



Universidade Estadual de Campinas
Faculdade de Engenharia Elétrica e Computação
Departamento de Engenharia de Sistemas

Um Modelo de Previsão de Carga por Barramento

Por: **Ricardo Menezes Salgado**

Orientador: **Prof Dr. Takaaki Ohishi**

Co-Orientadora: **Profa. Dra. Rosangela Ballini**

Dissertação de Mestrado apresentada à Faculdade de Engenharia Elétrica e Computação (FEEC) da Universidade Estadual de Campinas (UNICAMP) para o preenchimento dos pré-requisitos parciais para a obtenção do Título de MESTRE em Engenharia Elétrica.

Banca Examinadora:

Prof. Dr. Wagner da Silva Lima (EEE-UFG).

Prof. Dr. Secundino Soares Filho (FEEC-UNICAMP).

Prof. Dr. Fernando A. Campos Gomide (FEEC-UNICAMP).

Campinas, Julho de 2004.

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

Sa32m Salgado, Ricardo Menezes
Um Modelo de Previsão de Carga por Barramento /
Ricardo Menezes Salgado. --Campinas, SP: [s.n.], 2004.

Orientadores: Takaaki Ohishi, Rosangela Ballini.
Dissertação (Mestrado) - Universidade Estadual de Campinas,
Faculdade de Engenharia Elétrica e de Computação.

1. Análise de séries temporais - Processamento de dados
2. Redes Neurais (Computação). 3. Sistemas de energia elétrica -
Distribuição de carga. 4. Algoritmos difusos. I. Ohishi, Takaaki.
II. Ballini, Rosangela III. Universidade Estadual de Campinas.
Faculdade de Engenharia Elétrica e de Computação.
VI. Título

Resumo

Na operação de um sistema de energia elétrica, uma etapa importante é a determinação da programação da operação diária, a qual determina um plano de produção de energia elétrica para o(s) próximo(s) dia(s) para cada uma das unidades geradoras do sistema, geralmente em base horária ou de meia hora. Esta programação é utilizada pela operação em tempo real do sistema como uma referência operativa, e por isso é importante que a solução proposta assegure uma operação adequada do sistema. Para avaliar o impacto de um dado programa de operação sobre o sistema de transmissão, é necessário que se conheça a distribuição da carga ao longo da rede, pois o carregamento nas linhas de transmissão e transformadores depende da demanda de carga em cada barramento (ponto de entrega de energia elétrica). Num contexto de planejamento da operação diária é necessário conhecer a carga em cada barramento em cada intervalo de tempo considerado na programação. Ou seja, faz-se necessário uma previsão de carga de curto prazo por barramento.

O principal objetivo desta dissertação foi desenvolver um modelo de previsão de carga diária ativa, em base horária, por barramento. Dois tipos de metodologias foram implementadas: metodologias de previsão individual (MPI) que trata cada barramento de forma isolada e metodologias de previsão agregada (MPA) na qual a previsão é feita uma única vez para um dado conjunto de barramentos. O modelo agregado visa diminuir a necessidade da realização de previsões para cada barramento, propondo para isto, a realização de uma única previsão de forma agregada; este modelo é composto de três fases: (i) fase de agregação - onde as cargas dos barramentos são agregadas; (ii) fase de previsão - que é realizada através da série agregada em (i); (iii) fase de desagregação - onde a previsão é distribuída através dos barramentos agregados. Para agregar os barramentos utilizou-se técnicas de agrupamento de dados. A metodologia de previsão individual atende aos barramentos que apresentam pouca similaridade no seu perfil de demanda, quando comparados a outros barramentos, desta forma a sua agregação resulta em altos erros. Para os barramentos que apresentam alta similaridade com outros barramentos, as metodologias agregadas foram eficientes na previsão, proporcionando um menor número de previsões e resultados de boa qualidade. Os dados utilizados para testar os modelos são dados reais medidos em um sistema de transmissão e sub-transmissão do nordeste brasileiro.

Abstract

In the operation of an electric power system, an important stage is the determination of the daily operation programme, which determines a plan of electric power production for the following day(s) for each of the generating units of the system, usually on an hourly or a half-hourly basis. This programme is used by the system's real time operation as an operational reference, and therefore it is important that the proposed solution should assure an appropriate operation of the system. To evaluate the impact of any given operation programme on the transmission system, the distribution of the load along the net must be known, since the loading in the transmission lines and transformers depends on the load demand in each bus (point of electric power delivery). In a daily operation planning context, it is necessary to know the load in each bus in each time interval considered in the programme. In other words, a short-term load forecast per bus is necessary.

The main goal of this work was to develop a daily active load forecast model, on an hourly basis, per bus. Two types of methodologies were implemented: individual forecast methodology (MPI) that addresses each bus in an isolated way and aggregated forecast methodology (MPA) in which the forecast is made a single time for a given group of buses. The aggregate model seeks to reduce the need of forecasts being made for each bus, proposing for that a single forecast in an aggregated way. This model is composed of three stages: (i) aggregation phase - where the loads of the buses are joined; (ii) forecast phase - which is made through the aggregated series in (i); (iii) disaggregation phase - where the forecast is distributed through the joined buses. Two data clustering techniques were used for aggregating the buses. The individual forecast methodology considers the buses that present little similarity in their demand profile, when compared to the other buses, thus their aggregation results in high error rates. For the buses that present high similarity to other buses, the aggregated methodologies were efficient in the forecast, providing a smaller number of forecasts and results of good quality. The database used to test the models features real data measured in the transmission system and subtransmission of the Brazilian northeast region.

Este trabalho é dedicado ao Senhor
Jesus Cristo *ao qual pertence toda*
Glória, Louvor e Adoração hoje e sempre.

*Este projeto contou com o suporte financeiro da Fundação
de Amparo à Pesquisa do Estado de São Paulo - FAPESP.*

Agradecimentos

“Não tenho palavras pra agradecer Tua bondade

Dia após dia me cercas com fidelidade

Nunca me deixes esquecer

Que tudo o que tenho, tudo o que sou

O que vier a ser vem de Ti Senhor

Dependo de Ti, preciso de Ti

Sozinho, nada posso fazer

Descanso em Ti, espero em Ti

Sozinho, nada posso fazer...”

Vem de Ti/Diante do Trono.

Como esta canção, que é uma linda oração a Deus, diz eu não tenho palavras para agradecer a bondade do nosso Senhor Deus em minha vida, que por sua excelente vontade, proporcionou os meios necessários para que esta pesquisa pudesse ser realizada.

No mesmo tom de gratidão, agradeço à meus pais José e Enaíde que são co-autores de todas as minhas conquistas e sempre estarão presentes em minha vida. Agradeço também ao meu irmão Marcelo que em momentos difíceis foi um ponto de apoio fundamental. Agradeço a todos os membros da minha família pelas orações e incentivo.

À minha noiva Juliana expresso a minha profunda gratidão pois, com todo o seu carinho e amor, ela esteve presente em todas as horas me confortando e sempre me auxiliando. À família Russano pelos momentos alegres e pelas orações.

Aos professores Takaaki Ohishi e Rosangela Ballini pela amizade, paciência e orientação na execução e no desenvolvimento desta pesquisa. Aos amigos do COSE e DENSIS pelo companheirismo.

A Eng^a. Maria da Conceição Guedes Alcoforado (ONS/NE) pela concessão do banco de dados utilizado nesta pesquisa.

À FAPESP pelo suporte financeiro e o apoio necessário na execução deste projeto.

... Muito obrigado a todos!!!

*“A lei do Senhor é perfeita, e refrigera a alma; o testemunho do Senhor é fiel, e dá sabedoria aos simples”. **Salmo 19:7***

Sumário

Resumo	p. iii
Abstract	p. iv
Notação	p. xiii
Lista de Figuras	p. xiv
Lista de Tabelas	p. xvi
1 Introdução	p. 18
1.1 Revisão Bibliográfica	p. 19
1.2 Objetivos da Dissertação	p. 24
1.3 Organização da Dissertação	p. 25
2 Descrição do Problema e Apresentação dos Dados	p. 26
2.1 O Problema da Previsão de Carga por Barramento	p. 26
2.2 Apresentação dos Dados	p. 29
2.2.1 Análise dos Barramentos	p. 30
2.2.1.1 Normalização dos Dados	p. 30
3 Metodologias de Previsão	p. 36
3.1 Metodologia de Médias Móveis	p. 36

3.2	Metodologia Baseada em Rede Neural	p. 37
3.2.1	Histórico das Redes Neurais Artificiais	p. 37
3.2.2	Modelo de um Neurônio Artificial	p. 38
3.2.3	Estruturas de Redes Neurais Artificiais	p. 40
3.2.4	Tipos de Treinamento	p. 40
3.2.5	Redes Perceptron Multicamadas - MLP	p. 41
4	Metodologias de Agrupamento	p. 43
4.1	Introdução	p. 43
4.2	Definição	p. 44
4.2.1	Dificuldades no Problema de Agrupamento	p. 45
4.3	Mapas Auto-Organizáveis de Kohonen (SOM)	p. 45
4.4	Definição SOM	p. 46
4.4.0.1	Algoritmo de treinamento	p. 46
4.5	Técnica de Poda do Mapa Autorganizável de Kohonen - PSOM	p. 48
4.5.1	Definição - PSOM	p. 48
4.6	Fuzzy C-Means	p. 51
4.6.1	Noções Gerais	p. 51
4.6.2	Algoritmo FCM	p. 52
4.7	Resultados Obtidos	p. 54
4.7.1	Agrupamento de um dia	p. 55
4.7.2	Agrupamento de uma semana	p. 57
4.7.3	Agrupamento de 44 dias	p. 58
4.8	Conclusões	p. 60

5	Modelos de Previsão e Resultados	p. 61
5.1	Introdução	p. 61
5.2	Modelos de Previsão	p. 62
5.2.1	Modelos de Previsão Individual (MPI)	p. 62
5.2.2	Modelos de Previsão Agregada (MPA)	p. 67
5.2.2.1	Agregação dos Barramentos	p. 68
5.2.2.2	Previsão da Demanda Agregada	p. 69
5.2.2.3	Cálculo das Cargas Individuais - Desagregação	p. 70
5.2.3	Configuração das Metodologias de Previsão	p. 74
5.3	Resultados Alcançados	p. 75
5.3.1	Modelos de Previsão Individual - MPI	p. 76
5.3.2	Modelos de Previsão Agregada - MPA	p. 78
5.3.2.1	Previsão por Fator de Participação - PFP	p. 79
5.3.2.2	Rede Treinada com Perfil/Da - PRTDa	p. 80
5.3.2.3	PMDDa - Previsão Média Diária-Da	p. 81
5.3.3	Análise dos Resultados	p. 83
5.3.3.1	Avaliação: MPA	p. 83
5.3.3.2	Avaliação: MPA $\times$ MPI	p. 86
5.4	Conclusões	p. 88
6	Conclusão e Trabalhos Futuros	p. 90
6.1	Conclusões Gerais	p. 90
6.2	Trabalhos Futuros	p. 92
	Referências	p. 93

Apêndice A – MLP e Backpropagation p. 97

A.1	Noções Gerais	p. 97
A.1.1	Algoritmo de Aprendizado para uma rede MLP	p. 98
A.1.1.1	Derivada da Função de Ativação	p. 104
A.1.2	Taxa de Aprendizagem	p. 104
A.1.3	Inicialização dos Pesos	p. 105
A.1.4	Critério de Parada	p. 105
A.1.5	Generalização	p. 106
A.1.6	Problemas com o Algoritmo de treinamento	p. 106

Apêndice B – Tabelas - MAE p. 108

B.1	Metodologias de Previsão Individual	p. 108
B.2	Metodologias de Previsão Agregada	p. 109

Notação

- ▷ MPI: Modelos de Previsão Individual;
- ▷ PCD: Previsão da Curva de Carga Diária;
- ▷ PMD: Previsão da Carga Média Diária;
- ▷ MPA: Modelos de previsão agregada;
- ▷ PFP: Previsão via Fator de Participação;
- ▷ $Da(g, d, h)$: Demanda agregada no grupo g no dia d e na hora h ;
- ▷ $Da(g, d, h)_{prev}$: Demanda agregada prevista no grupo g no dia d e na hora h ;
- ▷ PRTDa: Previsão via Rede Treinada com Perfil da Da;
- ▷ PMDDa: Previsão da Média Diária - DA;
- ▷ $C(i, d, h)$: Carga real do barramento i na hora h do dia d ;
- ▷ $Tx_{(d)}$: Taxa de variação entre as média prevista e a média do dia escolhido como perfil;
- ▷ $Mp_{(d)}$: Média prevista no dia d ;
- ▷ $Mv_{(p)}$: Média verificada do dia indicado como perfil p ;
- ▷ $Cp(i, d, h)$: Carga prevista no barramento i no dia d na hora h ;
- ▷ $P(i, d, h)$: Participação do barramento i em relação a Da no dia d e na hora h ;
- ▷ $P(i, d, h)_{prev}$: Participação revista no barramento i no dia d e na hora h ;
- ▷ MAPE: Erro Percentual Médio Absoluto;
- ▷ MAE: Erro Médio Absoluto;
- ▷ FCM: *Fuzzy C- Means*;
- ▷ PSOM: *Pruning Technique Kohonen Self-Organizing Feature Map*;

Lista de Figuras

1	Sistema Interligado de Transmissão - Brasil/Fonte: ONS.	p. 28
2	Sistema de Potência - Dividido em Barramentos.	p. 29
3	Motivação para Padronização das Curvas.	p. 31
4	Perfil dos Barramentos 1 a 3.	p. 33
5	Perfil dos Barramentos 4 a 6.	p. 33
6	Perfil dos Barramentos 7 a 9.	p. 34
7	Perfil dos Barramentos 10 a 12.	p. 34
8	Perfil dos Barramentos 13 a 15.	p. 35
9	Modelo de neurônio Artificial.	p. 39
10	Arquiteturas da rede de Kohonen: (i)Arquitetura bidimensional (ii)Arquitetura unidimensional.	p. 46
11	Agrupamentos FCM 19/09/2001 - Quarta-Feira.	p. 56
12	Agrupamentos PSOM 19/09/2001 - Quarta-Feira.	p. 56
13	Desvio Padrão nos Grupos (fixando barramento 1).	p. 60
14	Curva diária - Barramento 1.	p. 63
15	Curva Média do Barramento 1.	p. 65
16	Porcentagem $\times$ Carga do barramento 12.	p. 71
17	Comparação dos MPI.	p. 78
18	Eficiência dos MPI.	p. 78
19	MPA - Fator de participação.	p. 80

20	Desempenho Metodologia - PRTDa.	p. 82
21	Desempenho metodologia PMDDa.	p. 82
22	Comparação dos MPA.	p. 84
23	Comparação - Barramentos com baixa demanda.	p. 85
24	Eficiência dos MPA.	p. 86
25	Comparação MPI $\times$ MPA.	p. 87
26	Eficiência MPI $\times$ MPA.	p. 87
27	Eficiência dos modelos - Porcentagem.	p. 88
28	Arquitetura MLP com duas camadas intermediárias.	p. 97
29	Saída do Neurônio j	p. 99
30	Saída do Neurônio Intermediário.	p. 102

Lista de Tabelas

1	Análise estatística.	p. 32
2	Agrupamentos FCM 19 a 25/09/2001.	p. 57
3	Agrupamentos PSOM 19 a 25/09/2001.	p. 58
4	Agrupamentos de Barramentos.	p. 59
5	Grupos de Barramentos (Fixando Barra 1).	p. 60
6	Vetor de Ajuste da RNA - PCD	p. 63
7	Vetor de Ajuste da RNA - PMD	p. 65
8	Vetor de Ajuste da RNA - $Da_{(prev)}$	p. 69
9	Vetor de Ajuste da RNA - PFP	p. 72
10	Arquitetura Rede Neural.	p. 75
11	MAPE Diário - PCD.	p. 76
12	MAPE Diário - PMD.	p. 77
13	MAPE Diário - PFP/Médias.	p. 79
14	MAPE Diário - PFP/RNA.	p. 80
15	MAPE Diário - PRTDa.	p. 81
16	MAPE Diário - PMDDa.	p. 83
17	MAE Diário - PCD.	p. 108
18	MAE Diário - PMD.	p. 109
19	MAE Diário - PFP/Médias.	p. 109
20	MAE Diário - PFP/RNA.	p. 110

21	MAE Diário - PRTDa.	p. 110
22	MAE Diário - PMDDa.	p. 111

1 *Introdução*

O acompanhamento da carga em um sistema de potência é exigência básica nos processos de operação. Para fornecer energia elétrica com boa qualidade, de maneira segura e econômica, uma empresa de energia elétrica precisa dispor de mecanismos que possibilitem a resolução de vários problemas de nível técnico e operacional. Com o propósito de obter melhorias no planejamento e controle do sistema, novas ferramentas de otimização e controle são desenvolvidas e aplicadas, tendo resultados relevantes na redução de custos. Em concordância com este objetivo, o conhecimento do comportamento da carga futura é o primeiro pré-requisito para um planejamento seguro e confiável do sistema de energia elétrica.

Um programa de operação de um sistema hidrotérmico estabelece a produção média de cada unidade geradora para cada intervalo de tempo do próximo dia. Obviamente, a produção total programada em cada intervalo de tempo deverá atender a demanda estimada para o respectivo intervalo de tempo. Especificamente, no contexto do planejamento da operação de curto prazo de sistemas de energia elétrica, a previsão da carga é importante na elaboração do programa de operação dos próximos dias em análises de segurança e estabilidade, pois erros na previsão da carga podem ter consequências sérias em termos da eficiência e segurança do sistema (aumento dos custos, não atendimento da demanda). Neste contexto, uma previsão de boa qualidade é essencial, pois visa um melhor atendimento da demanda, aumentando a eficiência e a confiabilidade do sistema.

Tradicionalmente, o programa de operação é baseado na previsão de carga global, isto é, o programa de operação se preocupa em atender a demanda total em cada intervalo de tempo. Entretanto, como o programa de operação é uma referência para o controle em tempo real do sistema, torna-se importante avaliar o impacto que o plano de produção de energia elétrica para o próximo dia terá sobre o sistema de transmissão. Para se realizar esta avaliação é necessário conhecer a demanda de carga por barramento (nós da rede de transmissão e de sub-transmissão do sistema de energia elétrica). Como na grande maioria dos casos não se dispõe da previsão por

barramento, é comum utilizar perfis típicos de demanda por patamar (carga média, leve e pesada) baseado em dados históricos. Em relação às análises de segurança e de estabilidade, nas quais se avalia o impacto que as alterações no sistema de transmissão têm sobre o desempenho do sistema, o conhecimento da demanda por barramento é também essencial. O interesse pela previsão é justificado na medida em que a mesma é um subsídio relevante à tomada de decisão.

1.1 Revisão Bibliográfica

A previsão de carga é um típico problema de previsão de séries temporais e já é um tema de pesquisa e desenvolvimento tradicional na área de sistemas de potência. Inicialmente, os modelos desenvolvidos eram baseados em modelos estatísticos (Gross e Galiana, 1987) sendo esta a técnica dominante até meados da década de 80.

Com o surgimento das redes neurais artificiais e a popularização das mesmas a partir da década do 80, a maioria dos modelos propostos para previsão de carga vem utilizando esta técnica, sendo uma das primeiras aplicações de sucesso das redes neurais (Ho et al., 1992).

Os modelos de Redes Neurais Artificiais (RNAs) têm mostrado ser uma poderosa ferramenta na classificação e reconhecimento de padrões. Inspiradas em sistemas biológicos relacionados com o cérebro humano, RNAs são capazes de aprender e generalizar com a experiência de dados passados. Atualmente, RNAs têm sido utilizadas para uma grande variedade de aplicações em diferentes áreas de negócios, indústria e ciência (Zhang et al., 1998).

Um grande número de aplicações de RNAs são vistos na área de previsão, uma vez que as RNAs possuem características bastante atraentes, tais como:

- Não requerem *a priori* informações sobre a modelagem do sistema a ser previsto, pois aprendem a partir de exemplos e, assim, capturam relações funcionais entre os dados mesmo se tais relações são dificilmente descritíveis.
- Têm a capacidade de generalizar. Após a fase de aprendizagem, quando já possuem algum conhecimento sobre as características do conjunto de dados, as RNAs conseguem inferir na população restante mesmo se dados amostrais contêm ruídos nas informações.
- São aproximadores universais, isto é, podem aproximar qualquer função contínua com uma dada precisão .

- É uma ferramenta não-linear, diferente de algumas técnicas utilizadas para a previsão tais como os modelos de Box-Jenkins.

A idéia de utilizar RNAs em previsão não é nova, sendo que a primeira aplicação data de 1964, realizada por Hu (Zhang et al., 1998). A utilização de redes neurais para a previsão de séries temporais foi proposta por vários autores, (Lapedes e Farber, 1987), (Weigend et al., 1990), mostrando a viabilidade de utilizar estes modelos.

Desde então, diversas estruturas de RNAs vêm sendo propostas baseadas principalmente nos modelos perceptron multi-camadas, mapas de Hopfield e redes auto-organizáveis de Kohonen. Em problemas de previsão, os modelos mais utilizados são redes estáticas multi-camadas, treinadas através do algoritmo de retropropagação, ou *backpropagation* (Rumelhart et al., 1986), cuja base é o método do gradiente descendente (Maier e Dandy, 1998).

(Lapedes e Farber, 1987) aplicaram redes neurais multicamadas (MLP) para a previsão de duas espécies de séries temporais caóticas geradas por um processo determinístico não-linear. Os resultados indicaram que os modelos MLP podem “aprender”, simulando sistemas não-lineares, com excelentes características quando comparadas a métodos tradicionais.

(Tang et al., 1991) compararam modelos de MLP e Box & Jenkins, usando séries de tráfego de passageiros de vôo internacional, vendas de carros domésticos e vendas de carros importados nos EUA. Os resultados mostraram que o desempenho do modelo de Box & Jenkins foi superior ao método de RNAs para uma previsão a curto prazo, no entanto para a previsão a longo prazo, os modelos baseados em redes neurais apresentaram um melhor desempenho.

(Lachtermacher e Fuller, 1995) mostram uma proposta híbrida para a aplicação de redes neurais em análise de séries temporais. A metodologia de Box & Jenkins foi utilizada como um procedimento para explorar importantes relações nos dados da série e estas informações foram usadas para definir uma estrutura inicial para a rede neural, reduzindo, assim, os parâmetros a serem estimados e os dados exigidos. Dois tipos de séries foram estudadas, séries estacionárias e não-estacionárias. Observou-se que o modelo de redes neurais, para a maioria das séries testadas, obteve um desempenho igual ou superior quando comparados aos modelos de Box & Jenkins.

(Park et al., 1991) utilizaram uma MLP para previsão de carga, tendo como entrada da RNA dados climáticos (temperatura). Park propôs uma RNA para aprender a relação entre as temperaturas e a carga. Para calcular a carga prevista, a RNA interpola os dados carga-temperatura num conjunto de treinamento. Para isto o problema de previsão foi subdividido em três sub-problemas,

isto é, foram desenvolvidas três redes neurais especializadas, uma para prever o pico de carga, uma para prever total de carga do dia e outra para o perfil horário da carga durante o dia, para horizontes de 24 horas à frente.

Dependendo do tipo de problema ao qual são submetidas, as RNAs têm apresentado um desempenho considerado superior aos métodos estatísticos utilizados para o mesmo fim. Os pesquisadores (Subramanian et al., 1993), desenvolveram uma pesquisa onde compararam as redes neurais com determinados métodos estatísticos de classificação. Concluíram, nas diversas circunstâncias estudadas, incluindo pequenos tamanhos de amostras e funções de maior complexidade, que as RNAs apresentaram melhores soluções.

(Lu et al., 1993) compararam a eficácia das RNAs para a previsão com horizontes de uma hora à frente com a previsão com horizonte de 24 horas à frente, analisando se o modelo de redes neurais é dependente da região de atuação da concessionária de energia elétrica. Para isto, foram utilizadas redes *feedforward* MLP para fazer a previsão 1 e 24 horas à frente durante a semana, utilizando como dados de entrada, carga, temperatura, hora do dia e dia da semana.

Para obter uma rede que se adapte às mudanças no comportamento da carga com o decorrer do tempo, (Dillon et al., 1991) propuseram um modelo de rede neural adaptativa para previsão de carga a curto-prazo uma hora à frente. Neste modelo, treina-se a rede neural com os dados de carga dos últimos 20 dias para prever a carga do dia 21.

(Peng et al., 1992) propuseram uma estratégia para seleção dos padrões de treinamento da rede utilizando a medida da mínima distância para identificar os padrões de carga dos dias semelhantes. Para previsão eles usaram um algoritmo com uma combinação de termos lineares e não-lineares através de uma rede neural apropriada para fazer a previsão horária com horizonte de 24 horas à frente.

(Mohammed et al., 1995) propuseram um algoritmo que considera a sazonalidade da carga atual e treina diversas redes especializadas em cada estação e em cada dia da semana. O treinamento é feito com os dados do mesmo período do ano anterior e corrigidos de acordo com os novos dados obtidos através de uma adaptação diária, semanal e mensal.

As RNAs e técnicas tradicionais para predição de séries temporais estão sendo comparadas há algum tempo (Tang et al., 1991). Grande parte dos estudos comparativos realizados atualmente utilizam o algoritmo *backpropagation*, ou algumas de suas variações. Tecnicamente, este modelo já foi utilizado com grande sucesso em diversas áreas de conhecimento, dentre elas

sistemas de potência, previsões econômicas e análise de séries temporais (Weigend et al., 1991a), (Weigend et al., 1991b).

Comparações feitas entre redes neurais com outras abordagens ainda não apresentam resultados concretos. Alguns autores evidenciam que as RNAs tem melhor desempenho que os métodos tradicionais (Weigend et al., 1991a), (Weigend et al., 1991b), enquanto outros acreditam no resultado oposto (Tang et al., 1991). As explicações para tais contradições podem ser descritas por diferentes fatores, como (Ballini, 2000):

- Estrutura da RNA;
- Tipo de série (estacionária ou não-estacionária);
- Tamanho da RNA versus número de entradas da série.

Atualmente um novo método está sendo usado para unir as vantagens de RNA e lógica nebulosa, resultando em um sistema integrado, onde a aprendizagem e o poder computacional dos modelos de RNA e a capacidade de representação e raciocínio da lógica nebulosa são combinados. Estes sistemas, chamados de redes neurais nebulosas (*redes neurofuzzy*) vêm sendo aplicados sucessivamente para sistemas de identificação não-linear, controle de processos, reconhecimento de padrões, previsão de carga (Ballini, 2000). Os sistemas nebulosos são sistemas computacionais desenvolvidos a partir das teorias de conjuntos nebulosos e da lógica nebulosa (fuzzy), visando aproveitar de forma pragmática o potencial destas teorias. Estes sistemas processam conhecimento e inferem conclusões a partir dos padrões apresentados (Pedrycz e Gomide, 1998).

Modelos de previsão, genericamente categorizados como aprendizagem de seqüências, constituem um dos mecanismos essenciais de aprendizagem humana e de máquina, desempenhando um importante papel em condicionamento clássico, controle, resolução de problemas e raciocínio (Anderson, 1995). Modelos de previsão são um importante componente de aprendizagem em vários domínios de sistemas inteligentes, tais como, planejamento, raciocínio, inferência, processamento de voz, reconhecimento de voz, mercado financeiro, controle adaptativo e previsão de séries temporais (Zhang et al., 1998).

Um tipo de modelo que provou ser útil no trabalho com séries temporais é o Modelo Oculto de Markov (*Hidden Markov Model - HMM*). A partir das características de uma série temporal existente, é possível a um HMM gerar uma série que se aproxime da real e melhorá-la sucessivamente. Para isso, é necessário descobrir tais características intrínsecas da série.

Em (Oates et al., 2000) tal procura é feita por um agrupamento de dados, cuja métrica de similaridade utilizada é denominada *Dynamic Time Warping (DTW)*. Como resultado, obtém-se um conjunto de seqüências, isto é, os dados da série que por terem características similares, pertencem a um mesmo grupo. Considerando a condição de que cada seqüência que pertence a um grupo foi gerada pelo mesmo HMM, os sucessivos agrupamentos terão como objetivo encontrar os HMMs que melhor ajustam as seqüências iniciais.

Outro exemplo da utilização dos HMMs é encontrado em (Shalizi et al., 2002). Diferente da abordagem anterior, os autores não consideram previamente características da arquitetura causal do processo markoviano, mas sim inferem tais propriedades do conjunto de dados. Em outras palavras, o objetivo é encontrar a estrutura dinâmica, causal e intrínseca ao processo, a fim de extrair todos os padrões que contém algum poder preditivo.

Já (Last et al., 2001) propuseram uma metodologia para a identificação dos atributos preditivos mais importantes e extração de um conjunto de regras de associação nebulosas que podem ser usadas para realizar previsões do comportamento da série temporal no futuro.

A ferramenta para agrupar dados com características semelhantes pode ser utilizada também em conjunto com as RNAs (Khalil et al., 2001). A ferramenta de agrupamento de dados pode ser testada no problema de previsão de carga, considerando a identificação de características da série que influenciarão na estimativa de valores futuros. Uma metodologia que vem sendo largamente usada por vários autores no que diz respeito à agrupamento são os Mapas Auto-Organizáveis de Kohonen desenvolvido por (Kohonen, 1982) (Kaski e Kohonen, 1995) (Kohonen, 2000).

(Francelin e Vieira, 1994) mostram que a Rede Auto-Organizável de Kohonen é eficaz para agrupar dados de séries temporais que apresentam características semelhantes.

(Lima e Ohishi, 1999) desenvolveram um classificador neural baseado em Mapas Auto-Organizáveis de Kohonen (SOM) para análise e classificação dos perfis de curvas de carga com influência climática.

Para encontrar um padrão de carga, (Ho et al., 1992) desenvolveram uma rede neural baseada nos Mapas Auto-Organizáveis de Kohonen para identificar os dias com padrões de carga similares, agrupando-os em classes. Uma característica especial dos mapas auto-organizáveis é que eles são capazes de identificar um novo tipo de padrão de carga antes que os operadores possam reconhecer o padrão do novo dia.

(Erkmen e Özökomen, 1993) desenvolveram uma rede neural híbrida. A primeira rede tem aprendizagem do tipo não supervisionada e serve para classificar os padrões de carga a partir de um histórico de carga horária. A outra rede, supervisionada, realiza a previsão horária uma hora à frente a partir do padrão já pré-estabelecido anteriormente.

Nos últimos anos, diferentes abordagens têm sido propostas para o problema de previsão e, entre elas, os modelos baseados em redes *neurofuzzy* têm apresentado resultados de excelente qualidade, despontando como uma alternativa para o problema de previsão de séries temporais (Mastorocostas et al., 1999).

(Last et al., 2001) propuseram uma metodologia para a identificação dos atributos preditivos mais importantes e extração de um conjunto de regras de associação nebulosas que podem ser usadas para realizar previsões do comportamento da série temporal no futuro.

Todos os trabalhos acima apresentados trataram somente da previsão de uma série temporal. Em particular, no caso da previsão de carga por barramento não foi encontrada nenhuma publicação.

(Salgado et al., 2004a) e (Salgado et al., 2004b) aplicaram o algoritmo Fuzzy C-means (FCM) para o agrupamento de curvas de demanda num determinado dia da semana. Os resultados demonstram que o algoritmo dividiu o conjunto de barramentos em grupos caracterizados por barramentos com perfis de consumo similares. Os grupos formados tendem a se manter ao longo do tempo.

1.2 Objetivos da Dissertação

O estudo proposto neste trabalho tem como objetivo geral a aplicação de modelos de previsão de séries temporais para prever o consumo de carga diária ativa nos barramentos do sistema elétrico para um período de 24 horas à frente.

A metodologia proposta baseia-se nos modelos neurais, em simbiose com algoritmos de agrupamento, em particular, foram investigadas arquiteturas de redes neurais para previsão de carga e algoritmos de agrupamento *fuzzy* e neural.

O desempenho dos modelos é avaliado fazendo-se comparações entre as previsões obtidas e os dados reais. As medidas de comparação utilizadas para avaliar os modelos serão os *erro*

percentual médio absoluto (MAPE) e o erro médio absoluto (MAE).

1.3 Organização da Dissertação

Este trabalho encontra-se dividido nos seguintes Capítulos:

- No Capítulo 2 são mostrados alguns conceitos sobre séries temporais, que possibilitaram uma análise estatística da série temporal de cada barramento. Esta análise tem o objetivo de possibilitar a extração de informações relevantes que possam caracterizar o comportamento da série de cada barramento;
- O Capítulo 3 expõe, de forma teórica, os modelos de previsão que serão utilizados neste trabalho para resolver o problema de previsão por barramento;
- O Capítulo 4 descreve os fundamentos e metodologias utilizadas para o agrupamento dos dados. O algoritmo PSOM proposto por (Castro, 1998) e o algoritmo FCM introduzido por (Bezdek, 1981) serão os enfocados como modelos de agrupamento. Os resultados e as comparações entre as metodologias de agrupamento são mostradas no fim deste capítulo;
- A forma como foi realizada a previsão agregada é apresentada no Capítulo 5 bem como os resultados e as comparações entre os modelos de previsão propostos.
- O Capítulo 6 apresenta uma análise crítica e uma avaliação de todos os métodos propostos e as contribuições provenientes deste estudo. Também são colocadas sugestões para continuidade deste tema de pesquisa;
- No apêndice A as redes neurais multi-camadas são detalhadas e também são feitos alguns comentários sobre o algoritmo de aprendizado *backpropagation*.
- No apêndice B são apresentadas as tabelas contendo os desempenho dos modelos de previsão propostos no Capítulo 5 medidos em termos de erro absoluto médio.

2 *Descrição do Problema e Apresentação dos Dados*

2.1 O Problema da Previsão de Carga por Barramento

Na operação de um sistema de energia elétrica, uma etapa importante é a programação da operação diária, a qual determina um plano de produção de energia elétrica para o(s) próximo(s) dia(s) para cada uma das unidades geradoras do sistema. Esta programação será utilizada pela operação em tempo real do sistema como uma referência operativa, e por isso é importante que a solução proposta assegure uma operação adequada do sistema. Ou seja, todos os requisitos operacionais devem ser detalhadamente considerados, entre os quais a análise do desempenho do sistema de transmissão é um dos mais importantes, pois é o meio de transporte da energia produzida até os centros de consumo.

Tradicionalmente, o programa de operação é baseado na previsão de carga global, isto é, o programa de operação se preocupa em atender a demanda total em cada intervalo de tempo. Entretanto, como o programa de operação é uma referência para o controle em tempo real do sistema, torna-se importante avaliar o impacto que o plano de produção de energia elétrica para o próximo dia terá sobre o sistema de transmissão. Para se realizar esta avaliação é necessário conhecer a distribuição geográfica da carga; isto é, conhecer a demanda de carga por barramento (nós da rede de transmissão e sub-transmissão do sistema de energia elétrica).

Quando fala-se em previsão de carga por barramento surge a seguinte questão: *“Como realizar uma previsão rápida que possa atender a todos os barramentos de forma satisfatória?”*

De acordo com as técnicas de previsão descrita no Capítulo 1 pode-se realizar a previsão aplicando um tratamento individual em cada barramento do sistema. Este procedimento é viável quando trabalha-se em um sistema com poucos barramentos. Já em um sistema com um

número elevado de barramentos, o tratamento individual é menos viável pois tal tarefa demandaria um grande tempo na obtenção das previsões. Observando, na Figura 1 o mapa do Sistema de Transmissão do SIN (Sistema Interligado Nacional) verifica-se que existem um grande número de barramentos, aumentando o grau de dificuldade na obtenção das previsões.

Desta forma, uma primeira alternativa para realizar a previsão de carga por barramento é fazer a previsão de forma agregada. A utilização desta previsão agregada visa proporcionar a realização de um número menor de previsões e, desta forma, reduzir o tempo e o esforço computacional que é exigido no tratamento individual de cada barramento.

Neste raciocínio, percebe-se que a previsão de carga por barramento é um problema extremamente complexo envolvendo vários fatores e dependendo de conhecimento especializado para determinar as características de cada centro consumidor.

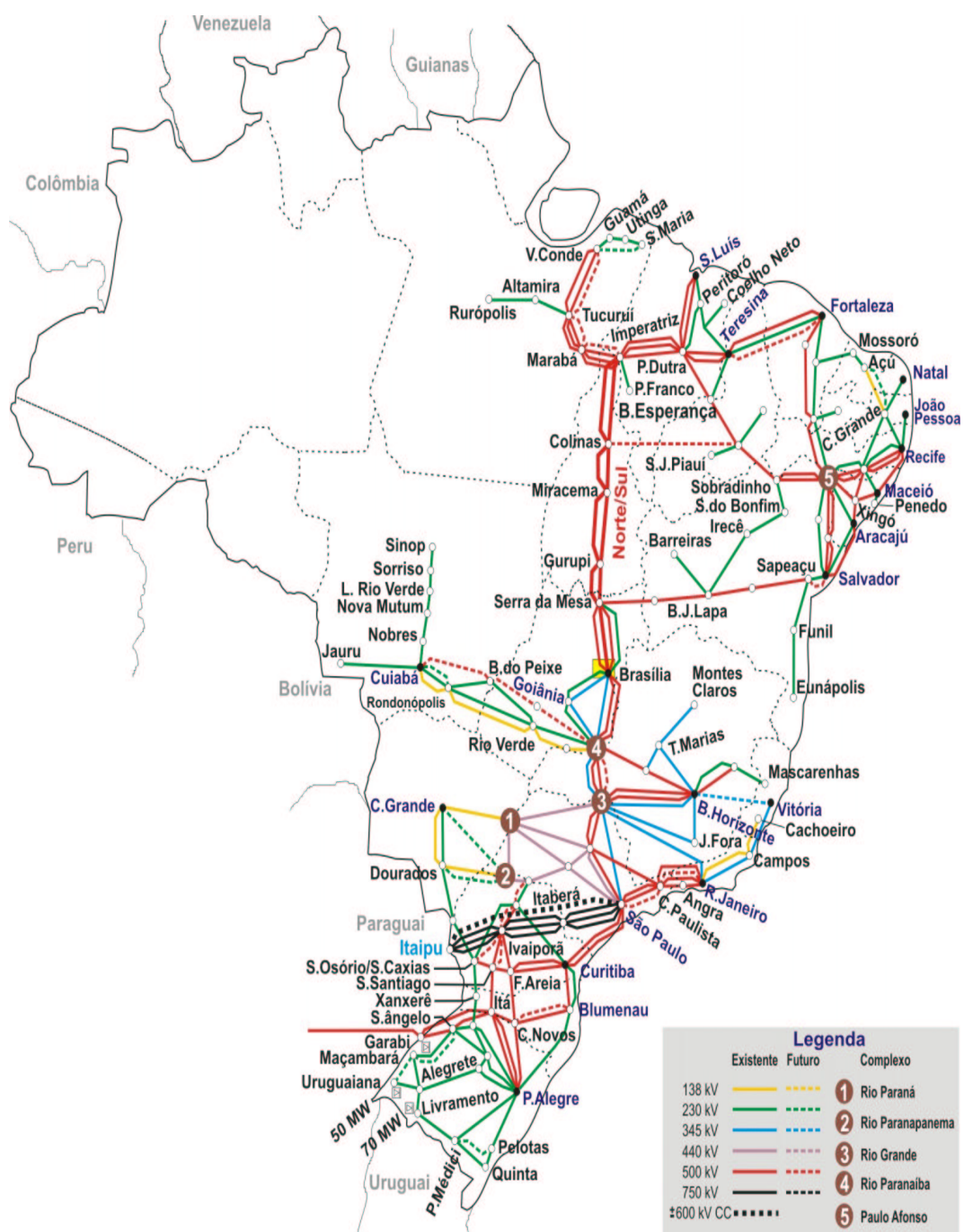


Figura 1: Sistema Interligado de Transmissão - Brasil/Fonte: ONS.

A Figura 2 ilustra um sistema de potência com seus respectivos barramentos. Os pontos de medição em cada barramento registram a potência (MW/h) consumida em cada um, sendo este registro utilizado como uma referência para calibrar os modelos de previsão.

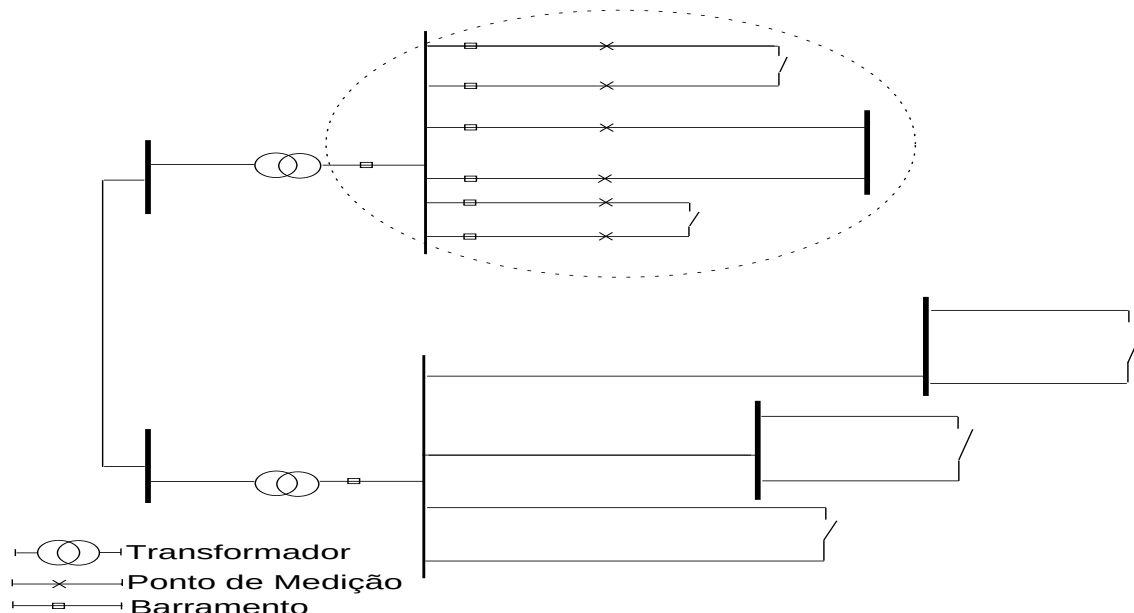


Figura 2: Sistema de Potência - Dividido em Barramentos.

A grande importância na previsão de carga por barramento se encontra no fato desta previsão proporcionar um conhecimento da carga (MW/h) na rede de distribuição facilitando o planejamento energético e proporcionando maior segurança e estabilidade nas operações em tempo real do sistema elétrico. De acordo com (Handschin e Dörnmann, 1988) o conhecimento da carga por barramento é uma referência fundamental no planejamento da expansão do sistema.

2.2 Apresentação dos Dados

Análise de séries temporais têm como objetivo básico a avaliação do fenômeno sob consideração, para se obter conclusões em termos estatísticos, bem como, sumarizar propriedades estatísticas e caracterizar seu comportamento (Ballini, 2000).

Conhecer o conjunto de dados que será utilizado na fase de ajuste e validação dos modelos de previsão é fundamental em qualquer etapa do desenvolvimento de uma pesquisa. A análise prévia dos dados pode fornecer informações relevantes que serão levadas em consideração

em etapas futuras no ajuste de um modelo de previsão.

2.2.1 Análise dos Barramentos

A análise dos barramentos têm sua relevância no sentido de identificar comportamentos, padrões e informações que estão presentes na série temporal de cada barra. Um conjunto de 15 barramentos, numerados de 1 a 15, são utilizados para testar e validar os modelos implementados nos Capítulos posteriores.

Neste trabalho, foram consideradas séries diárias de consumo de energia, isto é, por barramento, pertencentes ao sistema elétrico da região nordeste do Brasil.

As medições que representam os barramentos foram feitas no período de 01/08/2001 a 04/10/2001 (período do racionamento de energia). Tais barras possuem tensão variando de 38 a 230 KV. As Figuras 4 a 8 mostram as cargas (MW/h) observadas no período de medição.

2.2.1.1 Normalização dos Dados

As séries de carga de cada barra serão denotadas por $\{C(i, d, h), d = 1, \dots, 53; h = 1, \dots, 24; i = 1, \dots, 15\}$ tal que i denota o barramento do sistema, d o índice dos dias e h o número de horas. Como o enfoque deste trabalho é realizar previsões com base horária, entende-se que cada dia compreende 24 horas.

Para efetuar a análise detalhada dos barramentos, as séries temporais de cada barramento foram padronizadas (colocadas em um intervalo padrão, isto é, em p.u.). O objetivo de realizar a padronização de dados de barramentos é possibilitar uma análise conjunta de barramentos que apresentam tanto perfil de consumo, como níveis de carga distintos. A Figura 3 mostra claramente o objetivo da padronização das curvas de carga.

Para efetuar esta conversão as seguintes etapas foram realizadas:

Etapa 1: Cálculo da média $\hat{\mu}_{(i,d,h)}$ e variância $\hat{\sigma}_{(i,d,h)}^2$ para cada barramento usando as seguintes equações:

$$\hat{\mu}_{(i,d,h)} = \frac{1}{n} \sum_{d=1}^n C(i, d, h) \quad (2.1)$$

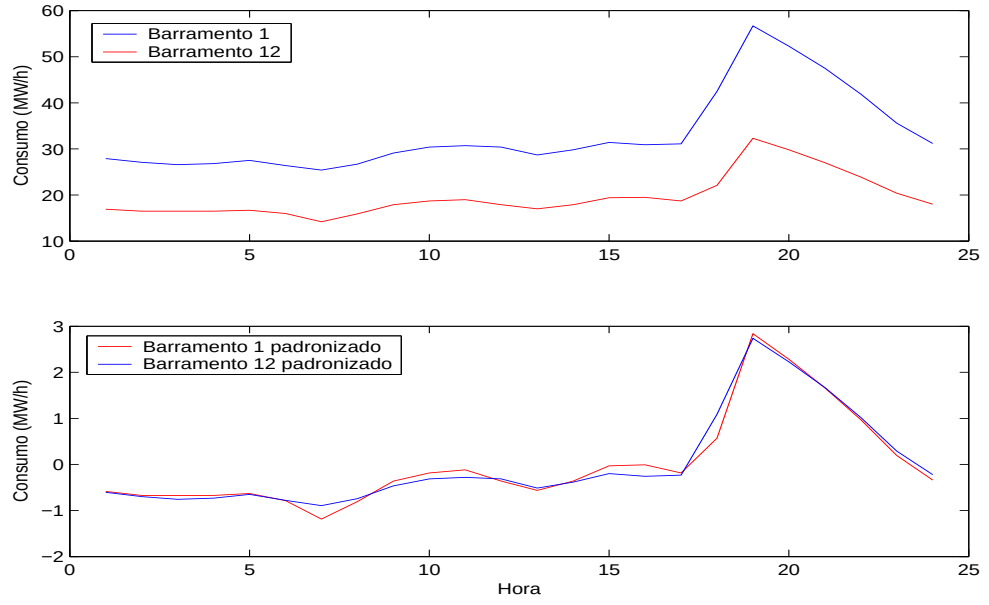


Figura 3: Motivação para Padronização das Curvas.

$$\hat{\sigma}_{(i,d,h)}^2 = \frac{1}{n-1} \sum_{d=1}^n (C(i,d,h) - \hat{\mu}_{(i,d,h)})^2 \quad (2.2)$$

Etapa 2: Nesta fase a série temporal de cada barramento é padronizada através da seguinte equação:

$$\hat{C}(i,d,h) = \frac{C(i,d,h) - \hat{\mu}_{(i,d,h)}}{\hat{\sigma}_{(i,d,h)}} \quad (2.3)$$

A Tabela 1 mostra valores máximo, mínimo, médio e o desvio padrão diários para os dados sem padronização e para cada um dos 15 barramentos. Nota-se que alguns barramentos apresentam grandes variações em termos de desvio padrão enquanto outras apresentam variações menores. A última coluna da Tabela 1 apresenta o desvio padrão dos elementos padronizados, com isto é possível detectar quais elementos possuem um comportamento com maior ou menor regularidade.

Tabela 1: Análise estatística.

Barramentos	Dados Sem Padronização (MW)				
Índice	<i>Máximo</i>	<i>Mínimo</i>	<i>Média</i>	<i>Desvio Padrão</i>	<i>Desvio Padrão(p.u.)</i>
1	65,50	22,10	35,60	2,42	0,02
2	2,80	0,70	1,39	0,11	0,03
3	75,0	16,80	49,03	4,42	0,04
4	8,70	4,50	7,20	0,71	0,05
5	160,30	83,10	117,64	10,89	0,05
6	259,50	122,40	190,29	20,19	0,06
7	13,40	3,20	9,10	1,41	0,08
8	22,60	8,50	14,82	1,52	0,06
9	261,50	89,50	179,57	21,41	0,06
10	114,70	53,50	77,72	8,56	0,07
11	53,10	13,10	33,22	5,04	0,08
12	37,30	11,70	20,92	1,86	0,03
13	4,00	1,10	2,27	0,23	0,05
14	55,80	17,80	32,23	5,43	0,09
15	10,60	1,80	4,81	1,06	0,09

Através das Figuras 4 a 8 pode-se ver o comportamento de cada barramento bem como a sua padronização obtida através da equação (2.3). Os barramentos apresentados nestas figuras apresentam comportamentos que possuem perfis distintos¹; algumas barras apresentam perfis tipicamente residencial (com demanda de ponta ocorrendo no período das 17:00h às 22:00h), outras comercial (com a carga maior no período comercial) e outras industrial (com a demanda constante ao longo do dia).

Nas Figuras 4 e 7, nota-se que os perfis de consumo dos barramentos 1, 2 e 12 apresentam grande regularidade, mostrando um comportado consumo tipicamente residencial. Isto também pode ser comprovado através dos desvios padrões em p.u. (Tabela 1) que são baixos em relação aos demais barramentos, o que confirma o comportamento com pouca variabilidade ao longo do tempo. Logo os perfis de carga dos barramentos 1, 2 e 12 apresentam um comportamento previsível e regular.

Seguindo o mesmo tipo de análise, pode-se observar que os barramentos 10, 11 (na Figura 7) e 14, 15 (na Figura 8), apresentam grande variabilidade, ou seja, os desvios padrões associados a estes perfis são mais elevados, segundo uma tendência de consumo tipicamente comercial; tendo um comportamento irregular e, portanto, mais difícil de prever.

¹Os perfis típicos de consumo estão de acordo com a demanda de carga do sistema elétrico do Brasil.

Na Figura 5, pode-se notar que o barramento 4 mostra uma curva de consumo constante ao longo do dia, isto reflete um perfil tipicamente industrial.

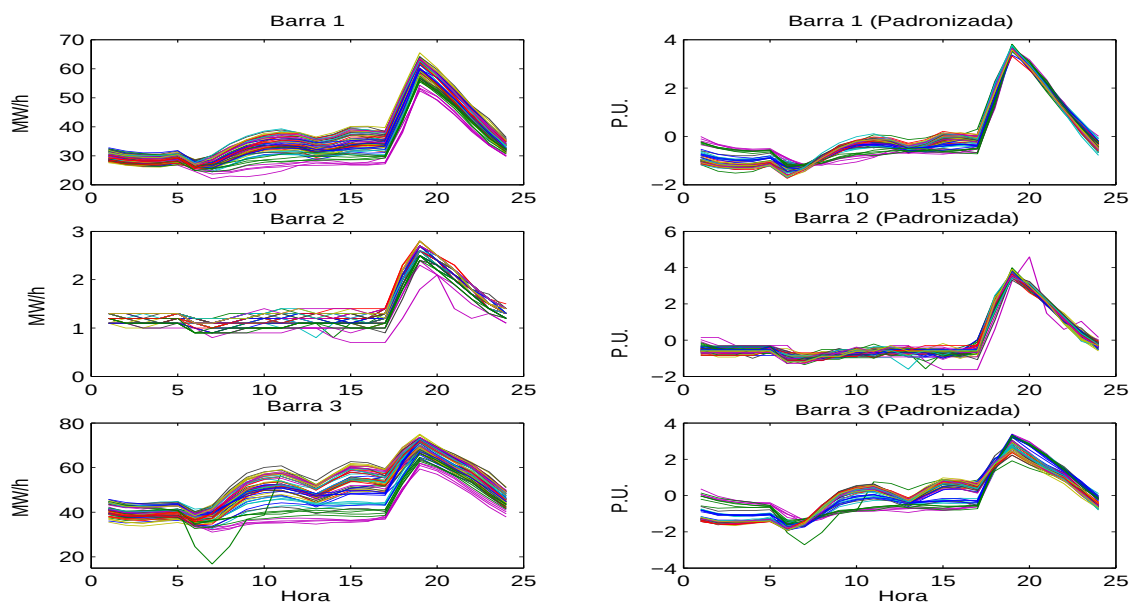


Figura 4: Perfil dos Barramentos 1 a 3.

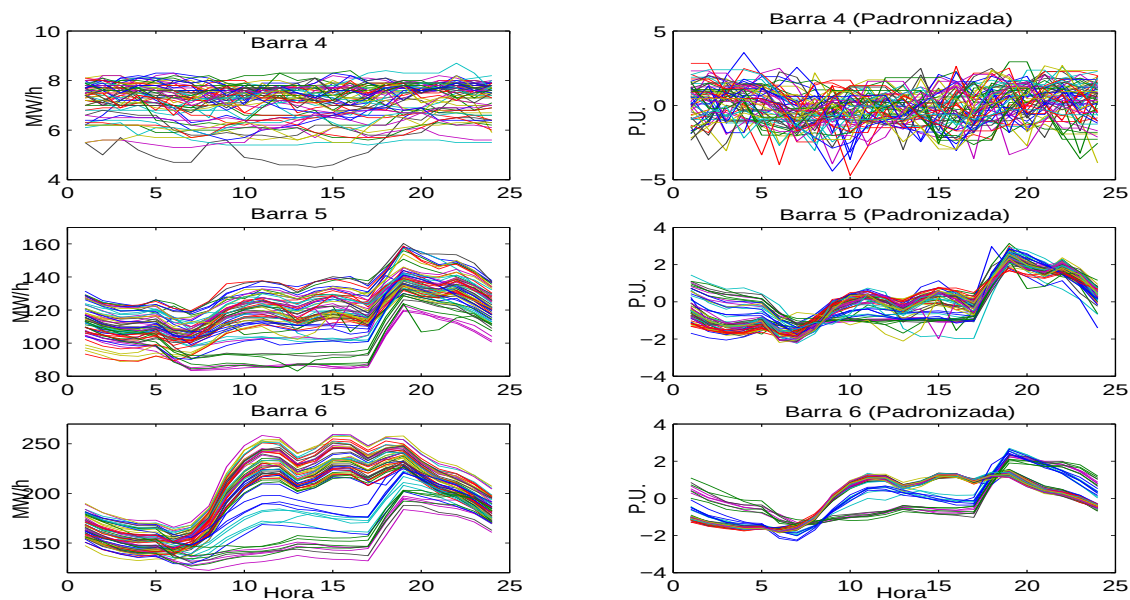


Figura 5: Perfil dos Barramentos 4 a 6.

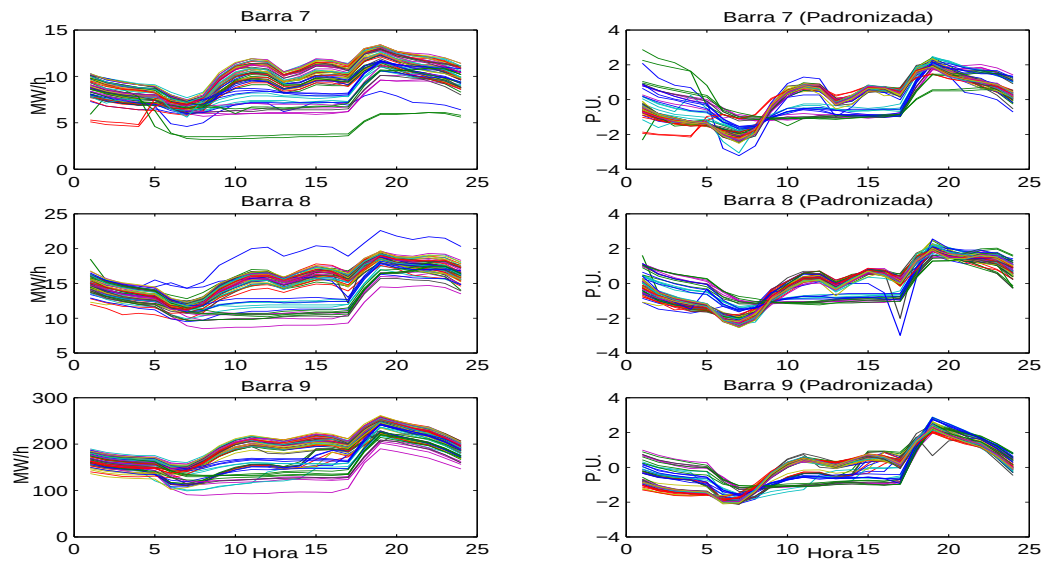


Figura 6: Perfil dos Barramentos 7 a 9.

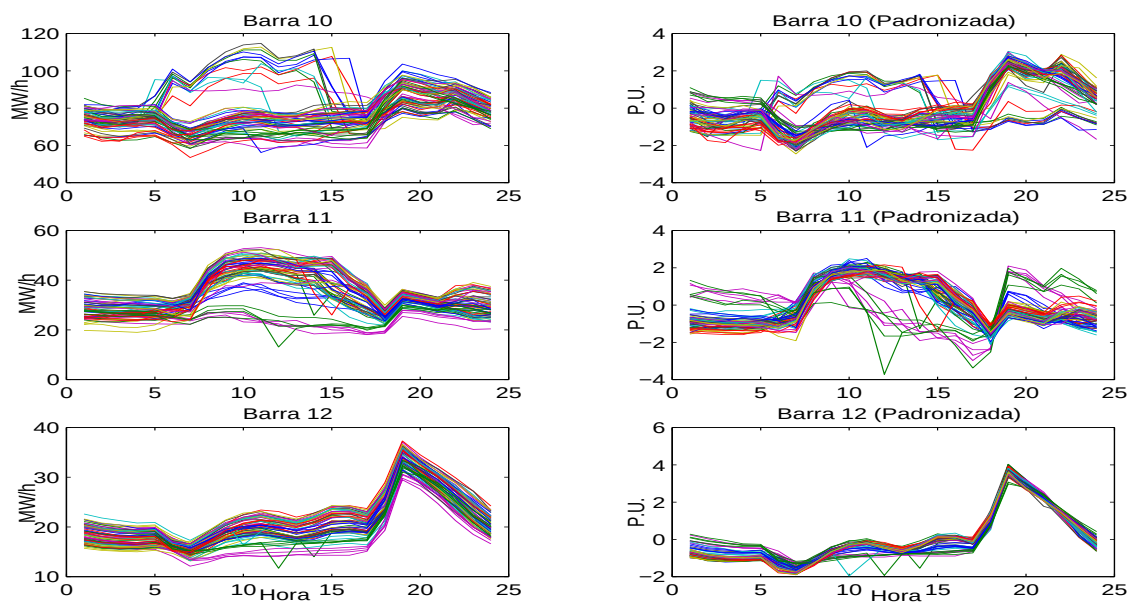


Figura 7: Perfil dos Barramentos 10 a 12.

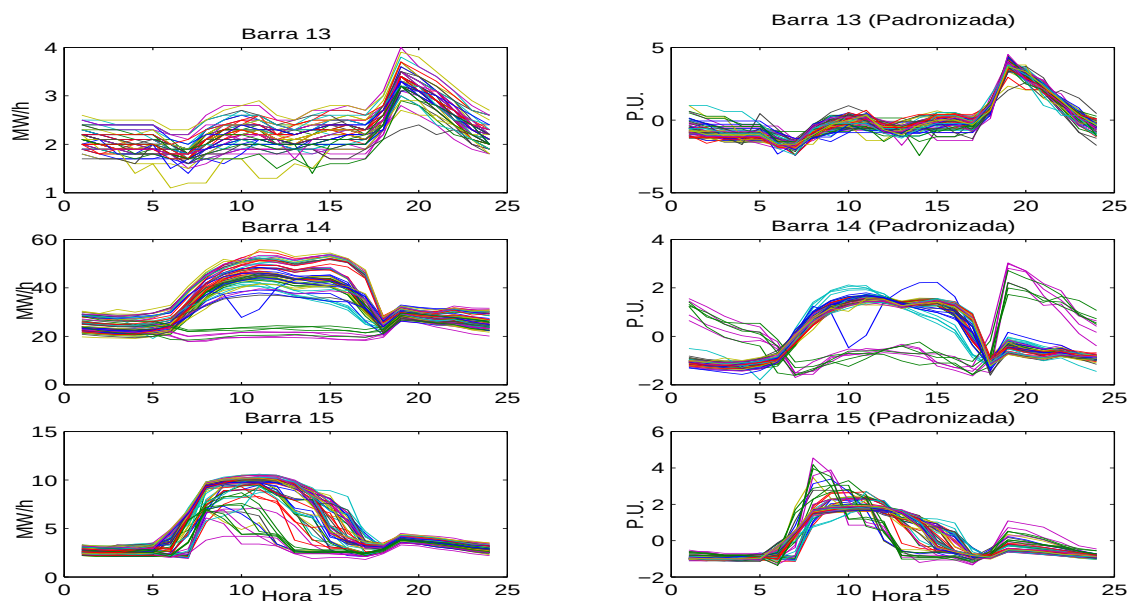


Figura 8: Perfil dos Barramentos 13 a 15.

Com a análise estatística apresentada na Tabela 1, pôde-se notar que determinados barramentos são mais regulares e outros menos, apresentando consumos máximo, mínimo e médios distintos. Em relação aos desvios padrões nota-se que existem barramentos que apresentam tipos de consumo variados, o que induz a modelos de previsão pouco precisos.

De uma forma geral, observando as Figuras 4 a 8 nota-se que existem vários tipos de barramentos que atendem consumidores distintos. O fato das curvas de consumo serem distintas alimenta a necessidade da realização de previsores que sejam capazes de enxergar tais distinções e proporcionar resultados acurados.

3 *Metodologias de Previsão*

O presente capítulo apresenta as duas metodologias de previsão de séries temporais utilizados neste trabalho, o primeiro baseado em médias (Mda) e o segundo com base nas redes neurais. As metodologias estão tratadas, de maneira objetiva, procurando situá-las no contexto abordado neste trabalho.

3.1 Metodologia de Médias Móveis

O método mais simples para realização da previsão de carga numa dada hora h é repetir a carga obtida na hora $(h - 1)$, supondo que o resultado neste instante seja conhecido. Outro modelo simplificado para prever a carga na hora h é dado através da média aritmética de cargas verificadas na mesma hora em dias anteriores.

Uma abordagem considerada neste trabalho foram as médias móveis. Como o próprio nome indica, o método das médias móveis é constituído por sucessivas médias movendo-se ao longo da série, sendo que cada observação usada para computar a média tem o mesmo peso de ponderação. O número de termos usados em uma média móvel determinará o grau de regularidade resultante. Assim, quanto mais termos se usam na média, tanto mais regularizada se torna a ocorrência.

Médias móveis para um período escolhido de extensão n consistem em uma série de médias aritméticas, calculadas ao longo do tempo, de modo que cada média aritmética seja calculada para uma sequência de valores observados que tenham extensão n . A este modelo denominou-se Mda.

Matematicamente, o Mda pode ser descrito da seguinte forma:

$$Cp(i, d, h) = \frac{1}{n} \sum_{s=1}^n C(i, (d-s), h) \quad (3.1)$$

Sendo que:

- $Cp(i, d, h)$: carga prevista no barramento i no dia d para a hora h ;
- $C(i, d, h)$: matriz com históricos de cargas verificadas de um determinado barramento i ;
- n : número de dias usados para calcular a média.

3.2 Metodologia Baseada em Rede Neural

As redes neurais artificiais são largamente utilizadas no problema de previsão de séries temporais, motivo pelo qual adotou-se esta metodologia para resolver o problema de previsão de carga por barramento.

Segundo (Braga et al., 1998) as redes neurais artificiais (RNAs) são sistemas paralelos distribuídos, compostos por unidades de processamento simples que calculam determinadas funções matemáticas. Geralmente, essas unidades são conectadas por canais de comunicação (sinapses) aos quais estão associados determinados pesos.

3.2.1 Histórico das Redes Neurais Artificiais

A história das RNAs é relativamente recente. Conta com pouco mais de meio século, tomando como pioneiros os trabalhos dos neurofisiologistas Donald Hebb e Karl Lashley, por volta de 1940, quando as primeiras simulações foram feitas. (McCulloch e Pitts, 1943) propuseram um modelo simplificado de neurônios biológicos. O modelo baseia-se no fato de que, em dado instante de tempo, o neurônio ou está ativo ou está inativo. O trabalho de McCulloch e Pitts propiciou uma rápida disseminação do entendimento de possíveis modelos neurais, atraindo e influenciando muitos pesquisadores.

F. Rosenblatt, em seu livro *Principles of Neurodynamics*, em 1958, forneceu várias idéias a respeito dos perceptrons, que são modelos de neurônios baseados nos conceitos de McCulloch e Pitts. Uma das principais idéias de Rosenblatt (Rosenblatt, 1958) foi a elaboração da

arquitetura *backcoupled Perceptron* e do algoritmo *backcoupled error correction*, que é capaz de adaptar os pesos de um conjunto de entradas de uma unidade de associação de acordo com uma saída desejada.

B. Widrow e M. Hoff desenvolveram o ADALINE (ADaptive LINear Element) e o MADALINE (Many ADALINE) Perceptron como dispositivos práticos para resolver tarefas de reconhecimento de padrões. Os modelos ADALINE e MADALINE usam saídas analógicas (contínuas) ao invés das binárias originalmente propostas por McCulloch e Pitts (Widrow e Hoff, 1960) e (Widrow e Hoff, 1962).

Minsky e Papert (Minsky e Papert, 1969) fizeram um estudo cuidadoso desses algoritmos e publicaram, em 1969, o livro *Perceptrons*. Este trabalho provou que uma rede formada de uma única camada de neurônios, independente do algoritmo de aprendizagem, é capaz de resolver o problema de associação de padrões apenas quando os conjuntos são linearmente separáveis. Estes resultados e observações feitas por Minsky e Papert foram devastadores, e a abordagem conexionista ficou relegada a um plano secundário durante toda a década de 70 e início da década de 80.

O interesse nas pesquisas de redes neurais ressurgiu após a publicação dos trabalhos do físico John Hopfield em 1982. Em (Hopfield, 1982) um modelo neural recorrente foi proposto. Este modelo busca atingir um estado de energia mínima, garantindo a estabilidade da rede.

A impotência das redes perceptron na resolução do problema de associação de padrões para um conjunto de padrões não-linear foi eliminada por Rumelhart, Hinton e Williams (Rumelhart et al., 1986). A solução encontrada foi a Regra Delta Generalizada, mais conhecida como Algoritmo de Correção de Erros de Retropropagação, em 1986, para redes perceptron de várias camadas de neurônios, com entradas e saídas analógicas.

3.2.2 Modelo de um Neurônio Artificial

Um neurônio é a unidade fundamental de processamento de informação de uma rede neural (Haykin, 1994). A Figura 9 mostra o modelo de um neurônio artificial.

Sendo $x_0 = b_k$, $x_1, \dots, x_p$ os sinais de entrada; $w_{k0}, w_{k1}, \dots, w_{kp}$ os pesos sinápticos do neurônio k ; $f(\cdot)$ a função de ativação e y_k é a saída do neurônio k .

A arquitetura de um neurônio artificial apresenta três elementos básicos:

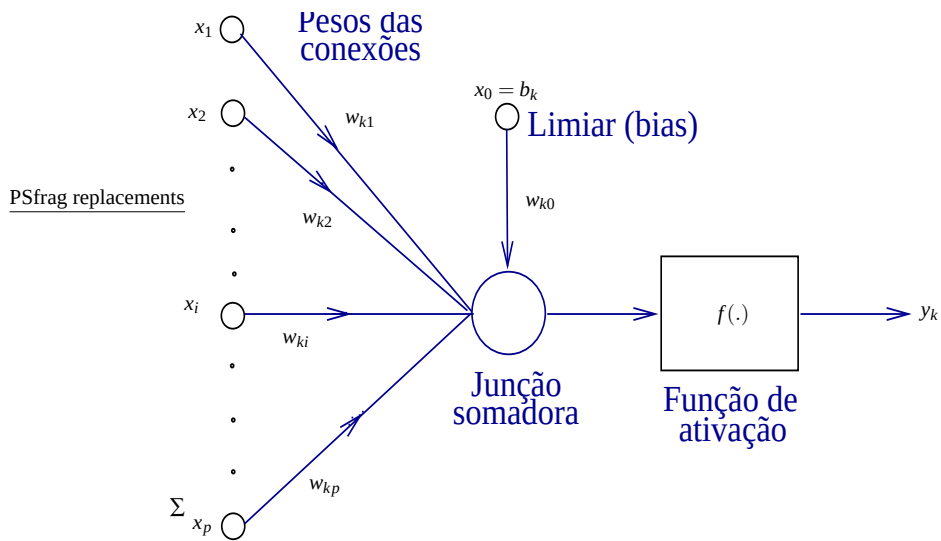


Figura 9: Modelo de neurônio Artificial.

1. Um conjunto de sinapses ou conexões de entrada, sendo cada entrada ponderada por um peso sináptico. Sendo assim, um sinal x_j na entrada da sinapse j conectada ao neurônio k é multiplicado pelo peso sináptico w_{kj} . Quando uma entrada de “bias” está presente (entrada x_0 na Figura 9), o peso sináptico correspondente é denominado peso de *polarização*.
2. Uma função soma, responsável pela combinação aditiva dos sinais de entrada, ponderados pelos respectivos pesos das sinapses do neurônio.
3. Uma função de ativação, geralmente não-linear, representando um efeito de saturação na ativação de saída y_k do neurônio.

Matematicamente, o modelo de neurônio ilustrado na Figura 9 pode ser descrito por:

$$v_k = \sum_{j=0}^p w_{kj} x_j \quad (3.2)$$

$$y_k = f(v_k), \quad (3.3)$$

Sendo v_k o nível de ativação interna do neurônio k . A função de ativação f mais utilizada é a função sigmoideal. Esta função é monotonicamente crescente e apresenta propriedades assintóticas e de suavidade. Um exemplo de função sigmoideal também chamada *função logística* é definida por:

$$f(v) = \frac{1}{1 + e^{-av}} \quad (3.4)$$

Sendo a o parâmetro de inclinação da função sigmoidal.

3.2.3 Estruturas de Redes Neurais Artificiais

As várias estruturas de redes neurais podem ser agrupadas, em duas classes: não-recorrentes e recorrentes.

As RNAs não-recorrentes são aquelas que não possuem realimentação de suas saídas para suas entradas e por isso são também ditas “estáticas”. A estrutura das RNAs ditas não-recorrentes podem ser formadas por uma ou mais camadas de neurônios artificiais.

As RNAs recorrentes são estruturas mais gerais, que contém realimentação das saídas para as entradas, sendo suas saídas determinadas pelas entradas atuais e pelas saídas anteriores. Da mesma forma, que as estruturas não recorrentes, estruturas podem estar organizadas em camadas ou não. Quando apresentam camadas, estas redes podem possuir interligações entre neurônios da mesma camada e entre camadas não consecutivas, gerando interconexões bem mais complexas que as RNAs não-recorrentes.

Neste trabalho, foi utilizado uma estrutura de rede multicamadas não-recorrente para a previsão das séries temporais de carga por barramento.

3.2.4 Tipos de Treinamento

Uma das propriedades das RNAs é a habilidade de aprender. Ao invés de especificar todos os detalhes de um aprendizado, tem-se a possibilidade de treinar uma rede para fazer este aprendizado. Isto significa que se pode tratar problemas onde regras apropriadas são muito difíceis de se conhecer a priori. O objetivo do treinamento de uma RNA é fazer com que a aplicação de um conjunto de entradas produza um conjunto de saídas. Cada conjunto de entrada ou saída é chamado de vetor. O treinamento é realizado pela apresentação dos vetores de entradas, enquanto os pesos da rede são ajustados de acordo com um procedimento de treinamento pré-determinado.

Existem dois procedimentos de treinamento:

1. Supervisionado

O algoritmo supervisionado necessita de um par de vetores composto pelo vetor de entrada e saída desejada ou vetor alvo. O procedimento de treinamento funciona da seguinte forma: os pesos são inicializados e o vetor de entrada é apresentado à rede. A saída da rede é calculada e comparada com o correspondente vetor alvo. O erro encontrado é realimentado através da rede e os pesos são atualizados de acordo com um determinado algoritmo, a fim de minimizar este erro. Este processo de treinamento é repetido até que o erro mínimo de treinamento seja obtido ou seja alcançado um número máximo de iterações.

2. Não-Supervisionado

O algoritmo não supervisionado não requer que as saídas desejadas sejam conhecidas. O processo de treinamento extrai as propriedades estatísticas do conjunto de treinamento agrupando os vetores similares em classes.

Existe uma grande variedade de algoritmos de aprendizado, tanto para a forma supervisionada, como para a não-supervisionada. Entre os algoritmos supervisionados, um dos mais difundidos é o algoritmo de retropropagação do erro (*backpropagation*) e entre os não-supervisionados o algoritmo de Kohonen. Ambos algoritmos são utilizados neste trabalho.

3.2.5 Redes Perceptron Multicamadas - MLP

As arquiteturas do tipo MLP constituem os modelos neurais artificiais mais utilizados e conhecidos atualmente. Tipicamente, esta estrutura consiste de um conjunto de unidades sensoriais que formam uma camada de entrada, uma ou mais camadas intermediárias e uma camada de saída. Os sinais de entrada são propagados camada a camada pela rede.

As redes do tipo MLP têm sido utilizadas com sucesso para a solução de vários problemas envolvendo altos graus de não-linearidade. Um método amplamente utilizado para ajuste dos parâmetros é o algoritmo de retropropagação do erro. Este algoritmo é baseado numa regra de aprendizagem que “corrige” o erro durante o treinamento (Haykin, 1994). Basicamente, o processo de retropropagação do erro é constituído de duas fases: uma fase de propagação do sinal funcional (*feedforward*) e uma de retropropagação do erro (*backpropagation*). Na fase de propagação, os vetores de dados são aplicados às unidades de entrada, e seu efeito se propaga pela rede, camada a camada. Finalmente, um conjunto de saídas é produzido como resposta da rede. Durante esta fase,

os pesos das conexões são mantidos fixos. Na retropropagação do erro, por outro lado, os pesos são ajustados de acordo com uma regra de correção do erro. Especificamente, a resposta da rede em um instante de tempo é subtraída da saída desejada para produzir um sinal de erro. Este sinal de erro é propagado da saída para a entrada, camada a camada, originando o nome “retropropagação do erro”.

O anexo A apresenta o detalhamento da estrutura de rede multicamadas e do algoritmo de retropropagação do erro.

4 *Metodologias de Agrupamento*

Metodologias de agrupamento são técnicas capazes de agrupar os elementos de um dado conjunto em grupos, de modo que os elementos de um mesmo grupo apresentem alguma similaridade. Estas metodologias vêm sendo empregadas em mineração de dados e reconhecimento de padrões (Cardoso, 2003).

Neste trabalho, as metodologias de agrupamento foram utilizadas para identificar barramentos com padrões de demanda diária similares. Os grupos determinados pelos algoritmos de agrupamento servirão de base para o processo de previsão de carga agregada apresentado no Capítulo 5.

Existem diversos métodos de agrupamentos difundidos na literatura. Neste capítulo serão apresentadas as duas metodologias empregadas para tratar o problema de agrupamento de barramentos; o método baseado nas redes de Kohonen e o método *fuzzy c-means*.

4.1 Introdução

Técnicas de agrupamento são empregadas para análise exploratória de dados, auxiliando na resolução de problemas de agrupamento. O objetivo do processo de agrupamento é separar um conjunto de dados em diferentes grupos de modo que os elementos de um dado grupo apresentem um certo grau de similaridade entre eles. O uso dessa técnica é apropriado quando se conhece pouco ou nada sobre a estrutura de um conjunto de dados (Backer, 1995).

Atualmente, diversos modelos de agrupamento de dados são utilizados no tratamento de séries temporais, sendo a maioria destes híbridos (Bose e Mahapatra, 2001). Segundo (Haykin, 1994) o mapa auto-organizável de Kohonen leva a uma representação organizada e estável de padrões apresentados no espaço de entrada.

O objetivo dos algoritmos de agrupamento é identificar duas ou mais coleções de dados que formam uma estrutura de grupo e que possam ser associadas a classes. Em sistemas cujos dados contém um grande número de atributos ao particionar a base de dados em um conjunto de grupos, proporciona-se um mecanismo de sumarização e compactação de informação (Cardoso, 2003).

Neste trabalho, o interesse pelas técnicas de agrupamento é obter o agrupamento de um conjunto de barramentos analisando o comportamento histórico de demanda diária de carga de cada barramento. O objetivo do agrupamento é determinar quais são os barramentos do sistema que possuem perfis de consumo semelhantes.

As metodologias aplicadas ao agrupamento de barramentos neste trabalho são: o algoritmo *Pruning Technique Kohonen Self-Organizing Feature Map - PSOM* proposto por (Castro e VonZuben, 1999) e o algoritmo *Fuzzy C-means - FCM* introduzido por (Bezdek, 1981). Ambos são detalhados no decorrer deste Capítulo.

4.2 Definição

A definição formal do problema de agrupamento encontrada em (Hruschka e Ebecken, 2001) é:

Considerando o agrupamento de um conjunto de n objetos $X = \{X_1, X_2, \dots, X_n\}$ onde cada $X_i \in \mathbb{R}^p$ é um vetor de p valores reais que dimensionam as características do objeto, estes devem ser agrupados em grupos disjuntos $G = \{G_1, G_2, \dots, G_k\}$, sendo k o número de grupos, de forma que as seguintes condições sejam atendidas:

1. $G_1 \cup G_2 \cup \dots \cup G_k = X$;
2. $G_i \neq \emptyset, \forall i, 1 \leq i \leq k$;
3. $G_i \cap G_j = \emptyset, \forall i \neq j, 1 \leq i, j \leq k$;

Por essas condições, um objeto não pode pertencer a mais de um grupo (grupos disjuntos) e cada grupo tem que ter ao menos um objeto. (Cole, 1998) ainda acrescenta que o valor de k pode ser desconhecido. Se k é conhecido, o problema é referido como o problema de *k-clustering*.

4.2.1 Dificuldades no Problema de Agrupamento

Encontrar o melhor agrupamento para um conjunto de objetos não é uma tarefa simples. (Hruschka e Ebecken, 2001) destacam que o problema de encontrar o melhor agrupamento é NP-completo e não é computacionalmente possível encontrá-lo, a não ser que o número de objetos n e o número de grupos k sejam pequenos, visto que o número de partições distintas em que se pode dividir n objetos em k grupos aumenta aproximadamente como:

$$\frac{k^n}{k!}$$

(Cole, 1998) mostra que existem cerca de 2.436.684.974.110.751 maneiras possíveis para agrupar 25 objetos em 5 grupos. Se o número de grupos é desconhecido, é necessário somar todas as partições possíveis para cada número de grupos entre 1 e 5 o que fornece um valor superior a 4×10^{18} partições possíveis. Para n objetos particionados em k grupos tem-se:

$$N(n, k) = \frac{1}{k!} \left(\sum_{i=0}^k (-1)^i \binom{k}{i} (k-i)^n \right) \quad (4.1)$$

Para agrupar os dados de 1 a k grupos tem-se:

$$\sum_{i=n}^1 N(n, k) \quad (4.2)$$

(Ankerst et al., 1999) escrevem que existem algumas razões para explicar porque o desempenho dos algoritmos de agrupamento é complexo. Primeiro, quase todos os algoritmos de agrupamento requerem valores para os parâmetros de entrada que são difíceis de determinar, especialmente para conjuntos de dados do mundo real contendo objetos com muitos atributos. Segundo, os algoritmos são muito sensíveis a estes parâmetros, podendo produzir partições muito diferentes para um mesmo conjunto de dados, dependendo dos ajustes destes parâmetros.

4.3 Mapas Auto-Organizáveis de Kohonen (SOM)

A rede de kohonen é uma rede onde os neurônios (ou unidades) vizinhos competem entre si. Nesta seção, apresenta-se características da rede proposta por (Kohonen, 1982).

4.4 Definição SOM

Os SOM fazem parte de um grupo de redes neurais normalmente chamadas de redes baseadas em competição, ou simplesmente redes competitivas (Cho, 1997) (Kaski e Kohonen, 1995). Estas redes combinam competição com a estratégia de um aprendizado não supervisionado para fazer o ajuste dos pesos, baseado nas semelhanças dos padrões de entrada. O principal objetivo dos mapas auto-organizáveis de Kohonen é agrupar os dados de entrada que são semelhantes entre si formando classes ou agrupamentos.

Em uma rede de classificação há uma unidade para cada componente do vetor de entrada. Cada unidade de saída representa um agrupamento, limitando o número de grupos ao número de unidades de saída. A Figura 10 apresenta as arquiteturas típicas de um SOM.

Durante o processo de treinamento, a rede determina a unidade de saída que está mais próxima do vetor dos padrões de entrada, e tal unidade é chamada de “unidade vencedora”. A unidade vencedora tem o seu vetor de pesos ajustados de acordo com o algoritmo de aprendizagem. As unidades dentro uma certa vizinhança ξ da “unidade vencedora” têm o seus pesos ajustados da mesma forma que a unidade vencedora.

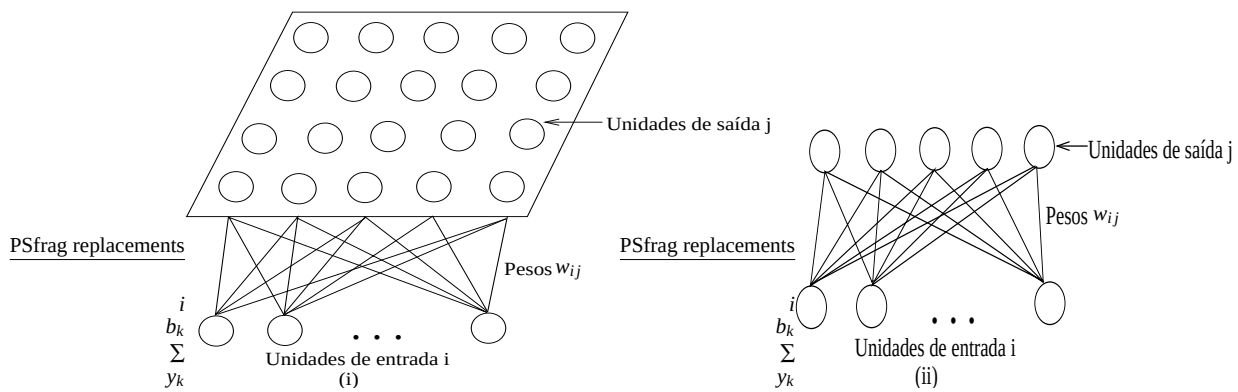


Figura 10: Arquiteturas da rede de Kohonen: (i)Arquitetura bidimensional (ii)Arquitetura unidimensional.

4.4.0.1 Algoritmo de treinamento

O primeiro passo do algoritmo é inicializar os valores dos pesos das conexões w_{ij} ($i = 1, \dots, n$ e $j = 1, \dots, m$) e uma vizinhança N_c . De acordo com (Kohonen, 2000) os pesos das conexões devem ser inicializados com valores pequenos e aleatórios. A vizinhança N_c é um fa-

tor importante no momento da atualização dos pesos. Neste algoritmo, a taxa de aprendizagem decresce lentamente com o tempo.

No processo de treinamento, N_c é decrementado após um determinado número de iterações, até que N_c seja igual a zero. Assim, diz-se que ocorreu uma iteração quando o vetor de padrões $(x_1, x_2, \dots, x_n)$ tiver sido apresentado uma vez.

O procedimento detalhado dentro de uma iteração acontece da seguinte maneira: depois de especificado os pesos iniciais das conexões, o valor de ativação de cada nó de saída pode ser computado. O neurônio j com o máximo valor de ativação é selecionado como a “unidade vencedora”, ou seja, o neurônio vencedor é o mais próximo do vetor de entrada segundo uma medida de distância. Esta unidade e as respectivas unidades vizinhas são atualizadas. O procedimento acima é repetido para cada padrão de entrada. Quando todos os n vetores de entrada tiverem sido apresentadas, diz-se que uma iteração está completa (Castro, 1998). Um pseudo-código do algoritmo SOM, é apresentado no Quadro 1 (Castro, 1998):

Quadro 1. Pseudo-código: algoritmo SOM.

1. Inicialização e definições de parâmetros

*Inicialize os pesos w_{ij} , $j = 1, \dots, m$ e $i = 1, \dots, n$
 Defina os parâmetros de vizinhança.
 Defina os parâmetros de aprendizagem.*

2. Enquanto a condição de parada não é satisfeita, faça:

2.1. Para cada j calcule:

$$2.1.1. D(j) = \operatorname{argmin} \{ \|w_{ji} - x_i\| \}.$$

2.1.3. Para $\forall j \in N_c(j)$, e $\forall i$:

$$w_{ij}(\text{novo}) = w_{ij}(\text{antigo}) + \alpha [x_i - w_{ij}(\text{antigo})].$$

2.2. Atualize a taxa de aprendizagem.

2.3. Reduza o raio de vizinhança.

4.5 Técnica de Poda do Mapa Autorganizável de Kohonen - PSOM

Esta seção apresenta a descrição do método de poda desenvolvido por (Castro e VonZuben, 1999), para as redes auto-organizáveis de Kohonen. (Castro e VonZuben, 1999) desenvolveram um algoritmo simplificado cujo objetivo principal é a eliminação das unidades de saída excedentes do método SOM, permitindo assim a determinação de uma arquitetura mínima capaz de representar o conjunto de dados.

4.5.1 Definição - PSOM

O PSOM é uma técnica capaz de podar as unidades que não são relevantes na arquitetura de um SOM, isto é, eliminar unidades que nunca vencem na competição. Este método tem por objetivo reduzir a dimensão topológica do mapa gerado pelo algoritmo de treinamento original, e suas aplicações têm ênfase em problemas de agrupamentos de padrões.

Quando unidades irrelevantes de uma rede neural artificial são eliminadas duas perguntas básicas surgem:

- Qual unidade será podada?
- Qual é o comportamento da rede (o que fazer) depois de feita a poda?

A meta é achar e podar as unidades de saída que não apresentam um número mínimo, pré-especificado, de padrões. Uma medida de agrupamento CM é definida para avaliar o grau representativo de cada unidade de saída. Essas unidades na qual a medida de agrupamento CM é inferior a um limiar pré-especificado ξ são retiradas uma-a-uma, juntamente com todas as suas conexões.

Por outro lado, se uma determinada unidade de saída possui uma medida de agrupamento CM muito alta, isto é, seja superior a um valor pré-definido ψ , um termo de penalização τ pode ser adicionado à medida de agrupamento desta unidade forçando o CM desta unidade a diminuir. Desta forma, aumenta-se a probabilidade de outras unidades estarem aptas a classificar parte do conjunto evitando excessos no processo de poda.

Após avaliada a medida de agrupamento de cada unidade de saída e feita a poda de uma destas unidades, alguns fatores devem ser considerados. Geralmente, os critérios de parada dos mapas auto-organizáveis são o limite de iterações ou o valor mínimo da taxa de aprendizagem α , já que esta decresce durante o processo de aprendizagem. O procedimento de poda deve ser um procedimento retardado, ou seja, a rede é treinada da maneira usual durante algumas iterações, e assim que o mapa apresente uma topologia pré-definida, após esta etapa inicia-se a eliminação de unidades cujo CM é pequeno. Então, uma unidade é retirada em fase adiantada do treinamento e o valor reduzido da taxa de aprendizagem permitirá pequenos ajustes no vetor de pesos. Este argumento evidencia a necessidade de uma reinicialização do valor do passo α e do raio de vizinhança N_c sempre que uma unidade for retirada (Castro, 1998).

A poda implica na retirada da unidade e reinicialização dos parâmetros de treinamento, tendo como condições iniciais, por exemplo, o estado anterior do processo. A condição de parada sugerida é um valor mínimo α_{min} para a taxa de aprendizagem α .

O parâmetro ξ é um valor de porcentagem, isto é, se uma unidade não faz a representação de pelo menos $\xi\%$ dos dados então, esta unidade é uma candidata para poda.

Para conjunto de dados pequeno, o número inicial de unidades de saída da rede pode ser tomado como sendo o número de padrões disponíveis. Esta escolha pode ser explicada pelo fato que, em um conjunto de dados com n vetores, não pode haver mais que n grupos diferentes. Se o conjunto de dados é muito grande, esta estratégia não é recomendada devido ao alto esforço computacional aplicado ao processo de treinamento.

Algumas das desvantagens do método proposto são a quantidade de parâmetros a serem arbitrados e um processo que inicia com uma arquitetura grande e tenta reduzi-la durante a adaptação.

O Quadro 2 descreve o algoritmo de treinamento PSOM.

Quadro 2. Pseudo-código: algoritmo PSOM.

Seja X , o vetor de entrada de Dados.

1. Inicialização e definição dos parâmetros:

Inicialize os pesos w_{ij} aleatoriamente, $j = 1, \dots, m$ e $i = 1, \dots, n$.

Definir a vizinhança N_{c0} e a taxa de aprendizagem α_0 .

Definir os parâmetros de penalização τ , poda ξ e avaliação do CM ψ .

Definir o número mínimo de saídas $m_{\min}$, retardo β e o N_{epocas} .

2. Enquanto o critério de parada não for satisfeito, faça:

2.1. Para cada j determine:

$$2.1.1 \ D(j) = \operatorname{argmin} \{ \|w_{ij} - x_i\| \}$$

2.1.2 Incrementar o fator CM correspondente a j .

2.1.3. $\forall j \in N_c(j)$, e $\forall i$:

$$w_{ij}(\text{novo}) = w_{ij}(\text{antigo}) + \alpha [x_i - w_{ij}(\text{antigo})]$$

2.2 Atualize a taxa de aprendizagem α .

2.3 Se $(m > m_{\min})$ e $(N_{\text{epocas}} > \beta)$

2.3.1. Poda a unidade que classifica menos que $\xi\%$ amostras.

2.3.2. Restaure $\alpha = \alpha_0$ e $N_c = N_{c0}$.

2.3 Caso contrário mantenha todas as unidades.

2.4 Se $CM(j) > \psi$, $j = 1, \dots, N$. puna-o com τ .

3 Decresça N_c e α .

4.6 Fuzzy C-Means

Nesta seção, será apresentada a metodologia *Fuzzy C-Means* proposta por (Bezdek, 1981). Este algoritmo de agrupamento de dados nebuloso baseia-se em um modelo não linear que agrupa dados com características similares de acordo com um procedimento iterativo de otimização de uma função objetivo que representa um critério de partição, ponderado pelos graus de pertinência dos dados aos respectivos grupos.

4.6.1 Noções Gerais

Dado um conjunto finito de dados X , o problema de agrupamento consiste em atribuir classes aos elementos de maneira que essas classes identifiquem grupos naturais em X . Essas classes devem formar uma partição de X ; indivíduos que pertencem ao mesmo bloco (classe) da partição estão totalmente relacionados (são indistinguíveis) e indivíduos que pertencem a blocos distintos não estão relacionados. Em outras palavras, indivíduos estão totalmente relacionados entre si ou não estão relacionados. No entanto, em muitas aplicações práticas uma partição rígida pode ser muito restritiva e inviável, devido, muitas vezes, à imprecisão dos dados.

Segundo (Bezdek e Pal, 1992), a imprecisão nos dados pode surgir de diversas fontes. Por exemplo, erros em instrumentos ou ruídos no experimento podem levar a valores parcialmente confiáveis de determinados atributos. Em alguns casos, o custo envolvido na extração de valores muito precisos de um atributo pode ser alto. Em outros casos, pode ser difícil decidir quais são os atributos mais relevantes que caracterizam um elemento. Por essas razões, torna-se conveniente o uso de variáveis lingüísticas e limitações para descrever os valores de atributos, em vez de tentar fornecer uma representação numérica exata para os dados com valores incertos dos atributos.

A incerteza no processo de agrupamento de dados pode provir também da sobreposição de várias classes. Geralmente, assume-se que um elemento pertence a apenas uma classe, o que nem sempre se verifica em domínios de dados reais, nos quais elementos pertencem a mais de uma classe, com diferentes graus de pertinência.

Uma alternativa viável para o agrupamento de dados com informações imprecisas é a utilização de conjuntos *fuzzy*. Segundo (Klir e Yuan, 1995), conjuntos *fuzzy* podem ser utilizados em pelo menos dois níveis para o problema de agrupamento:

- No nível de atributos, para representar os elementos do conjunto como um vetor de graus de pertinência, sendo que cada um desses graus representa o grau de “posse” do atributo em questão, por parte desses elementos;
- No nível de agrupamento, para representar a pertinência desses elementos às classes, bem como para prover uma estimativa das informações incompletas em termos de valores de pertinência.

Com a utilização da abordagem *fuzzy*, o problema passa então a ser caracterizado como um problema de agrupamento *fuzzy*, cujo objetivo é a obtenção de uma partição *fuzzy* em um conjunto de dados X . Em ambos os casos, permite-se diferentes graus de relacionamento entre elementos do conjunto, sendo que um elemento pode pertencer a mais de uma classe.

4.6.2 Algoritmo FCM

O algoritmo *Fuzzy C-Means* - FCM, é um dos mais atrativos para o agrupamento de dados nebulosos, devido principalmente à sua simplicidade e eficiência. Este algoritmo é baseado na otimização iterativa de uma função objetivo que traduz um critério de partição ponderado por graus de pertinência dos dados aos respectivos grupos.

O algoritmo FCM, apresentado no Quadro 3, utiliza os seguintes conceitos e notações: Seja $X = \{x_1, x_2, \dots, x_p\}$ um conjunto finito de dados, genericamente chamados de pontos de $\Re^n$, tal que $\Re^n$ é um espaço Euclidiano n -dimensional; c o número de grupos, sendo $2 \leq c \leq p$; U a matriz de pertinência $c \times p$, onde u_{ik} , com $1 \leq i \leq c$ e $1 \leq k \leq p$, o grau de pertinência do ponto x_k ao grupo i . Uma c -partição de X é definida por:

$$M_{fc} = \left\{ U \mid u_{ik} \in [0, 1] \forall i, k; \sum_{i=1}^c u_{ik} = 1, \forall k; 0 < \sum_{k=1}^p u_{ik} < p, \forall i \right\} \quad (4.3)$$

O algoritmo FCM procura agrupar os dados através de um procedimento que tem como objetivo determinar o mínimo da função:

$$J_m(U, V) = \sum_{k=1}^p \sum_{i=1}^c u_{ik}^m d_{ik}^2, U \in M_{fc}, 1 < m < \infty \quad (4.4)$$

sendo $d = \|x_k - v_i\|_A$ a distância entre x_k e v_i , $\|\cdot\|$ é uma norma induzida por um produto interno,

como por exemplo $\|x\|_A^2 = x^T A x$, tal que A é uma matriz $n \times n$ definida positiva, $V = \{v_1, v_2, \dots, v_p\}$ é o conjunto de centros do grupo, representado por uma matriz $c \times n$, sendo que cada $v_i \in \mathfrak{R}^n$, $1 \leq i \leq c$ é chamado de centro do i -ésimo grupo e m é o fator que define o grau de nebulosidade da partição nebulosa do sistema.

Quadro 3. Pseudo-código: algoritmo FCM.

Dado X , escolher o número de grupos $2 \leq c \leq p$, o parâmetro $m > 1$, o critério de parada $\varepsilon > 0$ e o número máximo de iterações $lmax$.

1. Inicializar $U^{(0)}$ e o contador de iterações $l = 1$.
2. Calcular os c centros dos grupos $\{v_1^l, v_2^l, \dots, v_p^l\}$ usando $U^{(l)}$, com a seguinte equação:

$$v_i^l = \frac{(\sum_{k=1}^p (u_{ik}^{l-1})^m x_k)}{(\sum_{k=1}^p (u_{ik}^{l-1})^m)}$$

3. Atualizar $U^{(l)}$ com o seguinte procedimento:

- 3.1. Para $1 \leq k \leq p$

- 3.1.1. Se $\|x_k - v_i^l\|^2 > 0$

$$u_{ik}^l = \left[\sum_{j=1}^c \left(\frac{\|x_k - v_i^l\|^2}{\|x_k - v_j^l\|^2} \right)^{\frac{1}{m-1}} \right]^{-1}, i = 1, 2, \dots, c.$$

- 3.1.2. Se $\|x_k - v_i^l\|^2 = 0$

$$u_{ik}^l = 1, 1 < i < c.$$

4. Calcular $\Delta = \|U^l - U^{l-1}\| = \max_{i,j} |u_{ij}^l - u_{ij}^{l-1}|$

- 4.1. Se $\Delta > \varepsilon$ ou $l < lmax$,
 $l = l + 1$ e voltar ao passo 2.

- 4.2. Caso contrário parar.
-

4.7 Resultados Obtidos

O agrupamento do conjunto de barramentos é relevante no sentido de determinar quais são os barramentos do sistema que apresentam um perfil de consumo semelhante no decorrer de um espaço de tempo.

Tradicionalmente, é usual considerar três classes de consumidores: residencial, comercial e industrial. Porém, como em cada uma das classes o perfil de demanda horário pode não ser único então foram testadas classificações considerando o número de grupos maiores que três. Nesta pesquisa, serão apresentados os resultados do algoritmo de agrupamento configurado para segregar o conjunto inicial de barramentos em 4 subconjuntos (ou 4 grupos). De um modo geral, notou-se que quando o número de grupos é menor, o grau de similaridade dos grupos tende a ser menor, e por outro lado, um número muito grande de grupos tende a formar grupos com poucos elementos.

A equação (4.5) apresenta a medida de similaridade utilizada a qual é dada pelo desvio das barramentos do grupo em relação a uma curva média dos barramentos:

$$Md = \frac{1}{n} \sum_{i=1}^n |(\|b(i)\| - \|mc(g)\|)| \quad (4.5)$$

Sendo que:

- Md : medida de distância;
- $b(i)$: barramento pertencente a um dado grupo g ;
- mc : curva média das barramentos do grupo g ;
- n : número de barramentos no grupo g .

Os algoritmos de agrupamento foram testados com os dados descritos no Capítulo 2 e os resultados alcançados apresentam-se na seguinte seqüência:

1. Agrupamento de um dia

Neste estudo, o conjunto de barramentos obtido no dia 19/09/2001 foi submetido aos algoritmos de agrupamento e o objetivo foi verificar o resultado dos agrupamentos para apenas um dia da semana.

2. Agrupamento de uma semana

Nesta etapa, o conjunto de dados utilizados para testar os algoritmos foram os dados da semana do dia 19 a 25/09/2001. Para cada dia desta semana o conjunto de barramentos foi submetido aos algoritmos de agrupamento. O objetivo foi verificar se os agrupamentos tendem a se manter ao longo da semana.

3. Agrupamento de 44 dias realizado no período de 01/06/2001 a 14/07/2001.

Neste último estágio, foram utilizados os dados dos dias de 01/06 a 14/07/2001, totalizando 44 dias corridos. Em cada um dos 44 dias, o conjunto de barramentos foi submetido aos algoritmos de agrupamento para verificar se a formação dos agrupamentos apresentava grandes variações ao longo do tempo. O objetivo desta análise foi verificar se os agrupamentos formados eram regulares ao longo do tempo.

4.7.1 Agrupamento de um dia

Como dito anteriormente, inicialmente o agrupamento foi feito em um dia apenas. Para o dia analisado foi montado uma única série temporal contendo as séries históricas de cada um dos 15 barramentos.

As Figuras 11 e 12 mostram os agrupamentos formados pelo algoritmo FCM e PSOM para o dia 19/09/2001 (quarta-feira), respectivamente.

Pode-se notar observando as Figuras 11 e 12 que no grupo 1 ambos os algoritmos apresentaram o mesmo agrupamento, evidenciando que os barramentos deste grupo possuem comportamento similar e independente do método de agrupamento, pois tais barramentos estarão sempre no mesmo agrupamento. Houve variações entre os agrupamentos nos grupos 2 e 3, onde as curvas foram agrupadas ou no grupo 2 ou no grupo 3 dependendo do algoritmo utilizado.

Através da medida de similaridade Md , pode-se notar que o algoritmo PSOM apresentou um desempenho melhor já que para o grupo 3 a medida Md foi menor que no algoritmo FCM. O que pode se comprovado pelas Figuras 11 e 12.

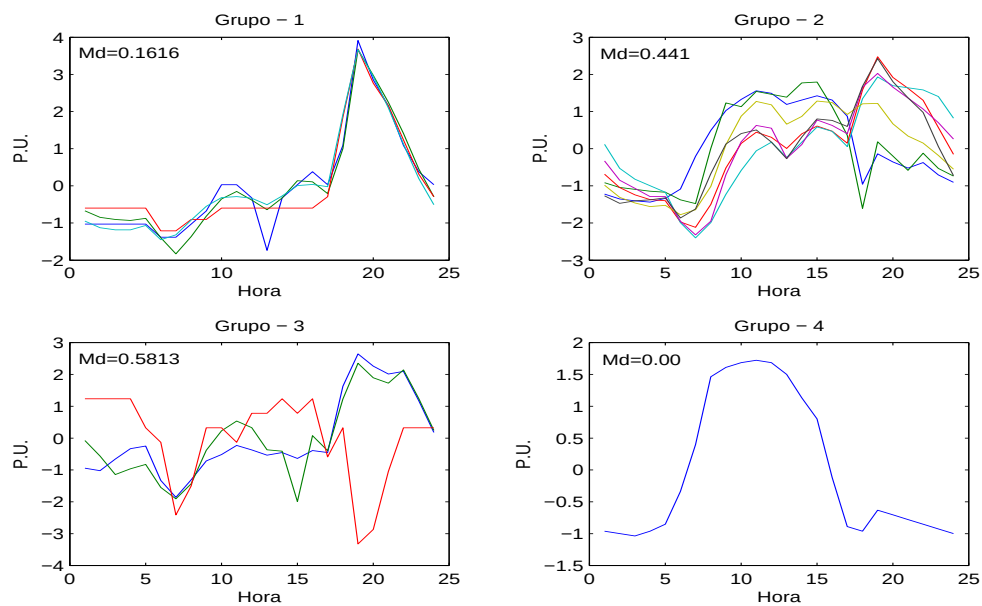


Figura 11: Agrupamentos FCM 19/09/2001 - Quarta-Feira.

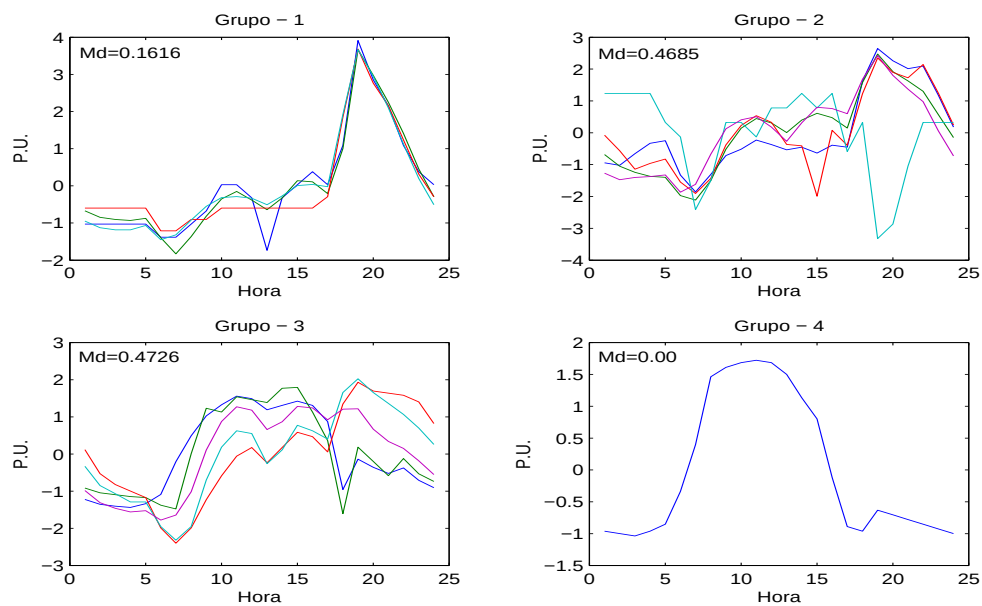


Figura 12: Agrupamentos PSOM 19/09/2001 - Quarta-Feira.

Com relação à medida Md nota-se que, em ambas as metodologias, o grupo 1 possui menor valor, quando comparados aos outros grupos, o que mostra que as curvas deste grupo estão mais próximas em relação à curva média. Visualmente falando, este grupo é mais regular que os demais e, logicamente, possui o menor valor Md como mostrado nas Figuras. O valor Md nos demais grupos é maior, mostrando que as curvas são menos similares, o grupo 4 apresenta $Md = 0,00$ porque este grupo contém apenas uma curva.

4.7.2 Agrupamento de uma semana

Os algoritmos foram testados e configurados para fornecerem os agrupamentos durante uma semana, no período de 19/09/2001 a 25/09/2001. Para cada dia desta semana foi montado uma única série temporal contendo as séries histórias de cada um dos 15 barramentos dia-a-dia.

As Tabelas 2 e 3 mostram os agrupamentos obtidos através do FCM e PSOM, e percebe-se que alguns agrupamentos se mantêm ao longo dos dias em ambos os algoritmos no entanto, na maioria dos casos, os algoritmos apresentam diferenças na formação dos grupos.

Nota-se que os agrupamentos formados nos finais de semana não possuem grande divergência com relação aos subconjuntos formados durante a semana, mostrando uma regularidade dos barramentos ao longo da semana.

Tabela 2: Agrupamentos FCM 19 a 25/09/2001.

	Grupo - 1	Grupo - 2	Grupo - 3	Grupo - 4
Quarta - 19/09	1 2 12 13	3 6 7 8 9 11 14	4 5 10	15
Quinta - 20/09	1 2 12	3 6 7 9 11 13 14	4 5 8 10	15
Sexta - 21/09	1 2 12 13	3 6 7 8 9	4 5 10	11 14 15
Sábado - 22/09	1 2 12	3 9 13	4 5 6 7 8 10	11 14 15
Domingo - 23/09	1 2 7 12	3 8 9 13	4 5 6 10 11 14	15
Segunda - 24/09	1 2 12	3 7 9 11 13	4 5 6 8 10	14 15
Terça - 25/09	1 2 12	3 6 7 8 9 13 14	4 5 10 11	15
Md (p.u.)	0,17	0,34	0,52	0,05

Da aplicação dos algoritmos de agrupamento aos dados de carga por barramento disponíveis, não se observou uma superioridade de nenhum dos métodos. Embora os resultados não sejam iguais, os grupos obtidos foram bastante similares.

É importante destacar que os algoritmos não chegam a um mesmo agrupamento, e um mesmo algoritmo pode agrupar os dados de maneira distinta devido à inicialização aleatória dos parâmetros em ambos os métodos utilizados neste trabalho.

Tabela 3: Agrupamentos PSOM 19 a 25/09/2001.

	Grupo - 1	Grupo - 2	Grupo - 3	Grupo - 4
Quarta - 19/09	1 2 12 13	3 4 5 9 10	6 7 8 11 14	15
Quinta - 20/09	1 2 12	3 4 5 6 7 8 9 10 13	11 14	15
Sexta - 21/09	1 2 12 13	3 4 5 6 7 8 9 10	11 14	15
Sábado - 22/09	1 2 3 12 13	4 5 6 7 8 9 10	11 14	15
Domingo - 23/09	1 2 12 13	4 5 6 8 9 10 11 13 14	7	15
Segunda - 24/09	1 2 12 13	3 4 5 6 7 8 9 10	11 14	15
Terça - 25/09	1 2 12	3 6 7 9 13	4 5 8 10 11 14	15
Md (p.u.)	0,17	0,38	0,26	0,0

Com relação à medida Md nota-se que ao longo da semana o grupo 1 possui menor valor na metodologia PSOM. No algoritmo FCM nota-se que o grupo 4 apresentou menor Md mas deve-se levar em conta que durante alguns dias da semana o barramento 15 ou estava isolado ou era agrupado juntamente com os barramentos 11 e 14. Entretanto estes barramentos (11 e 14) apresentam um grande desvio padrão quando comparados aos demais (Tabela 1, Capítulo 2). Desta forma, assim como no PSOM o grupo 1 possui maior regularidade que os demais grupos.

O valor Md nos demais grupos é maior, mostrando que as curvas são menos similares, o grupo 4 apresenta $Md = 0,00$ porque este grupo contém apenas uma curva. O resultados dos agrupamentos podem ser comparados com os resultados das Tabelas e Figuras do Capítulo 2.

4.7.3 Agrupamento de 44 dias

O algoritmo PSOM foi utilizado para formar os grupos das curvas de carga durante um período de 44 dias, 01/06/2001 a 14/07/2001, em cada dia deste período foi montado uma única série temporal contendo as séries histórias de cada um dos 15 barramentos dia-a-dia. A Tabela 4 mostra o número de vezes que determinados barramentos aparecem juntos uns aos outros após o agrupamento realizado nos 44 dias.

Na Tabela 4 estão relacionados os barramentos que foram agrupados no período descrito acima. Nota-se que, fixando o barramento 1 e observando ao longo de sua linha pode-se ver o número de vezes que os barramentos foram associados ao barramento fixo. Por exemplo, o barramento 2 é agrupado juntamente com o barramento 1 vinte e uma vezes (isto é, em 21 dias) durante os 44 dias, por outro lado o barramento 3 em nenhum dia foi agrupado juntamente com o barramento 1.

Pode-se observar que alguns barramentos sempre são agrupados juntos com outros,

indicando que estes sempre se comportam de maneira similar. Já outros barramentos aparecem menos vezes juntos, mostrando uma menor similaridade ao longo dos 44 dias.

Ainda na Tabela 4 nota-se a formação de diferentes grupos em função do número de vezes que os barramentos foram agrupados juntos. Por exemplo observando a linha do barramento 1, pode-se formar um primeiro grupo com os barramentos 1, 10, 12 e 13, pois estes foram agrupados juntos durante os 44 dias. O segundo grupo pode ser formado acrescentando-se ao primeiro grupo o barramento 2, pois este é agrupado junto 21 vezes. Desta forma, os outros grupos foram formados acrescentando-se nos grupos os barramentos menos similares ao barramento 1.

Tabela 4: Agrupamentos de Barramentos.

Barramentos	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	0	21	0	0	3	4	5	11	0	44	0	44	44	0	0
2	21	0	3	10	26	27	27	34	2	21	1	21	21	0	0
3	0	3	0	34	13	12	13	7	42	0	21	0	0	1	0
4	0	10	34	0	23	22	23	16	32	0	16	0	0	0	0
5	3	26	13	23	0	39	37	34	11	3	2	3	3	1	0
6	4	27	12	22	39	0	38	37	10	4	2	4	4	1	0
7	5	27	13	23	37	38	0	35	12	5	2	5	5	0	0
8	11	34	7	16	34	37	35	0	5	11	1	11	11	0	0
9	0	2	42	32	11	10	12	5	0	0	23	0	0	2	0
10	44	21	0	0	3	4	5	11	0	0	0	44	44	0	0
11	0	1	21	16	2	2	2	1	23	0	0	0	0	18	5
12	44	21	0	0	3	4	5	11	0	44	0	0	44	0	0
13	44	21	0	0	3	4	5	11	0	44	0	44	0	0	0
14	0	0	1	0	1	1	0	0	2	0	18	0	0	0	31
15	0	0	0	0	0	0	0	0	0	0	5	0	0	31	0

A Tabela 5 apresenta todos os grupos formados quando fixou-se o barramento 1. Para cada grupo calculou-se o desvio padrão. Como era esperado, nota-se que à medida em que os grupos aumentam (isto é, são acrescentados barramentos menos similares) a variabilidade também aumenta, ou seja, as cargas que compõem o grupo tornam-se menos similares. A Figura 13 mostra o comportamento do desvio padrão que cresce à medida que os grupos aumentam.

Pelas Tabelas 4 e 5, nota-se que os barramentos analisados possuem um comportamento regular ao longo de um grande período de tempo mostrando que estes não apresentam grandes variações de perfil. Existem alguns elementos menos regulares e que sempre vão estar agrupados em grupos diferentes dependendo do dia. Este tipo de comportamento está intimamente ligado ao tipo de barramento e a classe de consumidores que ele atende (residencial, industrial, comercial).

Tabela 5: Grupos de Barramentos (Fixando Barra 1).

	Grupo 1	Grupo 2	Grupo 3	Grupo 4	Grupo 5	Grupo 6
Barramento - N ^o vezes	1 - 44X	1 - 44X	1 - 44X	1 - 44X	1 - 44X	1 - 44X
	10 - 44X	10 - 44X	10 - 44X	10 - 44X	10 - 44X	10 - 44X
	12 - 44X	12 - 44X	12 - 44X	12 - 44X	12 - 44X	12 - 44X
	13 - 44X	13 - 44X	13 - 44X	13 - 44X	13 - 44X	13 - 44X
	***	2 - 21X	2 - 21X	2 - 21X	2 - 21X	2 - 21X
	***	***	8 - 11X	8 - 11X	8 - 11X	8 - 11X
	***	***	***	7 - 5X	7 - 5X	7 - 5X
	***	***	***	***	6 - 4X	6 - 4X
	***	***	***	***	***	5 - 3X
<i>Md (p.u.)</i>	0,0138	0,0245	0,0661	0,0699	0,0707	0,1254

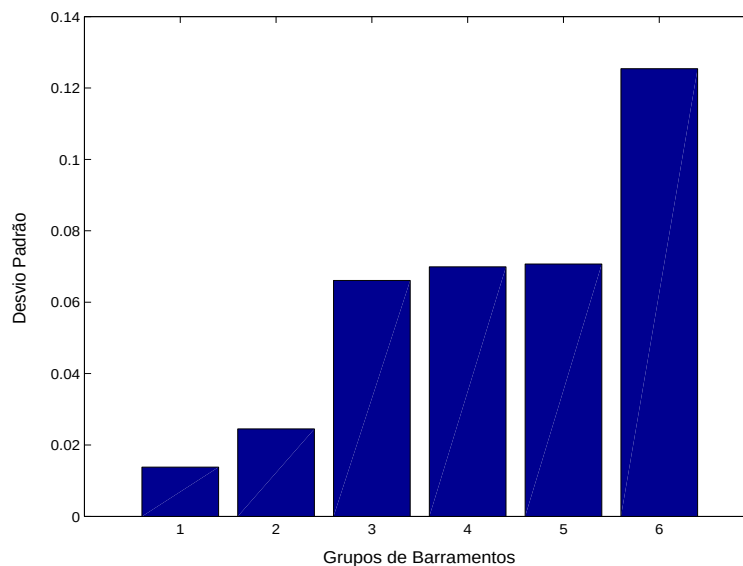


Figura 13: Desvio Padrão nos Grupos (fixando barramento 1).

4.8 Conclusões

Pelos resultados acima pôde-se concluir que as metodologias de agrupamento são eficientes na identificação de barramentos com perfis de demanda de carga semelhantes. Evidentemente, mesmo estando ligada ao histórico de cargas dos barramentos, a medida de distância *Md* foi eficiente para verificação dos grupos mais regulares.

No que diz respeito às técnicas de agrupamento FCM e PSOM, percebe-se que quando aplicadas ao problema de agrupamento as duas metodologias apresentaram resultados similares, pois os agrupamentos obtidos foram poucos divergentes na comparação das duas técnicas.

5 *Modelos de Previsão e Resultados*

As metodologias de previsão apresentadas no Capítulo 3 juntamente com as técnicas de agrupamento, expostas no Capítulo 4, serão utilizadas para o desenvolvimento dos modelos de previsão de carga por barramento. Assim, o objetivo principal deste capítulo é introduzir os modelos de previsão de carga por barramento, bem como mostrar os resultados obtidos quando tais modelos são aplicados a dados de um sistema real.

5.1 Introdução

Quando se fala em previsão de carga por barramento a questão que surge é a seguinte: “*Como realizar uma previsão de boa qualidade e rápida que atenda a todas os barramentos do sistema?*” Esta pergunta tem várias possibilidades de respostas, pois existem várias técnicas de previsão que podem ser aplicadas à previsão por barramento; entretanto, este problema de previsão é um problema bastante complexo.

A demanda de cada barramento constitui uma série temporal e a partir desta pode-se estimar sua carga futura através de modelos de previsão de carga tradicionais. O inconveniente desta alternativa é o grande número de previsões necessárias devido à quantidade de barramentos existentes num sistema de transmissão/distribuição de energia elétrica.

Neste cenário há em princípio duas alternativas de previsão as metodologias de previsão individuais (isto é, a previsão é realizada para cada barramento), e as metodologias agregadas, os quais fariam uma previsão baseada na carga agregada, e a partir desta previsão seriam obtidas as cargas individuais. A partir destas duas abordagens é possível gerar muitas variações que vão desde os casos de previsão individual semanal, até a de previsão agregada pontual.

5.2 Modelos de Previsão

Como comentado anteriormente, para resolver o problema de previsão de carga por barramento foram desenvolvidas duas metodologias de previsão:

- **Metodologias de Previsão Individual (MPI)** - O modelo individual trata a série temporal de cada barramento separadamente. O objetivo da alternativa individual é ajustar um modelo de previsão para cada barramento. O fator negativo desta alternativa é o fato de serem necessárias uma previsão para cada barramento e, neste caso, se o sistema possui um grande número de barramentos estas previsões necessitarão de um alto custo computacional, pois seria necessário ajustar um previsor para cada barramento. O fator positivo é que os modelos são específicos para cada barramento de modo que os seus resultados tendem a ser melhores quando comparados a modelos agregados.
- **Metodologias de Previsão Agregada (MPA)** - O modelo agregado trabalha com uma série temporal agregada formada através da agregação de um determinado grupo de barramentos com característica similares. Para encontrar a série agregada é necessário saber como realizar o cálculo da demanda agregada; para resolver este problema as técnicas de agrupamento são de grande utilidade, pois identificam os barramentos com perfis de demandas diárias similares e a partir deste conhecimento pode ser realizado o cálculo da demanda agregada. Após o cálculo da demanda agregada efetua-se uma previsão desta demanda agregada e a partir desta previsão inicia-se um processo de cálculo das cargas individuais dos barramentos que foram utilizados para a formação da demanda agregada. O inconveniente do modelo agregado é que o processo de agregação não é indicado para todas os barramentos do sistema, podendo em alguns casos gerar previsões com erros elevados quando comparados a previsão individual do mesmo barramento; por outro lado um fator positivo é o fato de ser possível a realização de uma previsão por barramento de forma mais rápida.

5.2.1 Modelos de Previsão Individual (MPI)

Na categoria individual foram implementados dois modelos de previsão, de forma que seus resultados foram tomados como referencial para avaliação do desempenho dos modelos agregados. Neste sentido implementou-se dois modelos de previsão individual: PCD (previsão da curva diária) e PMD (previsão da média diária).

Previsão da curva diária - (PCD):

Neste modelo, o objetivo é encontrar a curva de consumo diária num horizonte futuro. Para isto deve-se prever as cargas horárias de um dia para obter a curva desejada. A Figura 14 mostra o perfil da curva diária para o barramento 1.

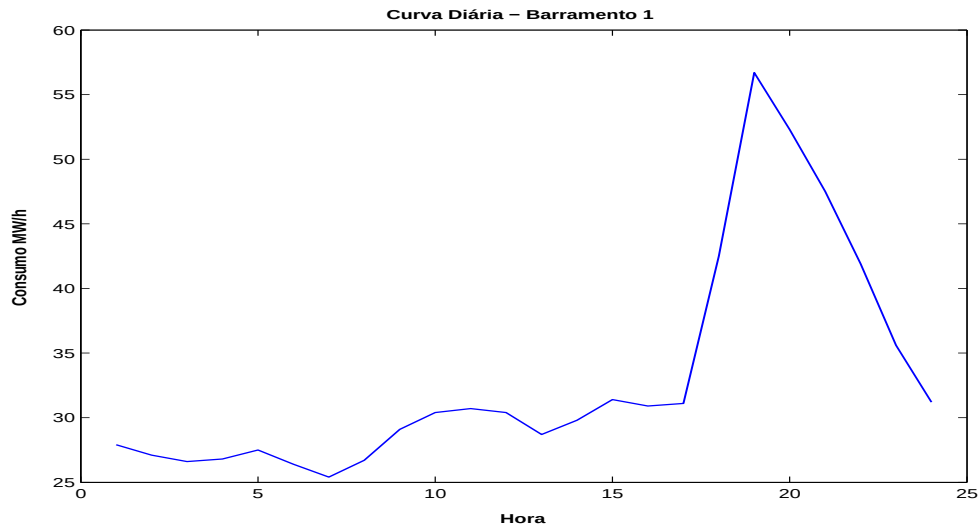


Figura 14: Curva diária - Barramento 1.

Com este objetivo, utilizou-se uma rede neural MLP do tipo *perceptron* com 1 camada intermediária. A rede foi treinada utilizando o seguinte vetor de entradas:

Tabela 6: Vetor de Ajuste da RNA - PCD

Entrada da Rede	$C_{(d-192)}$	$C_{(d-168)}$	$C_{(d-144)}$	$C_{(d-120)}$	$C_{(d-96)}$	$C_{(d-72)}$	$C_{(d-48)}$	$C_{(d-24)}$	$\widehat{C}_{(d)}$
-----------------	---------------	---------------	---------------	---------------	--------------	--------------	--------------	--------------	---------------------

Sendo que:

- $\widehat{C}_{(d)}$: carga estimada pela rede (demanda desejada para para o dia d na mesma hora h);
- $C_{(d-i*24)}$ $i : 1, \dots, 8$: cargas utilizadas como entrada na rede (demanda no dia $[d - (i \times 24)]$ na mesma hora h).

Ajustou-se uma rede para cada hora do dia. Como entrada foram utilizados 8 padrões para ajuste e um para saída desejada. Tais padrões representam 8 dias anteriores ao dia desejado d

na a mesma hora h . Este procedimento é repetido 24 vezes para cada hora do dia, totalizando 24 previsões para obter a curva diária prevista.

Previsão da média diária - (PMD):

Este modelo trabalha com a previsão da média diária. Neste caso não há a necessidade de realizar uma previsão da curva diária, mas apenas da média do dia pois a partir da previsão da média diária obtém-se a curva diária de consumo. A Figura 15 mostra o comportamento da série de carga média diária no período de 1/8/2001 a 4/10/2001. Pode-se perceber que a carga média diária tem um comportamento sazonal (de ciclo semanal) com uma tendência crescente. Abaixo detalha-se o procedimento utilizado para o cálculo da carga média diária.

Cálculo da Carga Média Diária

A carga média diária é calculada de forma que a curva diária possa ser representada por um valor numérico. Adotou-se a média aritmética das 24h do dia por ser um valor que possibilita a obtenção do perfil da curva diária de forma simples (Kadowaki et al., 2002). Para calcular a carga média diária, para o barramento i , utilizou-se a seguinte equação:

$$M(i, d, h) = \left(\frac{1}{24} \sum_{h=1}^{24} C(i, d, h) \right) \quad (5.1)$$

Sendo que:

- $M(i, d, h)$: média diária no barramento i na hora h do dia d ;
- $C(i, d, h)$: carga verificada no barramento i na hora h do dia d .

Na Figura 15, pode-se visualizar a curva de carga média obtida através da aplicação da equação 5.1 na série de consumo do barramento 1.

Como o objetivo deste modelo de previsão é obter a curva diária prevista, os seguintes passos são executados:

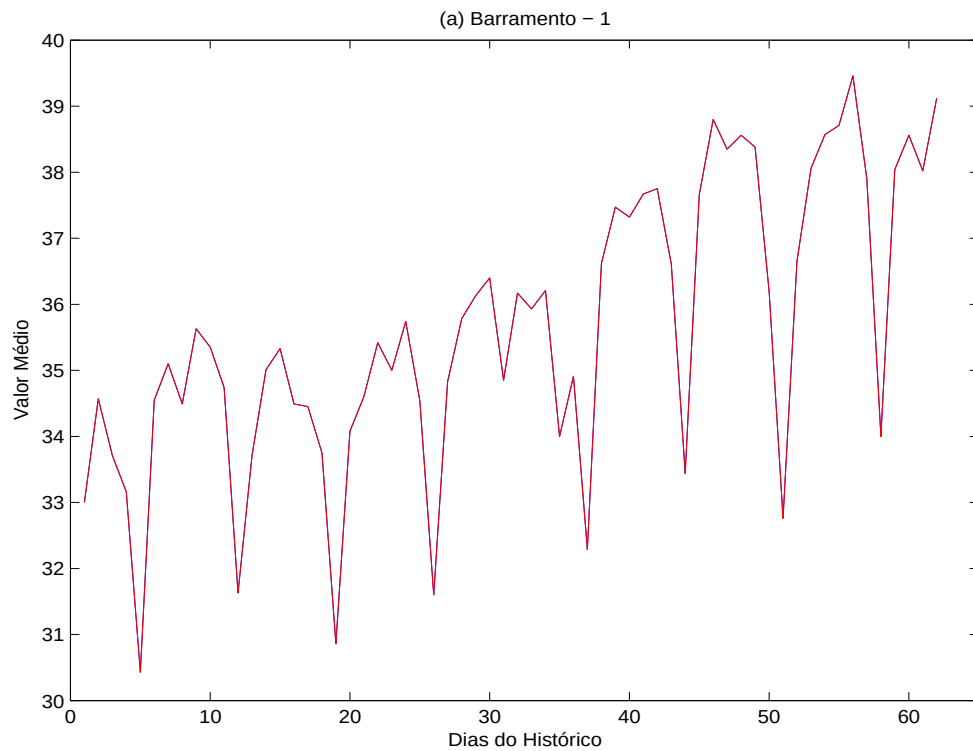


Figura 15: Curva Média do Barramento 1.

1- Previsão da Carga Média Diária

Efetua-se a previsão da carga média para um determinado dia. Por exemplo, supondo que o dia a ser previsto é uma terça-feira, deve-se obter a previsão da carga média desta terça-feira.

Para prever a média diária, utilizou-se uma rede neural do tipo MLP *perceptron* com 1 camada intermediária. A rede foi treinada utilizando o seguinte vetor de entradas:

Tabela 7: Vetor de Ajuste da RNA - PMD

Entrada da Rede	$M_{(d-3)}$	$M_{(d-2)}$	$M_{(d-1)}$	$\widehat{M}_{(d)}$
-----------------	-------------	-------------	-------------	---------------------

Sendo que:

- $\widehat{M}_{(d)}$: saída desejada no ajuste da rede (carga média desejada);
- $M_{(d-i)}$ $i : 1, \dots, 3$: padrões utilizados no ajuste da rede.

2- Escolha do Perfil da Curva de Carga Diária

No passo anterior, calcula-se uma estimativa da carga média diária e é necessário, agora, obter a curva de carga diária a qual dá a estimativa da carga a cada hora do dia. O perfil diário, em geral, é fortemente influenciado pelo dia da semana, pela temperatura e pelas condições climáticas.

Em relação ao dia da semana, há basicamente quatro perfis típicos: sábado, domingo, segunda-feira, e os dias úteis (terça a sexta-feira). O perfil da segunda-feira se diferencia dos demais dias úteis na madrugada e no começo da manhã, pois nesse período o seu comportamento se aproxima mais a um domingo. Quanto à temperatura, a sua influência pode tanto ser em termos de variação da carga média diária, pois em dias mais quentes a demanda média também tende a ser maior; como também em termos do perfil diário onde a influência da temperatura pode ser significativa, principalmente no período da tarde devido a utilização de aparelhos condicionadores de ar. Finalmente, as condições climáticas também influenciam o perfil de consumo, pois nos dias nublados e chuvosos a carga de iluminação tende a ser maior.

Neste trabalho, o dia escolhido como perfil ao dia previsto, foi o mesmo dia da semana anterior ao dia a ser previsto (no exemplo, o perfil é a terça-feira da semana que antecede a terça-feira a ser prevista).

3- Correção da Carga Média Diária

Observa-se na Figura 15 que há uma tendência de crescimento na carga média diária devido a este fator, obtém-se um fator de correção no valor da carga média para proporcionar maior precisão no cálculo da curva de carga diária.

Para obter a curva diária do dia a ser previsto, calcula-se uma taxa de variação entre a carga média prevista e a carga média do dia selecionado como perfil, de acordo com a seguinte equação:

$$Tx_{(d)} = \frac{Mp(d) - Mv(p)}{Mp(d)} \quad (5.2)$$

Sendo que:

- $Tx_{(d)}$: taxa de variação entre as cargas médias no dia d e do dia escolhido como perfil p ;
- $Mp(d)$: carga média prevista no dia d ;

- $Mv(p)$: carga média verificada no dia utilizado como perfil p .

4- Obtenção da Curva de Carga Diária

A curva de carga diária prevista (a curva de terça-feira prevista como no exemplo) é obtida através da equação (5.3).

$$Cp(i, d, h) = C(i, d, h)[1 + Tx_{(d)}] \quad (5.3)$$

Sendo que:

- $Cp(i, d, h)$: carga prevista no barramento i no dia d na hora h ;
- $C(i, d, h)$: carga verificada do barramento i na hora h do dia escolhido como perfil p ;
- $Tx_{(d)}$: taxa de variação entre as médias no dia d e do dia escolhido como perfil p .

Como o modelo PMD é um modelo individual o procedimento acima deve ser repetido para todos os dias a serem previstos e, sucessivamente, para todos os barramentos do sistema.

5.2.2 Modelos de Previsão Agregada (MPA)

Na grande maioria dos casos, a demanda global apresenta um comportamento bastante regular, isto é, o perfil de carga de uma segunda-feira, por exemplo, é muito semelhante à demanda da segunda-feira de uma semana anterior sendo a regularidade mantida ao longo da semana. Quando chega-se ao nível dos barramentos, observa-se uma menor regularidade, pois a demanda em um dado barramento é definida pelo comportamento do mercado ao qual este ponto de suprimento de energia atende. Assim, a variação no comportamento da carga dos barramentos está relacionada com o tipo de carga e, portanto, depende dos hábitos de seus consumidores.

O objetivo da previsão agregada é evitar a previsão individualizada das cargas por barramento, e para isto a idéia é realizar uma única previsão para um dado conjunto de barramentos. Neste contexto, os algoritmos de agrupamentos são essenciais pois viabilizam um processo de agrupamento para barramentos com perfis de consumo semelhantes de forma rápida e confiável (Salgado et al., 2004a) e (Salgado et al., 2004b).

O modelo de previsão agregada trabalha com um histórico agregado. Este histórico pode ser obtido através da agregação dos barramentos em uma única série temporal. A metodologia de previsão agregada pode ser resumida nos seguintes passos:

- (i) Agrupamentos dos barramentos e cálculo da série temporal agregada - (Agregação);
- (ii) Previsão da série temporal agregada;
- (iii) Cálculo das curva de carga individual a partir da previsão agregada - (Desagregação).

O objetivo da etapa (i) é agrupar os barramentos com comportamentos similares. Para cada grupo, as cargas são agrupadas formando uma série temporal agregada;

Na etapa (ii) é realizada uma previsão para estimar a demanda agregada de cada grupo. Esta previsão é baseada na série temporal agregada;

A última etapa determina a curva de consumo diário para cada barramento a partir da previsão da carga global do grupo.

5.2.2.1 Agregação dos Barramentos

No processo de agregação o principal ponto é como agrupar os barramentos. Neste sentido os métodos de agrupamento são relevantes, pois tais técnicas possibilitam agrupar barramentos que possuem um perfil de consumo semelhante ao longo do tempo.

Forma de Agregação dos Barramentos

Para se calcular a demanda agregada a informação com relação aos agrupamentos formados é essencial. Neste trabalho, optou-se por utilizar os dados de demandas por barramento do dia anterior ao dia a ser previsto para realizar os agrupamentos. Desta forma, se a previsão for feita para uma quinta-feira (20/09/2001), o dia usado para agrupar os barramentos será a quarta-feira (19/09/2001), isto é, foi utilizado o dia anterior ao dia a ser previsto.

Nos agrupamentos obtidos o barramento 15 sempre foi agrupado individualmente (observar Tabela 2, Capítulo 4, Seção 4.7.2). Assim, este barramento deve receber um tratamento individual. Por este motivo tal barramento não foi considerado nos resultados dos modelos de previsão agregada.

Cálculo da Carga Agregada - Agregação

A equação 5.4, abaixo, é utilizada para obter a demanda agregada no grupo g no dia d e na hora h . O processo de agregação é definido como sendo o somatório das cargas dos barramentos agregados em cada intervalo dia/hora.

$$Da(g, d, h) = \sum_{i=1}^n C(i, d, h) \quad (5.4)$$

Sendo que:

- $Da(g, d, h)$: Demanda agregada no grupo g no dia d e na hora h ;
- n : Número de barramentos pertencentes ao grupo g ;
- $C(i, d, h)$: Carga dos i barramentos pertencentes ao grupo g no dia d e na hora h .

5.2.2.2 Previsão da Demanda Agregada

A partir do cálculo da demanda agregada, o próximo passo foi realizar uma previsão da demanda agregada ($Da_{(prev)}$). Com o objetivo de prever a demanda agregada, utilizou-se uma rede neural do tipo *perceptron* com 1 camada intermediária. A rede é treinada utilizando o seguinte vetor de entradas:

Tabela 8: Vetor de Ajuste da RNA - $Da_{(prev)}$

Entrada da Rede	$C_{(d-192)}$	$C_{(d-168)}$	$C_{(d-144)}$	$C_{(d-120)}$	$C_{(d-96)}$	$C_{(d-72)}$	$C_{(d-48)}$	$C_{(d-24)}$	$\widehat{C}_{(d)}$
-----------------	---------------	---------------	---------------	---------------	--------------	--------------	--------------	--------------	---------------------

Sendo que:

- $\widehat{C}_{(d)}$: saída desejada da rede (demanda agregada desejada para para o dia d na hora h);
- $C_{(d-i*24)}$ $i : 1, \dots, 8$: carga de entrada na rede (demanda agregada no dia $[d - (i \times 24)]$ na mesma hora h).

Assim, ajustou-se uma rede para cada hora com um total de 8 entradas imediatamente anteriores à entrada a ser ajustada.

A previsão da demanda agregada é um fator fundamental e será utilizada nos modelos de previsão agregada descritos a seguir.

5.2.2.3 Cálculo das Cargas Individuais - Desagregação

Após a previsão da demanda agregada ($Da_{(prev)}$), a próxima etapa foi a determinação das cargas individuais para cada barramento do grupo. Com o objetivo de obter a carga individual de cada barramento foram desenvolvidas três alternativas de desagregação. Tais alternativas representam modelos que possuem princípios distintos para o cálculo das cargas individuais. Os processos de desagregação desenvolvidos estão detalhados abaixo e são classificados da seguinte forma:

1.- Previsão via Fator de Participação - (PFP);

A idéia básica do modelo dos fatores de participação é utilizar a participação que cada barramento em relação à demanda agregada. Desta forma, se os barramento 1, 2 e 3 são agregados, e supondo que a contribuição de cada um é 25%, 25% e 50% respectivamente no calculo da demanda agregada, então informação é usada para a desagregação dos barramentos; isto é, efetua-se a previsão da série de participações e utiliza-se a $Da_{(prev)}$ para efetuar a desagregação com base nos fatores previstos. O inconveniente deste modelo é a necessidade da realização de previsões para a série de participações de carga de cada barramento e, neste caso, volta-se ao problema encontrado nos modelos de previsão individual.

A série histórica de participações é determinada para cada barramento. A equação (5.5) mostra o cálculo desta participação, e a Figura 16 mostra o comportamento da série de participações em relação a Da e $Da_{(prev)}$ para o barramento 12 na hora 1 do dia 01/08/01 a 30/09/01. Como pode-se ver a série de participações (porcentagens) tem um comportamento regular, o que motivou a sua utilização no processo de desagregação. Através da Figura 16 pode-se notar que a medida que a carga cresce sua porcentagem tende a crescer.

$$P(i, d, h) = \frac{C(i, d, h)}{Da(g, d, h)} \times 100 \quad (5.5)$$

Sendo que:

- $P(i, d, h)$: participação do barramento i no grupo g em relação à demanda agregada no dia d

e na hora h ;

- $C(i, d, h)$: carga do barramento i no dia d e na hora h ;
- $Da(g, d, h)$: demanda agregada no grupo g no dia d e na hora h .

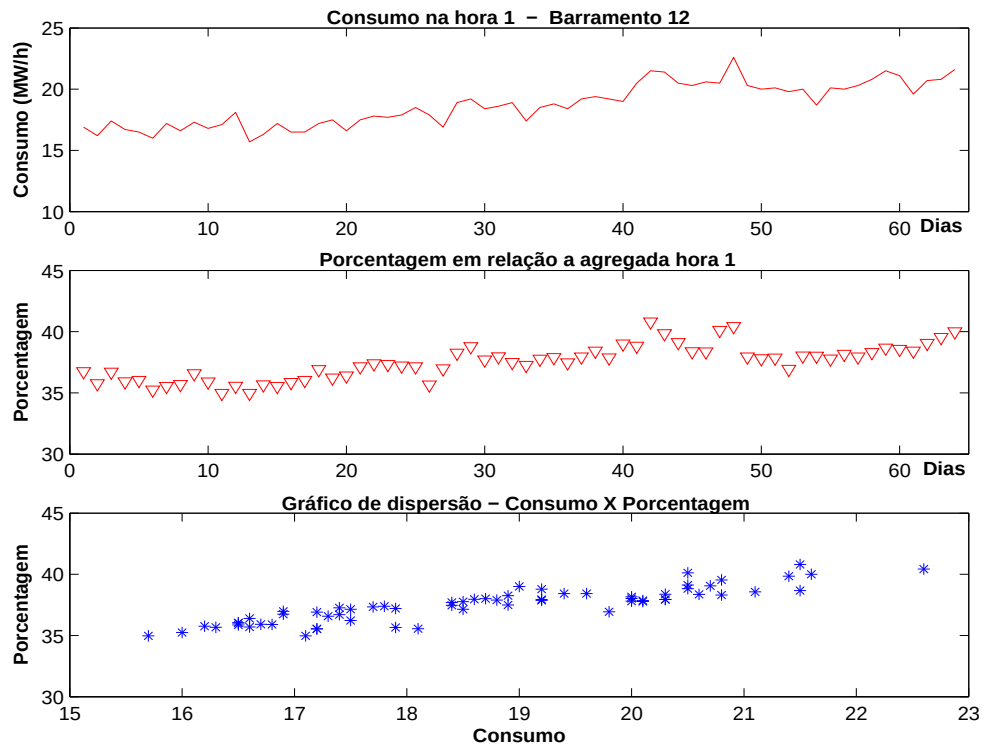


Figura 16: Porcentagem $\times$ Carga do barramento 12.

Desta forma, para cada barramento uma nova série temporal é formada contendo a participação em relação à demanda agregada. Para encontrar a previsão para cada barramento, o primeiro passo é realizar uma previsão utilizando a série de participações. Tal previsão pode ser feita por qualquer técnica de previsão existente na literatura. Neste trabalho, esta previsão da participação foi feita utilizando duas alternativas descritas no Capítulo 3:

- **Previsão via Médias Móveis - Mda;**

A estimativa de participação via média aritmética é obtida calculada através da média aritmética das cargas dos últimos dias ao dia estimado para a mesma hora.

• **Previsão via Redes Neurais.**

Neste caso, o fator de participação é dado por uma rede que estima a participação de cada barramento em relação à demanda agregada. A rede neural utilizada foi ajustada de acordo com as seguintes características: Uma rede neural MLP do tipo *perceptron* com 1 camada intermediária foi ajustada, de acordo com seguinte vetor de entradas:

Tabela 9: Vetor de Ajuste da RNA - PFP

Entrada da Rede	$P_{(d-192)}$	$P_{(d-168)}$	$P_{(d-144)}$	$P_{(d-120)}$	$P_{(d-96)}$	$P_{(d-72)}$	$P_{(d-48)}$	$P_{(d-24)}$	$\widehat{P}_{(d)}$
-----------------	---------------	---------------	---------------	---------------	--------------	--------------	--------------	--------------	---------------------

Sendo que:

- ▷ $\widehat{P}_{(d)}$: participação desejada da rede (participação estimada no barramento i no dia d e na hora h);
- ▷ $P_{(d-i*24)} \ i : 1, \dots, 8$: participação utilizada na entrada da rede (participação estimada no barramento i no dia $[n - (i \times 24)]$ na mesma hora h).

Assim, ajustou-se a rede para encontrar o fator de participação para cada dia e para cada hora. Após a previsão do fator de participação, vem o processo de cálculo das cargas individuais para os barramentos. O valor desta carga, em cada barramento, é obtido através da equação:

$$Cp(i, d, h) = P(i, d, h)_{(prev)} \times Da(g, d, h)_{(prev)} \quad (5.6)$$

Sendo que:

- ▷ $Cp(i, d, h)$: carga prevista no barramento i do grupo g no dia d e na hora h ;
- ▷ $P(i, d, h)_{(prev)}$: participação prevista no barramento i no dia d e na hora h ;
- ▷ $Da(g, d, h)_{(prev)}$: demanda agregada prevista do grupo g no dia d e na hora h .

Se o modelo de previsão adotado for um modelo simples como o modelo baseado em médias, esta alternativa torna-se mais viável. Se o modelo for um modelo mais complexo, exigindo maior esforço computacional, como um modelo neural, esta alternativa acaba se tornando menos viável pois será necessário ajustar uma rede para cada barramento, exigindo assim um grande tempo de processamento caso o sistema de transmissão possua um grande número de barramentos. Neste caso, é mais vantajoso realizar diretamente a previsão individual.

2.- Previsão via Rede Treinada para previsão da Da - (PRTDa);

Este modelo parte do princípio que a demanda agregada apresenta as características dos barramentos que foram utilizados para sua composição. Assim, se a Da representa o perfil dos barramentos do seu grupo, conseqüentemente o modelo ajustado para a previsão da Da será capaz de prever as cargas individuais de cada barramento sem a necessidade de um novo ajuste. Novamente, vale ressaltar que este procedimento apresentará resultados coerentes se a Da apresentar um comportamento similar aos barramentos do seu grupo. Seguindo este raciocínio, treinou-se uma rede neural utilizando a série temporal da Da e após o ajuste da rede o processo de desagregação é feito da seguinte maneira:

Utiliza-se uma rede neural treinada para obter a previsão da Da. Após o treinamento, aplica-se a rede ajustada para previsão dos barramentos. Desta forma não é necessária a realização de previsões para todos os barramentos, já que o modelo é ajustado apenas uma única vez para previsão em cada grupo.

A rede neural utilizada para prever o perfil da demanda agregada foi ajustada hora-a-hora de acordo com o vetor de entradas apresentado na Tabela 8.

3.-Previsão da Média Diária da Da - (PMDDa).

Este modelo trabalha com a previsão da média diária na Da e, como detalhado no modelo individual PMD, a partir da previsão da média da Da obtém-se a curva diária de consumo em cada barramento.

Assim como no modelo de previsão individual PMD, neste modelo agregado também é necessário efetuar o cálculo da média diária para realizar a previsão. A diferença das duas abordagens (PMD e PMDDa) é que no caso do PMDDa, trabalha-se com a série temporal da demanda agregada nas previsões.

Analogamente ao modelo individual PMD, o cálculo do valor da carga média, bem como o ajuste da rede neural para prever a média, são feitos do mesmo modo no modelo PMDDa.

Para efetuar o processo de desagregação a taxa de variação calculada para encontrar a curva média para demanda agregada é propagada através dos barramentos que fazem parte do grupo no qual a demanda está associada. Neste caso, também pressupõem-se que o comportamento da demanda foi influenciado pelo comportamento dos barramentos que compõem o seu grupo. A equação (5.7) é utilizada para propagar a taxa de variação da Da através dos barramentos que

compõem o seu grupo:

$$Cp(i, d, h) = C(i, d, h)[1 + Tx(d)_{(g)}] \quad (5.7)$$

Sendo que:

- ▷ $Cp(i, d, h)$: carga prevista no barramento i do grupo g no dia d e na hora h ;
- ▷ $C(i, d, h)$: carga verificada no barramento i do grupo g no dia perfil p e na hora h ;
- ▷ $Tx(d)$: taxa de variação da Da no grupo g entre os dia d e o dia perfil p .

5.2.3 Configuração das Metodologias de Previsão

Os modelos de previsão apresentados no Capítulo 3 foram aplicados aos modelos descritos anteriormente. Cada modelo foi ajustado de acordo com os procedimentos descritos abaixo, Procurando-se obter o melhor rendimento na escolha dos parâmetros de entrada dos modelos.

Modelo Baseado em Médias Móveis - Mda

No modelo Mda, utilizou-se como padrão $n = 3$ (isto é, três últimos valores anteriores ao valor a ser previsto) para estimar a carga consumida num instante t . Assim, a carga prevista num dia d foi dada pelo cálculo da média aritmética dos 3 dias anteriores ao dia d na mesma hora h .

Rede Neural Artificial

As RNAs utilizada são redes perceptron multi-camadas, com uma única camada intermediária utilizando o algoritmo de retropropagação do erro para treinamento. A taxa de aprendizagem foi de 0,9 e o termo momento constante foi de 0,6. A arquitetura ($A \times B \times C$) significa: rede *feedforward* com 3 camadas, tendo A neurônios na camada de entrada, B neurônios na camada intermediária e C neurônios na camada de saída. A Tabela 10 mostra a arquitetura das redes usadas nos modelos de previsão:

Tabela 10: Arquitetura Rede Neural.

Modelos	Arquitetura
PCD	$8 \times 10 \times 1$
PMD	$1 \times 10 \times 1$
PFP	$8 \times 10 \times 1$
PRTDa	$8 \times 10 \times 1$
PMDDa	$1 \times 10 \times 1$
$Da_{(prev)}$	$8 \times 10 \times 1$

5.3 Resultados Alcançados

Para o treinamento dos modelos utilizou-se o histórico de dados do período de 01/08/01 até 19/09/01 e para o período de testes utilizou-se o histórico de 20/09/2001 até 26/09/2001. Para avaliar e comparar o desempenho dos modelos, a medida utilizada foi o *erro percentual médio absoluto (MAPE)* e o *erro médio absoluto (MAE)* entre as cargas observadas e as cargas estimadas na previsão. O MAPE e o MAE são definidos como segue:

$$MAPE(i, d) = \frac{100}{24} \sum_{h=1}^{24} \frac{|C(i, d, h) - Cp(i, d, h)|}{C(i, d, h)}; \quad (5.8)$$

$$MAE(i, d) = \frac{1}{24} \sum_{h=1}^{24} |C(i, d, h) - Cp(i, d, h)|. \quad (5.9)$$

Sendo que:

- $C(i, d, h)$: carga verificada no barramento i no dia d e na hora h ;
- $Cp(i, d, h)$: carga estimada, no barramento i no dia d e na hora h .

Nas metodologias abordadas neste trabalho utilizou-se preferencialmente o MAPE como medida de desempenho e comparação entre os previsores. Desta forma, as tabelas do MAE são apresentadas apenas no anexo B, e no presente Capítulo são mostrados somente as tabelas e figuras com os valores do MAPE.

5.3.1 Modelos de Previsão Individual - MPI

As metodologias de previsão individuais foram testadas durante o período de uma semana. Ambos os modelos foram submetidos aos mesmos dados nas etapas de ajuste e conseqüentemente o mesmo vetor de teste foi aplicado para avaliar as metodologias. Abaixo seguem os resultados e as análises quantitativas e qualitativas acerca do desempenho de cada uma das técnicas desenvolvidas.

As Tabelas 11 e 12 mostram o desempenho do PCD e PMD, respectivamente. Em ambos os modelos observa-se que os erros médios de previsão variam muito dependendo do barramento, e de um modo geral os barramentos com menores desvios padrão apresentam também menores erros de previsão. Destes testes também não se observa a predominância de um dos métodos; aproximadamente 50% dos barramentos, o modelo PCD apresentou melhores resultados e nos outros 50% o melhor método foi o PMD.

Para a metodologia PCD treinou-se uma RNA para cada hora de dia ao longo de uma semana. Sendo feita uma previsão pontual de base horária. No modelo PMD, treinou-se uma RNA para prever apenas a média do dia e a partir da previsão desta média encontra-se a curva diária. Logo, na metodologia PCD foram feitas 24 previsões para cada dia e na metodologia PMD foi feita apenas uma previsão para cada dia.

Tabela 11: MAPE Diário - PCD.

-	qui	sex	sáb	dom	seg	ter	qua	-
Barramentos	20/set	21/set	22/set	23/set	24/set	25/set	26/set	Erro Médio
1	1,07	1,62	2,00	1,45	1,56	1,56	1,48	1,53%
2	3,96	4,80	3,69	5,13	7,00	6,63	7,25	5,49%
3	2,76	2,80	2,52	2,38	2,04	4,62	3,79	2,99%
4	4,03	2,81	2,33	2,46	10,49	18,28	9,83	7,18%
5	1,50	1,86	5,92	5,52	2,29	1,95	2,63	3,10%
6	2,14	2,56	1,92	5,40	2,53	2,15	3,26	2,85%
7	3,97	3,64	2,59	13,36	3,01	3,16	2,06	4,54%
8	2,41	1,96	3,33	5,53	2,36	3,04	1,91	2,93%
9	1,39	0,90	1,74	3,21	1,15	2,02	1,09	1,64%
10	2,50	3,56	1,44	1,85	2,84	3,01	3,38	2,65%
11	4,07	3,66	2,41	6,73	6,86	8,14	2,81	4,95%
12	4,70	4,88	4,31	2,41	7,25	5,01	2,89	4,49%
13	9,65	7,54	11,65	10,56	8,90	7,50	10,20	9,43%
14	9,80	8,14	9,83	9,87	5,09	7,16	9,36	8,46%
15	4,05	9,16	6,32	4,54	8,00	5,28	6,12	6,21%

De acordo com a Tabela 1, no Capítulo 2, pode-se notar que o barramento 1 possui um desvio padrão baixo, enquanto que o barramento 4 possui um alto desvio padrão; espera-se que os barramentos com baixo desvio padrão tenham baixos erros de previsão. Os erros de previsão expressos na Tabela 11 e na Figura 17 mostram que, na maioria dos casos, os barramentos com elevado desvio padrão apresentam um altos erros de previsão. Assim, para os barramentos com menor desvio padrão o PMD obteve melhores resultados e, para os mais variáveis o PCD apresentou melhor comportamento. Isto se deve ao fato da previsão via PCD ser pontual (hora-a-hora).

Tabela 12: MAPE Diário - PMD.

-	qui	sex	sáb	dom	seg	ter	qua	
Barramentos	20/set	21/set	22/set	23/set	24/set	25/set	26/set	Erro Médio
1	1,01	0,92	2,09	1,30	1,94	1,29	1,66	1,46%
2	2,65	4,31	3,06	2,41	1,73	3,91	4,15	3,17%
3	1,61	1,23	1,26	1,81	1,34	2,06	1,63	1,56%
4	4,84	5,34	4,25	6,86	5,16	31,68	6,84	9,28%
5	3,52	4,46	3,14	4,62	1,99	4,53	4,48	3,82%
6	2,32	2,31	1,63	2,83	1,87	1,97	3,19	2,30%
7	3,99	3,58	2,22	25,49	13,47	3,48	3,79	8,00%
8	6,85	7,79	9,73	3,06	1,65	4,25	8,18	5,93%
9	1,38	1,49	1,01	2,40	1,44	1,86	2,76	1,76%
10	6,22	5,02	5,85	6,17	7,50	4,58	5,61	5,85%
11	7,41	5,74	3,64	5,85	7,28	8,11	4,67	6,10%
12	1,74	1,44	0,95	3,64	4,05	5,42	1,59	2,69%
13	14,70	6,59	7,20	5,89	4,96	9,46	6,65	7,92%
14	5,33	11,85	5,51	3,41	5,21	13,89	7,64	7,55%
15	3,35	5,43	5,00	12,73	12,34	8,72	5,02	7,51%

Os modelos de previsão individuais são indicados para barramentos que apresentam características especiais e por este motivo não podem ser agrupadas. Apesar dos resultados das metodologias individuais serem superiores aos resultados das agregadas estes métodos apresentam limitações quanto à sua aplicação; o grande inconveniente dos MPI é a necessidade de realizar uma previsão para cada barramento gerando um alto custo computacional.

Deve-se destacar que para ambos os casos a mesma estrutura de rede neural foi aplicada para todos os barramentos. Desta forma, acredita-se que para cada barramento haja modelos específicos mais adequado, os quais apresentariam erros bem menores dos que os obtidos. Para isto, porém, seria necessário trabalhar especificamente com o histórico de cada barramento para determinar o modelo mais adequado.

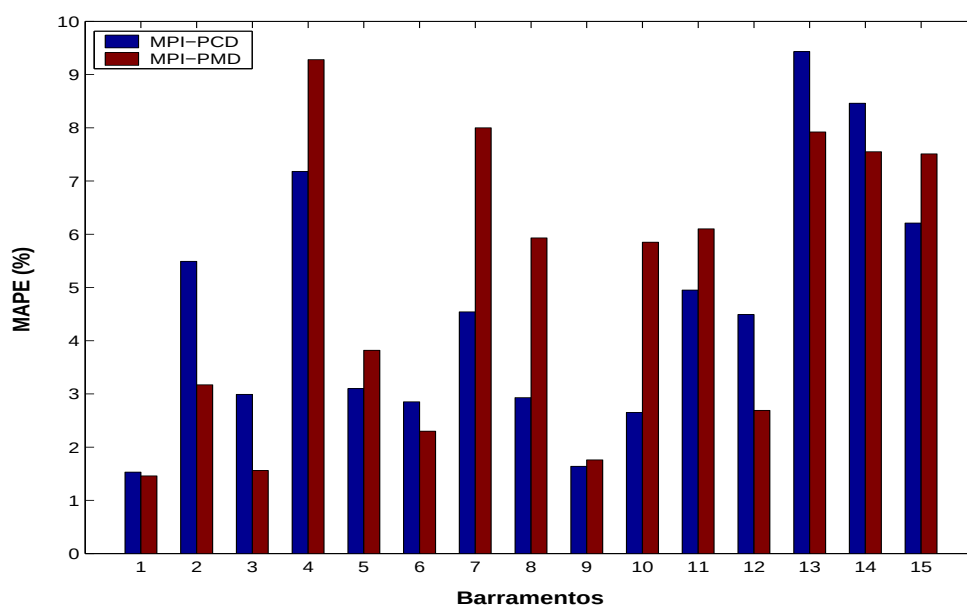


Figura 17: Comparação dos MPI.

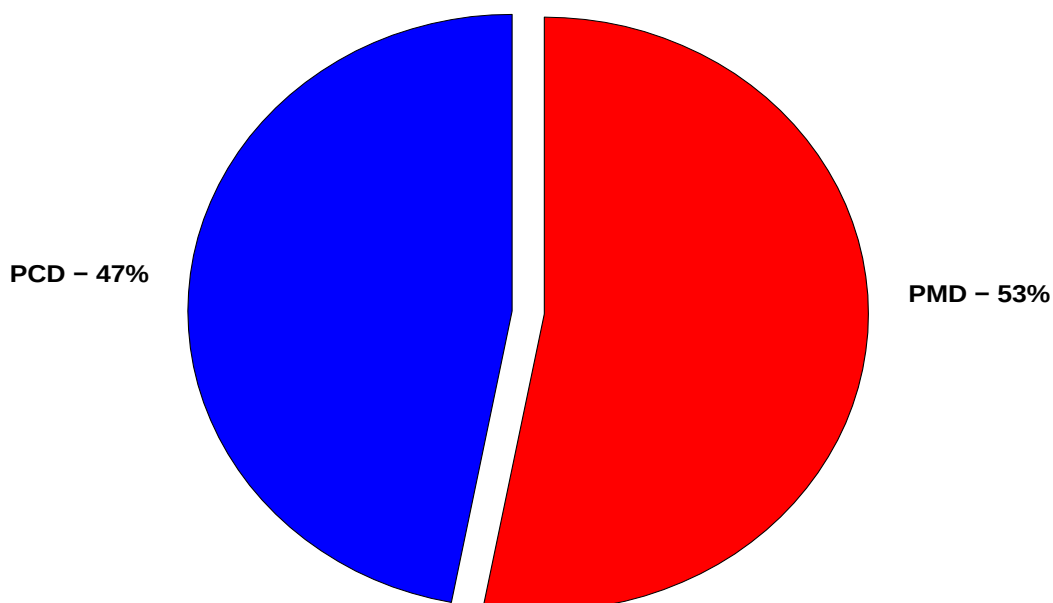


Figura 18: Eficiência dos MPI.

5.3.2 Modelos de Previsão Agregada - MPA

Os modelos de previsão agregados foram aplicados a dados reais do sistema elétrico para verificar e avaliar o desempenho e a robustez das metodologias.

5.3.2.1 Previsão por Fator de Participação - PFP

A metodologia PFP foi testada e as Tabelas 13 e 14 mostram os resultados expressos em termos dos erros. Para cálculo do fator de participação foram utilizados as técnicas de previsão baseada nas médias (PFP/Médias) e em RNA (PFP/RNA).

Observando as Tabelas 13 e 14, de um modo geral, os resultados obtidos através de fatores de participação apresentam maiores erros do que a previsão individual (PCD). Para o caso do fator de participação ser calculado pelas médias, em todos os casos os resultados foram piores do que para a previsão individual (PCD). Porém, em alguns casos, quando o fator de participação é calculado por uma rede neural para os barramentos com maiores cargas, os erros foram melhores do que os obtidos pelo modelo individual. Portanto para estes barramentos o modelo PFP/RNA é uma alternativa interessante.

Observando a Tabela 1 e as Figuras 4 a 8, no Capítulo 2, nota-se que alguns barramentos possuem uma carga muito baixa, na ordem de 2 a 10 MW/h aproximadamente fazendo com que o fator de participação em relação a Da seja consideravelmente baixo quando comparado a outros barramentos do sistema. Este fator implica que um pequeno erro na previsão desta participação resulta em um erro extremamente elevado na aplicação do processo de desagregação utilizando fatores de participação.

Tabela 13: MAPE Diário - PFP/Médias.

-	qui	sex	sáb	dom	seg	ter	qua	-
Barramentos	20/set	21/set	22/set	23/set	24/set	25/set	26/set	Erro Médio
1	1,88	1,03	6,51	10,95	10,65	5,74	2,00	5,54%
2	1,44	2,12	5,59	14,04	8,93	6,08	3,52	5,96%
3	2,53	1,88	8,71	14,29	13,32	4,29	1,97	6,71%
4	3,43	4,38	3,40	6,35	7,95	41,13	26,91	13,36%
5	1,98	1,25	5,31	13,15	13,82	4,93	2,61	6,15%
6	1,6	1,34	15,23	20,10	17,99	4,14	3,26	9,09%
7	3,5	1,72	15,34	17,44	16,31	6,35	4,79	9,35%
8	2,11	1,66	5,71	16,70	13,94	5,09	4,16	7,05%
9	1,32	1,47	11,67	15,20	15,64	1,56	2,19	7,01%
10	2,22	2,33	2,93	5,30	11,07	7,64	3,34	4,98%
11	3,81	8,19	5,33	41,77	19,48	10,19	5,24	13,43%
12	2,69	2,30	4,93	9,35	10,25	6,69	1,53	5,39%
13	4,33	4,41	5,55	3,70	14,98	4,97	7,18	6,45%
14	4,66	7,19	7,25	41,68	25,94	4,58	12,29	14,80%

Tabela 14: MAPE Diário - PFP/RNA.

-	qui	sex	sáb	dom	seg	ter	qua	-
	20/set	21/set	22/set	23/set	24/set	25/set	26/set	Erro Médio
1	2,60	5,62	6,37	2,79	3,14	3,79	2,56	3,84%
2	4,44	3,81	3,24	74,45	19,52	4,02	9,36	16,98%
3	1,77	1,55	1,69	2,34	5,13	3,79	3,18	2,78%
4	3,36	3,19	3,31	44,82	63,53	8,45	26,32	21,85%
5	1,82	1,93	1,76	2,67	7,03	2,23	2,02	2,78%
6	2,82	2,35	2,13	2,40	5,73	3,50	3,62	3,22%
7	42,72	58,78	62,89	67,46	4,29	39,73	59,15	47,86%
8	4,25	2,09	3,86	4,64	7,22	1,49	3,97	3,93%
9	2,00	1,04	2,84	3,64	5,43	3,75	1,99	2,96%
10	2,13	2,07	2,26	2,37	5,55	2,81	3,76	2,99%
11	3,40	2,77	2,09	4,48	7,47	8,79	6,40	5,06%
12	2,38	1,72	1,75	2,10	2,94	3,25	1,91	2,29%
13	2,93	92,71	4,72	64,35	81,15	83,80	90,87	60,08%
14	2,84	2,50	8,33	14,65	5,37	5,08	6,58	6,48%

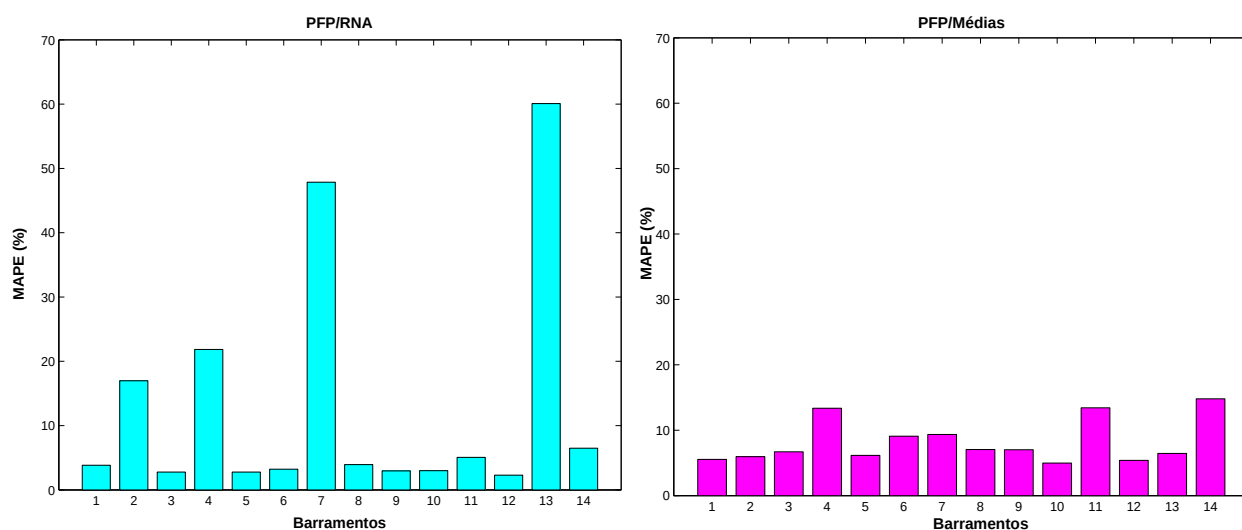


Figura 19: MPA - Fator de participação.

5.3.2.2 Rede Treinada com Perfil/Da - PRTDa

Para esta metodologia, conforme exposto anteriormente, uma rede neural é ajustada para previsão da demanda agregada. Esta mesma rede é utilizada para fazer a previsão de cada barramento do grupo. Isto é, não existe um processo de desagregação, pois a previsão de cada barramento é realizada utilizando a própria série história, através da rede neural ajustada para previsão

da demanda agregada.

Neste teste, foram utilizados os grupos obtidos através do processo de agrupamento mostrados no Capítulo 4. Aqui nota-se obviamente a influência da homogeneidade dos elementos de cada grupo, pois para o Grupo 1, compostos pelos barramentos 1, 2, 12; dois dos barramentos apresentam melhores resultados do que o caso individual (PCD). Nos demais grupos, a grande maioria dos casos a previsão via PRTDa é inferior à previsão individual (PCD).

Tabela 15: MAPE Diário - PRTDa.

-	qui	sex	sáb	dom	seg	ter	qua	-
Barramentos	20/set	21/set	22/set	23/set	24/set	25/set	26/set	Erro Médio
1	2,2	2,48	2,57	4,25	8,88	3,5	2,31	3,74%
2	2,89	1,8	4,07	4,33	15,23	4,12	3,88	5,19%
3	4,94	3,67	1,46	4,91	3,31	5,37	3,57	3,89%
4	2,83	3,33	3,61	4,58	5,09	7,99	12,93	5,77%
5	3,67	1,35	2,2	3,6	3,86	6,06	8,83	4,22%
6	2,14	4,49	1,38	7,9	3,84	7,86	2,24	4,26%
7	3,02	4,85	2,44	7,39	8,89	9,18	8,14	6,27%
8	2,07	1,8	2,79	9,84	3,8	5,35	3,89	4,22%
9	2,5	3,42	1,28	6,16	3,83	8,05	2,16	3,91%
10	2,31	1,7	2,01	3,21	3,37	3,81	5,08	3,07%
11	8,37	8,28	6,29	5,61	21,89	12,65	9,2	10,33%
12	2,82	2,39	4,57	5,63	8,99	6,06	2,78	4,75%
13	6,84	5,57	5,68	5,84	5,51	6,71	6,11	6,04%
14	19,45	9,72	14,02	15,63	44,54	10	4,1	16,78%

5.3.2.3 PMDDa - Previsão Média Diária-Da

Nesta metodologia, conforme exposto anteriormente, a previsão é realizada com base na previsão da carga média diária calculada para a Da de cada grupo. Após a previsão da carga média diária obtém-se as curvas diárias dos barramentos. A Tabela 16 mostra os erros de previsão para esta metodologia.

Neste caso, também é forte a influência da homogeneidade dos elementos de cada grupo. Novamente, dois dos três barramentos do Grupo 1, foram melhores do que a previsão individual, e da mesma forma dois barramentos do Grupo 2.

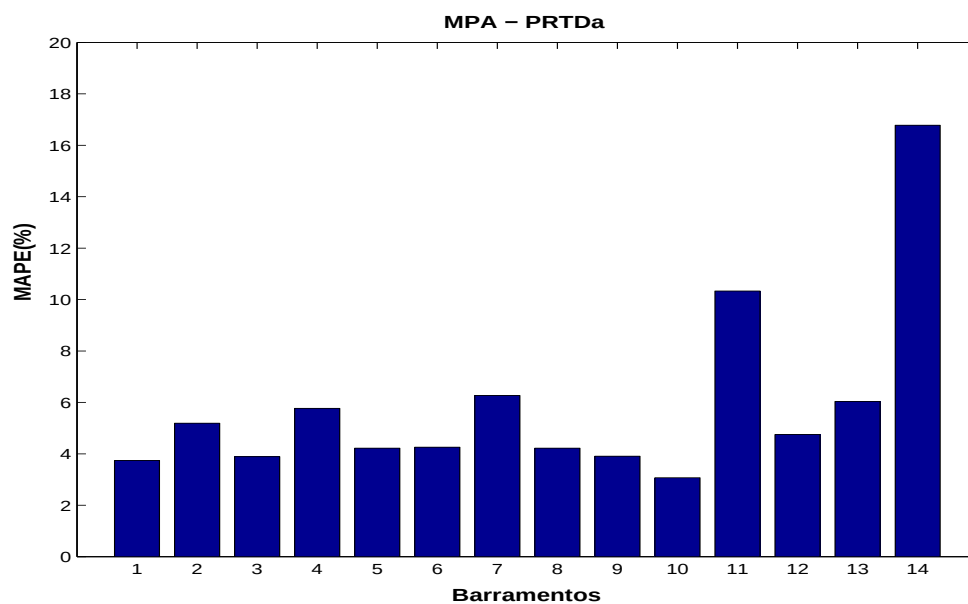


Figura 20: Desempenho Metodologia - PRTDa.

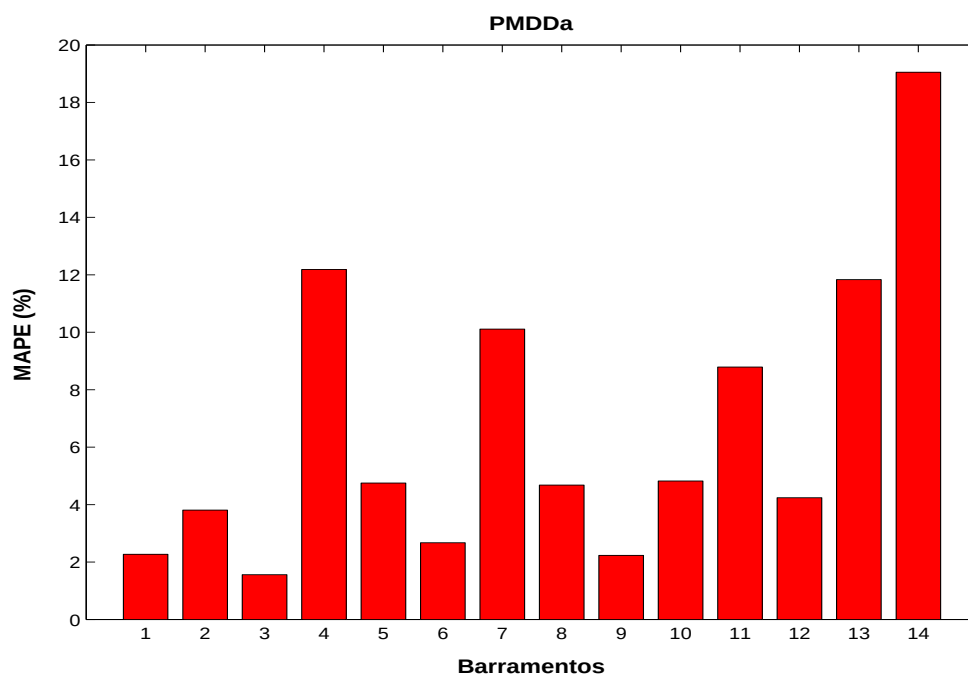


Figura 21: Desempenho metodologia PMDDa.

Tabela 16: MAPE Diário - PMDDa.

-	qui	sex	sáb	dom	seg	ter	qua	-
Barramentos	20/set	21/set	22/set	23/set	24/set	25/set	26/set	Erro Médio
1	1,88	1,14	6,11	1,38	2,49	1,23	1,67	2,27%
2	2,24	4,23	5,92	2,48	3,98	3,74	4,09	3,81%
3	1,61	1,23	1,26	1,81	1,34	2,06	1,63	1,56%
4	4,95	7,65	6,08	11,01	7,30	22,60	25,65	12,18%
5	4,59	5,19	5,01	5,67	3,73	4,18	4,84	4,75%
6	3,05	3,10	1,61	3,22	1,99	3,11	2,61	2,67%
7	4,38	4,08	2,40	39,60	9,02	7,65	3,61	10,11%
8	3,99	6,31	11,65	3,08	1,65	3,47	2,57	4,68%
9	3,40	2,83	1,04	2,39	1,57	2,16	2,18	2,23%
10	4,52	3,81	2,66	4,21	5,25	6,80	6,51	4,82%
11	15,53	5,97	4,55	6,03	6,80	11,94	10,72	8,79%
12	3,64	3,27	7,25	3,66	5,11	5,17	1,59	4,24%
13	19,63	13,69	7,90	10,78	5,59	13,32	11,91	11,83%
14	20,18	28,18	23,67	17,39	9,30	15,20	19,47	19,05%

5.3.3 Análise dos Resultados

5.3.3.1 Avaliação: MPA

A Figura 22 mostra os resultados da previsão, utilizando os MPA para todos os barramentos. Nota-se que não há um modelo que seja o melhor para todos os barramentos. Observa-se que os barramentos com elevado desvio padrão apresentam os maiores erros em todos os modelos. Nota-se também que para os barramentos com baixa carga, alguns modelos apresentaram elevados erros.

Os erros dos MPA tendem a ter erros maiores que os erros dos MPI, haja visto que trata-se de uma previsão via demanda agregada, e esta metodologia tende a apresentar erros superiores à metodologia individual, já que o conjunto de barramentos estudados possui barramentos com comportamento irregular.

Outro fator relevante é que existe um acúmulo de erros nos modelos MPA já que nos modelos MPA primeiro é realizada uma previsão da Da e, posteriormente, utiliza-se esta Da prevista para efetuar a previsão de cargas individuais. Desta forma, os erros de previsão da Da podem influenciar no desempenho do processo de cálculo das cargas individuais, afetando o processo de desagregação e conseqüentemente o desempenho dos modelos de previsão agregada.

Como o objetivo da previsão de carga é fornecer dados de demanda para a análise do

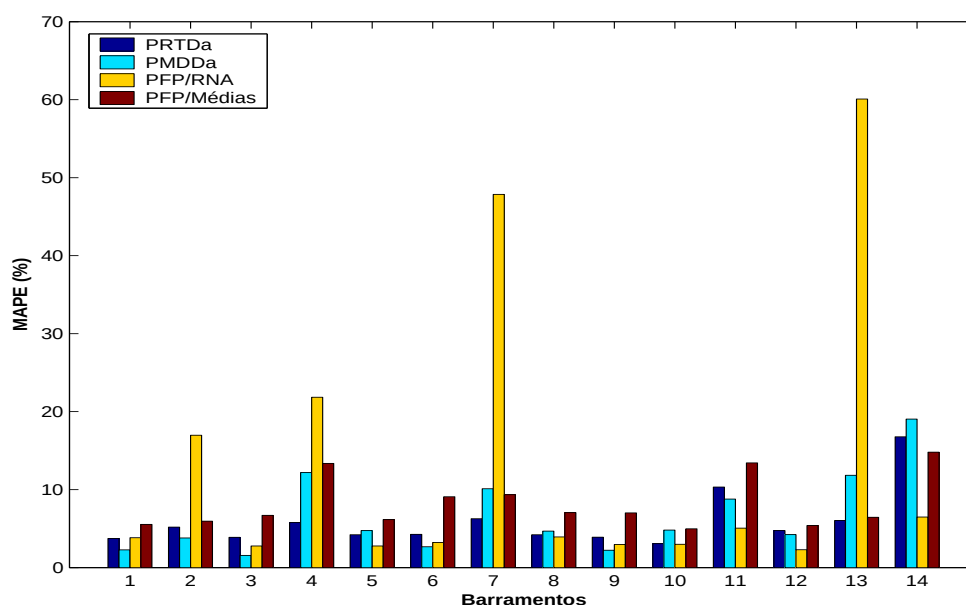


Figura 22: Comparação dos MPA.

carregamento do sistema de transmissão, e as demandas de barramento de reduzido valor têm pouca influência sobre o carregamento nas linhas de transmissão, o erro nas previsões das demandas destes barramentos tem uma influência muito pequena na análise de segurança. Desta forma, para este tipo de barramento é recomendado realizar a previsão com o valor médio das cargas verificadas anteriormente (Média das Cargas Anteriores - Mda). A Figura 23 mostra que a média se mostrou adequada quando comparado aos modelos agregados e apresenta um baixo custo computacional; outra solução para tais barramentos seria a utilização de um modelo individual como descrito anteriormente. Nas Figuras 23 (a) e 23 (b) pode-se observar erro dos barramentos 2, 4, 7 e 13 que possuem baixa carga.

Na Figura 24, pode-se ver a eficiência dos modelos de MPA para todos os barramentos. Nota-se que o modelo PFP teve desempenho superior em 43% dos barramentos enquanto que o PMDDa e o PRTDa foram superiores em 36% e 21%, respectivamente. O desempenho do PFP é superior quando comparado às demais metodologias, porque como dito anteriormente a forma de previsão do PFP é o mesmo utilizado nos modelos individuais, isto é, existe necessidade de realizar uma previsão individual para cada barramento.

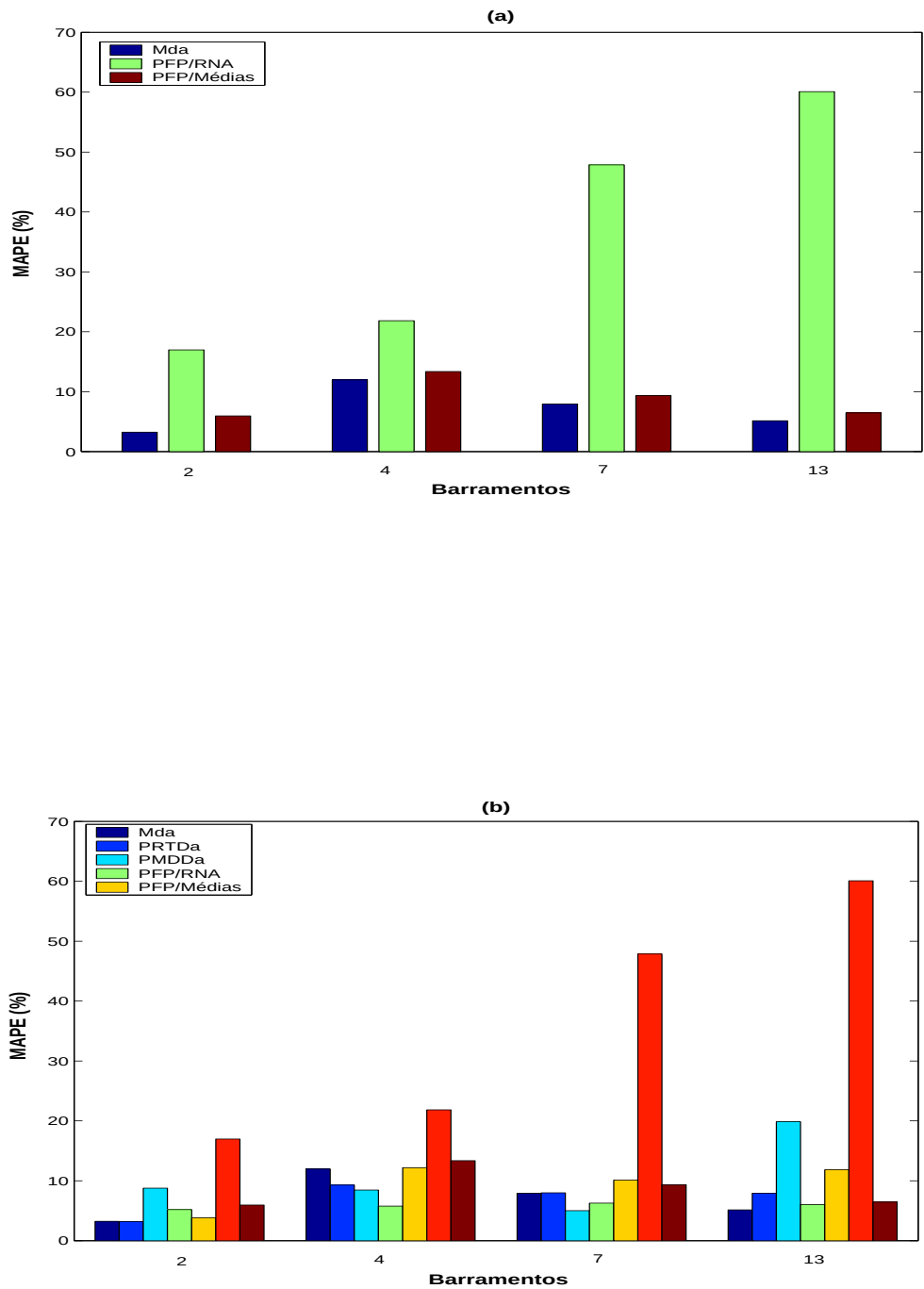


Figura 23: Comparação - Barramentos com baixa demanda.

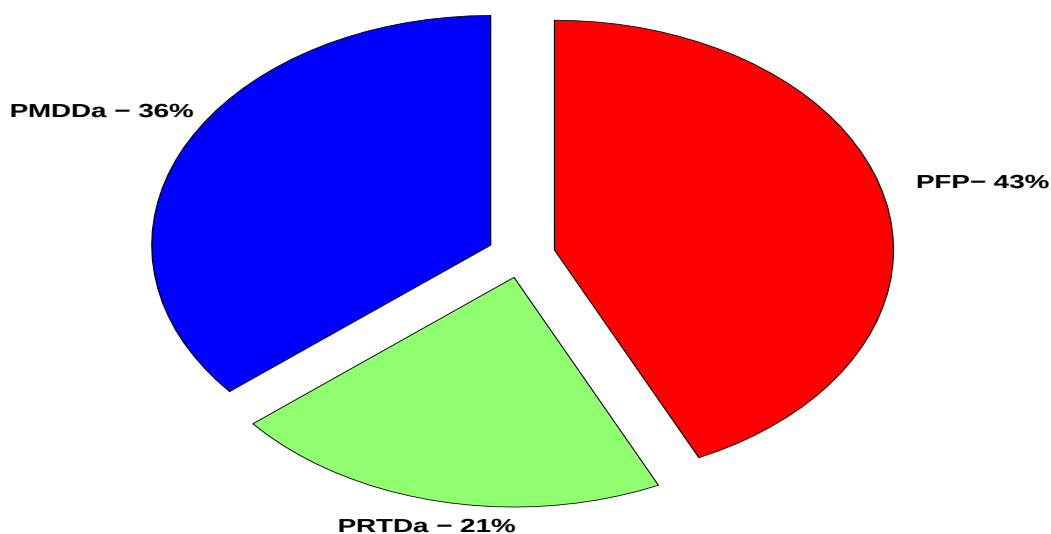


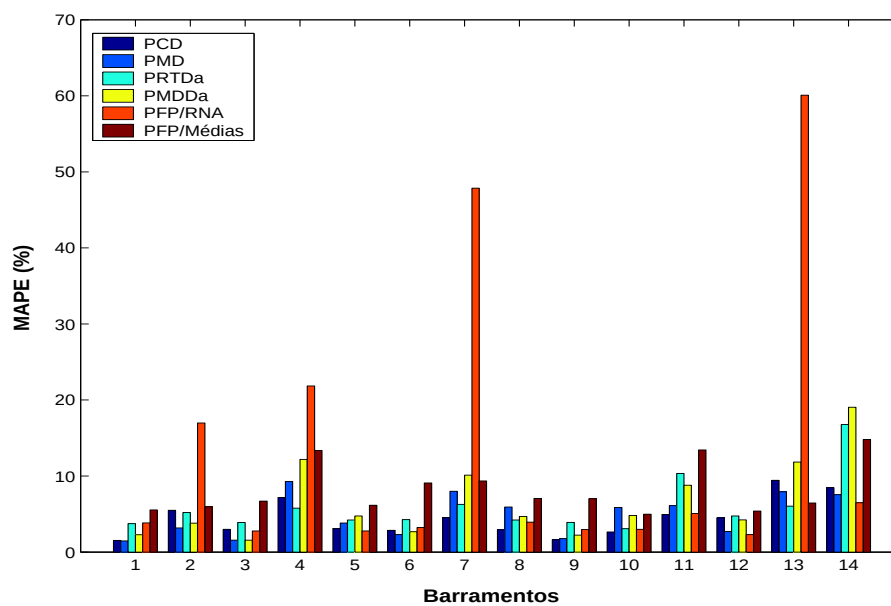
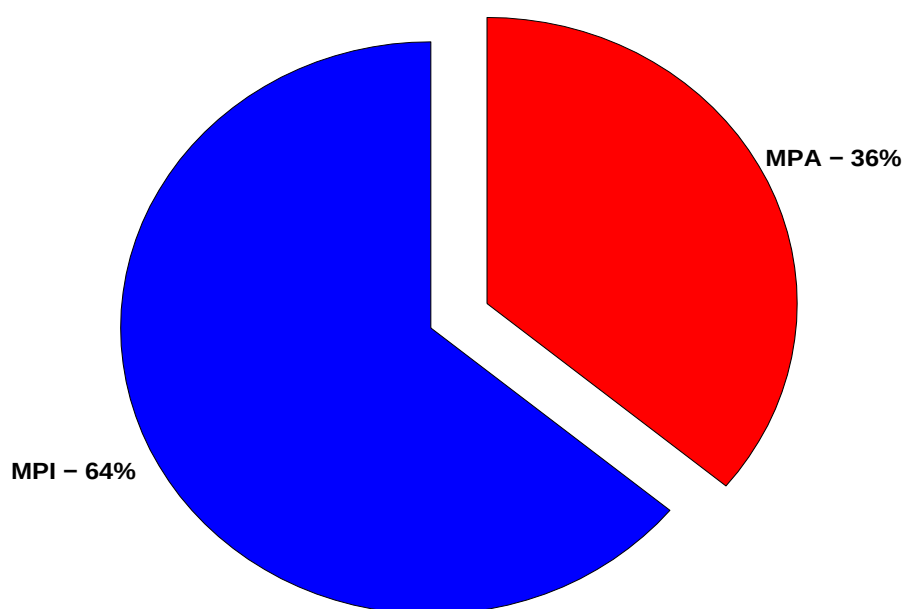
Figura 24: Eficiência dos MPA.

5.3.3.2 Avaliação: MPA $\times$ MPI

Com a Figura 25 visualiza-se os erros de todas as metodologias de previsão desenvolvidas neste trabalho. Como era esperado, os MPI apresentaram melhores resultados na maioria dos barramentos. Entretanto, os MPA apresentaram resultados satisfatórios quando comparados às metodologias de previsão individual. Para os barramentos com elevada carga e desvio padrão baixo e médio, os modelos PMDDa e PFP/RNA foram os melhores e nos outros casos estiveram próximos dos melhores.

Ainda na Figura 25 nota-se que os MPA são compatíveis com os MPI com erros próximos em quase todos os casos analisados, mostrando que a metodologia de previsão agregada, proposta neste trabalho, é eficaz na resolução do problema proposto.

Nas Figuras 26 e 27 estão apresentadas a eficiência dos MPA e MPI, em termos de porcentagem. Neste caso, os MPI foram superiores em 64% dos barramentos analisados. Entretanto, pela Figura 25 os erros entre os MPI e MPA não estão muito divergentes confirmando que a metodologia agregada mostrou-se eficaz na resolução do problema de previsão de carga para um conjunto considerável de barramentos.

Figura 25: Comparação MPI $\times$ MPA.Figura 26: Eficiência MPI $\times$ MPA.

Na Figura 27, pode-se notar que os modelos de previsão individual demonstraram melhor desempenho (PMD - 29% e PCD - 36%). Por outro lado, os modelos agregados tiveram rendimento inferior aos individuais, como já era esperado (PFP - 21% e PRTDa - 14%).

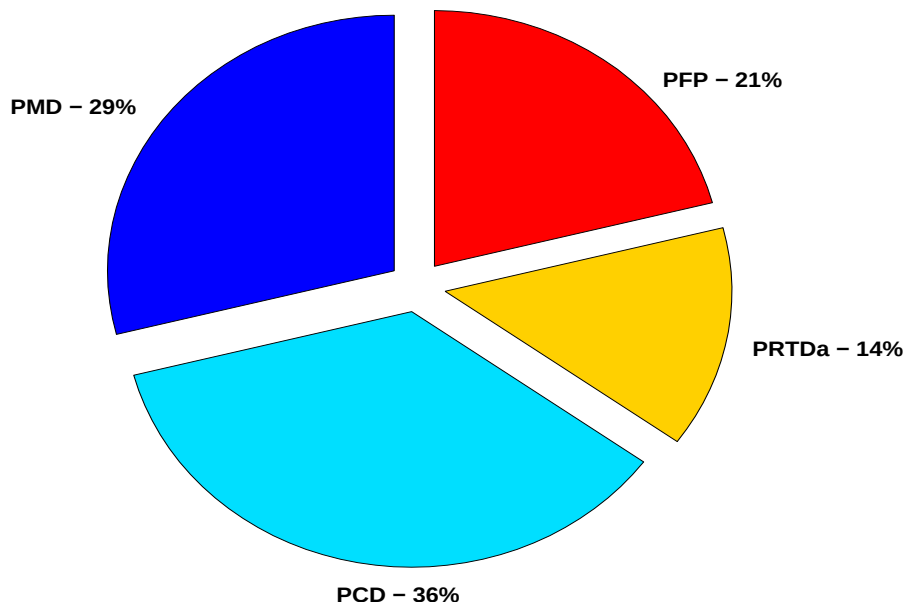


Figura 27: Eficiência dos modelos - Porcentagem.

5.4 Conclusões

No problema de previsão por barramento foram comparados os modelos de previsão individual e agregado. Notou-se que os MPI foram mais eficientes em grande parte dos barramentos. Os resultados dos MPA quando comparados aos MPI, de uma forma geral não apresentaram erros muito maiores, tendo em alguns barramentos (3, 4, 5, 13, e 14) erros inferiores ao modelo MPI. Desta forma, os MPA propostos foram eficientes na resolução do problema de previsão de carga por barramento, pois apresentam um menor número de previsões e resultados aceitáveis quando comparados aos MPI.

Através deste estudo observou-se que os resultados da previsão estão intimamente ligados ao valor do desvio padrão dos barramentos analisados conseqüentemente, não se pode tratar todos os barramentos da mesma forma. Desta forma, podemos classificar os barramentos em 2 tipos de conjuntos de barramentos que podem ser tratados especificamente:

1. **Barramentos Agregáveis** - são aqueles que apresentam:

- Desvio padrão baixo. - $\text{DesvPad (Pu)} = 0.02$ - Barramentos 1 e 12.
- Desvio padrão médio - $\text{DesvPad (Pu)} = 0.05$ - Barramentos 3, 5, 6, 8, 9 e 10.

2. **Barramentos Não-Agregáveis** - são aqueles que apresentam:

- Desvio padrão alto - $\text{DesvPad (Pu)} = 0.09$ - Barramentos 11, 14 e 15.
- Carga reduzida - Barramentos 2, 4, 7 e 13.

6 *Conclusão e Trabalhos Futuros*

Este capítulo apresenta as conclusões gerais e as contribuições deste trabalho, bem como as sugestões para futuras pesquisas nesta área de previsão de carga por barramento.

6.1 **Conclusões Gerais**

O presente trabalho apresentou a construção de um método de previsão de carga ativa por barramento num período de 24h à frente (curto prazo). Este tema tem relevância no sentido de colaborar para o planejamento do sistema elétrico de forma a viabilizar o conhecimento regional da carga, auxiliando na operação do sistema.

O modelo para a previsão de carga por barramento desenvolvido nesta pesquisa foi constituído de duas metodologias: **previsão individual** e **previsão agregada**.

O modelo de previsão individual considera a série temporal de carga para cada barramento isoladamente. Desta forma, foi necessário o ajuste de um modelo de previsão para cada um dos barramentos sendo assim, realizadas um grande número de previsões, proporcionando um alto custo computacional.

Já a previsão de carga agregada é um problema mais complexo pois a previsão depende de 3 fases:

- 1- os barramentos devem ser agrupados em grupos com comportamentos similares;
- 2- calcula-se a demanda agregada e faz-se uma previsão desta carga agregada;
- 3- determina-se a carga individual baseada na previsão da carga agregada.

Neste processo, as etapas mais críticas são as de agrupamento e do cálculo das cargas

individuais.

Neste trabalho, para resolver a fase 1 foram utilizados dois algoritmos de agrupamento: *Pruning Technique Kohonen Self-Organizing Feature Map - PSOM* e *Fuzzy C-means - FCM*. Ambos algoritmos mostraram-se eficientes no agrupamento de barramentos com perfis de comportamentos similares. Evidentemente, o grau de similaridade dos elementos de cada grupo formado por estes algoritmos depende do histórico de cargas. Porém, de um modo geral, obteve-se grupos com alto grau de similaridades, outros médios e outros grupos com barramentos pouco similares. Para estes últimos grupos, as previsões agregadas apresentaram erros elevados, pois a demanda agregada foi pouco similar às demandas de cargas individuais. Portanto, para estes barramentos deve-se dar um tratamento específico.

Os algoritmos de agrupamento se mostraram úteis no processo de previsão por barramento, pois as previsões por grupos se mostraram eficientes e estes algoritmos também possibilitaram identificar os barramentos que necessitam de tratamento específico.

Um dos grandes problemas dos Modelos de Previsão Agregada - MPA está no processo de cálculo das cargas individuais nos barramentos (desagregação). Assim neste trabalho foram estudadas três alternativas para efetuar este cálculo: Previsão via Fator de Participação - **PFP**, Previsão da Média Diária da Demanda Agregada - **PMDDa** e Previsão via Rede Treinada para Demanda Agregada - **PRTDa**. Destes métodos de previsão via demanda agregada, as 3 abordagens obtiveram desempenho equivalente, sendo que nenhum se mostrou eficiente para o tratamento de todos os barramentos do conjunto analisado.

No problema de previsão, quando compara-se os modelos de previsão individual com os agregados notou-se que os MPI foram mais eficientes em grande parte dos barramentos. Entretanto, os resultados dos MPA apesar de inferiores quando comparados aos MPI, não apresentaram grandes variações tendo, em alguns barramentos, erros inferiores ao modelo MPI.

Desta forma, pode-se concluir que os MPA propostos foram eficientes na resolução do problema de previsão de carga por barramento, pois apresentam um menor número de previsões e resultados aceitáveis quando comparados aos MPI.

Através deste estudo observou-se três tipos de grupos de barramentos que podem ser tratados especificamente:

- **Conjunto de barramentos agregados:** pertenceram a essa classe todos os barramentos que

fazem parte de grupos com alto grau de similaridade;

- **Conjunto de barramentos com carga reduzida** : pertenceram a essa classe os barramentos que apresentam uma demanda de carga muito reduzida. Para essas barras a sugestão é efetuar a previsão através da média aritmética da carga verificada recentemente;
- **Conjunto de barramentos especiais**: pertenceram a essa classe os barramentos que apresentaram um comportamento atípico, como por exemplo, barramentos que alimentam exclusivamente uma dada indústria, ou por exemplo, barramentos em zonas rurais cujas cargas são muito influenciadas pelo período do ano.

Do ponto de vista de modelagem, os modelos de previsão descritos neste trabalho apresentaram-se altamente promissores para esta aplicação. Entretanto, tais metodologias podem ser aprimoradas e especializadas visando a obtenção de resultados mais eficientes.

6.2 Trabalhos Futuros

O primeiro passo a ser tomado futuramente é testar e verificar a robustez dos métodos propostos nesta pesquisa em um sistema de transmissão e sub-transmissão de grande porte, ou seja, onde existam uma grande quantidade de barramentos.

Em termos de trabalhos futuros, os modelos de previsão por barramento podem ser aperfeiçoados no seguinte sentido:

1. Um estudo detalhado das séries temporais de cada barramento, utilizando um algoritmo de filtragem, buscando eliminar todos os dados não relevantes, para fornecer entradas mais consistentes aos modelos;
2. Pesquisar novas formas de agregação dos barramentos;
3. Aplicar outros modelos de previsão de séries temporais;
4. O desenvolvimento de um sistema de previsão de carga por barramento amigável, que contenha uma interface gráfica de modo que o usuário final possa interagir e tomar decisões auxiliado pelo sistema.

Referências

- Anderson, J. (1995). *Learning and Memory*, John Wiley and Sons, New York.
- Ankerst, M., Breunig, M. e Kriegel, H. (1999). Optics: Ordering points to identify the clustering structure, *Proceedings of the ACM SIGMOD Conference on Management of Data* pp. 49–60.
- Backer, E. (1995). *Computer Assisted Reasoning in Cluster Analysis*, Prentice Hall, New York.
- Ballini, R. (2000). *Análise e previsão de vazões utilizando modelos de séries temporais, redes neurais e redes neurais nebulosas*, PhD thesis, Unicamp.
- Bezdek, C. J. e Pal, S. K. (1992). *Fuzzy Models for Pattern Recognition*, IEEE Press, New York.
- Bezdek, J. (1981). *Pattern Recognition with Fuzzy Objective Function Algorithms*, Plenum Press, New York.
- Bose, I. e Mahapatra, R. (2001). Business data minig - a machine learning perspective, *Information & Management* **39**: 211–225.
- Braga, A., Carvalho, A. P. e Ludemir, T. (1998). Fundamentos de redes neurais artificiais, *11a Escola de Computação/Universidade Federal do Rio de Janeiro*. .
- Cardoso, G. C. (2003). *Modelo de previsão baseado em agrupamento e base de regras nebulosas*, Master's thesis, Unicamp.
- Castro, L. N. (1998). *Análise e síntese de estratégias de aprendizado para redes neurais artificiais*, Master's thesis, FEEC-Unicamp, Campinas.
- Castro, L. e VonZuben, F. (1999). An improving pruning technique with restart for the kohonen self-organizing feature map, *In Proceedings of the IEEE International Joint Conference on Neural Networks* **3**: 1916–1919.
- Cho, S.-B. (1997). Self-organizing map with dynamical node splitting: Application to handwritten digit recognition, *Neural Computation* **9**: 1345–1355.
- Cole, R. M. (1998). *Clustering with genetic algorithms*, Master's thesis, Department of Computer Science, University of Western Australia.
- Dillon, T. S., Sestito, S. e Leung, S. (1991). Short-term load forecasting using an adaptive, *Eletrical Power and Energy Systems* **13**(4): 186–192.

- Erkmen, I. e Özökmen (1993). An hybrid neural network for short-term load forecasting, *Proceedings of IEEE/NTUA Athens Power Tech Conference: Planning Operation and Control of Today's Electric Power Systems II*: 811–815.
- Francelin, R. A. e Vieira, L. R. (1994). A rede de kohonen, *Technical report*, Universidade de São Paulo-IMCMSC.
- Gross, G. e Galiana, F. (1987). Short-term load forecasting, *Proceedings of the IEEE* **75**(12): 1558–1573.
- Handschin, E. e Dörnmann, C. (1988). Bus load modelling and forecasting, *IEEE Transactions on Power Systems* **3**(2): 627–633.
- Haykin, S. (1994). *Neural Networks, A Comprehensive Foundation*, Macmillan Publishing Company.
- Ho, K., Hsu, Y. e Yang, C. (1992). Short-term load forecasting using a multi-layer neural network with an adaptative learning algorithm, *IEEE Transactions on Power Systems* **7**: 141–149.
- Hopfield, J. (1982). Neural networks and physical systems with emergent collective computational abilities, *Proceedings of the National Academy of Sciences of the U.S.A* **79**: 2554–2558,.
- Hruschka, E. R. e Ebecken, N. F. (2001). A genetic algorithm for cluster analysis, *IEEE Transactions on Evolutionary Computation* .
- Kadowaki, M., Ohishi, T., Soares, S. e Ballini, R. (2002). Short term load forecasting using a neurofuzzy network model, *Congresso Brasileiro Automática* **1**: 365–370.
- Kaski, S. e Kohonen, T. (1995). Exploratory data analysis by the self-organizing map: Structures of welfare and poverty in the world, *Proceedings of the Third International Conference on Neural Networks in the Capital Markets* pp. 498–507.
- Khalil, M., Panu, U. e Lennox, W. (2001). Groups and neural networks based streamflow data infilling procedures, *Journal of Hydrology* **241**: 153–176.
- Klir, G. J. e Yuan, B. (1995). *Fuzzy Sets and Fuzzy Logic Theory and Applications*, Prentice Hall.
- Kohonen, T. (1982). Self-organized formation of topologically correct feature maps, *Biological Cybernetics* **43**: 59–69.
- Kohonen, T. (2000). *Self-Organizing Maps*, 3rd edn, Springer-Verlag.
- Lachtermacher, G. e Fuller, J. D. (1995). Backpropagation in time-series forecasting, *Journal of Forecasting* **14**(4): 381–393.
- Lapedes, A. e Farber, R. (1987). Nonlinear signal processing using neural networks: Prediction and system modelling, *Technical report*, Los Alamos National Laboratory.
- Last, M., Klein, Y. e Kandel, A. (2001). Knowledge discovery in time series databases, *IEEE Transactions on Systems, Man and Cybernetics - Part B: Cybernetics* **31**(1): 160–169.

- Lima, W. S. (1996). *Previsão de carga no período de demanda de ponta utilizando redes neurais artificiais*, Master's thesis, FEEC-Unicamp, Campinas.
- Lima, W. S. e Ohishi, T. (1999). Mapas auto-organizáveis não-paramétricos para análise da influência climática em curvas de carga, *Seminário Nacional de Produção e Transmissão de Energia Elétrica* **14**: 35–62.
- Lu, C. N., Wu, H. T. e Vemuri, S. (1993). Neural network based short-term load forecasting, *IEEE Transactions on Power Systems* **8**(1): 336–342.
- Maier, H. e Dandy, G. (1998). Understanding the behaviour and optimising the performance of back-propagation neural networks: an empirical study, *Environmental Modelling and Software* **13**: 179–191.
- Masterocostas, P., Theocharis, J. e Bakirtzis, A. (1999). Fuzzy modeling for short term load forecasting using the orthogonal least squares method, *IEEE Transactions on Power Systems* **14**(1).
- McCulloch, W. e Pitts, W. (1943). A logical calculus of ideas immanent in nervous activity, *Bulletin of Mathematical Biophysics* **5**: 115–133.
- Minsky, P. L. e Papert, S. A. (1969). *Perceptrons*, MIT - Press, Cambridge, MA.
- Mohammed, O., Park, D., Merchant, R. Dinh, T., Tong, C., Azzem, A., Farah, J. e C., D. (1995). Practical experiences with an adaptive neural network short-term load forecastin system, *IEEE Transactions on Power Systems* **10**(1): 254–265.
- Oates, T., Firofu, L. e Cohen, P. (2000). Using dynamic time warping to bootstrap hmm-based clustering of time series, *Sequence Learning - Paradigms, Algorithms and Applications*.
- Park, D. C., El-Sharkawi, A., M. e Marks II, R. J. (1991). Eletric load forecasting using an artificial neural network, *IEEE Transactions on Power Systems* **6**(2): 442–449.
- Pedrycz, W. e Gomide, F. (1998). *An Introduction to Fuzzy sets: Analysis and Design*, MIT Press, Cambridge, M.A.
- Peng, T. M., Hubele, N. F. e Karady, G. G. (1992). Advancement in the application of neural networks for short-term load forecasting, *IEEE Transactions on Power Systems* **7**(1): 250–257.
- Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain, *Psychological Review* **65**: 386–408.
- Rumelhart, D., Hinton, G. e Willians, R. (1986). *Learning internal representation by error propagation, Parallel Distributed Processing: Explorations in the Microstruture of Cognition*, Vol. 1, MIT Press, Cambridge.
- Salgado, R. M., Ohishi, T. e Ballini, R. (2004a). Agrupamento de curvas de demanda, *Proceedings of CBA - Congresso Brasileiro de Automática*.

- Salgado, R. M., Ohishi, T. e Ballini, R. (2004b). Clustering bus load curves, *Proceedings of Power Systems Conference Expositon PSCE- IEEE* .
- Shalizi, C., Shalizi, K. e Crutchfield, J. (2002). Pattern discovery in time series, part i: theory, algorithm, analysis, and convergence, *Technical report*, Santa Fe Institute.
- Subramanian, V., Hung, M. e Hu, M. (1993). An experimental evaluation of neural networks for classification, *Computers & Operations Research* **20**(7): 769–782.
- Tang, Z., Almeida, C. e Fishwick, P. A. (1991). Time series forecasting using neural networks vs. Box-Jenkins methodology, *Simulation* **57** 5: 303–310.
- Weigend, A. S., Rumelhart, D. e Huberman, B. A. (1990). Predicting the future: A connectionist approach, *International Journal Neural System* **1**:3.
- Weigend, A. S., Rumelhart, D. e Huberman, B. A. (1991a). Generalization by weight-elimination application to forecasting, advances in neural, *Information Processing Systems* **3**: 875–882.
- Weigend, A. S., Rumelhart, D. e Huberman, B. A. (1991b). Generalization by weight-elimination applied to currency of neural networks, *Proceedings of the IEEE* **1**: 837–841.
- Widrow, B. e Hoff, N. (1960). Adaptative switching circuits, *IRE WESCON Convencion Record* pp. 66–104.
- Widrow, B. e Hoff, N. (1962). Generalization and information storage in networks of adaline neurons, *Self - Organizing Systems* pp. 435–461.
- Zhang, Q., Patuwo, B. e Hu, M. (1998). Forecasting with artificial neural networks: the state of the art, *International Journal of Forecasting* **14**: 35–62.

APÊNDICE A – MLP e Backpropagation

Neste capítulo apresenta-se de maneira detalhada os conceitos teóricos que fundamentam as MLP e o algoritmo de retropropagação do erro.

A.1 Noções Gerais

A figura 28 apresenta uma arquitetura do tipo MLP com duas camadas intermediárias. A rede está conectada a todas as outras unidades (neurônios) na camada anterior. O fluxo de sinais através da rede é feito da camada de entrada para a camada de saída.

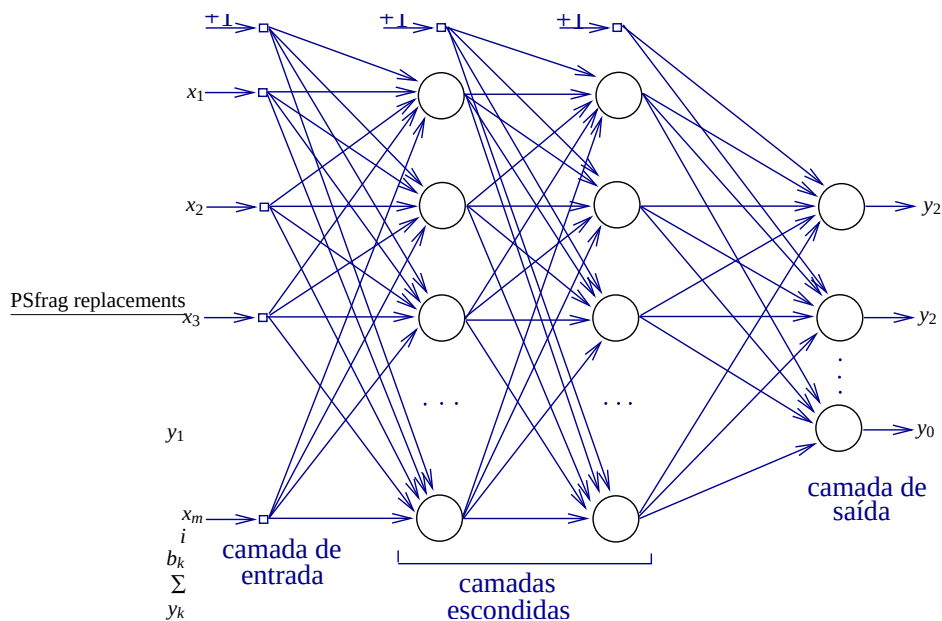


Figura 28: Arquitetura MLP com duas camadas intermediárias.

A camada de entrada geralmente é composta por neurônios sensoriais, ou seja, unidades que não modificam os sinais externos, apenas os distribuem para a primeira camada intermediária.

As unidades de saída constituem a camada de saída da rede e as demais unidades constituem as camadas intermediárias. Cada unidade intermediária ou de saída é responsável por duas tarefas:

- Calcular o sinal na saída da unidade, que geralmente é expresso como uma função não-linear do sinal de entrada e pesos sinápticos associados;
- Calcular uma estimativa instantânea do vetor gradiente, que é necessário para a retro-propagação do erro através da rede;

A.1.1 Algoritmo de Aprendizado para uma rede MLP

O algoritmo de retro-propagação consiste em dois passos de computação: o processamento direto e o processamento reverso. No processamento direto, uma entrada é aplicada à rede neural e seu efeito é propagado pela rede, camada a camada. Durante o processamento direto, os pesos da rede permanecem fixos. No processamento reverso, um sinal de erro calculado na saída da rede é propagado no sentido reverso, camada a camada, e ao final deste processo os pesos são ajustados de acordo com uma regra de correção de erro. A seguir, apresentam-se o algoritmo de retropropagação.

O sinal de erro na saída do neurônio j na iteração n é definido por:

$$e_j(n) = s_j(n) - y_j(n), \quad (\text{A.1})$$

sendo que $s_j(n)$ é a resposta desejada para o neurônio j da camada de saída.

O erro quadrático para o neurônio j é definido como sendo $\frac{1}{2}e_j^2(n)$. A soma dos erros quadráticos da rede é então definida por:

$$\varepsilon(n) = \frac{1}{2} \sum_{j \in R} e_j^2(n) \quad (\text{A.2})$$

em que R é o conjunto que contém todos os neurônios da camada de saída da rede.

Seja N o número total de padrões de treinamento. O erro quadrático médio é dado por:

$$E_{cust} = \frac{1}{N} \sum_{n=1}^N \varepsilon(n) \quad (\text{A.3})$$

A soma dos erros quadráticos e, portanto o erro quadrático médio, é função dos pesos sinápticos da rede. Para um dado conjunto de treinamento, E_{cust} representa uma função de custo, isto é, uma medida de desempenho da rede neural. O objetivo do processo de treinamento é minimizar E_{cust} .

Neste trabalho é considerado o método de treinamento em que os pesos são ajustados *padrão-a-padrão*, ou seja, os pesos da rede são ajustados de acordo com o erro computado para *cada* padrão apresentado à rede.

A figura 29, mostra o neurônio de saída j , alimentado pelas ativações de todos os neurônios da camada imediatamente anterior. O nível de ativação interno do neurônio j é dado por:

$$v_j(n) = \sum_{i=0}^p w_{ji}(n) y_i(n) \quad (A.4)$$

em que p é o número total de entradas (excluindo a polarização) aplicadas ao neurônio j , ou seja, o número de neurônios da camada imediatamente anterior. Portanto a ativação $y_j(n)$ do neurônio j é dada por

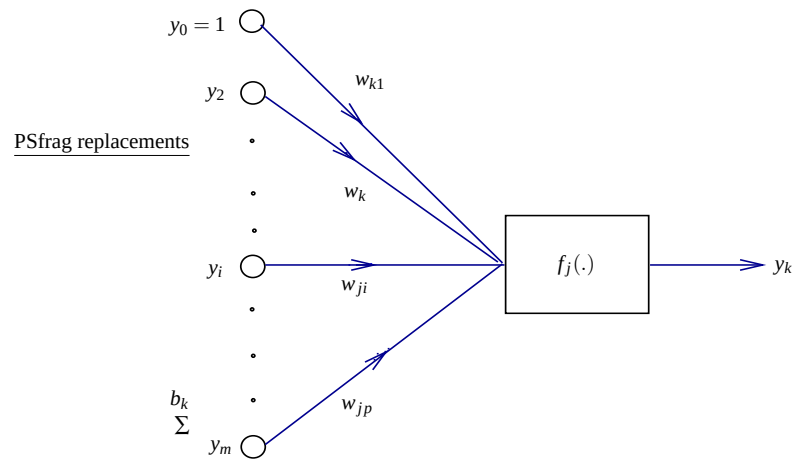


Figura 29: Saída do Neurônio j .

$$y_j(n) = f_j(v_j(n)) \quad (A.5)$$

Para minimizar o erro quadrático médio, é necessário determinar $\partial \varepsilon(n) / \partial w_{ji}(n)$, isto é,

o gradiente instantâneo. Aplicando a regra da cadeia, o gradiente é expresso como sendo:

$$\frac{\partial \varepsilon(n)}{\partial w_{ji}(n)} = \frac{\partial \varepsilon(n)}{\partial e_j(n)} \frac{\partial e_j(n)}{\partial y_j(n)} \frac{\partial y_j(n)}{\partial v_j(n)} \frac{\partial v_j(n)}{\partial w_{ji}(n)} \quad (\text{A.6})$$

Derivando a equação (A.2) em relação a $e_j(n)$, obtém-se,

$$\frac{\partial \varepsilon(n)}{\partial e_j(n)} = e_j(n) \quad (\text{A.7})$$

Derivando a equação (A.1) em relação a $y_j(n)$, obtém-se,

$$\frac{\partial e_j(n)}{\partial y_j(n)} = -1 \quad (\text{A.8})$$

Da mesma forma, derivando a equação (A.5) em relação a $v_j(n)$, obtém-se,

$$\frac{\partial y_j(n)}{\partial v_j(n)} = f'_j(v_j(n)) \quad (\text{A.9})$$

Finalmente, derivando a equação (A.4) em relação a $w_{ji}(n)$ obtém-se,

$$\frac{\partial v_j(n)}{\partial w_{ji}(n)} = y_i(n) \quad (\text{A.10})$$

Substituindo as equações (A.7) a (A.10) na equação (A.6), obtém-se,

$$\frac{\partial \varepsilon(n)}{\partial w_{ji}(n)} = -e_j(n) f'_j(v_j(n)) y_i(n) \quad (\text{A.11})$$

que é a equação da derivada do erro instantâneo $\varepsilon(n)$ em relação ao peso sináptico $w_{ji}(n)$ do neurônio j da camada de saída.

Note que a equação (A.11) corresponde a um componente do vetor gradiente do erro, cujos elementos representam a derivada parcial de $\varepsilon(n)$ em relação a todos os pesos da rede neural, arranjados em uma ordem fixa, mas arbitrária.

Definindo o gradiente local $\delta_j(n)$ como sendo:

$$\delta_j(n) = -\frac{\partial \varepsilon(n)}{\partial e_j(n)} \frac{\partial e_j(n)}{\partial y_j(n)} \frac{\partial y_j(n)}{\partial v_j(n)} = e_j(n) f'_j(v_j(n)) \quad (\text{A.12})$$

a equação (A.11) pode ser reescrita da seguinte maneira,

$$\frac{\partial \varepsilon(n)}{\partial w_{ji}(n)} = -\delta_j(n) y_i(n) \quad (\text{A.13})$$

Para minimizar o erro, aplica-se aos pesos uma correção proporcional ao oposto do gradiente do erro, ou seja, na direção oposta ao do gradiente. Esta regra de ajuste pode ser expressa:

$$w_{ji}(n+1) = w_{ji}(n) - \eta(n) \frac{\partial \varepsilon(n)}{\partial w_{ji}(n)} \quad (\text{A.14})$$

em que $\eta(n)$ é o passo do algoritmo na iteração n .

A equação (A.14) pode ser reescrita da seguinte forma:

$$w_{ji}(n+1) = w_{ji}(n) + \Delta w_{ji}(n), \quad (\text{A.15})$$

sendo que a expressão $\Delta w_{ji}(n)$ é conhecida como Regra Delta para ajuste dos pesos.

Pelas equações (A.12) e (A.13), nota-se que um fator fundamental para o cálculo do gradiente do erro é o sinal do erro $e_j(n)$. Deve-se então considerar dois casos distintos:

Caso I: O Neurônio j é um Neurônio na camada de saída da rede

Se o neurônio j é um neurônio de saída, então a saída desejada $d_j(n)$ para o neurônio é conhecida. Desta forma, a equação (A.1) pode ser usada para calcular o sinal de erro $e_j(n)$ associado ao neurônio j , com $s_j(n) = d_j(n)$. Com o erro $e_j(n)$ determinado, o cálculo de $\delta_j(n)$ é facilmente obtido através da equação (A.12).

Caso II: Neurônio j é um Neurônio pertencente a uma camada escondida

Como o neurônio j é um neurônio escondido, então não há nenhuma saída desejada pré-especificada para o neurônio. Desta forma, o sinal de erro deve ser calculado através dos sinais de erro de todos os neurônios aos quais o neurônio escondido está diretamente conectado. A figura 30, mostra o neurônio j como um neurônio escondido da rede. Pela equação (A.12), pode-se redefinir o gradiente local $\delta_j(n)$ para o neurônio escondido j como sendo:

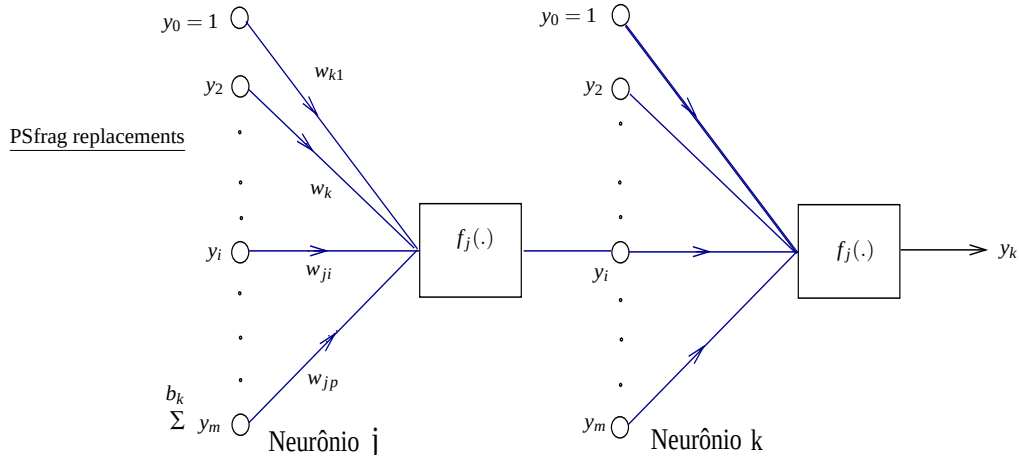


Figura 30: Saída do Neurônio Intermediário.

$$\delta_j(n) = -\frac{\partial \varepsilon(n)}{\partial y_j(n)} \frac{\partial y_j(n)}{\partial v_j(n)} = -\frac{\partial \varepsilon(n)}{\partial y_j(n)} f'_j(v_j(n)) \quad (\text{A.16})$$

A derivada parcial $\partial \varepsilon(n) / \partial y_j(n)$, é calculada através dos seguintes passos: para o neurônio k pertencente à camada de saída figura 30, o erro é dado por:

$$\varepsilon(n) = \frac{1}{2} \sum_{k \in C} e_k^2(n) \quad (\text{A.17})$$

Derivando a equação (A.17) em relação a $y_j(n)$, obtém-se,

$$\frac{\partial \varepsilon(n)}{\partial y_j(n)} = \sum_k e_k \frac{\partial e_k(n)}{\partial y_j(n)} \quad (\text{A.18})$$

Usando a regra da cadeia para calcular $\partial e_k(n) / \partial y_j(n)$, pode-se reescrever a equação (A.18) da seguinte forma;

$$\frac{\partial \varepsilon(n)}{\partial y_j(n)} = \sum_k e_k(n) \frac{\partial e_k(n)}{\partial v_k(n)} \frac{\partial v_k(n)}{\partial y_j(n)} \quad (\text{A.19})$$

O erro na saída do neurônio k é dado por:

$$e_k(n) = d_k(n) - y_k(n) = d_k(n) - f_k(v_k(n)) \quad (\text{A.20})$$

sendo $v_k(n)$ é o nível de ativação interna do neurônio k . Logo,

$$\frac{\partial e_k(n)}{\partial v_k(n)} = -f'_k(v_k(n)) \quad (\text{A.21})$$

Da figura 30, pode-se ver que o nível de ativação interna do neurônio k é dado por:

$$v_k(n) = \sum_{j=0}^q w_{kj}(n)y_j(n) \quad (\text{A.22})$$

em que q é o número total de entradas (excluindo a polarização) aplicadas ao neurônio k .

Derivando a equação A.1 em relação a $y_j(n)$ obtém-se,

$$\frac{\partial v_k(n)}{\partial y_j(n)} = w_{kj}(n) \quad (\text{A.23})$$

Desta forma, efetuando as devidas substituições (A.19) encontra-se a derivada parcial desejada:

$$\begin{aligned} \frac{\partial \varepsilon(n)}{\partial y_j(n)} &= -\sum_k e_k(n) f'_k(v_k(n)) w_{kj}(n) = \\ &= -\sum_k \delta_k(n) w_{kj}(n) \end{aligned} \quad (\text{A.24})$$

sendo que, na segunda linha, usou-se a definição de gradiente local $\delta_k(n)$ dada na equação (A.12),

Finalmente, substituindo a equação (A.24) na equação (A.16), obtem-se a expressão para o gradiente local $\delta_j(n)$ para o neurônio escondido j :

$$\delta_j(n) = f'_j(v_j(n)) \sum_k \delta_k(n) w_{kj}(n). \quad (\text{A.25})$$

O termo $f'_j(v_j(n))$ na equação (A.25) depende apenas da derivada da função de ativação associada ao neurônio j . O termo restante, isto é, o somatório sobre k , depende de dois conjuntos de parâmetros. O primeiro conjunto de parâmetros $\delta_k(n)$, exigem o conhecimento dos sinais de erro $e_k(n)$, para todos os neurônios localizados na camada imediatamente posterior à camada em que se encontra o neurônio j , e que estão diretamente conectados ao neurônio j figura 30. O segundo conjunto de parâmetros, $w_{kj}(n)$, consiste nos pesos sinápticos associados a estas conexões.

A.1.1.1 Derivada da Função de Ativação

Pelas equações (A.12) e (A.25), nota-se que é necessário conhecer a derivada da função de ativação $f(\cdot)$ para calcular os gradientes locais δ^1 . Uma das funções mais utilizadas é a função logística, definida na equação (3.4). Para um neurônio j , tem-se:

$$y_j(n) = f_j(v_j(n)) = \frac{1}{1 + e^{-v_j(n)}}$$

sendo que $v_j(n)$ é o nível de ativação interna do neurônio j .

Uma característica da função logística é que sua derivada pode ser expressa através da ativação $y_j(n)$, a saber:

$$f'_j(v_j(n)) = y_j(n)[1 - y_j(n)].$$

A função logística é limitada, assumindo valores no intervalo (0, 1).

A.1.2 Taxa de Aprendizagem

A taxa de aprendizagem η é uma constante de proporcionalidade, que assume valores no intervalo (0, 1) e proporciona a adaptação dos pesos.

Um alto valor de η produzirá uma rápida aprendizagem, resultando em grandes alterações nos pesos. Entretanto, existe o risco da rede tornar-se instável, podendo oscilar ao redor da solução e provocar a saturação dos neurônios (Ballini, 2000).

De outra maneira, se a taxa de aprendizagem assumir valores pequenos, as mudanças nos pesos sinápticos serão menores de uma iteração para a outra e as trajetórias no espaço definido pelos pesos serão mais suaves. Tal benefício é obtido com um alto custo computacional no processo de treinamento, pois este torna-se mais lento por requerer um número maior de iterações até o processo de convergência ser alcançado (Ballini, 2000).

Um método simples que aumenta a velocidade da aprendizagem é obtido através da modificação da Regra Delta (equação (A.15)) adicionando um termo *momentum* como proposto

¹A diferenciabilidade é uma propriedade fundamental que deve ser apresentada pela função de ativação.

por (Rumelhart et al., 1986).

$$\Delta w_{ji}(n) = \alpha \Delta w_{ji}(n-1) + \eta \delta_j(n) y_i(n) \quad (\text{A.26})$$

sendo α um número no intervalo $[0, 1)$, chamado constante *momentum*.

O termo *momentum* relaciona as alterações dos pesos sinápticos na iteração n com as alterações realizadas na iteração $(n-1)$. A utilização desta constante dificulta a mudança de tendência na atualização dos pesos como também pode ser usada para reduzir a possibilidade mínimos locais ocorrerem (Ballini, 2000).

A.1.3 Inicialização dos Pesos

O primeiro passo do algoritmo de retro-propagação é, a inicialização dos pesos da rede, o que corresponde à definição de um ponto inicial da superfície de erro. Uma boa escolha inicial dos pesos é fundamental para um bom desempenho do algoritmo de retro-propagação.

Como usualmente não se tem nenhuma informação que possa ser diretamente empregada na inicialização dos pesos da rede, um método usualmente utilizado é inicializar os pesos aleatoriamente, com distribuição uniforme sobre um pequeno intervalo em torno de zero (Haykin, 1994).

A.1.4 Critério de Parada

O critério de parada do algoritmo de retro-propagação não é bem definido e a convergência não é garantida. Um critério de parada é interromper o treinamento após um número fixo de iterações, mas esta forma não é recomendável por não levar em conta as informações sobre o estado do processo de aprendizagem (Ballini, 2000). Assim deve-se buscar um critério de convergência que considere a existência de mínimos locais. (Haykin, 1994) considera tal propriedade é o fato de que a função objetivo (erro quadrático médio) é estacionária no ponto $w = w^*$, onde w^* é um vetor peso que denota o mínimo, podendo ser local ou global. Este critério de parada pode ser definido da seguinte forma: o algoritmo de retro-propagação converge se a variação do erro quadrático médio de uma época para outra assumir um valor suficientemente pequeno.

Um outro critério de parada é testar, após cada iteração, o desempenho de generaliza-

ção do algoritmo. O treinamento é encerrado quando o desempenho de generalização apresentar tendência de piora.

A.1.5 Generalização

Generalização é uma das mais importantes características de redes neurais, podendo ser influenciada por três fatores:

1. O tamanho e a eficiência do conjunto de treinamento, que deve ser muito bem escolhido, pois deve expressar todas as características relevantes acerca do problema no universo considerado;
2. A arquitetura da rede neural: números de camadas e neurônios intermediários;
3. A complexidade do problema a ser tratado.

Uma rede neural apresenta uma boa generalização se, fornecidos os conjuntos de dados de treinamento e teste, a relação entre os dados de entrada/saída calculada pela rede já treinada está correta ou tão próxima quanto desejável para os padrões pertencentes ao conjunto de teste (Ballini, 2000).

A.1.6 Problemas com o Algoritmo de treinamento

Apesar do grande sucesso das aplicações do algoritmo backpropagation e da sua enorme popularidade, muitos problemas ainda o acompanham. Entre eles, os que mais se destacam são (Lima, 1996):

Período de treinamento: O treinamento pode demorar um longo período, principalmente para problemas complexos, sendo que não há garantias que depois deste tempo o treinamento tenha sido um sucesso. O termo *momentum* pode ser uma alternativa para diminuir o tempo de treinamento, assegurando um maior grau de estabilidade para a rede. Foi proposto por Rumelhart, Hinton e Williams em 1986 (Rumelhart et al., 1986) e se mostrou bastante eficiente em alguns casos.

Mínimos locais: O treinamento é um processo de otimização por gradiente, e sofre dos males inerentes a estes processos. Um dos problemas é que como E_{cust} é não-linear podem existir

mínimos locais. O algoritmo de retropropagação não garante que o erro mínimo seja alcançado durante o processo de treinamento, mas apenas algum erro local mínimo.

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

APÊNDICE B – Tabelas - MAE

As Tabelas abaixo mostram a eficiência dos de previsão de carga por barramento medidos em termos de erro absoluto médio - MAE.

B.1 Metodologias de Previsão Individual

As Tabelas 17 e 18 mostram os erros absolutos médios obtidos pelos modelos de previsão da curva diária (PCD) e previsão da média diária (PMD), respectivamente, ao longo da semana de 19 a 26/09/01.

Tabela 17: MAE Diário - PCD.

-	qui	sex	sáb	dom	seg	ter	qua	-
Barramentos	20/set	21/set	22/set	23/set	24/set	25/set	26/set	Erro Médio
1	0,40	0,58	0,69	0,45	0,55	0,56	0,53	0,54
2	0,06	0,07	0,05	0,06	0,08	0,08	0,10	0,07
3	1,59	1,55	1,26	1,01	0,99	2,44	2,00	1,55
4	0,30	0,22	0,18	0,19	0,73	1,02	0,55	0,45
5	1,88	2,31	7,05	5,71	2,73	2,40	3,32	3,63
6	4,85	5,02	3,46	8,95	4,86	4,26	6,91	5,47
7	0,44	0,36	0,25	1,10	0,29	0,25	0,21	0,41
8	0,40	0,30	0,53	0,74	0,37	0,46	0,30	0,44
9	2,86	1,67	3,12	5,02	2,16	3,96	2,15	2,99
10	2,02	2,91	1,14	1,36	2,18	2,41	2,75	2,11
11	1,36	1,38	0,81	1,61	2,69	3,42	1,02	1,76
12	1,11	1,08	0,86	0,47	1,45	1,11	0,64	0,96
13	0,21	0,18	0,24	0,22	0,20	0,18	0,25	0,21
14	3,02	2,68	2,96	2,17	1,37	2,91	3,87	2,71
15	0,22	0,42	0,32	0,16	0,33	0,28	0,30	0,29

Tabela 18: MAE Diário - PMD.

-	qui	sex	sáb	dom	seg	ter	qua	-
Barramentos	20/set	21/set	22/set	23/set	24/set	25/set	26/set	Erro Médio
1	0,38	0,33	0,72	0,40	0,69	0,46	0,59	0,51
2	0,04	0,06	0,04	0,03	0,02	0,05	0,06	0,04
3	0,93	0,68	0,63	0,77	0,65	1,09	0,86	0,80
4	0,36	0,41	0,32	0,52	0,36	1,77	0,38	0,59
5	4,40	5,55	3,74	4,78	2,37	5,58	5,66	4,58
6	5,26	4,53	2,94	4,69	3,59	3,90	6,76	4,52
7	0,44	0,35	0,21	2,10	1,30	0,28	0,39	0,72
8	1,13	1,21	1,54	0,41	0,26	0,65	1,30	0,93
9	2,84	2,77	1,81	3,75	2,70	3,65	5,45	3,28
10	5,03	4,10	4,64	4,55	5,75	3,66	4,56	4,61
11	2,48	2,16	1,23	1,40	2,85	3,41	1,70	2,18
12	0,41	0,32	0,19	0,71	0,81	1,20	0,35	0,57
13	0,32	0,16	0,15	0,12	0,11	0,23	0,16	0,18
14	1,64	3,90	1,66	0,75	1,40	5,65	3,16	2,59
15	0,18	0,25	0,25	0,45	0,51	0,46	0,25	0,33

B.2 Metodologias de Previsão Agregada

As Tabelas 19 e 20 mostram os erros absolutos médios obtidos pelos modelos de previsão de fator de participação via médias (PFP/Médias) e via redes neurais (PFP/RNA) na semana de 19 a 26/09/01.

Tabela 19: MAE Diário - PFP/Médias.

-	qui	sex	sáb	dom	seg	ter	qua	-
Barramentos	20/set	21/set	22/set	23/set	24/set	25/set	26/set	Erro Médio
1	0,68	0,4	2,23	3,59	4,05	2,18	0,81	1,99
2	0,02	0,03	0,08	0,17	0,13	0,09	0,05	0,08
3	1,38	1,06	4,25	6,18	7,30	2,21	1,02	3,34
4	0,25	0,33	0,26	0,49	0,54	2,29	1,52	0,81
5	2,32	1,56	6,12	12,79	16,99	5,94	3,27	7,00
6	3,21	2,76	27,46	30,36	39,89	8,47	7,06	17,03
7	0,33	0,18	1,34	1,24	1,78	0,55	0,46	0,84
8	0,31	0,26	0,87	1,87	2,21	0,77	0,58	0,98
9	2,47	2,91	19,81	22,25	31,99	3,11	4,09	12,38
10	1,82	1,9	2,43	3,71	8,48	5,93	2,55	3,83
11	1,24	3,46	1,85	9,73	8,01	3,92	1,97	4,31
12	0,54	0,49	0,98	1,76	2,32	1,48	0,34	1,13
13	0,1	0,1	0,12	0,08	0,35	0,12	0,17	0,15
14	1,44	2,18	2,22	8,38	10,80	1,7	5,02	4,53

Tabela 20: MAE Diário - PFP/RNA.

-	qui	sex	sáb	dom	seg	ter	qua	-
Barramentos	20/set	21/set	22/set	23/set	24/set	25/set	26/set	Erro Médio
1	0,94	2,18	2,18	0,91	1,19	1,44	1,04	1,38
2	0,06	0,05	0,05	0,90	0,28	0,06	0,13	0,23
3	0,97	0,87	0,82	1,01	2,81	1,95	1,65	1,38
4	0,24	0,24	0,25	3,46	4,32	0,47	1,49	1,33
5	2,13	2,41	2,03	2,60	8,64	2,69	2,53	3,16
6	5,66	4,84	3,84	3,63	12,71	7,16	7,84	6,03
7	4,03	6,15	5,49	4,80	0,47	3,44	5,68	4,30
8	0,62	0,33	0,59	0,52	1,14	0,23	0,55	0,55
9	3,74	2,06	4,82	5,33	11,11	7,48	3,72	5,22
10	1,75	1,69	1,87	1,66	4,25	2,18	2,87	2,30
11	1,11	1,17	0,73	1,04	3,07	3,38	2,41	1,62
12	0,48	0,37	0,35	0,40	0,67	0,72	0,42	0,48
13	0,07	2,10	0,10	1,39	1,90	2,02	2,15	1,38
14	0,88	0,76	2,55	2,95	2,24	1,89	2,69	1,99

As Tabelas 21 e 22 mostram os erros absolutos médios obtidos pelo modelo de previsão via rede treinada (PRTDa) e pelo modelo de previsão da média diária (PMDDa) da demanda agregada na semana de 19 a 26/09/01.

Tabela 21: MAE Diário - PRTDa.

-	qui	sex	sáb	dom	seg	ter	qua	-
Barramentos	20/set	21/set	22/set	23/set	24/set	25/set	26/set	Erro Médio
1	0,80	0,96	0,88	1,39	3,38	1,33	0,94	1,38
2	0,04	0,03	0,06	0,05	0,22	0,06	0,06	0,07
3	2,69	2,07	0,71	2,12	1,81	2,77	1,85	2,00
4	0,21	0,25	0,28	0,35	0,35	0,44	0,73	0,37
5	4,30	1,68	2,54	3,50	4,75	7,30	11,06	5,02
6	4,29	9,25	2,49	11,93	8,51	16,08	4,85	8,20
7	0,28	0,51	0,21	0,53	0,97	0,8	0,78	0,58
8	0,3	0,28	0,43	1,10	0,60	0,81	0,54	0,58
9	4,68	6,77	2,17	9,02	7,83	16,05	4,03	7,22
10	1,89	1,39	1,67	2,25	2,58	2,96	3,88	2,37
11	2,72	3,50	2,18	1,31	9,00	4,87	3,46	3,86
12	0,57	0,51	0,91	1,06	2,03	1,34	0,62	1,01
13	0,16	0,13	0,12	0,13	0,13	0,16	0,14	0,14
14	6,01	2,95	4,29	3,14	18,54	3,71	1,67	5,76

Tabela 22: MAE Diário - PMDDa.

-	qui	sex	sáb	dom	seg	ter	qua	-
Barramentos	20/set	21/set	22/set	23/set	24/set	25/set	26/set	Erro Médio
1	0,72	0,40	2,15	0,43	0,86	0,44	0,60	0,80
2	0,03	0,06	0,09	0,03	0,05	0,05	0,05	0,05
3	0,93	0,68	0,63	0,77	0,65	1,09	0,86	0,80
4	0,37	0,59	0,47	0,84	0,49	1,26	1,45	0,78
5	5,71	6,46	5,98	5,89	4,47	5,14	6,11	5,68
6	6,85	6,17	2,90	5,37	3,96	6,01	4,96	5,18
7	0,49	0,41	0,21	3,20	0,81	0,66	0,33	0,87
8	0,66	0,98	1,79	0,41	0,26	0,52	0,36	0,71
9	6,98	5,37	1,85	3,69	2,76	4,01	3,85	4,07
10	3,68	3,13	2,10	3,13	4,02	5,42	5,27	3,82
11	5,22	2,04	1,60	1,47	2,81	5,14	4,12	3,20
12	0,75	0,72	1,57	0,72	1,04	1,10	0,35	0,89
13	0,42	0,32	0,17	0,23	0,13	0,33	0,29	0,27
14	6,65	8,77	7,06	3,72	2,67	6,26	7,82	6,14