo). Os resultados obtidos para a avaliação do teste de determinismo, e conseqüentemente os de previsão, vieram a sugerir que existe uma componente determinística nos sinais de dados de energia elétrica, mas que esta não está sempre presente. Ou seja, pode-se suspeitar que alguns conjuntos de dados (como foi feito neste estudo, com a divisão de janelas) podem ser previstos com uma boa precisão, mas que existem outros os quais não se pode esperar resultados semelhantes.

As previsões obtidas com as séries de resíduo, somadas ao ciclo médio, foram consideradas boas em relação à média (ou seja, melhores que se utilizasse apenas as médias destes dados para a previsão). No entanto, não se pode comparar o desempenho obtido neste trabalho ( $MAPE < 5\%$ ) com os na literatura (normalmente  $MAPE < 3\%$  (Reis (1999); Drezga and Rahman (1998, 1999))), devido ao fato dos dados terem características muito diferentes.

## 7.3 Sugestão para trabalhos futuros

Apesar de ser considerado um estudo de caso, os resultados apresentados neste trabalho sugerem que o teste de determinismo merece ser mais investigado em estudos de previsão de energia elétrica. Além disto, observou-se um comportamento diferente para a previsão de dados iniciada em dias e horários diferentes do testado (a previsão neste caso, iniciou após o treinamento de uma semana).

Assim sendo, como sugestão de trabalhos futuros, outras séries de carga devem ser testadas, bem como a realização de testes de previsões iniciando em dias e horários diferentes.

Os resultados da análise de classificação de padrões realizados no capítulo 4, sugerem a fazer estudos em sistemas com maior quantidade de dados me-

dados para se verificar a existência da correlação entre a classe social e o consumo de energia.

Uma avaliação da série original dos dados fornecidos pela CFLCL, conforme descrito neste trabalho usando a série de resíduos, também sugere um trabalho a ser pesquisado futuramente.



# Referências Bibliográficas

- Aguirre, A. and Aguirre, L. A. (1999). Modelos estatísticos e econométricos para a análise da variação sazonal das séries de preços - o caso dos preços do boi gordo no estado de são paulo. *Nova Economia*, 9(1):171–193.
- Aguirre, A. and Aguirre, L. A. (2000). Time series analysis of monthly beef cattle prices with nonlinear autoregressive models. *Journal of Applied Economics*, 32:265–275.
- Aguirre, L. A. (2000). *Introdução à Identificação de Sistemas*. Editora UFMG.
- Amaral, G. F. V. (2001). Uso de conhecimento a priori e redes neurais na identificação de sistemas dinâmicos não lineares. *Dissertação de Mestrado - Universidade Federal de Minas Gerais, PPGEE-UFMG*.
- Billings, S. A. (1980). Identification of nonlinear systems - a survey. *IEEE Proceedigs Pt. D*, 127(6):272–285.
- Braga, A. P., Ludemir, T. B., and Carvalho, A. C. P. L. F. (1998). *Fundamentos de Redes Neurais Artificiais*. 11a Escola de Computação.
- Burke, H., Rosen, D., and Goodman, P. (1995). Comparing the prediction accuracy of neural networks and other statistical models for breast cancer survival. *Neural Information Processing Systems 7*.
- Caminhas, W. M., Pereira, Tavares, H., and Gomide, F. (2001). Detecção de falhas em sistemas dinâmicos: Abordagem baseada em redes neurofuzzy com aprendizado on-line. *V Congresso Brasileiro de Redes Neurais*, pages 613–618.
- Caminhas, W. M., Pereira, M. F., Aguirre, L. A., Tavares, H., and Gomide, F. (1997). Detecção e identificação de falhas de instrumentação num sistema

- de tanques interativos usando rede neurofuzzy. *III Congresso Brasileiro de Redes Neurais*, pages 201–206.
- Cassini, S. C. C. (1999). Estimação recursiva de características estáticas não lineares utilizando modelos polinomiais NARMAX. *Dissertação de Mestrado - Universidade Federal de Minas Gerais*.
- Charytoniuk, W. and Chen, M. (2000). Very short-term load forecasting using artificial neural networks. *IEEE Transactions on Power Systems*, 15(1):263–268.
- Chen, S., Billings, S., and Luo, W. (1989). Orthogonal least squares methods and their application to non-linear system identification. *International Journal of Control*, 50(5):1873–1896.
- Chow, T. W. S. and Leung, C. T. (1996). Nonlinear autoregressive integrated neural network model for short-term load forecasting. *IEE Proceedings transactions*, 143(5):500–506.
- Chowdhury, F. N. and Aravena, J. L. (1998). A modular methodology for fast fault detection and classification in power systems. *IEEE Transactions on Control Systems Technology*, 6(5):623–633.
- Cybenko, G. (1988). Continuous valued neural networks with two hidden layers are sufficient. *Technical Report, Department of Computer Science*.
- Denker, J. a. a. (1996). Neural network recognizer for handwritten zip code digits. *Neural Network World*, 6(3):241–249.
- Drezga, I. and Rahman, S. (1998). Input variable selection for ann-based short-term load forecasting. *IEEE Transactions on Power Systems*, 13(4):1238–1244.
- Drezga, I. and Rahman, S. (1999). Short-term load forecasting with local ann predictors. *IEEE Transactions on Power Systems*, 14(3):844–850.
- Engbert, R. (2000). Testing for nonlinearity: the role of surrogate data. *Chaos, Solitons and Fractals*, 13:79–84.
- Fritzke, B. (1991). Kohonen feature maps and growing cell structures - a performance comparison. *International Computer Science Institute*, pages 123–130.

- Fukushima, K. (1975). Cognitron: a self-organising multilayered neural network. *Biol. Cybern.*, 23:121–134.
- Gomes, M. E. D. (2001). Técnicas de sistemas dinâmicos não lineares na análise da modulação autonômica da variabilidade da frequência cardíaca. *Tese de doutorado - Universidade Federal de Minas Gerais*.
- Gomes, M. E. D., Souza, A. V. P., Guimarães, H. N., and Aguirre, L. (2000). Investigation of determinism in heart rate variability. *CHAOS*, 10(2):398–410.
- Haykin, S. (1999). *Neural Networks - A Comprehensive Foundation*, volume second edition. Prentice Hall.
- Haykin, S. and VanVeen, B. (1999). *Sinais e Sistemas*. Bookman.
- Hebb, D. O. (1949). The organization of behavior. *Wiley*.
- Henrique, H. M., Lima, E. L., and Seborg, D. E. (2000). Model structure determination in neural network models. *Chemical Engineering Science*, 55:5457–5469.
- Hippert, H. S., Pedreira, C. E., and Souza, R. C. (2001). Neural networks for short-term load forecasting: A review and evaluation. *IEEE Transactions on Power Systems*, 16(1):44–55.
- Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective properties. *Proc. Nat. Acad. Sci.*, 79:2555–8.
- Howell, J. (1993). Kohonen networks: Self-ordering maps. <http://nips.djvuzone.org>.
- Ilic, M., Shanktze, P., and Visudhiphan, P. (2001). Electricity troubles in california: Who's next? *IEEE Spectrum*, pages 11–13.
- Jen-Lun, Y. and Terrence, L. F. (1993). Forecasting demand for electric power. <http://nips.djvuzone.org>, pages 739–746.
- Khotanzad, A., Afkhami-Rohani, R., Lu, T. L., Abaye, A., Davis, M., and Maratukulan, D. J. (1997). Annstlf - an neural-network-based electric load forecasting system. *IEEE Transactions on Neural Networks*, 8(4):835–846.
- Khotanzad, A., Afkhami-Rohani, R., and Maratukulan, D. J. (1998). Annstlf - artificial neural network short-term load forecaster - generation three. *IEEE Transactions on Power Systems*, 13(4):1413–1422.

- Kohonen, T. (1974). An adaptative associative memory principle. *IEEE Transactions on Computers*, 23:444–445.
- Kohonen, T. (1982). Self-organized formation of topologically correct feature maps. *Biological Cybernetics*, 43.
- Kohonen, T. (1989). Self-organization and associative memory. *Springer-Verlag*, 3 edition.
- Lamedica, R., Prudenzi, A., Sforma, M., Caciotta, M., and Cencelli, V. O. (1996). A neural network based technique for short-term forecasting of anomalous load periods. *IEEE Transactions on Power Systems*, 11(4):1749–1756.
- Leontaritis, I. J. and Billings, S. A. (1985a). Input-output parametric models for nonlinear systems part I: Deterministic nonlinear systems. *Int. J. Control*, 41(2):303–328.
- Leontaritis, I. J. and Billings, S. A. (1985b). Input-output parametric models for nonlinear systems part II: Stochastic nonlinear systems. *Int. J. Control*, 41(2):329–344.
- Liangyue, C., Yiguang, H., Hanzhang, Z., and Shuhui, D. (1995). Predicting economic time series using a nonlinear deterministic technique. *Computational Economics*, pages 149–178.
- Lima, S. T., Rodrigues, D., and Aguirre, L. (2002). Técnicas de análise de séries temporais não-lineares aplicadas ao problema de previsão de carga em sistemas elétricos. *Relatório PPGEE/MACSIN/01-02, Universidade Federal de Minas Gerais*.
- Liu, K., Subbarayan, S., Shoults, R. R., Manry, M. T., Kwan, C., Lewis, F. L., and Naccarino, J. (1996). Comparison of very short term load forecasting techniques. *IEEE Transactions on Power Systems*, 11(2):877–882.
- Ljung, L. (1987). *System Identification, theory for the user*. New Jersey: Prentice Hall.
- Makansi, J. (2001). California’s electricity crisis rooted in many failings. *IEEE Spectrum*, pages 24–27.

- Mc Culloch, W. S. and Pitts, W. (1943). A logical calculus of the ideas immament in nervous activity. *Bulletin of Mathematical Biophysics*, 5:115–133.
- Minsky, M. and Papert, S. (1969). Perceptrons: an introduction to computational geometry. *MIT Press*.
- Mori, H. and Kobayashi, H. (1996). Optimal fuzzy interference for short-term load forecasting. *IEEE Transactions on Power Systems*, 11(1):390–396.
- Oliveira, C. M., Machado, L. J. B., and Wazlawick, R. (2000). Modelo de rede neural artificial para previsão automática da carga ativa nodal em sistemas de energia elétrica. *XIII Congresso Brasileiro de Automação*, pages 1138–1143.
- Overbye, T. J. and Weber, J. D. (2001). Visualizing the electric grid. *IEEE Spectrum*, pages 52–57.
- Papoulis, A. (1991). *Probability, Random Variables and Stochastic Process*, volume 3rd edition. Mc Graw- Hill.
- Peng, T. M., Hubele, N. F., and Karady, G. G. (1993). An adaptative neural network approach to one-week ahead load forecasting. *IEEE Transactions on Power Systems*, 8(3).
- Pham, D., T. and Xing, L. (1995). *Neural Networks for Identification, Prediction and Control*, volume 2nd edition. Springer-Verlag.
- Reategui, E. and Campbell, J. A. (1994). A classification system for credit card transactions. *Proceedings of The Second European Workshop on Case-Based Reasoning*, pages 167–174.
- Reed, R. (1993). Prunning algorithms - a survey. *IEEE Transactions on Neural Networks*, 4(5):740–747.
- Reis, A. (1999). Pré-processamento de séries temporais para um previsor de carga elétrica a curto prazo via redes neurais artificiais. *Dissertação de Mestrado - Escola Federal de Engenharia de Itajubá*.
- Reis, A. and Silva, A. (2000). Pré-processamento de séries temporais para a previsão de carga a curto prazo via redes neurais. *XIII Congresso Brasileiro de Automação*, pages 2036–2041.

- Rich, E. and Knight, K. (1994). Inteligência artificial. *McGraw-Hill*, segunda edição.
- Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain. *Psychol. Rev.*, 65:386–408.
- Rumelhart, D. E., Hinton, G. E., and Williams, R. J. (1986). Learning representations by backpropagation errors. *Nature*, 323:533–536.
- Silva, A. P. A., Ferreira, C., Souza, A. C. Z., and Torres, G. L. (1997). A new constructive ANN and its application to electric load representation. *IEEE Transactions on Power Systems*, 12(4).
- Silva, A. P. A. and Moulin, L. S. (2000). Confidence intervals for neural network based short-term load forecasting. *IEEE Transactions on Power Systems*, 15(4).
- Sodré, E. and Mota, W. (2000). Um sistema neuro-fuzzy para análise de segurança da estabilidade de tensão. *XIII Congresso Brasileiro de Automação*, pages 2036–2041.
- Teive, R., Silva, E., and Gauche, E. (2000). Síntese de redes de transmissão de sistemas de energia elétrica-uma abordagem baseada em fluxos em redes e inteligência artificial. *XIII Congresso Brasileiro de Automação*, pages 765–770.
- Theiler, J., Eubank, S., Longtin, A., and Galdrikian, B. and Farmer, J. D. (1992). Testing for nonlinearity in time series: the method of surrogate data. *Physica D*, 58:77–94.
- Valença, M. and Ludermir, T. (1998a). Self-organizing modeling in forecasting daily river flows. *IEEE Spectrum*, pages 210–214.
- Valença, M. and Ludermir, T. (1998b). Uma nova rede neural polinomial com aplicação na previsão de vazões. *V Simpósio Brasileiro de Redes Neurais*, pages 273–278.
- Widrow, B. and Hoff, M. E. (1960). Adaptive switching circuits. *Institute of Radio Engineers, Western Electronic Show and Convention*.
- Zaded, L. A. (1965). Fuzzy sets. *Information and Control*, 8:338–353.