

DSEE

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



Aplicação de Redes Neurais
Artificiais na Previsão de Curto Prazo da
Carga Nodal de Sistemas de Energia Elétrica.

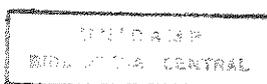
Autor: Francisco Eugenio de Andrade Leite

Orientador: Prof. Dr. André Luiz Morelato França

Dissertação apresentada à Faculdade de Engenharia Elétrica e de Computação da UNICAMP, como parte dos requisitos exigidos para a obtenção do título de Mestre em Engenharia Elétrica.

Campinas, Maio de 1997

Este exemplar corresponde à edição final da tese
elaborada por *Francisco Eugenio de Andrade Leite*
e aprovada pela Comissão
 julgadora em *05 / 05 / 1997*.
André Luiz Morelato França, Orientador



9708253

UNIDADE	BC
N.º CHAMADA	1/UNICAMP
	L5360
V	Es.
TÍTULO	BC/31078
PRD.	284197
C	<input type="checkbox"/>
D	<input checked="" type="checkbox"/>
PREÇO	R\$ 11,00
DATA	15/07/97
N.º CPD	

CM-0009883 0-6

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

L536a

Leite, Francisco Eugenio de Andrade

Aplicação de redes neurais artificiais na previsão de curto prazo da carga nodal de sistemas de energia elétrica / Francisco Eugenio de Andrade Leite.--Campinas, SP: [s.n.], 1997.

Orientador: André Luiz Morelato França.

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

1. Sistemas de energia elétrica - Controle. 2. Redes neurais (Computação). 3. Análise de séries temporais. 4. Redes elétricas - Análise. I. França, André Luiz Morelato. II. Universidade Estadual de Campinas. Faculdade de Engenharia Elétrica e de Computação. III. Título.

Lista de Abreviaturas

ADALINE - *Adaptive Linear Combiner*.

ARX - *Auto-regressive model with exogenous variables* (modelo linear de auto-regressão).

ARMAX - *Auto-regressive Moving Average model with exogenous variables* (modelo de auto-regressão com médias móveis).

ARIMAX - *Auto-regressive Integrated Moving Average model with exogenous variables* (modelo de auto-regressão de médias móveis integrado).

ANN - *Artificial Neural Network*.

ANNPS - *International Forum on Applications of Neural Networks to Power Systems*.

ANNSTLF - *Artificial Neural Network Short Term Load Forecasting* (previsão de carga de curto prazo por redes neurais artificiais).

ASHRAE - *American Society of Heating, Refrigerating and Air-Conditioning Engineers*.

CV - Coeficiente de Variação. (medida estatística).

CPFL - Companhia Paulista de Força e Luz.

FIR - *Finite Impulse Response Filter* (filtro de resposta finita ao impulso).

IEEE - *The Institute of Electrical and Electronic Engineers*.

ISAP - *Intelligent Systems Applications to Power*.

MA - *Moving Average model*.

MADALINE - *Multiple Adaptive Linear Combiner*.

MIMO - *Multiple Input Multiple Output* (modelo de múltiplas entradas e múltiplas saídas).

MISO - *Multiple Input Single Output* (modelo de múltiplas entradas e uma única saída).

MLP - *Multiple Layer Perceptron* (rede de perceptrons de múltiplas camadas).

NARX - *Nonlinear ARX model*

NARMAX - *Nonlinear ARMAX model*

NARIMA - *Nonlinear ARIMAX model*

NOE - *Nonlinear Output-Error model*

NFIR - *Nonlinear Finite Impulse Response model* (esse modelo não realiza auto-regressão, depende apenas de variáveis externas. Não confundir com FIR que representa um filtro).

STLF - *Short Term Load Forecasting.*

PCA - *Principal Component Analysis.*

PSCC - *International Power Systems Computation Conference.*

RBF - *Radial Basis Function.*

Resumo

Este trabalho apresenta três modelos NARX utilizando redes neurais artificiais para previsão de carga nodal ativa e reativa, tendo em vista a posterior previsão do estado do sistema através do uso de um algoritmo de fluxo de carga, com a finalidade de poder estudar a segurança de operação do sistema para um perfil de carga futuro.

Os dois primeiros modelos são implementados utilizando-se uma rede de perceptrons de múltiplas camadas estática e o terceiro modelo é implementado utilizando uma rede de perceptrons com filtros FIR em suas sinapses para permitir processamento temporal.

Para cada um dos modelos, diferentes tamanhos de conjunto de treinamento e de topologia da rede neural foram testados, bem como vários tipos de variáveis de entrada para o modelo NARX, a fim de verificar qual a melhor resposta em termos de precisão. Nos testes foram utilizados dados reais de carga de várias subestações.

Neste estudo, verificou-se que uma única rede neural estática usada para tratar todos os dias da semana apresentou o melhor resultado de previsão, sendo que a previsão de potência reativa apresentou um resultado pior do que a de potência ativa.

Observou-se ainda que o erro de previsão do ângulo de tensão nas barras apresenta elevada sensibilidade em relação ao erro de previsão da potência ativa e que o erro na previsão da tensão nodal (tanto magnitude quanto ângulo) apresenta pequena sensibilidade em relação ao erro de previsão da potência reativa.

Abstract

This work presents three artificial neural network NARX models for bus load active and reactive power forecasting. The predicted injections can be used as inputs to a load flow to obtain the forecasted bus states aiming to perform security analysis in a future operating point.

The first two models considered were implemented with a static Multilayer Perceptron Network (MLP) and the third model was implemented using a FIR neural network for temporal processing.

For each one of the above models, different sizes of training sets and neural network topologies were tested, as well as distinct choices of neural networks input variables, in order to check the accuracy of prediction. Measured real data concerning several substations were used to test the models.

As a result, the study shows that one single neural MLP network model for all days of the week give the least active and reactive power forecast error and that the active power prediction is more accurate than the reactive power one.

It was also observed that the state angle forecasting shows a high sensitivity to the active power prediction error and that the voltage (magnitude and phase) prediction error shows a low sensitivity to the reactive power prediction error.

Agradecimentos

- Prof. André Luiz Morelato França, pela orientação e apoio ao longo desse período em que trabalhamos juntos,
- aos professores do Departamento de Sistemas de Energia Elétrica, por seus esforços em manter e disponibilizar as excelentes condições de trabalho para os estudantes que utilizam os laboratórios desse departamento,
- aos amigos de Juiz de Fora, cuja convivência, auxílio e amizade foram fundamentais durante o período de realização deste trabalho,
- a todos os amigos do Laboratório de Sistemas de Energia Elétrica, pela convivência saudável e proveitosa que mantivemos,
- aos amigos Eduardo Nicola Ferraz Zagari e Prof. Fujio Sato, pelos esclarecimentos sobre os dados obtidos junto à CPFL e ao amigo Antônio César Baleeiro Alves pelo auxílio em relação as funções de análise de rede,
- à CPFL, cuja cessão de medições de potência em suas subestações tornou possível a realização deste trabalho e a seus técnicos e engenheiros que se prontificaram em fornecer e comentar estes dados, e
- ao CNPq, pelo apoio financeiro.

Dedico este trabalho:

a Deus, sem o qual nada se realizaria.

*a meus pais, Francisco e Licínia.
a minhas irmãs, Angélica e Fátima.
em memória de meu avô,
Eng. Odilon Pereira de Andrade.*

Conteúdo

1	Introdução	1
2	Aplicação de Redes Neurais em Previsão de Carga	7
2.1	Considerações Gerais	7
2.2	Técnicas Tradicionais Usadas em Previsão de Carga	11
2.2.1	Modelos baseados em Regressão	11
2.2.2	Modelos baseados em Processos Estocásticos	13
2.2.3	Caracterização de modelos ARMA	14
2.2.4	Composição de Modelos ARMA e Regressivos	16
2.2.5	Modelos ARIMA	18
2.3	Comentários Conclusivos	19
2.4	Redes Neurais	19
2.4.1	Breve Histórico	19
2.4.2	Fundamentos	22
2.5	Redes Neurais para Identificação de Séries Temporais	29
2.5.1	Redes NARX para Processamento Temporal	32

2.6	Aplicação de Redes Neurais à Previsão de Carga	38
2.6.1	Particionamento do Conjunto de Treinamento	39
2.6.2	Modelos de Regressão Utilizados	41
2.6.3	Estratégias de Aprendizado	42
2.6.4	Fatores Causais e Topologia da Rede	44
2.6.5	Modelagem de Dias Especiais e Tratamento de Valores Espúrios	45
3	Características da Curva de Carga Nodal	46
3.1	Análise dos Dados de Carga Nodal Ativa e Reativa	46
3.1.1	Tratamento de Erros Grosseiros	46
3.1.2	Caracterização da Carga das Subestações	47
4	Testes e Resultados	69
4.1	Considerações Iniciais	69
4.2	Modelo M	70
4.2.1	Algoritmo de Previsão	72
4.3	Modelo L	77
4.4	Modelo F	81
4.5	Descrição dos Resultados	84
4.5.1	Resultados com Modelo M	86
4.5.2	Comentários sobre os resultados do modelo M	94
4.5.3	Resultados com Modelo L	96
4.5.4	Comentários sobre os resultados do modelo L	99

4.5.5	Resultados com Modelo F	100
4.5.6	Comentários sobre os resultados do modelo F	109
4.5.7	Comparações de Resultados	110
4.5.8	Gráficos dos Resultados	113
5	Previsão do Estado Futuro	148
5.1	Considerações Iniciais	148
5.2	Efeito da Previsão de Carga no Estado	151
5.2.1	Análise do Erro de Previsão de Carga no Estado do Sistema IEEE14 . . .	152
6	Conclusões	160
6.1	Sugestão para Trabalhos Futuros	161
A	Descrição do IEEE14 barras	163

Lista de Tabelas

3.1	Subestação com carga tipicamente residencial	49
3.2	Subestação com grande consumidor público (repartição)	50
3.3	Subestação com carga equitativa entre consumidores	51
3.4	Subestação com carga tipicamente residencial e comercial	52
3.5	Subestação com carga tipicamente industrial	53
3.6	Subestação com carga tipicamente residencial e industrial (ind. leves)	54
3.7	Subestação com carga tipicamente residencial e industrial (ind. pesadas)	55
3.8	Subestação com Grande Consumidor Comercial	56
3.9	Subestação com predominância residencial e rural	57
3.10	Subestação com predominância comercial e industrial	58
3.11	Subestação com carga mista e grandes consumidores comerciais e industriais	59
4.1	Variáveis de entrada utilizadas no modelo M1 e M2	73
4.2	Valores base para potência ativa e reativa.	85
4.3	Tabela comparativa com simulações para o modelo M1. Potência Ativa.	88
4.4	Tabela comparativa com simulações para o modelo M1. Potência Reativa.	89

4.5	Tabela comparativa com simulações para o modelo M2. Potência Ativa.	90
4.6	Tabela comparativa com simulações para o modelo M2. Potência Reativa.	91
4.7	Tabela comparativa com simulações para o modelo M3. Potência Ativa.	92
4.8	Tabela comparativa com simulações para o modelo M3. Potência Reativa.	93
4.9	Tempos de execução em segundos para as diversas variações no modelo M.	94
4.10	Tabela comparativa com simulações para o modelo L3	97
4.11	Tabela comparativa com simulações para o modelo L4	98
4.12	Tempos de execução em segundos para as diversas variações no modelo L.	99
4.13	Tabela comparativa com simulações para o modelo F1. Potência Ativa.	101
4.14	Tabela comparativa com simulações para o modelo F1. Potência Reativa.	102
4.15	Tabela comparativa com simulações para o modelo F2. Potência Ativa.	103
4.16	Tabela comparativa com simulações para o modelo F2. Potência Reativa.	104
4.17	Tabela comparativa com simulações para o modelo F3. Potência Ativa.	105
4.18	Tabela comparativa com simulações para o modelo F3. Potência Reativa.	106
4.19	Tabela comparativa com simulações para o modelo F4. Potência Ativa.	107
4.20	Tabela comparativa com simulações para o modelo F4. Potência Reativa.	108
4.21	Tempos de execução em segundos para as diversas variações no modelo F.	109
4.22	Comparação entre os modelos que apresentaram melhor desempenho. Potência Ativa.	111
4.23	Comparação entre os modelos que apresentaram melhor desempenho. Potência Reativa.	112
5.1	Valor das injeções de potência na barra 2	152

5.2	Estado na barra 2	152
5.3	Relação entre o número da barra e o número da subestação.	153
5.4	Tabela com as potências base para o fluxo nos ramos	153
5.5	Estatísticas sobre o estado do sistema (IEEE14).	155
5.6	Estatísticas do fluxo de potência ativa (IEEE14).	156
5.7	Estatísticas do fluxo de potência reativa (IEEE14).	157

Lista de Figuras

1.1	Sistema de Potência fictício com duas áreas interligadas.	2
1.2	Perfil de carga.	4
2.1	Representação gráfica de um Perceptron.	20
2.2	Exemplo de Rede Neural de Múltiplas Camadas	23
2.3	Rede MLP representando modelo NARX.	30
2.4	Diagrama esquematizado de uma rede FIR $1 : 1 : 1 : 1 : 25 \times 5 \times 5$	37
2.5	Diferentes arquiteturas de redes neurais utilizadas em previsão	40
3.1	Potência ativa e reativa na subestação 40	49
3.2	Funções de autocorrelação (potência ativa e reativa) na subestação 40	49
3.3	Potência ativa e reativa na subestação 80	50
3.4	Funções de autocorrelação (potência ativa e reativa) na subestação 80	50
3.5	Potência ativa e reativa na subestação 221	51
3.6	Funções de autocorrelação (potência ativa e reativa) na subestação 221	51
3.7	Potência ativa e reativa na subestação 232	52
3.8	Funções de autocorrelação (potência ativa e reativa) na subestação 232	52

3.9	Potência ativa e reativa na subestação 40	53
3.10	Funções de autocorrelação (potência ativa e reativa) na subestação 379	53
3.11	Potência ativa e reativa na subestação 463	54
3.12	Funções de autocorrelação (potência ativa e reativa) na subestação 463	54
3.13	Potência ativa e reativa na subestação 572	55
3.14	Funções de autocorrelação (potência ativa e reativa) na subestação 572	55
3.15	Potência ativa e reativa na subestação 623	56
3.16	Funções de autocorrelação (potência ativa e reativa) na subestação 623	56
3.17	Potência ativa e reativa na subestação 668	57
3.18	Funções de autocorrelação (potência ativa e reativa) na subestação 668	57
3.19	Potência ativa e reativa na subestação 762	58
3.20	Funções de autocorrelação (potência ativa e reativa) na subestação 762	58
3.21	Potência ativa e reativa na subestação 840	59
3.22	Funções de autocorrelação (potência ativa e reativa) na subestação 840	59
3.23	Potência ativa global. Dia de Semana	61
3.24	Potência ativa subestação 40. Dia de Semana	62
3.25	Potência ativa subestação 232. Dia de Semana	62
3.26	Potência ativa subestação 379. Dia de Semana	63
3.27	Potência reativa subestação 40. Dia de Semana	63
3.28	Potência reativa subestação 232. Dia de Semana	64
3.29	Potência reativa subestação 379. Dia de Semana	64
3.30	Potência ativa global. Fim de Semana	65

3.31	Potência ativa subestação 40. Fim de Semana	66
3.32	Potência ativa subestação 232. Fim de Semana	66
3.33	Potência ativa subestação 379. Fim de Semana	67
3.34	Potência reativa subestação 40. Fim de Semana	67
3.35	Potência reativa subestação 232. Fim de Semana	68
3.36	Potência reativa subestação 379. Fim de Semana	68
4.1	Rede neural recorrente	74
4.2	Deslocamento da janela de tempo.	75
4.3	Diagrama esquematizado da rede FIR utilizada para previsão de carga	83
4.4	Potência ativa prevista, real e erro. Subestação 40. Modelo M1.	114
4.5	Potência reativa prevista, real e erro. Subestação 40. Modelo M1.	115
4.6	Gráfico da porcentagem de medidas em função do erro absoluto. Barra 40, potência ativa. Modelo M1	116
4.7	Gráfico da porcentagem de medidas em função do erro absoluto. Barra 40, potência reativa. Modelo M1	116
4.8	Potência ativa prevista, real e erro. Subestação 80. Modelo M1.	117
4.9	Potência reativa prevista, real e erro. Subestação 80. Modelo M1.	118
4.10	Gráfico da porcentagem de medidas em função do erro absoluto. Barra 80, potência ativa. Modelo M1	119
4.11	Gráfico da porcentagem de medidas em função do erro absoluto. Barra 80, potência reativa. Modelo M1	119
4.12	Potência ativa prevista, real e erro. Subestação 221. Modelo M1.	120
4.13	Potência reativa prevista, real e erro. Subestação 221. Modelo M1.	121

4.14	Gráfico da porcentagem de medidas em função do erro absoluto. Barra 221, potência ativa. Modelo M1	122
4.15	Gráfico da porcentagem de medidas em função do erro absoluto. Barra 221, potência reativa. Modelo M1	122
4.16	Potência ativa prevista, real e erro. Subestação 232. Modelo M1.	123
4.17	Potência reativa prevista, real e erro. Subestação 232. Modelo M1.	124
4.18	Gráfico da porcentagem de medidas em função do erro absoluto. Barra 232, potência ativa. Modelo M1	125
4.19	Gráfico da porcentagem de medidas em função do erro absoluto. Barra 232, potência reativa. Modelo M1	125
4.20	Potência ativa prevista, real e erro. Subestação 379. Modelo M1.	126
4.21	Potência reativa prevista, real e erro. Subestação 379. Modelo M1.	127
4.22	Gráfico da porcentagem de medidas em função do erro absoluto. Barra 379, potência ativa. Modelo M1	128
4.23	Gráfico da porcentagem de medidas em função do erro absoluto. Barra 379, potência reativa. Modelo M1	128
4.24	Potência ativa prevista, real e erro. Subestação 463. Modelo M1.	129
4.25	Potência reativa prevista, real e erro. Subestação 463. Modelo L4.	130
4.26	Gráfico da porcentagem de medidas em função do erro absoluto. Barra 463, potência ativa. Modelo M1	131
4.27	Gráfico da porcentagem de medidas em função do erro absoluto. Barra 463, potência reativa. Modelo L4	131
4.28	Potência ativa prevista, real e erro. Subestação 572. Modelo M1.	132
4.29	Potência reativa prevista, real e erro. Subestação 572. Modelo M1.	133
4.30	Gráfico da porcentagem de medidas em função do erro absoluto. Barra 572, potência ativa. Modelo M1	134

4.31 Gráfico da porcentagem de medidas em função do erro absoluto. Barra 572, potência reativa. Modelo M1	134
4.32 Potência ativa prevista, real e erro. Subestação 623. Modelo M1.	135
4.33 Potência reativa prevista, real e erro. Subestação 623. Modelo M1.	136
4.34 Gráfico da porcentagem de medidas em função do erro absoluto. Barra 623, potência ativa. Modelo M1	137
4.35 Gráfico da porcentagem de medidas em função do erro absoluto. Barra 623, potência reativa. Modelo M1	137
4.36 Potência ativa prevista, real e erro. Subestação 668. Modelo F2.	138
4.37 Potência reativa prevista, real e erro. Subestação 668. Modelo M1.	139
4.38 Gráfico da porcentagem de medidas em função do erro absoluto. Barra 668, potência ativa. Modelo F2	140
4.39 Gráfico da porcentagem de medidas em função do erro absoluto. Barra 668, potência reativa. Modelo M1	140
4.40 Potência ativa prevista, real e erro. Subestação 762. Modelo M1.	141
4.41 Potência reativa prevista, real e erro. Subestação 762. Modelo M1.	142
4.42 Gráfico da porcentagem de medidas em função do erro absoluto. Barra 762, potência ativa. Modelo M1	143
4.43 Gráfico da porcentagem de medidas em função do erro absoluto. Barra 762, potência reativa. Modelo M1	143
4.44 Potência ativa prevista, real e erro. Subestação 840. Modelo M1.	144
4.45 Potência reativa prevista, real e erro. Subestação 840. Modelo M1.	145
4.46 Gráfico da porcentagem de medidas em função do erro absoluto. Barra 840, potência ativa. Modelo M1	146
4.47 Gráfico da porcentagem de medidas em função do erro absoluto. Barra 840, potência reativa. Modelo M1	146

5.1	Diagrama de blocos para o problema da análise de segurança com previsão de carga.	150
5.2	Sistema para exemplo de previsão de estado.	151
5.3	Erro na potência ativa e reativa e sua relação com o erro no ângulo e na tensão. Subestação 80	158
5.4	Erro na potência ativa e reativa e sua relação com o erro no ângulo e na tensão. Subestação 379	159

Aplicação de Redes Neurais Artificiais na Previsão de Curto Prazo da Carga Nodal de Sistemas de Energia Elétrica.

Francisco Eugenio de Andrade Leite
DSEE-FEE - UNICAMP

2 de junho de 1997

Capítulo 1

Introdução

A presente dissertação tem como objetivo analisar modelos de previsão de carga nodal, bem como sua influência no estado futuro de sistemas de energia elétrica. Embora a motivação inicial deste trabalho seja a aplicação da previsão para a análise de segurança, outras aplicações em sistemas de potência serão brevemente abordadas nesse capítulo.

O objetivo de uma empresa de energia elétrica é fornecer eletricidade aos seus consumidores de uma forma confiável e econômica. Para isso, é necessário um sistema que possa transportar a energia das unidades geradoras até o consumidor final (residencial, comercial, industrial, etc), percorrendo um longo caminho constituído pelas linhas de transmissão, subestações e linhas de distribuição urbanas.

Muito embora os requisitos básicos para transmissão de energia apresentados acima possam ser atendidos por uma única empresa operando sozinha, é uma prática comum que as diversas empresas geradoras e fornecedoras de energia elétrica se interconectem de forma a melhorar a qualidade do fornecimento de energia, bem como torná-lo mais econômico. A figura 1.1 mostra um sistema de energia hipotético com oito nós (cada nó equivale a uma subestação ou barramento), com duas áreas operando conjuntamente, interligadas pelas linhas entre os nós 3 e 6. As áreas podem representar uma empresa com duas regiões de operação ou duas empresas distintas operando de forma interligada.

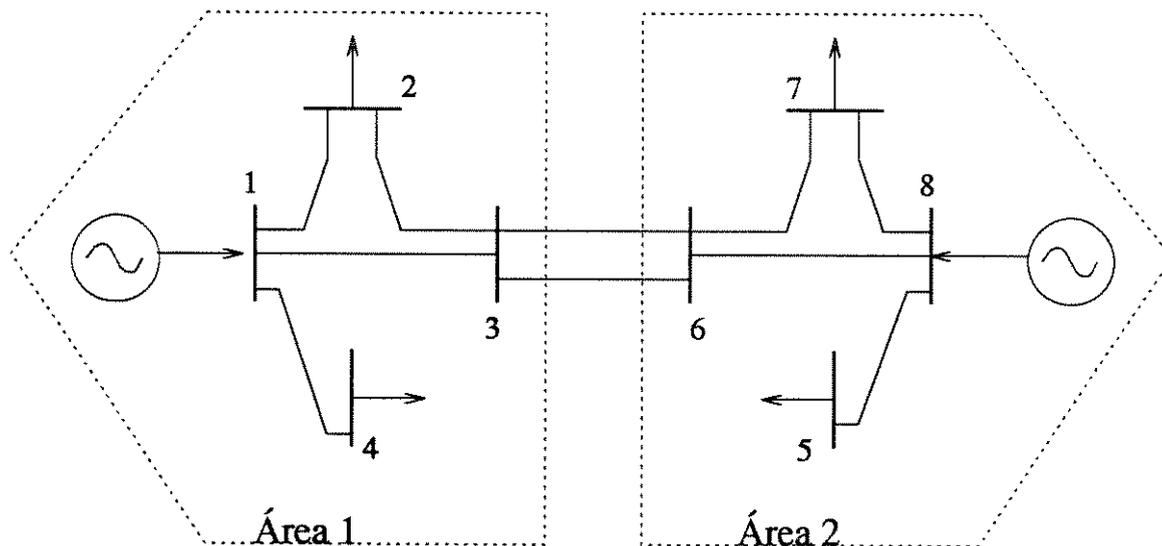


Figura 1.1: Sistema de Potência fictício com duas áreas interligadas.

O objetivo da operação interligada é alcançar uma gama de benefícios econômicos e operacionais, tais como dividir o custo da carga do sistema entre as diversas unidades geradoras, evitar investimentos elevados em novas unidades, minimizar o efeito de uma contingência¹ qualquer no sistema, manter a frequência e a tensão do sistema dentro de limites previamente especificados, entre outros.

De forma a realizar o complexo conjunto de tarefas acima, cada empresa possui um centro de controle e operação responsável por realizar estas tarefas, sendo que normalmente este centro se comunica com os centros de operação de outras empresas ou áreas e com o centro de operação interligada. Esses centros devem realizar um conjunto de funções básicas, a saber [Horton, 1987]:

1. Armazenamento de dados
2. Aplicativos para monitoração da segurança: estimador de estados, análise de contingências, fluxo de carga ótimo, fluxo de carga, análise de estabilidade transitória, previsão de carga e simulações de treinamento.
3. Controle e despacho da geração: controle automático da geração, despacho de carga e intercâmbio/compra de energia entre áreas.
4. Manutenção de dados: serviços de recuperação dos dados para todas as funções.
5. Aquisição e processamento de alarmes.

¹Uma contingência representa a falta parcial ou total de uma unidade geradora, linha de transmissão, transformador ou carga.

6. Interface homem-máquina.

Para uma visão geral de todas as funções acima, bem como um exemplo de integração das mesmas em ambiente computacional, o leitor pode se referir a [Zagari, 1996]. Nesta seção serão tratadas apenas as funções dos itens 2 e 3 as quais se relacionam diretamente com a previsão de carga.

A monitoração da segurança é por assim dizer a parte inteligente do sistema, uma vez que através de suas funções é possível verificar se o sistema se encontra em estado de operação normal (veja [Monticelli, 1983]) com carga e limites de geração/transmissão atendidos ou em situação de emergência (neste caso, ocorreu uma contingência no sistema). Para várias dessas funções é necessário conhecer o comportamento futuro do sistema. Dessa forma, a previsão de carga passa a ser uma necessidade em um centro de controle, se há o desejo de mantê-lo operando de forma econômica e com maior grau de confiabilidade. Os itens seguintes explicam de forma sucinta a relação entre a previsão de carga e as funções dos itens 2 e 3 [Debs, 1988, Wood, 1996].

Despacho de geração

Esta é, economicamente, a mais importante aplicação de previsão de carga. O despacho de geração pode ser definido de seguinte maneira [Wood, 1996]:

Dado o conjunto de N unidades geradoras disponíveis em uma empresa ou agregado delas, qual é o subconjunto S de unidades capaz de atender a carga com o menor custo possível.

O exemplo abaixo procura esclarecer a importância do uso da previsão de carga e sua relação com o despacho de geração. Suponha que o perfil da carga para um determinado intervalo de tempo em um dia da semana seja aquele mostrado na figura 1.2, e que a empresa (ou agregado delas) possua quatro unidades geradoras:

- G1 500 MW
- G2 400 Mw
- G3 200 Mw
- G4 100 Mw

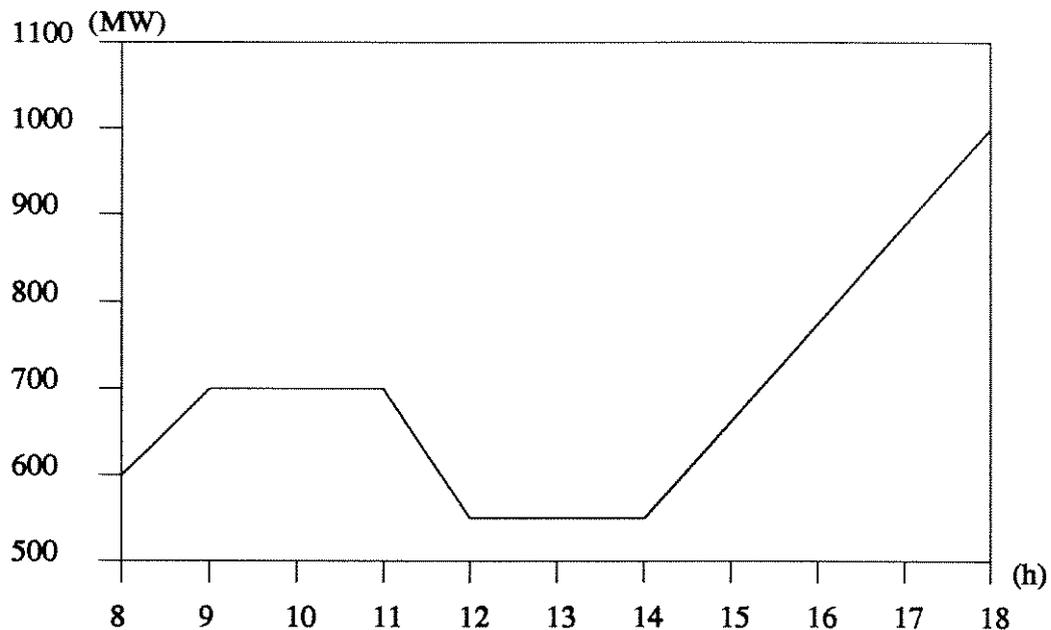


Figura 1.2: Perfil de carga.

Se a empresa mantiver as 4 unidades funcionando constantemente, então a produção total de energia equivalerá a 1200 MW, mas conforme se vê na figura 1.2 a carga total que necessita ser atendida não chega a esse valor. Além do mais, o período do dia em que é necessária maior geração é um intervalo pequeno próximo às 18:00 hs. Claro está que o ideal seria manter funcionando apenas o número de unidades geradoras necessárias para suprir a carga em um determinado horário. Porém, isso só é possível se houver um conhecimento a priori da curva de carga, o que é obtido através da previsão. Deve-se notar que quanto menor o erro de previsão melhor será aproveitada a capacidade de geração.

Muito embora o raciocínio acima tenha sido empregado para a carga total do sistema, o mesmo é válido se for necessário conhecer a carga de determinados nós de sistema com o objetivo de poder comprar, vender ou intercambiar energia entre empresas ou consumidores com o menor prejuízo possível. Se o consumo futuro for conhecido, então menores serão as chances de comprar energia que não será utilizada, ou vender uma quantidade de energia, cujo custo de geração seja muito elevado. Geralmente o intervalo de previsão para esse problema abrange o período de um dia até uma semana.

Análise de segurança

Este problema é de certa maneira uma extensão do problema do despacho de geração e pode ser descrito da seguinte forma:

Dado um conjunto S de unidades geradoras em funcionamento e que as restrições de operação do sistema são atendidas no instante atual de funcionamento, ou

seja, o sistema opera em seu estado normal, deseja-se saber se a evolução atual e futura da carga em cada barramento do sistema será atendida no evento de uma contingência, ou se uma carga futura pode provocar uma contingência².

Embora esse problema tenha uma certa relação com o anterior, diferencia-se deste no aspecto da importância em assegurar a segurança estática e dinâmica do sistema. A segurança estática está relacionada com a verificação do comportamento do estado permanente do sistema pós-contingência e a análise de segurança dinâmica verifica o comportamento da transição entre a pré-contingência e a pós-contingência. Correntemente, apenas a função de análise estática é implementada com sucesso em centros de controle, já a análise de segurança dinâmica ainda não se encontra totalmente operacional [El-Sharkawi, 1996]. O intervalo de previsão para este problema é de 1 a 24 horas.

De um modo geral os modelos de previsão para análise de segurança tendem a ser mais simples do que os modelos para despacho de carga, sendo escolhidos para serem estudados neste trabalho conforme descrito no início deste capítulo.

Estimação de estado [Couto, 1994, Silva, 1993]

A estimação de estados é uma função importante em um centro de operação, pois ela é responsável por verificar a confiabilidade dos dados obtidos pelo sistema de aquisição e corrigi-los quando for necessário ou gerar pseudo-valores no caso de não haver medições em alguma parte do sistema. Embora seja comum não se fazer a previsão de estado, é possível dotar algoritmos de estimação com esta capacidade. Neste caso, podem ser feitos três tipos de previsão: previsão de carga, estado ou medição. Cada uma delas possuindo suas vantagens e aplicações dentro do problema de estimação. Aqui os benefícios alcançados são os seguintes: validação de dados, rapidez de processamento do algoritmo de estimação e melhora na qualidades das estimativas.

Fluxo de carga ótimo e controle automático de geração (CAG) [Debs, 1988]

Para este, caso a literatura pesquisada não cita o uso de previsão. Porém, a mesma é potencialmente útil para as duas funções, uma vez que o conhecimento da tendência do comportamento da carga permite ao CAG acompanhar a taxa de variação da mesma e, portanto, adequar a geração a ela. Em relação ao fluxo de carga ótimo, a previsão de carga para barramento dentro de um intervalo de 10 a 60 minutos é útil para determinar o nível ótimo de geração e assim auxiliar a função do CAG.

Em síntese, procurou-se neste capítulo mostrar um quadro das aplicações de previsão de carga em análise de sistemas de potência. No capítulo 2 são apresentadas as formas pelas quais modelos para previsão de carga são elaborados e a vantagem em se utilizar redes neurais com essa função. No capítulo 3 mostra-se o comportamento da curva de carga nodal ativa e reativa das subestações cujos dados foram utilizados no presente estudo, além de se comparar as diferenças

²Esse evento é difícil de acontecer em um sistema de grande porte, pois ele é planejado para atender aquela carga. Contudo, no caso do sistema estar operando em condição de alerta/emergência esse fator passa a ter maior importância.

entre elas e problemas decorrentes. No capítulo 4 são descritos os modelos estudados neste trabalho e suas potencialidades, bem como resultados estatísticos do uso desses modelos para previsão considerando os dados reais. No capítulo 5 analisa-se a obtenção da previsão de estado (magnitude e fase da tensão) a partir da previsão de carga nodal, utilizando-se um algoritmo de fluxo de carga que é o ponto de partida para se realizar a análise de segurança. Por fim, o capítulo 6 apresenta as conclusões deste trabalho.

Capítulo 2

Aplicação de Redes Neurais em Previsão de Carga

2.1 Considerações Gerais

Como se pode perceber da leitura de artigos que fazem um revisão geral sobre o assunto [El-Sharkawi, 1996, Gross, 1987], a grande maioria dos trabalhos citados na literatura trata da previsão da carga ativa total de um sistema de potência, que é a soma das demandas individuais em todos os seus nós. No caso de carga nodais, há poucas publicações referentes ao problema [Abu-El-Magdad, 1981, Debs, 1988, Falcão, 1986, Handschin, 1988, Keyhani, 1983], sendo que destas, apenas [Falcão, 1986] faz referência à previsão de carga nodal reativa. Assim as considerações feitas a seguir sobre os fatores genéricos que afetam a carga devem ser tomadas com relação à carga total do sistema ou à carga ativa de um determinado nó. Parte dessas considerações também podem ser aplicadas à modelagem da carga reativa, como a influência de fatores passados da carga, fatores temporais e efeitos aleatórios. Contudo, não se encontrou um estudo que mostre uma relação entre a carga reativa e fatores externos, tais como clima ou crescimento econômico, da mesma forma que há para a carga ativa.

A carga total do sistema e a carga nodal ativa e reativa podem ser modeladas para fins de previsão de carga a curto prazo (que neste trabalho será também denominada de STLF, “Short Term Load Forecasting”), levando-se em consideração alguns fatores, classificados segundo as categorias abaixo:

Fatores Econômicos

Esses fatores influenciam bastante a carga do sistema, já que eles determinam a carac-

terística populacional, industrial, agrícola e conseqüentemente do consumo de energia de uma determinada área. Isto determina de forma marcante a característica da curva de carga, conforme será observado no capítulo 3. Todavia, tais fatores apresentam constante de tempo bastante longas e geralmente não são considerados explicitamente na previsão de carga de curto prazo.

Fatores Temporais

Esses fatores são aqueles que mais influenciam as características da carga e podem ser separados em três componentes.

- Efeitos Sazonais: estão relacionados com a estação do ano, pelo fato da demanda de energia no inverno ser geralmente diferente da do verão. Neste caso as mudanças na característica da curva de carga ocorrem em geral gradualmente. Contudo, mudanças relacionadas com o início ou fim de horário de verão, início de período escolar, feriados e atividades eventuais (shows ou jogos de futebol, por exemplo) podem mudar a característica da curva de carga de forma bastante significativa no dia do evento. (É desejável que o modelo de STLF possa lidar com estes tipos de eventos com a menor perda possível de precisão).
- Ciclo Semanal e Diário: está relacionado com o ciclo de trabalho e descanso de uma dada região e definem marcadamente o padrão da carga.
- Feriados: percebe-se que há uma diminuição no consumo de energia nesses dias, sendo que os dias antecedentes e precedentes podem também ter seus padrões de carga afetados (tomando-se como referência a forma do padrão da carga para um dia “normal” da semana).

Alguns exemplos de variáveis que expressam os fatores acima são:

- Hora do dia, para um determinado dia da semana. Geralmente é representada através de números binários.
- Indicador de dia da semana. Pode ser representada através de números binários, podendo, porém, assumir valores decimais ou fracionários.
- Indicador de evento especial ou feriado. Modelos de carga têm incorporado essa característica através de módulos específicos [Bakirtzis, 1996, Papalexopoulos, 1994], em vez de considerá-la como uma variável.
- Para o caso de previsão de carga nodal em áreas industriais pode ser relevante considerar variáveis que levem em conta diferentes tipos de atividades industriais.

Fatores Climáticos

As condições meteorológicas são responsáveis por variações significativas na carga, principalmente quando a empresa fornecedora de energia abrange uma extensa área geográfica ou atende grandes centros, onde o consumo de energia em aparelhos de ar condicionado é grande. Cargas nodais também podem ser influenciadas por fatores climáticos [Chen, 1992], inclusive a carga representada por um único consumidor [Ohlsson, 1994].

Em diversos modelos citados na literatura que utilizam variáveis climáticas, a temperatura é sempre levada em consideração. Porém, no caso de se considerar a previsão de curto prazo de uma carga nodal de consumo preponderantemente residencial ou comercial, uma variável tão importante quanto a temperatura é a indicação de intensidade luminosa e/ou nebulosidade [Abu-El-Magdad, 1981, Ohlsson, 1994], já que esse fator afeta o uso de energia elétrica para iluminação de forma quase instantânea, ao passo que a variação na temperatura demora mais para influenciar o comportamento da carga. Outros fatores climáticos podem, ainda, ser considerados, como por exemplo: velocidade do vento e umidade do ar. Cabe notar que a consideração de variáveis climáticas em um modelo de STLF depende em grande parte do grau de influência das mesmas sobre a carga. Dessa forma, não há como saber a priori quais variáveis devem ser consideradas ou de que forma elas serão consideradas. Para isso, é necessário fazer um estudo de correlação com o comportamento da carga. Abaixo, são dados alguns exemplos de variáveis que podem ser utilizadas:

- temperatura no momento da previsão.
- temperatura máxima e mínima no dia da previsão.
- temperatura algumas horas antes da previsão.
- média da temperatura em dias anteriores.
- indicador de intensidade luminosa.
- indicador de nebulosidade.
- umidade relativa no dia da previsão.
- indicador de velocidade do vento

Mesmo que se tenha um bom modelo para descrever a influência das variáveis climáticas na carga, é grande a dificuldade de incluir tal efeito nos esquemas de previsão de carga, pois isso depende diretamente do nível de precisão das previsões de tempo e de temperatura, que se sabe serem não muito confiáveis. No caso de cargas nodais isto é ainda mais difícil, pois a área abrangida por uma subestação é muito pequena. Por essas razões, a influência da temperatura e de outras variáveis climáticas não será considerada neste trabalho.

Comportamento Passado da Carga

Este fator está relacionado com o estudo estatístico do comportamento da carga através da análise das funções de autocorrelação e autocorrelação parcial, as quais mostram que valores de carga, ocorridos no mesmo instante de tempo, em dias diferentes se correlacionam, conforme se pode observar nos gráficos de autocorrelação apresentados no Capítulo 3. É, em conjunto com a temperatura, uma das variáveis mais utilizadas, senão a mais utilizada. Geralmente eles são utilizados da seguinte forma:

- certo número de valores discretos da carga passada, representados por atrasos no tempo, que possuem maior correlação com a carga presente.
- medidas dos 24 valores de carga correspondentes a um ou dois dias antes do dia de previsão.

Fatores Aleatórios

Neste item estão incluídos os efeitos causados por eventos distintos dos descritos anteriormente. O sistema elétrico está continuamente sujeito a distúrbios aleatórios, provocados por pequenas perturbações ou por variações na carga devido à mudança de operação de alguma grande indústria, e que podem causar variações imprevisíveis no padrão da carga. No caso de se estar prevendo a carga de uma barra com consumo preponderantemente industrial, esse fato traz uma complicação extra na modelagem da carga.

As variáveis apresentadas acima são apenas exemplos da forma como os vários fatores que influenciam a carga podem ser considerados, podendo haver outros tipos de variáveis, que dependem da técnica de previsão a ser adotada, quer seja regressão, modelos estocásticos, redes neurais ou outras utilizadas na previsão de carga¹.

Uma vez apresentados os fatores que influenciam a carga pode-se elaborar o seguinte modelo geral de carga ativa, que é dado pela equação em variáveis discretas abaixo:

$$P_L(k) = B(k) + W(k) + S(k) + v(k) \quad (2.1)$$

onde: k é o tempo no domínio discreto, $P_L(k)$ é a carga ativa total do sistema ou da barra, $W(k)$ é a parte da carga influenciada pela temperatura, $B(k)$ é a carga típica ou carga base, $S(k)$ é a componente da carga devido à ocorrência de eventos especiais e $v(k)$ é a componente devido à variações aleatórias.

A equação (2.1), escrita em sua forma algébrica, é empregada comumente em técnicas de previsão baseadas em regressão, análise de séries temporais e variáveis no espaço de estado. Ao se considerar técnicas baseadas em inteligência artificial, da qual redes neurais é um ramo, a

¹Os exemplos de variáveis apresentados são típicos daqueles modelos utilizados para fazer previsão através de redes neurais, sendo que algumas delas também são utilizadas por outras técnicas.

equação (2.1) deve ser entendida em um sentido mais qualitativo, representando a decomposição do modelo da carga em módulos responsáveis pela identificação de certas dependências ou características específicas, tais como a relação entre temperatura e carga ou previsão de cargas para feriados. Porém, isso só tem sentido se a rede tiver uma estrutura modular ou for uma rede neural fuzzy. No caso, da rede neural ser constituída por um modelo único para o comportamento da carga, a equação (2.1) perde o sentido para essa modelagem.

Para dar uma visão geral da evolução do tema, serão introduzidas no restante deste capítulo algumas técnicas tradicionais de previsão de carga, bem como os modelos baseados em redes neurais.

2.2 Técnicas Tradicionais Usadas em Previsão de Carga

2.2.1 Modelos baseados em Regressão

Esses são os modelos mais simples existentes e dependem em grande parte de intervenção humana, muito embora haja algoritmos robustos e automatizados, ideais para operação “on-line”.

O modelo mais simples consiste em, dada uma base de dados com medições de carga, obter um conjunto de valores de carga típicos que melhor represente o comportamento de carga que foi mais observado, ou cuja probabilidade de ocorrência é maior, e utilizar este valor para fazer a previsão. Os valores da curva de carga típica, obtidos da forma descrita acima, podem ser utilizados para se fazer a previsão da carga diária, semanal e, até mesmo, mensal. É claro que esse modelo apresentará erros de previsão elevados em situações em que há desvio no comportamento normal da carga, devido, por exemplo, a uma elevação de temperatura.

Uma evolução desse modelo consiste em extrair não apenas uma curva de carga típica, mas sim várias curvas de carga típicas, dependendo da condição de operação do sistema. Dessa maneira, o comportamento da curva de carga passa a ser descrito pela equação (2.2):

$$z(t) = \sum_{i=1}^N \alpha_i f_i(t) + v(t) \quad (2.2)$$

onde α_i são os coeficientes que ponderam os valores das curvas típicas de carga dados pelas funções f_i e $v(t)$ é um ruído branco.

Nesse sentido, uma técnica para previsão de carga nodal ativa pode ser elaborada, tomando-se as curvas de carga típicas de cada consumidor conectado a um determinado barramento e

fazendo com que a carga deste nó seja a soma ponderada da carga dos consumidores a ele conectados [Handschin, 1988]. Assim, se a curva de carga típica de diversos tipos de consumidores (industriais, públicos, residenciais, etc) for conhecida, pode-se levantar a carga dos barramentos de toda uma área e não apenas de um único nó, já que os barramentos de uma área possuem uma determinada combinação de consumidores. Esse fato é demonstrado pela equação (2.3):

$$\underline{z}(t) = H \cdot \underline{x} + \underline{v} \quad (2.3)$$

onde \underline{x} representa a curva de carga típica dos diversos tipos de consumidores, mais uma possível correção para efeitos ocasionados por alguma variação climática, h_{ij} (elemento da matriz H) é o peso que a carga do consumidor j representa no barramento i e \underline{z} é o vetor de carga da área e \underline{v} é um ruído branco.

Esse é um método bastante interessante para a previsão de carga ativa nodal, embora aproximado e envolvendo grande quantidade de dados. Entretanto, estudos precisariam ainda ser feitos para verificar a possibilidade de aplicar esse método na previsão de potência reativa.

O modelo descrito pela equação (2.2) pode tomar uma forma mais complexa [Christiaanse, 1971, Papalexopoulos, 1990, Strasser, 1990], onde f_i não é mais uma curva de carga típica, mas sim elementos funcionais caracterizados por funções senoidais, polinomiais, ou lineares, as quais representam os efeitos temporais da carga (periodicidade, por exemplo) ou efeitos climáticos. Os valores α_i são os coeficientes de regressão, calculados pelo método dos mínimos quadrados ou mínimos quadrados ponderados, com o objetivo de minimizar a função:

$$J = \sum_{j=1}^M (\hat{z}_j - z_j)^2 \quad (2.4)$$

onde M é o número de amostras da série, \hat{z}_j é o valor estimado e z_j o valor medido.

A equação (2.4) tem a seguinte versão ponderada ($0 < \beta < 1$):

$$J = \sum_{j=1}^M \beta^j (z_j - \hat{z}_j)^2 \quad (2.5)$$

A ponderação é introduzida com a finalidade de dar maior importância aos valores mais novos de medição com relação aos mais antigos.

Cabe notar que os coeficientes α_i podem ser recalculados em intervalos regulares, como por

exemplo diariamente [Christiaanse, 1971], para levar em consideração efeitos recentes da carga, ou até mesmo em intervalos de um ano [Papalexopoulos, 1990]. Eles também podem ser recalculados adaptativamente [Baumann, 1993, Strasser, 1990] através de um filtro de Kalman, a fim de corrigir a previsão caso os valores do erro se tornem elevados, fazendo com que o algoritmo tenha características dinâmicas, permitindo acompanhar a evolução da carga de maneira bastante precisa, inclusive na ocorrência de variações bruscas. Um problema que pode ocorrer ao se usar filtros de Kalman é a possibilidade de haver instabilidade numérica no algoritmo [Fan, 1994].

De um modo geral, os métodos regressivos apresentam resultados bastante satisfatórios em condições normais de operação, sendo que eles têm a vantagem de serem estruturalmente simples e os parâmetros do modelo poderem ser atualizados sem grande esforço. Contudo, esses modelos apresentam o problema [Papalexopoulos, 1994] de não captar convenientemente as relações não lineares e estocásticas existentes entre as variáveis regressoras e a carga, além de não serem flexíveis em situações de rápida variação de carga, muito embora esta dificuldade possa ser em parte resolvida utilizando um filtro de Kalman se a estrutura do modelo assim o permitir.

2.2.2 Modelos baseados em Processos Estocásticos

Esses modelos consideram que a variação de carga é um processo estocástico, relacionado com o comportamento recente da carga, clima e efeitos aleatórios. Eles podem ser de dois tipos básicos: modelos ARMA (“Autoregressive Moving Average”) e modelos de variável de estado [Gross, 1987]. Dentro do modelo ARMA existem ainda dois sub-modelos: aqueles que modelam a carga segundo uma componente determinística, semelhante ao apresentado pela equação (2.2), mais uma componente devido à presença de efeitos aleatórios da carga e os modelos ARMA puros. Em ambos os casos o efeito de variáveis climáticas, basicamente temperatura, são considerados na modelagem.

Da mesma forma que na seção anterior, esse tipo de modelagem será descrita neste trabalho, de maneira mais geral, não só por motivos históricos, mas porque esse é um dos modelos mais empregados em previsão de carga e pelo fato da previsão de carga por redes neurais se utilizar das análises feitas na elaboração dos modelos acima citados. O leitor interessado em conhecer melhor esses métodos deve se referir ao excelente livro clássico escrito por Box e Jenkins [Box, 1976]. Outro texto mais simples e que apresenta a modelagem por processos estocásticos de forma bastante prática é apresentado por Nazem [Nazem, 1988].

Antes de se comentar a forma pela qual modelos tipo ARMA são utilizados na previsão de carga, é necessário caracterizar a forma pela qual eles são representados, o que será feito logo a seguir.

2.2.3 Caracterização de modelos ARMA

Pode-se demonstrar, através do teorema da decomposição de Wold [Nazem, 1988], que um processo estocástico estacionário é formado por uma soma de componentes auto-determinísticas e uma componente de médias móveis (média ponderada de ruídos brancos) de ordem infinita. Assim, observa-se que modelos ARMA são uma representação linear de um processo estocástico estacionário. A conceituação de um processo estacionário, por si só, já é bastante complexa. Porém, uma condição de processo estacionário conhecida por estacionaridade por covariância² é mais simples de ser verificada [Gupta, 1971]. Nessa conceituação, a qual será utilizada no restante deste trabalho, um processo estocástico é dito ser estacionário se as seguintes condições forem observadas:

1. a média:

$$\bar{y}_i = \frac{1}{T} \sum_{k=i+1}^{i+T} y(k) \quad (2.6)$$

2. e a variância:

$$\sigma_i = \frac{1}{T} \sum_{k=i+1}^{i+T} (y(k) - \bar{y}(k))^2 \quad (2.7)$$

para $i = 1, 2, \dots$ são independentes de i , ou seja, a média e a variância são constantes. T é o período da série e $y(k)$ são os valores observados da série temporal.

O modelo ARMA é na verdade uma composição de dois modelos básicos para séries temporais bastante importantes: modelo AR (“Autoregressive”) e modelo MA (“Moving Average”) [Box, 1976].

- Em um modelo AR(p), o valor presente do processo, $\tilde{y}(k)$, é dado por um arranjo linear de p valores passados de $\tilde{y}(k)$ mais um ruído branco Gaussiano $a(k)$, sendo p a ordem do modelo:

$$\tilde{y}(k) = \phi_1 \tilde{y}(k-1) + \phi_2 \tilde{y}(k-2) + \dots + \phi_p \tilde{y}(k-p) + a(k) \quad (2.8)$$

onde $\tilde{y}(k) = y(k) - \mu$ e μ a média em torno da qual o processo varia. Observe que \tilde{y} é obtido através de uma regressão usando valores prévios da mesma variável, daí o nome “autoregressivo”.

Introduzindo-se o operador deslocamento $By(k) = y(k-1)$ tem-se:

$$\phi(B) = 1 - \phi_1 B - \dots - \phi_p B^p$$

²Também chamado de processo fracamente estacionário [Keyhani, 1983]

e a equação (2.8) fica:

$$\phi(B)\tilde{y}(k) = a(k) \quad (2.9)$$

- Em um modelo MA(q), de ordem q, o valor presente \tilde{y} é definido por uma combinação linear de q valores passados da variável aleatória $a(k)$:

$$\tilde{y}(k) = a(k) - \theta_1 a(k-1) - \dots - \theta_q a(k-q) \quad (2.10)$$

Pode-se então escrever que:

$$y(k) = \theta(B)a(k) \quad (2.11)$$

onde $\theta(B) = 1 - \theta_1 B - \dots - \theta_q B^q$ é definido como operador MA.

Combinando-se os modelos AR e MA a equação do modelo ARMA assume então a forma:

$$\phi(B)y(k) = \theta(B)a_k \quad (2.12)$$

Em geral, séries temporais reais não apresentam estacionaridade, podendo a não estacionaridade se manifestar na média, na variância, ou em ambas. Pode ainda acontecer de a série apresentar um comportamento periódico como é o caso da carga de sistemas de energia. De qualquer maneira, uma série não estacionária não pode ser modelada diretamente por um modelo ARMA, sendo necessário fazer transformações na mesma de modo a torná-la estacionária. Isso é feito diferenciando-se a série original $y(k)$, obtendo-se a série estacionária $z_a(k)$ da seguinte maneira:

$$z_a(k) = y(k) - y(k-1) = (1-B)y(k) = \nabla y(k)$$

Havendo a necessidade de diferenciar a série mais de uma vez, tem-se a equação abaixo:

$$z_{a^d}(k) = \nabla^d y(k)$$

onde d é a ordem da diferenciação. A equação acima se refere a um modelo não sazonal. Para o caso de uma série sazonal, como a curva de carga, onde é possível verificar a repetição de um determinado padrão ao longo do tempo, torna-se necessário fazer uma diferenciação sazonal, conforme mostram as equações abaixo [Debs, 1988]:

$$z_b(k) = y(k) - y(k-24) = (1-B^{24})y(k) = \nabla_{24}y(k)$$

$$z_c(k) = y(k) - y(k - 168) = (1 - B^{168})y(k) = \nabla_{168}y(k)$$

A diferenciação feita para obter a série z_a elimina a não estacionaridade na média. A diferenciação das séries z_b e z_c é feita para eliminar periodicidades diárias ou semanais, no caso de se tratar de medições de carga elétrica horária. Uma composição das diferenciações acima também é possível, como por exemplo:

$$z(k) = (1 - B)(1 - B^{24})y(k) = \nabla\nabla_{24}y(k)$$

O modelo ARMA diferenciado recebe o nome especial ARIMA (“Autoregressive Integrated Moving Average”), que pode ser equacionado da seguinte maneira:

$$\phi(B)\nabla^d\nabla_S^D z(k) = \theta(B)a(k) \quad (2.13)$$

onde S corresponde à sazonalidade da série e D ao grau de diferenciação sazonal.

É possível transformar o modelo ARIMA em um modelo equivalente de variáveis de estado [Campo, 1987, Debs, 1988]. Contudo, é necessária uma modelagem bastante robusta, a fim de que o modelo no espaço de estados não contenha um excessivo número de variáveis. Cabe notar que os modelos que incorporam em sua representação efeitos do clima apresentam resultados mais satisfatórios [Campo, 1987, Hagan, 1987].

2.2.4 Composição de Modelos ARMA e Regressivos

Este tipo de modelo é a composição de uma série determinística ou periódica mais uma série estocástica na forma mostrada pela equação (2.14). No caso desse modelo, a periodicidade da carga elétrica é incorporada na componente determinística não sendo assim necessário utilizar um modelo ARIMA.

$$y(k) = y_p(k) + y_s(k) \quad (2.14)$$

na equação, $y_p(k)$ é a componente periódica representada por uma forma semelhante à da equação (2.2). O processo $y_s(k)$ é um processo estocástico não-estacionário. A não estacionaridade de $y_s(k)$ está relacionada ao efeito da variação da carga ocasionado pela influência de alguma variável climática, tipicamente a temperatura, nesse tipo de modelo. De uma forma geral, a modelagem de $y_s(k)$ é feita decompondo-se o processo em uma série estacionária, representada por um modelo ARMA, mais uma série não estacionária, geralmente uma função polinomial ou

linear da temperatura. Modelos que não representam a temperatura não satisfazem a condição de estacionaridade, necessitando de uma constante readaptação dos parâmetros. Geralmente apresentam erros de previsão elevados em condições de variações climáticas bruscas.

De modo geral, $y_s(k)$ pode ser escrita na forma:

$$y_s(k) = \sum_{i=1}^n a_i y_s(k-i) + \sum_{l=1}^{n_u} \sum_{j_l=0}^{m_l} b_{j_l}^l u_l(k-j_l) + \sum_{h=1}^H c_h w(k-h) \quad (2.15)$$

onde n é a ordem do modelo AR, n_u é o número de variáveis climáticas do modelo, $u(k)$ é uma função não-linear que relaciona a carga com a variável climática e H é a ordem do modelo MA.

Portanto, com base nas equações (2.2) e (2.15), a equação para $y(k)$ assume a forma matricial dada por:

$$y(k) = \alpha^T \underline{f}(k) + \mathbf{a}^T \underline{y}_s + \mathbf{b}^T \underline{u} + \mathbf{c}^T \underline{w} \quad (2.16)$$

onde:

$$\alpha^T = [\alpha_1, \dots, \alpha_N]$$

$$\mathbf{a}^T = [a_1, \dots, a_n]$$

$$\mathbf{b}^T = [b_0^1, \dots, b_{m_{n_u}}^{n_u}]$$

$$\mathbf{c}^T = [c_1, \dots, c_H]$$

$$\underline{f}^T = [f_1, \dots, f_N]$$

$$\underline{y}_s^T = [y_s(k-1), \dots, y_s(k-n)]$$

$$\underline{u}^T = [u_1(k-1), \dots, u_{n_u}(k-m_{n_u})]$$

$$\underline{w}^T = [w_1, \dots, w_H]$$

Os parâmetros usados pela equação (2.16) podem ser calculados utilizando-se um estimador de máxima verossimilhança ou um estimador recursivo de mínimos quadrados ponderados [Fan, 1994, Galiana, 1974], entre outros. A frequência com que esses parâmetros são recalculados depende das condições de variação da carga. Se houver mudanças bruscas no comportamento os parâmetros podem ser recalculados a cada nova medida obtida pelo sistema de aquisição de dados. Caso contrário, os parâmetros são reestimados a cada dia ou semana.

Utilizando esta concepção de separar a componente determinística da componente estocástica da carga, [Keyhani, 1983] apresenta um modelo "on-line" para previsão de carga horária nodal ativa em que a carga total de uma área servida por uma empresa de energia elétrica é prevista e, então, através do uso de fatores de distribuição (calculados a cada hora, utilizando-se uma

equação recursiva) são obtidas medidas de potência ativa nodal. No caso desse modelo, a componente determinística não é uma composição de funções periódicas, mas sim uma curva de carga típica gerada através de um modelo ARMA.

2.2.5 Modelos ARIMA

Esses modelos são elaborados considerando que a série possui uma não estacionaridade, podendo a série apresentar também uma sazonalidade. Portanto, possuem uma modelagem do tipo apresentado pela equação 2.13, não possuindo componentes determinísticas na forma considerada nos modelos de regressão. Sabe-se, porém, que é possível estabelecer uma equivalência entre modelos ARIMA e modelos baseados em regressão sob certas condições [Abraham, 1983].

O efeito da temperatura pode ser incluído através de uma função de transferência linear ou não linear, sendo que neste caso a precisão do modelo tende a melhorar [Hagan, 1987]. Um exemplo de um modelo ARIMA que leva em conta a não estacionaridade na média e a periodicidade é mostrada na equação 2.17, sem levar em conta o efeito da temperatura.

$$\phi_p(B)\Phi_P(B^{24})\nabla^d\nabla_{24}^D z(k) = \theta_q\Theta_Q(B^{24})a(k) \quad (2.17)$$

onde $\Phi_P(B^{24})$ representa um modelo AR sazonal de ordem P e $\Theta_Q(B^{24})$ um modelo MA sazonal de ordem Q. Esses modelos correspondem aos modelos apresentados nas equações 2.8 e 2.10, respectivamente, exceto que B deve ser substituído por B^{24} [Nazem, 1988].

Utilizando-se de um modelo semelhante ao apresentado pela equação 2.17, [Falcão, 1986] e [Abu-El-Magdad, 1981] desenvolveram modelos para previsão de carga nodal ativa, com os parâmetros do modelo calculados segundo as equações de Yule-Walker, sendo que Falcão também desenvolveu um modelo para previsão de carga reativa, através de uma modelagem indireta utilizando o fator de potência por meio do conceito de espaço de estados. Nesses casos um modelo ARI multivariável foi utilizado, ou seja, ao invés de se ter um modelo para cada nó (um modelo de variável única) elabora-se um modelo que englobe todos os nós. Não se considerou a representação de médias móveis pelo fato da estimação dos parâmetros para essa componente ser não-linear e exigir algum esforço computacional. Cabe ressaltar que os dados usados por [Falcão, 1986] são resultantes de simulação e não de medidas reais, o que, de certa forma, não representa todas as condições de carga existentes no sistema, fazendo com que o modelo mostre resultados bastante favoráveis.

Por outro lado, [Abu-El-Magdad, 1981] testou seu modelo com valores reais de carga. Ele também desenvolveu um modelo baseado em espaço de estados, que tem como vantagem o fato de ter os seus parâmetros atualizados “on-line” e de não necessitar muita intervenção humana para sua elaboração. Uma desvantagem, porém, é a possibilidade de instabilidade numérica. Nenhum dos dois autores considerou o uso de variáveis climáticas.

2.3 Comentários Conclusivos

Muito embora modelos baseados em regressão e modelos ARMA incorporem tanto características sazonais da carga [Galiana, 1974, Hagan, 1987] quanto variáveis climáticas (estas variáveis podem não estar implícitas no modelo) e efeitos aleatórios, aparentemente eles não conseguem captar mudanças repentinas na carga, sendo também difícil incorporar a representação do dia da semana e de feriados, tornando-os pouco flexíveis [Papalexopoulos, 1990]. Outra desvantagem é o excesso de intervenção humana necessária na fase de identificação, podendo esse processo ser, contudo, parcialmente automatizado [Vemuri, 1981]. O uso de variáveis de estado também torna o processo de identificação mais simples, podendo seus parâmetros serem atualizados “on-line”, o que permite o acompanhamento de variações bruscas na carga [Baumann, 1993]. Neste caso pode haver instabilidade numérica, embora esse fato possa ser contornado [Fan, 1994]. Uma vantagem de modelos ARMA reside no fato de se poder medir analiticamente a adequação do modelo.

Deve-se ainda ressaltar que nada impede que métodos mais robustos e que façam uma regressão não-linear sejam preferíveis aos métodos lineares, apesar da modelagem não linear da temperatura em um modelo ARIMA melhorar seu erro de previsão e da facilidade de uso de modelos baseados em espaço de estados.

2.4 Redes Neurais

2.4.1 Breve Histórico

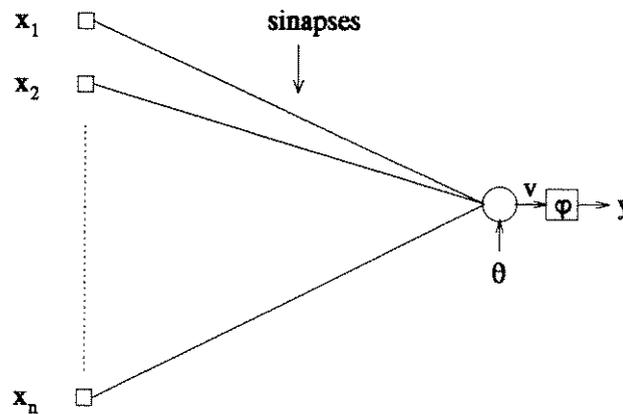
A seguir apresenta-se um breve histórico sobre a evolução da teoria de redes neurais até o desenvolvimento do algoritmo de Retropropagação (“Backpropagation Algorithm”) para o treinamento de redes neurais de múltiplas camadas. O objetivo deste histórico é apenas apresentar o motivo pelo qual se originou o estudo de redes neurais e sua importância, não correspondendo a um estudo minucioso da evolução dessa teoria. O leitor interessado no assunto encontrará uma boa referência em [Haykin, 1994]. Um texto mais simples, porém elucidativo, se encontra em [El-Sharkawi, 1996] e [Freeman, 1991].

Pode-se dizer que o interesse por redes neurais é bastante remoto e vem desde a época em que alguns filósofos como John Locke, George Berkeley e David Hume especulavam sobre o relacionamento entre espírito, memória e comportamento humano. Porém, foi com o desenvolvimento das Neurociências que o interesse pelo estudo de redes neurais e seus paradigmas computacionais se desenvolveu e chegou ao estágio avançado dos dias atuais. O primeiro passo significativo nesse sentido foi feito por McCulloch e Pitts em 1943, quando procuravam reproduzir o comportamento fisiológico do neurônio através de um modelo computacional simples [Freeman, 1991, Haykin, 1994]. Posteriormente, Rosenblatt em 1959 introduziu o Perceptron

(Fig. 2.1), baseando-se no trabalho de McCulloch, e com ele um modelo computacional capaz de classificar padrões linearmente separáveis, através de um algoritmo de correção do erro (“error-correction algorithm”) que permite o ajuste dos parâmetros do Perceptron utilizando o teorema da convergência. O Perceptron é capaz de responder apenas positiva ou negativamente a um determinado conjunto de sinais a ele aplicados. A cada sinal corresponde um valor, o qual é ponderado por um peso (parâmetro), representando a conexão sináptica entre neurônios. Dessa forma, pode-se formular um conjunto de equações que representam o comportamento do Perceptron, que são mostradas abaixo:

$$v = \sum_{i=1}^p w_i x_i - \theta \quad (2.18)$$

$$\varphi(v) = \begin{cases} 1 & \text{se } v > 0 \\ -1 & \text{se } v \leq 0 \end{cases} \quad (2.19)$$



x_1 , x_2 e x_n são os estímulos ou sinais de entrada do Perceptron

φ é a função de ativação

y é a saída do perceptron e θ é o valor de polarização

Figura 2.1: Representação gráfica de um Perceptron.

Paralelamente a esses resultados, Kolmogorov em 1957 provou um teorema importante para a teoria de funções e que bem mais tarde seria utilizado de uma forma modificada para provar a capacidade de redes neurais em representar funções. Contudo, não se tem notícia de que o

teorema de Kolmogorov e a teoria de rede neurais apresentassem alguma relação na época em que o mesmo foi elaborado.

Foi somente com o trabalho de Widrow-Hoff em 1960 [Haykin, 1994, Widrow, 1990] e sua regra delta que se elaborou um algoritmo para ajuste dos parâmetros do Perceptron, o qual resumidamente consiste dos seguintes passos:

- Inicialização: Faça

$$w_k(1) = 0 \quad (2.20)$$

- Filtragem:

para $n = 1, 2, 3, \dots$, calcule:

$$y(n) = \sum_{j=1}^p w_j(n)x_j(n) \quad (2.21)$$

$$e(n) = d(n) - y(n) \quad (2.22)$$

$$w_k(n+1) = w_k(n) + \eta e(n)x_k(n) \quad (2.23)$$

para $k = 1, 2, \dots, p$

onde, nas equações acima p é o número de padrões, n é o tempo discreto, η uma constante que controla a taxa de ajuste e $x(n)$ as instâncias que os padrões assumem ao longo do tempo. A equação (2.23) pode ser compreendida também como a evolução temporal da variável $x(n)$ e p a ordem de um preditor AR. Dessa forma fica claro que o algoritmo acima é uma realização dos modelos AR vistos na seção anterior, sendo que a mesma idéia pode ser estendida para um modelo ARMA.

Widrow utilizou o algoritmo acima para elaborar o ADALINE (elemento linear adaptativo), que consistia em uma máquina para classificação de padrões lineares. Na realidade, o ADALINE é conceitualmente semelhante ao Perceptron, exceto que os parâmetros do ADALINE são calculados de forma a minimizar o erro de saída do elemento linear no sentido dos mínimos quadrados, enquanto o objetivo do perceptron e seu algoritmo de convergência é o de adaptar os parâmetros até que os padrões apresentados ao neurônio estejam separados. Mais tarde, Widrow estendeu o seu trabalho para uma rede de ADALINES chamada MADALINE. Esses estudos lançaram as bases para o uso da Rede de Perceptrons, também conhecida por Rede de Múltiplas Camadas. Outros trabalhos seguiram-se aos de Widrow, porém um grande problema ainda persistia, - como adaptar os parâmetros dos neurônios da rede, pois até então não existia uma regra geral para

isso. Foi graças a Rumelhart [Rumelhart, 1986] e seu algoritmo de Retropropagação (Backpropagation), que havia sido também desenvolvido independentemente por Werbos em 1974, que se tornou possível calcular os parâmetros de uma rede de neurônios (Perceptrons) abrindo caminho para que as redes neurais ganhassem grande popularidade. Esse achado, em conjunto com uma variação do teorema de Kolmogorov, permitiu aplicar redes neurais em uma série de problemas em diversos campos da ciência e em particular ao problema da previsão de séries temporais. Cabe notar que, enquanto as redes de Perceptrons eram estudadas, outros tipos de redes neurais foram também desenvolvidos como as redes de Kohonen e de Hopfield.

2.4.2 Fundamentos

Nesta seção são apresentados os conceitos básicos sobre Redes Neurais de Múltiplas Camadas (MLP), bem como algoritmos de treinamento (adaptação) dos coeficientes (pesos) da rede. Como este trabalho trata apenas de redes MLP, as características de outros tipos de redes serão comentadas apenas quando for necessário.

A estrutura de uma rede genérica MLP de n camadas pode ser representada matematicamente na forma mostrada pela equação (2.24).

$$\underline{y} = F[W^n F[W^{n-1} \dots F[W^1 \underline{x} + \underline{\theta}^1] + \dots + \underline{\theta}^{n-1}] + \underline{\theta}^n] \quad (2.24)$$

onde W^i é a matriz de pesos da i -ésima camada e $\underline{\theta}^i$ é o seu vetor de valores de polarização. Cada nó (neurônio) possui um valor de polarização próprio. F é um operador não-linear $F(\underline{x}) = [\varphi^1, \varphi^2, \dots, \varphi^n]$ e φ^i são as funções de ativação para $i = 1, 2, \dots, n$.

A Figura 2.2, mostra uma rede neural de três camadas (duas camadas intermediárias e uma camada de saída). A primeira camada geralmente não é contada, pois cada nó representa apenas os valores dos elementos do vetor de entrada. Notar que os nós correspondentes aos valores de polarização foram omitidos. Os nós da rede são estruturas computacionais semelhantes ao Perceptron mostrado na Figura 2.1, exceto que a função de ativação φ é representada por uma função suave semi-linear do tipo sigmóide (A função tangente hiperbólica, usada neste trabalho, é um exemplo deste tipo de função)

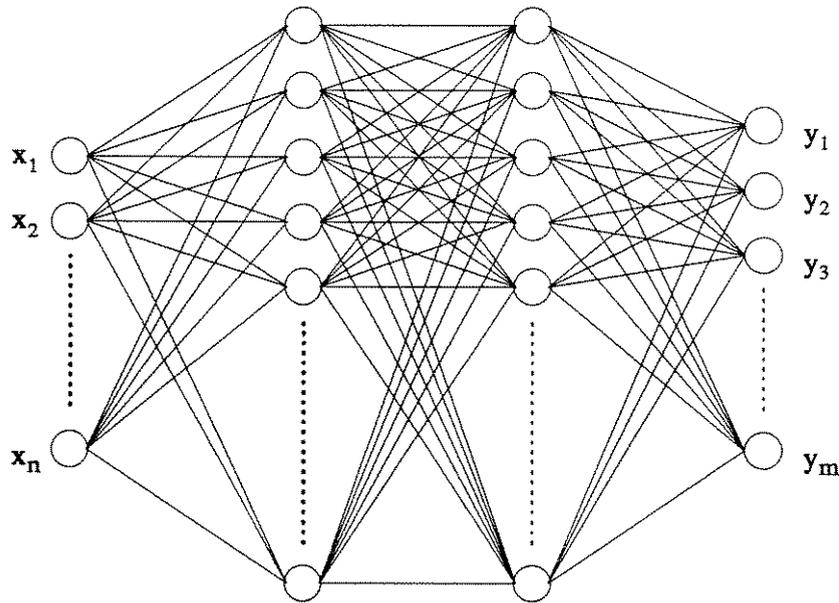


Figura 2.2: Exemplo de Rede Neural de Múltiplas Camadas

Observando-se a equação (2.24), percebe-se que esta é a expressão genérica de um regressor não-linear. Dessa forma, nota-se que uma rede MLP está relacionada à mesma classe de problemas de regressão não-linear, ou seja, interpolação e aproximação de funções, que pode ser enunciada de forma simplificada da seguinte maneira³:

Seja \underline{x} um vetor de variáveis independentes, \underline{y} um vetor de variáveis dependentes, ϵ um vetor de ruídos no processo e f uma função desconhecida tal que:

$$\underline{y} = f(\underline{x}) + \epsilon \quad (2.25)$$

Obtenha, então, a função \hat{f} e um conjunto de parâmetros ϕ de forma que $\hat{\underline{y}}$ gerado pela equação 2.26.

$$\hat{\underline{y}} = \hat{f}(\phi, \underline{x}) \quad (2.26)$$

esteja o mais próximo possível de \underline{y} , segundo algum critério de minimização, ou seja:

$$\| \hat{\underline{y}} - \underline{y} \| \leq \epsilon \quad (2.27)$$

ou escrevendo na forma de uma função custo (onde \mathbf{X} representa todo o conjunto de amostras) :

³Para um melhor detalhamento desse problema veja por exemplo [Von Zuben, 1996]

$$J(\phi) = \sum_{\mathbf{x} \in \mathbf{X}} \| f(\mathbf{x}) - \hat{f}(\phi, \mathbf{x}) \| \quad (2.28)$$

A capacidade de redes neurais em resolver o problema acima foi demonstrada por inúmeros autores [Haykin, 1994] e é resultado direto do Teorema Universal da Aproximação de Funções. Esse fato é importante porque a previsão de séries temporais, e por conseguinte da carga de sistemas elétricos, é uma extensão direta dele, conforme será visto mais adiante.

Muito embora o problema formulado pelas equações (2.25), (2.26) e (2.27) possa ser resolvido por qualquer regressor não-linear descrito na forma da equação (2.24), ou seja, utilizando um conjunto de funções não-lineares como polinômios, ou funções trigonométricas, estudos têm mostrado [Narendra, 1996] que o uso de redes neurais artificiais apresenta numerosas vantagens práticas em relação ao uso de outros regressores. Algumas das vantagens são: a necessidade de menos parâmetros, menor sensibilidade a ruídos e maior tolerância a erros. Por esse motivo redes MLP apresentam um grande potencial de aplicação prática.

Para se utilizar uma rede neural como aproximador de funções ou preditor, é necessário um algoritmo que possibilite o ajuste dos coeficientes da rede. Esse processo também é chamado de treinamento da rede neural, que consiste na minimização de uma função custo como a da equação (2.28), utilizando a norma euclídeana. O treinamento pode ser feito utilizando-se o algoritmo de Retropropagação comentado na seção anterior, que será apresentado a seguir em uma forma resumida, sem a demonstração da Regra da Cadeia Generalizada, que pode ser encontrada em [El-Sharkawi, 1996], [Haykin, 1994] ou [Rumelhart, 1986] por exemplo. Deve ficar claro que esse algoritmo permite apenas calcular os coeficientes ϕ da equação (2.26), dado que as funções não lineares \hat{f} estão previamente fixadas, ou seja a rede neural representa um modelo paramétrico. Contudo é possível [Von Zuben, 1996] desenvolver algoritmos em que \hat{f} não está fixada, utilizando métodos construtivos para redes neurais não paramétricas (\hat{f} e ϕ não são fixos). O uso desses métodos permite obter modelos de aproximação potencialmente mais eficientes do que os modelos paramétricos, mas não são considerados neste trabalho.

I Inicialização.

Uma configuração inicial deve ser escolhida e os pesos devem assumir pequenos valores gerados aleatoriamente.

II Apresentação dos pares de treinamento.

Um conjunto de pares de treinamento $[\underline{x}(1), \underline{x}(2), \dots, \underline{x}(t)]$ e $[\underline{d}(1), \underline{d}(2), \dots, \underline{d}(t)]$, onde \underline{x} são os vetores de entrada e \underline{d} são os vetores alvo, deve ser apresentado à rede a fim de que os valores dos pesos possam ser ajustados. Assim, para cada par de treinamento $[\underline{x}(k), \underline{d}(k)]$, com $k = 1, 2, \dots, t$, os seguintes cálculos devem ser feitos:

II.1 Fase de propagação.

Calcule os potenciais de ativação $v(k)$ e o sinal de saída da função de ativação $Y(k)$ de cada neurônio de uma camada, procedendo com a propagação do sinal de entrada $\underline{x}(k)$, camada por camada até a camada de saída. O valor do potencial de ativação ou nível de atividade interna v_j do neurônio j da camada l é calculado segundo a expressão:

$$v_j^l(k) = \sum_{i=0}^p w_{ji}^l y_i^{(l-1)}(k) + \theta_j$$

onde $y_i^{(l-1)}(k)$ é o sinal de saída do neurônio i da camada anterior ($l-1$) na iteração k , w_{ji} é o peso representando a conexão sináptica entre o neurônio i da camada ($l-1$) ao neurônio j da camada l e θ_j é o valor de polarização do neurônio j na camada l . Considerando a função tangente hiperbólica (utilizada neste trabalho) como função de ativação, então o sinal de saída do neurônio j da camada l para cada camada é dado por

(a) se o neurônio j estiver na camada de entrada ($l = 0$):

$$y_j^0(k) = x_j(k)$$

onde x_j é o j -ésimo elemento do vetor \underline{x}

(b) se o neurônio j estiver em uma camada intermediária: ($l = 1, \dots, L-1$).

$$y_j^l(k) = \tanh(v_j^l(k))$$

(c) se o neurônio j está na última camada ($l = L$), então:

$$y_j^L(k) = o_j(k)$$

onde $o_j(k)$ é o valor de saída da função de ativação da última camada, que no presente estudo foi considerada linear.

II.2 Cálculo do sinal de erro.

para os neurônios da última camada, calcule o sinal de erro:

$$e_j(k) = d_j(k) - o_j(k)$$

onde d_j é a j -ésima componente do vetor $\underline{d}(k)$.

II.3 Fase de retropropagação.

Calcule os valores dos δ da rede por retropropagação do sinal de erro, conforme as equações abaixo:

(a) para neurônio pertencente à camada L (função de ativação linear):

$$\delta_j^L(k) = e_j^L(k)$$

(b) para neurônio pertencente à camada intermediária $l = (L - 1, \dots, 1)$:

$$\delta_j^l(k) = (1 - (y_j^l(k))^2) \sum_z \delta_z^{l+1}(k) w_{zj}^{l+1}(k)$$

Ajuste o valor dos pesos da camada l de acordo com a regra delta generalizada

$$w_{ji}^l(k+1) = w_{ji}^l(k) + \alpha[w_{ji}^l(k) - w_{ji}^l(k-1)] + \eta \delta_j^l(k) y_i^{j-1}(k)$$

onde η é a taxa de aprendizado e α é uma constante chamada momentum, cujos valores devem ser heurísticamente arbitrados.

III Ciclo de treinamento.

Repita o item II até que um determinado critério de parada tenha sido atingido. Cabe notar que o algoritmo descrito acima apresenta uma forma de treinamento em que o ajuste dos pesos é feito por padrão. Há outra forma de ajuste conhecida por ajuste em batelada, no qual os pesos só são ajustados após todos os padrões terem sido apresentados.

Ao realizar o ajuste dos pesos de forma a minimizar a diferença entre a saída $o(k)$ da rede e o vetor $d(k)$, para todos os pares $[x(k), d(k)]$, diz-se que a rede constrói uma representação interna do mapeamento que relaciona o vetor $\underline{x} \in R^N$ ao vetor $\underline{d} \in R^M$. Essa é uma grande vantagem da rede neural, ao mesmo tempo que é uma desvantagem. Ao construir uma representação interna de um determinado mapeamento a rede neural toma para si a função de modelar a relação $[\underline{x}, \underline{d}]$, não havendo necessidade de intervenção humana ou de um conjunto de regras para modelagem. Por outro lado, não há como saber de que forma a rede neural constrói a representação do sistema que está modelando [Lee, 1992, Papalexopoulos, 1994], muito embora tal fato possa ser contornado utilizando-se Redes Neurais Fuzzy [Mori, 1994], ou escolhendo-se um conjunto apropriado de funções base [Von Zuben, 1996]. O uso de Redes Neuro-Fuzzy ou a escolha de funções base apropriadas permite utilizar informações prévias sobre as características da função f que se deseja identificar a fim de melhorar a flexibilidade do modelo de aproximação em representar f [Von Zuben, 1996].

A capacidade de uma rede neural construir representações, ou aprender, está intrinsecamente relacionada com a forma pela qual são feitas as conexões entre os neurônios, a topologia da rede, o ambiente na qual ela é utilizada e do paradigma de treinamento para a rede.

Há três tipos básicos de paradigmas de treinamento [Jain, 1996]:

- **Treinamento Supervisionado:**

Nesse tipo de treinamento para cada vetor de entrada é fornecido um vetor de saída contendo a resposta desejada, a qual a rede deve aprender. Esse é o tipo de treinamento

utilizado no algoritmo de retropropagação apresentado, no qual é utilizada uma regra de correção de erro. Há outras regras que permitem o treinamento supervisionado, como: “regra de Hebbian”, “regra de Boltzman” ou “competitive rule”. Neste trabalho será utilizado apenas o treinamento supervisionado com correção de erro, também chamado “Teacher-Forcing”.

- **Treinamento Não Supervisionado:**

Nesse tipo de treinamento não há necessidade de se apresentar à rede um vetor de saída desejado. Ela mesma é capaz de construir uma representação dos padrões de entrada, segundo as categorias em que os mesmos estão agrupados.

- **Treinamento Híbrido:**

É uma combinação dos dois tipos de treinamento apresentados acima.

Muito embora redes MLP sejam um poderoso método de regressão não-linear, algumas limitações ou dificuldades precisam ser consideradas quando de sua utilização. Essas dificuldades estão relacionadas basicamente às seguintes questões:

Critério de Minimização no Algoritmo de Treinamento

No uso do algoritmo de Retropropagação clássico a minimização da função custo é feita através do Método do Gradiente. Por ser linear esse método apresenta problemas de baixa taxa de convergência (no caso da rede ser sobredimensionada), além de geralmente convergir para soluções sub-ótimas ou ficar preso em um mínimo local. Uma forma de resolver esse problema de minimização é utilizar o Método de Newton ou suas derivações, uma vez que eles possuem uma taxa de convergência maior do que o Método do Gradiente. Contudo, seu uso não garante a convergência para uma solução ótima. Deve-se, ainda, considerar o esforço adicional de se trabalhar com a matriz hessiana, além da possibilidade de um mau condicionamento numérico.

Para atingir a solução ótima é necessário o uso de algoritmos de busca, tais como algoritmos genéticos, “simulated annealing” ou outros algoritmos desenvolvidos para esse problema [Shang, 1996], sob pena de um longo tempo de computação. O presente estudo faz uso apenas do Método do Gradiente, usando momento.

Número de Neurônios por Camada e Número de Camadas

O problema do número de neurônios a ser utilizado está relacionado a outro problema, descrito a seguir, que é a capacidade de generalização. Na prática não há uma maneira de se saber, a priori, qual o número ótimo de neurônios a utilizar, dado um conjunto de treinamento. A regra estabelecendo que $(2N_1 + 1)$ (N_1 é o número de neurônios existentes na camada de entrada) representa um número suficiente de neurônios, na maioria das vezes, fornece um valor superdimensionado. Contudo, se não for utilizado algum método para

redução do número de neurônios, verifica-se [Park, 1996] que uma rede superdimensionada oferece maior capacidade de generalização do que uma rede subdimensionada, desde que não se avance muito no processo de treinamento, ou seja, o número de iterações no algoritmo de Retropropagação seja reduzido. Por sua vez, não há forma de calcular esse valor e ele deve ser obtido por tentativa e erro através da técnica de validação cruzada⁴, a qual será comentada no item seguinte. Esse mesmo procedimento também é utilizado para a redução do número de neurônios na camada intermediária. De fato, alguns algoritmos [Cottrell, 1995, Satoh, 1994] para redução de neurônios incorporam essa técnica durante a fase de treinamento da rede. Um problema no uso desses algoritmos é que eles requerem o retreinamento da rede para verificar se não houve degradação da performance da mesma. Entretanto, há métodos que reduzem este problema [Park, 1996], porém, eles devem ser utilizados com técnicas de treinamento mais eficientes do que o método do gradiente. É possível e mais indicado utilizar algoritmos construtivos, que ao invés de retirar neurônios, acrescentam neurônios. Dessa forma, pode-se garantir que a rede, partindo apenas de um neurônio inicial, atinja um número ótimo de neurônios na camada intermediária [Von Zuben, 1996]. De qualquer forma, esses algoritmos se restringem às redes de apenas uma camada intermediária.

Outro problema está relacionado ao número de camadas a ser utilizado. Muito embora uma rede com apenas uma camada intermediária seja capaz de aproximar qualquer função, sabe-se que em certas situações o número de neurônios necessários para realizar esta tarefa é muito elevado, tornando o seu uso impraticável, sendo que uma rede com duas camadas pode ser utilizada de uma forma mais eficiente [Rush, 1993]. Porém a questão do número de camadas a ser utilizado permanece em aberto. Neste trabalho utilizou-se apenas redes de uma camada intermediária.

Por fim, convém salientar que o número de variáveis independentes (o que corresponde ao número de neurônios na camada de entrada) bem como o número de neurônios da camada intermediária deve ser o menor possível.

Capacidade de Generalização

A capacidade de generalização se refere à habilidade da rede neural fornecer um valor de saída que seja aceitável, segundo algum critério, quando um conjunto de dados não utilizados durante o treinamento lhe for apresentado. Esse fato está relacionado com os fatos anteriores, pois a capacidade de generalização depende da estimativa dos parâmetros da rede (pesos) e do número de neurônios utilizados [Sjöberg, 1995]. Para um número de neurônios ótimo, o nível de generalização também é ótimo. Contudo, se a rede é subdimensionada para o problema, então ela não generalizará bem e se a rede é superdimensionada, a capacidade de generalização será também afetada por haver um excesso de parâmetros na rede. Todavia, é possível melhorar a generalização parando o processo de minimização

⁴[Sjöberg, 1995] comenta que o uso excessivo dessa técnica pode levar a uma conclusão polarizada sobre a capacidade de generalização da rede, uma vez que o uso repetido do conjunto de dados para teste da rede, equivaleria a incorporá-lo no conjunto de treinamento.

antes que o mínimo seja atingido (no caso de uma rede com número de neurônios fixos). Buscar o mínimo, nesse caso, corresponde também a aumentar a variância do modelo. O número de amostras utilizadas para treinamento também influencia a capacidade de generalização. Em geral, um número maior de amostras tende a aumentar a capacidade de generalização, desde que o problema abordado seja estático.

Redes do tipo MLP são conhecidas por representarem um modelo estacionário, o qual é caracterizado por equações que não possuem memória, significando que a sua saída é função apenas da entrada corrente e não de valores passados. Contudo, o problema de séries temporais requer o uso de valores passados da série. Para isso, é necessário utilizar um modelo dinâmico de rede, que será objeto da próxima seção.

2.5 Redes Neurais para Identificação de Séries Temporais

Modelos de rede dinâmicos são sistemas que apresentam memória e geralmente as equações dos seus nós são descritas por equações diferenciais ou de diferenças, que são adequadas para descrever o comportamento de uma série temporal. Esses modelos podem ser classificados em três diferentes grupos: redes com linha de derivação de atraso (“tapped delay line”), redes com realimentação de saída e redes com realimentação de estado. A primeira categoria não apresenta recorrência (podendo ser usada de forma recorrente quando de sua utilização após o treinamento) e as outras duas sim. Uma vez que neste trabalho só foram utilizadas redes com linha de derivação por atraso e suas variações, os detalhes de implementação serão apresentados apenas para estes modelos. O leitor interessado em estudar modelos de redes recorrentes poderá utilizar as referências [Connor, 1994], [Narendra, 1990], [Nerrand, 1993], [Von Zuben, 1996] e [Williams, 1989].

Quando se considera o uso de redes MLP para identificação de séries temporais de valores $y(t)$, os elementos do vetor de entrada \underline{x} da rede neural passam a ser valores passados da série $u(t)$ (onde u é uma variável exógena (externa) da série $y(t)$) e, também, da própria série $y(t)$ ⁵. Nesse caso, \underline{x} passa a ser escrito na forma:

$$\underline{x}(t) = [y(t-1), y(t-2), \dots, y(t-p), u(t-1), u(t-2), \dots, u(t-m)] \quad (2.29)$$

O problema da aproximação, descrito na seção 2.7.2, passa a ser o seguinte:

Seja a série $y(t)$ representada por uma função não linear desconhecida f de valores passados $\underline{x}(t)$, mais um ruído branco $w(t)$:

⁵Por simplicidade, os modelos de identificação apresentados trataram apenas do caso monovariável do tipo SISO (single input single output). Porém, a extensão para o caso multivariável é simples.

$$y(t) = f(\underline{x}(t)) + w(t) \quad (2.30)$$

Deseja-se encontrar \hat{f} , representada por uma rede neural, que aproxime f , segundo uma parametrização $\phi(t)$, tal que:

$$\hat{y}(t) = \hat{f}(\phi(t), \underline{x}(t)) \quad (2.31)$$

O enunciado acima representa a formulação geral da identificação de uma série temporal que, para o caso linear, é apresentada na seção 2.2.3 através da equação (2.8). Na realidade, a função \hat{f} com $\underline{x}(t)$ definido segundo (2.29) representa um modelo NARX⁶ (modelo não linear autoregressivo com variáveis exógenas) da série $y(t)$, cuja representação por rede neural é esquematizada na Figura 2.3, onde o operador Z^{-1} indica atraso unitário.

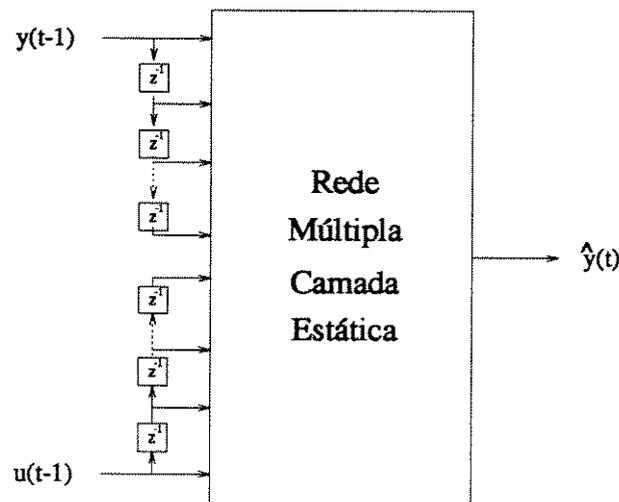


Figura 2.3: Rede MLP representando modelo NARX.

Abaixo segue uma caracterização dos possíveis modelos de identificação [Sjöberg, 1995] e [Von Zuben, 1996]:

- NFIR

Este modelo usa somente variáveis exógenas como entrada e é um modelo sempre estável. Pode requerer uma dimensão elevada para o vetor $\underline{x}(t)$.

⁶Os modelos ARX e ARMAX são um caso particular dos modelos NARX E NARMAX, para os quais a rede neural passa a ser representada por apenas um neurônio com uma função de ativação linear, conforme visto na seção 2.7.1 .

- NARX

Esse modelo usa como variáveis de entrada, tanto valores passados da saída do sistema quanto valores passados das variáveis exógenas. Ele tem a vantagem de ser simples e representar melhor a dinâmica do sistema se comparado ao modelo anterior (do ponto de vista da identificação, na forma apresentada pela equação (2.31), ele é considerado completo). Não há ainda a necessidade de uma representação explícita do ruído, a qual é incorporada na representação dinâmica. A sua desvantagem é a necessidade de uma dimensão elevada para o vetor $\underline{x}(t)$, caso se deseje representar o ruído adequadamente.

- NOE

Esse modelo é semelhante ao NARX diferenciando-se deste pelo fato de os valores passados da saída do sistema, utilizados na composição do vetor $\underline{x}(t)$, serem valores simulados da saída, em vez de valores medidos. Isto faz com que o modelo apresente realimentação da saída (recorrência), tornando sua modelagem mais complexa.

- NARMAX⁷

Este modelo é idêntico ao modelo NARX, possuindo uma modelagem separada para o ruído, o qual é representado pela equação $w(t) = y(t) - \hat{y}(t)$. Dessa forma, o vetor $\underline{x}(t)$ passa a conter valores passados da saída do sistema, de variáveis exógenas e do ruído. Por incorporar a descrição do ruído, a dimensão do vetor $\underline{x}(t)$ fica reduzida. Contudo, ele também requer realimentação, tornando o algoritmo de treinamento mais complexo.

Há a possibilidade de se elaborar modelos simplificados das estruturas NARX, NOE e NARMAX contendo algumas partes lineares, bem como construir modelos complexos, cuja estrutura é montada utilizando-se blocos das estruturas NFIR, NARX, NOE ou NARMAX. Esse fato é comum em previsão de carga.

Uma generalização das estruturas acima pode ser obtida por meio de um modelo por espaço de estados [Sjöberg, 1995], o qual é representado pelas equações abaixo:

$$s(t+1) = f(s(t), u(t), w(t)) \quad (2.32)$$

$$\hat{y}(t) = g(s(t), u(t)) + e(t) \quad (2.33)$$

⁷Esse modelo apresenta um variação chamada NBJ, na qual $x(t)$ é composto por valores passados de $u(t)$, $\hat{y}(t)$ (valor simulado), $w(t)$ e valores simulados do ruído.

onde $w(t)$ e $e(t)$ são respectivamente chamados de ruído do processo e ruído de medição. A diferença básica entre o modelo por espaço de estados e os outros modelos, chamados de entrada/saída, é que aqueles possuem flexibilidade em representar o sistema tanto através da equação (2.32), quanto da equação (2.33), enquanto que a flexibilidade dos modelos entrada/saída é representada apenas através da equação (2.33), uma vez que esta equação é previamente fixada e representa o vetor $\underline{x}(t)$ [Sjöberg, 1995].

Neste trabalho utilizou-se apenas o modelo NARX, devido as suas vantagens. Porém, esse modelo apresenta um problema estrutural, relacionado ao uso de redes MLP para a sua representação. Conforme visto na seção 2.4.2, esse tipo de rede não possui memória e, portanto, é capaz de aproximar uma série apenas no sentido espacial, o que resulta na necessidade de encontrar certas regularidades na série, a fim de utilizar essa técnica com eficiência. Isso implica em modelos às vezes complexos. Uma opção mais atraente é aproximar a dinâmica que gera a série temporal. Sendo assim, um outro tipo de rede neural capaz de representar o processamento temporal de um modelo NARX [Haykin, 1994] foi estudado e será apresentado a seguir.

2.5.1 Redes NARX para Processamento Temporal

Equação de Propagação

Redes FIR (“Finite Impulse Response”)⁸ são redes semelhantes às redes perceptron multicamada (MLP) do ponto de vista da topologia da rede, que é a mesma da figura (2.2). A diferença é que elas são capazes de reproduzir funções que variam no tempo ou séries não-estacionárias. Isso significa dizer que, dado um conjunto de valores passados da série, as redes FIR têm a capacidade de reproduzir valores futuros da série com razoável precisão, até um determinado instante de tempo futuro, a partir do qual a sua saída sofre uma deterioração em relação ao valor verdadeiro da série. Essa rede é capaz de construir relações internas em seus neurônios podendo modelar o que seria a lei de formação da série temporal. Em outras palavras a rede possui uma memória, que é implementada através de filtros FIR associados às sinapses da rede.

Assim, em vez dos pesos das sinapses serem constituídos por valores escalares, eles passam a ser vetores cujos elementos são os coeficientes do filtro e a camada de entrada passa a ser constituída por apenas um neurônio.

O nível interno parcial de atividade de um neurônio j da camada $c + 1$ em termos do sinal gerado pelo neurônio i da camada c é dado pela equação (2.34).

⁸Nesta seção são apresentados apenas as equações básicas acerca de redes FIR. Para uma leitura mais aprofundada o leitor deve consultar as referências [Haykin, 1994] e [Wan, 1993b].

$$s_{ji}(n) = F(n)x_i(n) \quad (2.34)$$

onde $F(n) = a_0 + a_1q^{-1} + \dots + a_Mq^{-M}$ representa a equação do filtro no domínio do tempo discreto. M é a ordem do filtro, a_i são os coeficientes do filtro e q^{-i} é um operador de atraso unitário. A função $x_i(n)$ é o sinal de saída do neurônio i da camada c , no instante n .

Expandindo a equação (2.34), obtém-se:

$$s_{ji}(n) = a_0x_i(n) + \dots + a_Mx_i(n - M) \quad (2.35)$$

Escrevendo na notação mais peculiar de redes neurais os coeficientes $a_l, l = 0, \dots, M$ passam a ser identificados por $w_{ji}(l)$. Aqui $w_{ji}(l)$ é o peso correspondente à sinapse, relacionada à derivação de atraso l , que liga o neurônio i da camada c ao neurônio j da camada $c + 1$. Assim, a equação (2.35) pode ser escrita como segue:

$$s_{ji}(n) = \sum_{l=0}^M w_{ji}(l)x_i(n - l) \quad (2.36)$$

Pode-se escrever (2.36) em notação vetorial na forma que segue abaixo:

$$s_{ji}(n) = \underline{w}_{ji}^T \cdot \underline{x}_i(n) \quad (2.37)$$

onde:

$$\underline{x}_i(n) = [x_i(n), x_i(n - 1), \dots, x_i(n - M)]^T \quad (2.38)$$

$$\underline{w}_{ji} = [w_{ji}(0), w_{ji}(1), \dots, w_{ji}(M)]^T \quad (2.39)$$

Conforme dito acima, $s_{ji}(n)$ refere-se ao nível parcial de atividade do neurônio j devido ao neurônio i . Dessa forma, considerando todos os p neurônios pertencentes à camada C , tem-se que o nível total de atividade do neurônio j e seu valor de saída são dados pelo par de equações (2.40) e (2.41):

$$v_j(n) = \sum_{i=1}^p \underline{w}_{ji}^T \underline{x}_i(n) - \theta_j \quad (2.40)$$

$$y_j(n) = \varphi(v_j(n)) \quad (2.41)$$

onde $\varphi(\cdot)$ é a função de ativação não linear e θ_j é valor de polarização do neurônio j .

As equações (2.40) e (2.41) representam a propagação do sinal da camada c para a camada $c + 1$.

Equações de Retropropagação

A equação para ajuste do peso em uma rede FIR tem a forma familiar:

$$\underline{w}_{ji}(n+1) = \underline{w}_{ji} + \eta \delta_j(n) \underline{x}_i \quad (2.42)$$

Note que a equação (2.42) é semelhante à equação de ajuste de pesos de uma rede MLP estática, exceto que agora w e x são vetores em vez de escalares. O índice j se refere ao neurônio que está na camada onde é feita a atualização dos pesos e i é o neurônio que se conecta ao neurônio j pela sinapse representada por \underline{w}_{ji} e, portanto, pertence à camada anterior à camada do neurônio j . Em (2.42) n é o instante de tempo e não o padrão apresentado à rede.

Para realizar a adaptação de \underline{w}_{ji} é necessário calcular o valor de $\delta_j(n)$, o que é feito seguindo os passos abaixo.

1. Se o neurônio j pertence à última camada.

$$\delta_j(n) = e_j(n) \varphi'(v_j(n)) \quad (2.43)$$

onde $e_j(n) = y_j(n) - d_j(n)$ é erro entre o valor calculado e o valor real, para o instante de tempo n .

2. Se o neurônio j pertence a uma camada intermediária (Na equação que segue supõe-se que o neurônio j se conecta a todos os neurônios da camada posterior à sua).

$$\delta_j(n) = \varphi'(v_j(n)) \sum_{m=1}^N \Delta_m^T(n) \cdot \underline{w}_{mj} \quad (2.44)$$

onde a camada $c + 1$ (a qual pertence o neurônio m) contém N neurônios, M é a ordem dos filtros presentes nas sinapses que conectam a camada c à camada $c + 1$, $\varphi'(\cdot)$ é a derivada de $\varphi(\cdot)$ e $\Delta_m(n) = [\delta_m(n), \delta_m(n + 1), \dots, \delta_m(n + M)]$.

Analisando a equação (2.44) percebe-se que ela é não-causal. Isso significa que não é possível calcular o valor de $\delta_j(n)$, uma vez que ele depende de valores futuros ainda não calculados. A fim de tornar (2.44) causal é necessário adicionar operadores de atraso na rede, o que matematicamente é representado por:

$$\delta_j(n - M) = \varphi'(v_j(n - M)) \sum_{m=1}^N \Delta^T(n - M) \cdot \underline{w}_{mj} \quad (2.45)$$

Dessa forma, tem-se:

1. Se o neurônio j pertence à camada de saída:

$$\underline{w}_{ji}(n + 1) = \underline{w}_{ji}(n) + \eta \delta_j(n) \underline{x}_i(n) \quad (2.46)$$

$$\delta_j(n) = e_j(n) \varphi'(v_j(n)) \quad (2.47)$$

2. Se o neurônio j pertence à camada intermediária:

$$\underline{w}_{ji}(n + 1) = \underline{w}_{ji}(n) + \eta \delta_j(n - cM) \underline{x}_i(n - cM) \quad (2.48)$$

$$\delta_j(n) = \varphi'(v_j(n - cM)) \sum_{m=1}^N \Delta_m^T(n - cM) \cdot \underline{w}_{mj} \quad (2.49)$$

Nas equações acima c é o número da camada, sendo que $c = 0$ indica a camada de saída. O produto de c por M indica o atraso total existente entre a camada de saída e a camada na qual se está fazendo o ajuste do pesos, para o caso de todas as camadas possuírem filtros de mesma ordem. No evento da ordem dos filtros variar de camada para camada, mas permanecendo a mesma para uma dada camada, substitui-se o valor de cM pelo valor [Wan, 1993b]:

$$cM \leftrightarrow \sum_{c=0}^{C-1} M^c \quad (2.50)$$

onde C é o número total de camadas e M^c indica a ordem M do filtro pertencente à camada c .

Algoritmo de Treinamento⁹

O algoritmo de treinamento de uma rede FIR é semelhante ao utilizado para treinamento de redes tipo MLP, exceto que antes de proceder ao treinamento da rede propriamente dito é necessário inicializar o estado da rede. Isso significa que os vetores \underline{x}_i e Δ devem ser inicializados com valores adequados. O procedimento para fazer isso, bem como o treinamento da rede, será exemplificado¹⁰ a seguir.

Suponha uma série temporal $S(n)$ com N pontos amostrados para treinamento $[S(0), \dots, S(N)]$ e uma rede FIR 1 : 1 : 1 : 1 (indica uma rede de quatro camadas, com um neurônio por camada), que pode ser esquematizada conforme mostra a Figura 2.4. A ordem¹¹ dos filtros por camada é $25 \times 5 \times 5$ (a ordem dos filtros correspondentes à primeira camada intermediária é 25 e das camadas subsequentes é 5). A camada de entrada não possui filtros.

Na figura 2.4 observa-se os vetores X_1 , X_2 , X_3 e X_4 , correspondentes aos vetores \underline{x}_i da equação 2.38. A inicialização destes vetores é feita como segue:

I Fase de propagação inicial.

Primeiro inicializam-se os valores de $X_1(l)$, onde l é uma derivação de atraso, até atingir a ordem do filtro da camada, ou seja, até que $x_1(25) = S(0)$. Então, propaga-se o sinal da camada de entrada para a primeira camada intermediária, e apenas para ela, até que o elemento $x_2(5)$ receba o valor calculado no instante 26, ou seja, $x_2(5) = O_2(26)$. Após isso, passa-se à propagação do sinal da camada de entrada para a primeira camada e daí para a segunda camada e apenas até essa camada. Continua-se com esse procedimento até que o valor de saída da rede $O_4(0)$ seja calculado, o que no exemplo ocorrerá no instante 36. Passa-se então para a fase II:

II Fase de retropropagação inicial.

Nesta fase ocorre a inicialização dos estados dos vetores Δ . De forma semelhante à fase I, procede-se primeiro ao cálculo do erro na camada de saída. Tendo-se calculado o valor de δ para a camada de saída, passa-se ao ajuste de peso dessa camada e apenas dela. Notar que não houve propagação do sinal de erro para as outras camadas, o que só acontecerá no instante em que o elemento 6 do vetor Δ possuir o valor $\delta(36)$. Após isso pode-se propagar o sinal de erro para a camada 3 e apenas para ela. Assim, enquanto o elemento 6 de Δ_3 não possuir um valor de δ calculado, somente serão ajustados os valores

⁹O algoritmo de treinamento apresentado aqui é baseado no código do programa para treinamento de uma rede FIR escrito pelo professor Eric A. Wan, o qual pode ser encontrado no site <http://www.cse.ogi.edu/~ericwan>.

¹⁰A rede apresentada aqui como exemplo é uma variação de rede utilizada pelo Professor Eric A. Wan no "Santa Fé Institute Time Series Prediction and Analysis Competition". Para maiores detalhes ver [Haykin, 1994] ou [Wan, 1993b].

¹¹A notação utilizada por Wan para identificar o número de neurônios em uma topologia e a ordem dos filtros é a inversa da mostrada nesse trabalho, a qual foi utilizada apenas para compatibilizar com a notação mais corrente de identificação de uma topologia de redes MLP.

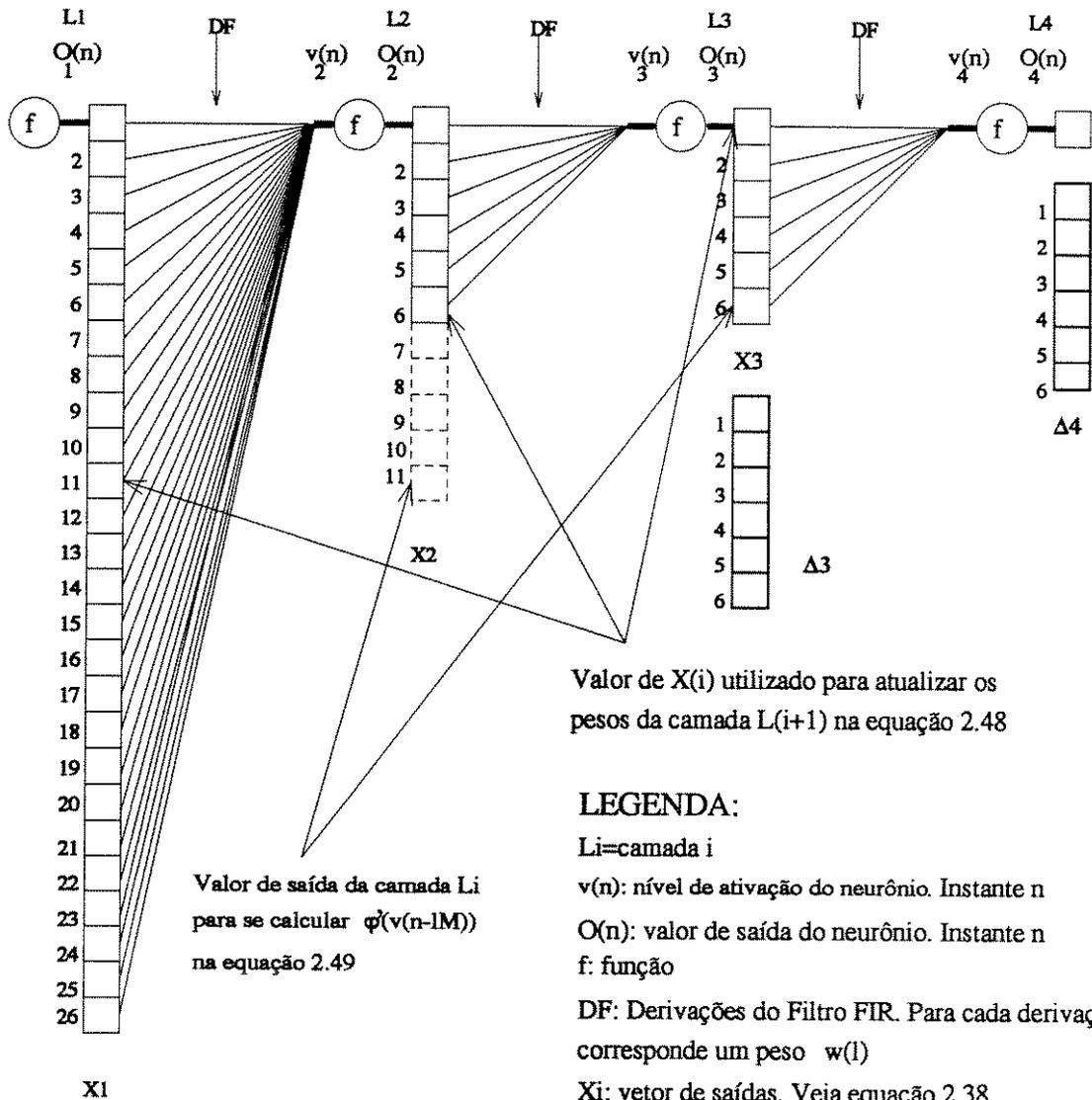


Figura 2.4: Diagrama esquematizado de uma rede FIR 1:1:1:1 25 x 5 x 5

de peso da camada de saída e da camada 3. Esse procedimento é repetido até que se possa calcular o valor de δ da primeira camada escondida. Tendo-se chegado a este ponto, entra-se na fase III.

III Fase de treinamento.

Esta fase é idêntica ao processo de treinamento de uma rede MLP, contendo os passos seguintes:

- (1) propagação do sinal pelas camadas.
- (2) cálculo do erro na camada de saída.
- (3) retropropagação do sinal pelas camadas.
- (4) ajuste do peso para cada camada.

Note que os valores da série $S(n)$ utilizados nas fases I e II acima não são utilizados nesta fase, já que a apresentação dos pontos tem uma seqüência temporal que é a mesma na qual os pontos de S foram amostrados.

IV Treinamento de compensação.

Tendo-se apresentado todos os N pontos da série durante uma época de treinamento, é necessário ainda prolongar o treinamento além de N , pelo fato do ajuste dos pesos depender apenas de saídas passadas apresentadas, conforme se observa na equações (2.46), (2.47), (2.48) e (2.49). Isso é feito tomando-se apenas os passos (2), (3) e (4) da fase III, fazendo-se com que o erro calculado na camada de saída seja igual a zero, passando então para os passos (3) e (4). Repete-se esse procedimento até que o último elemento de Δ_4 seja nulo. Após isso, repete-se os passos (3) e (4) desconsiderando, porém, a última camada. Continua-se assim até que o valor do último elemento do vetor Δ da segunda camada intermediária seja nulo, o que na ilustração apresentada corresponderia a $\Delta_3(6) = 0$.

2.6 Aplicação de Redes Neurais à Previsão de Carga

A utilização de redes neurais para previsão de carga de curto prazo (STLF) vem ganhando atenção nos últimos anos devido a algumas características já apresentadas, notadamente: capacidade de explorar a relação não linear entre a carga elétrica e variáveis independentes correlacionadas a ela, flexibilidade às mudanças rápidas no comportamento do sistema, habilidade de distinguir entre diferentes padrões de carga como dia-da-semana e fim-de-semana, entre outras. Resultados obtidos através de estudos e implementações demonstram que redes neurais apresentam um desempenho superior ao das técnicas convencionalmente empregadas em STLF [Baumann, 1993], [Chen, 1992], [El-Sharkawi, 1996], [Kreider, 1994] e [Papalexopoulos, 1990].

Dos tipos de redes neurais utilizados, o que mais se destaca é a rede MLP, treinada com o algoritmo de Retropropagação, utilizando o método do gradiente. Dessa forma, o modelo de



previsão mais empregado é do tipo NARX, usado em inúmeras variações, algumas das quais são mostradas na Figura 2.5. Na realidade, as diferentes formas de se utilizar redes neurais demonstram que não há uma técnica estabelecida para se determinar qual a melhor arquitetura de rede, tamanho do conjunto de treinamento, número de variáveis de entrada e outros parâmetros. A busca por um previsor de carga baseado em redes neurais (ANNSTLF) se torna então uma tarefa exaustiva de análise de vários fatores que afetam o desempenho da mesma. Um outro problema adicional se refere ao fato de que a série temporal que se deseja identificar, ou seja a carga elétrica nodal, é caracterizada por um comportamento não-estacionário, apresentando uma característica cíclica predominantemente diária e semanal, sendo necessário ao modelo de previsão reproduzir esse comportamento.

Nesta seção serão mostrados alguns resultados e regras gerais para a elaboração de um previsor de carga e que são citados na literatura. O leitor desejando obter uma lista mais completa de citação de artigos e trabalhos sobre ANNSTLF, bem como sua descrição, deve se referir a [El-Sharkawi, 1996]. É necessário ressaltar que a grande maioria dos trabalhos sobre ANNSTLF se referem à previsão da carga global¹² do sistema, sendo que alguns poucos se referem à previsão de carga nodal. [Chen, 1992] elaborou um previsor para uma carga tipicamente residencial, [Kreider, 1994] cita vários previsores para a demanda de energia elétrica em um prédio e [Germond, 1996] cita um previsor para carga elétrica regional.

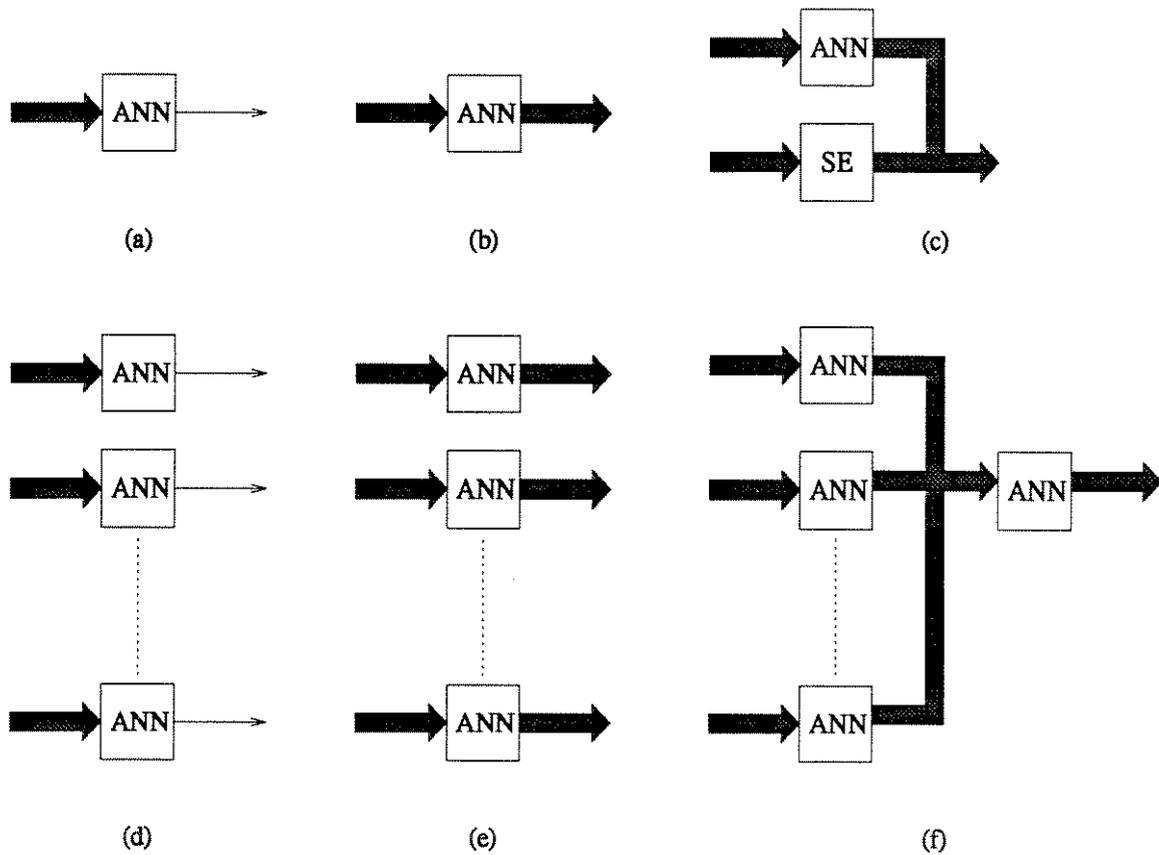
2.6.1 Particionamento do Conjunto de Treinamento

Conforme citado anteriormente, um dos primeiros detalhes notados ao desenvolver um modelo para previsão de carga é o fato da mesma apresentar um comportamento ciclo-estacionário com regularidades relacionadas ao ciclo horário, diário, semanal e até mesmo sazonal. Dessa característica infere-se que diferentes redes neurais podem ser usadas para processar informações referentes às diferentes regularidades encontradas. Esquema este adotado por vários autores [Connor, 1994, Djukanovic, 1995, Hsu, 1991a, Khotanzad, 1995, Lee, 1992, Peng, 1992, Mohammed, 1995, Srinivasan, 1995, Sharkawi, 1993]. Assim, obtido um determinado conjunto de dados para treinamento, pode-se escolher um subconjunto que contenha apenas algumas das regularidades citadas acima e utilizá-lo para treinar uma rede neural que será específica para esse conjunto. A forma pela qual um subconjunto é definido varia e pode ser baseada em conhecimento prévio do comportamento da carga adquirido pelo operador do sistema [Khotanzad, 1995, Lee, 1992] ou classificação de padrões feita por uma rede de Kohonen [Hsu, 1991a].

Pode-se, assim, caracterizar os tipos de particionamento na forma apresentada abaixo:

1. Subconjunto de treinamento igual ao conjunto de treinamento.

¹²A palavra carga mencionada no restante desta seção se refere à potência ativa. Não foi encontrado na literatura pesquisada nenhum trabalho que citasse a previsão de carga reativa por redes neurais.



a) Rede Neural de múltiplas entradas e uma saída (MISO).

b) Rede Neural de múltiplas entradas e múltiplas saídas (MIMO).

c) Rede Neural corrigida por um Sistema Especialista.

d) Rede MISO com particionamento.

e) Rede MIMO com particionamento.

f) Estrutura Hierárquica de Redes MIMO.

SE representa algum tipo de sistema especialista, uma rede neural para classificação, ou algum outro regressor.

Figura 2.5: Diferentes arquiteturas de redes neurais utilizadas em previsão

Nesse caso, utiliza-se apenas uma única rede neural para modelar a carga, sendo que as características cíclicas podem ser codificadas em variáveis representando o dia da semana, hora do dia, estação do ano, entre outras. [Bakirtzis, 1996], [Chen, 1992] e [Papalexopoulos, 1990].

2. Subconjunto de treinamento definido segundo conhecimento do operador ou programa de classificação (cabe notar que essa classificação pode ser feita por um sistema especialista.)

Dentro desta categoria, destacam-se as seguintes subcategorias:

- (a) Cada dia da semana representa um subconjunto:

Para esse caso uma rede é alocada para cada dia da semana [Bakirtzis, 1996] e [Khotanzad, 1995].

- (b) Separam-se os dias da semana em grupos:

A separação em grupos pode ser feita por meio de agrupamentos simples, como separar o final de semana dos dias de semana [Djukanovic, 1995], [Lee, 1992], e [Peng, 1992], ou por meio de regras pré-definidas [Mohammed, 1995].

- (c) Um sistema de classificação é utilizado para separar os dias que possuam padrões semelhantes e uma rede é então alocada para cada um desses padrões.

Para cada particionamento acima pode ser feito um subparticionamento, por exemplo, para cada hora do dia ou agrupamento de horas [Connor, 1994, Khotanzad, 1995, Lee, 1992, Mohammed, 1995]. É possível também tomar, dentro de um determinado grupo de dias pertencentes a um determinado particionamento, um subgrupo que contenha características semelhantes às do dia de previsão (é um passo posterior ao particionamento 2.b ou 2.c acima). O critério de semelhança pode ser dado, por exemplo, por uma medida de distância entre vetores de regressão dentro do grupo de dias e o vetor de regressão usado para realizar a previsão [Djukanovic, 1995, Peng, 1992]. Ressalta-se que para diferentes subparticionamentos o vetor de regressão \underline{x} pode ser composto por diferentes variáveis. [Connor, 1994, Lee, 1992, Khotanzad, 1995].

Não há ainda forma de se saber qual o melhor particionamento a ser adotado. Contudo, particionamentos do grupo 2b, com subparticionamento horário, têm sido utilizados com sucesso. [Khotanzad, 1995, Sharkawi, 1993]. Isso, porém, não desmerece estruturas construídas utilizando uma única rede [Bakirtzis, 1996, Papalexopoulos, 1990]. Comparações entre diferentes tipos de particionamentos podem ser encontradas em [Bakirtzis, 1996, Sharkawi, 1993].

2.6.2 Modelos de Regressão Utilizados

A grande maioria dos modelos de regressão utilizados, inclusive em centros de operação [Khotanzad, 1995] e [Papalexopoulos, 1990], são do tipo NARX com duas estruturas básicas:

- MIMO: Múltiplas entradas e múltiplas saídas.

Nessa estrutura a saída é um conjunto de valores de carga. Geralmente valores horários de carga correspondentes às 24 horas do dia. Na Figura 2.5 essa estrutura corresponde aos blocos de saída com setas grandes.

- MISO: Múltiplas entradas e uma única saída.

Nessa estrutura a saída é geralmente o valor horário da carga. Ela corresponde aos blocos de saída com setas pequenas na figura 2.5.

Modelos NARMAX e NFIR são também encontrados na literatura. No caso de modelos NARMAX [Connor, 1994, Czernichow, 1993, Mori, 1993] algumas comparações demonstram uma performance superior ao modelo NARX, ou, pelo menos, um vetor de regressão com dimensões reduzidas em relação ao modelo NARX. Porém, esses resultados não são conclusivos, pois há contra exemplos¹³.

De qualquer maneira, os modelos NARX, NFIR ou NARMAX encontrados na literatura são utilizados em diferentes tipos de estruturas de acordo com os particionamentos apresentados.

Modelos híbridos também são utilizados, como combinações de redes MLP com redes de Kohonen [Djukanovic, 1995, Hsu, 1991b], sistema especialista fuzzy [Srinivasan, 1995] ou [Kim, 1995], ou mesmo com modelos de regressão lineares [Satoh, 1994] (veja Figura 2.5.c). Encontram-se também referências ao uso de Redes Neurais Fuzzy [Bakirtzis, 1994, Mori, 1994]. A vantagem desse tipo de rede, na forma em que foram implementadas por seus autores, reside em tempos de treinamento inferiores se comparados à rede MLP e possibilidade de estabelecer uma melhor compreensão da relação existente entre as variáveis independentes e a variável dependente.

O uso de funções de base radial também é proposto [El-Sharkawi, 1996]. A vantagem desse modelo é sua rapidez (possui convergência linear), bem como a facilidade de realizar o particionamento de padrões durante uma das fases de seu treinamento [Rush, 1993], além de poder gerar índices de confiabilidade da previsão. Muito embora, isso também possa ser realizado por uma rede MLP [Nix, 1995].

2.6.3 Estratégias de Aprendizado

Duas estratégias de aprendizado são possíveis (bem como uma combinação das mesmas):

¹³Em uma competição realizada pela ASHRAE para previsão do consumo de energia elétrica, água e vapor de um prédio, os três primeiros vencedores utilizaram modelos NARX ou NFIR [Kreider, 1994].

1. Adaptação constante do coeficientes da rede (treinamento contínuo).
2. Cálculo dos coeficientes e posterior utilização, sem retreinamento.

Na primeira estratégia os pesos da rede são calculados continuamente (“on-line”) de forma a considerar a não-estacionaridade da carga [El-Sharkawi, 1996, Nerrand, 1993, Williams, 1989] e captar o comportamento mais recente da mesma, incorporando novos padrões que possam surgir [Djukanovic, 1995, Hsu, 1991a]. Dessa maneira, estabelece-se uma janela de tempo contendo os mais recentes padrões, a forma da janela pode ser retangular ou exponencial [El-Sharkawi, 1996, Nerrand, 1993]. Isso significa que apenas uma pequena base de dados com valores passados de carga e variáveis independentes é mantida, geralmente constituída de algumas semanas de medições [Djukanovic, 1995, El-Sharkawi, 1996]. Essa base de dados pode conter tanto algumas semanas de dados imediatamente anteriores aos dias de previsão quanto dados de carga de anos anteriores, desde que estes possuam alguma relevância para o modelo de previsão. Nesse caso, é comum tomar-se os dados de dias que correspondam aos dias para os quais se deseja fazer previsão.

Muitos autores têm utilizado essa técnica com bons resultados, inclusive superando os resultados obtidos com a segunda estratégia [Bakirtzis, 1996]. A frequência de retreinamento varia de diária a mensal¹⁴. Cabe mencionar que um dos modelos de previsão aqui referenciados e realmente implementado em centros de operação utiliza esta técnica [Khotanzad, 1995], sendo que outro modelo implementado utiliza a segunda técnica.

Na segunda estratégia estabelece-se o conjunto de treinamento e, então, os coeficientes da rede são calculados. Uma vez feito isso, os coeficientes são utilizados na fase de previsão sem modificação adicional [Papalexopoulos, 1990, Park, 1991, Sharkawi, 1993, Srinivasan, 1995].

Assim, o número de dias escolhidos para compor o conjunto de treinamento varia de acordo com a estratégia adotada:

- Pode-se escolher uma base de dados global, contendo um ou dois anos de dados. [Khotanzad, 1995] e [Mohammed, 1995]
- Pode-se escolher uma base de dados local, contendo dias pertencentes a algumas semanas anteriores à previsão, incluindo a possibilidade de tomar um conjunto de dias do ano anterior que inclui, também o dia correspondente ao dia para o qual se está fazendo a previsão.
- Combinação dos itens acima [Mohammed, 1995]

¹⁴Deve ficar claro que todo modelo de regressão para um série não estacionária necessita readaptar seus parâmetros. Nesse sentido, considerou-se, neste trabalho, que uma rede de treinamento contínuo é aquela em que o retreinamento é realizado no máximo mensalmente.

2.6.4 Fatores Causais e Topologia da Rede

Fatores causais correspondem aos tipos de variáveis que serão consideradas para compor o vetor de regressão e que, portanto, são capazes de explicar de acordo com algum critério estatístico o comportamento da carga. A primeira questão a ser levantada para definir as variáveis a utilizar na regressão se refere à disponibilidade delas, ou seja, caso se deseje montar um modelo que utilize uma variável correspondente à temperatura horária local, por exemplo, é necessário que essa medida esteja disponível e que seja confiável. Uma vez determinada uma lista de variáveis disponíveis, passa-se à fase de determinação das variáveis a serem utilizadas, escolhidas geralmente segundo algum critério estatístico ou conhecimento adquirido por um operador.

Normalmente os critérios estatísticos adotados são: análise da função de autocorrelação [Lee, 1992], análise da função de correlação cruzada e procedimentos adotados para a elaboração de modelos ARIMA [Chen, 1992]. Métodos estatísticos mais sofisticados para redução da dimensionalidade do vetor de regressão também são citados. Dentro dessa categoria pode-se destacar os seguintes métodos: métodos de poda [Cottrell, 1995], análise da componente principal (PCA) [Feuston, 1994], regularização [MacKay, 1994], entre outros.

O conhecimento empírico do operador também é importante na elaboração das variáveis que serão utilizadas, ou que devam ser consideradas, restringindo o universo de variáveis candidatas. O mesmo conhecimento pode também estabelecer que uma determinada variável (na maior parte das vezes relacionada ao clima) deve ser utilizada não na forma como é medida, mas sim após sofrer alguma transformação não-linear (pré-processamento) [Papalexopoulos, 1990].

Quanto à topologia usada, geralmente utilizam-se redes com uma única camada intermediária, sendo o número de neurônios determinado através do uso de métodos de validação cruzada, ou de métodos estatísticos de poda [Cottrell, 1995, Satoh, 1994]. Redes neurais de duas camadas também são citadas [Baumann, 1993, Cottrell, 1995, Hsu, 1991b]. [Kiarziz, 1994] mostra que a função de ativação adotada não influencia significativamente no resultado (para o caso de diferentes tipos de função sigmoidal). Contudo, o critério de normalização do vetor de entrada é importante para a taxa de convergência. [Peng, 1992] obteve bons resultados utilizando uma transformação do tipo apresentado abaixo, onde \bar{x} representa valor médio, σ_i o desvio padrão, i é a i -ésima componente do vetor \underline{x} e \tilde{x} a variável normalizada.

$$\tilde{x}_i = \frac{(x_i - \bar{x})}{\sigma_i}$$

2.6.5 Modelagem de Dias Especiais e Tratamento de Valores Espúrios

A previsão de carga para dias correspondentes a feriados, ou atividades incomuns, apresentam valores elevados de erro, uma vez que esses dias possuem uma curva de carga atípica. Para tratar esse problema é necessário uma correção do valor, a qual pode ser feita por meio de equações algébricas [Bakirtzis, 1996, Papalexopoulos, 1990], treinamento utilizando uma base de dados especiais [Djukanovic, 1995, Hsu, 1991a, Peng, 1992] ou algum sistema especialista, por exemplo, um sistema fuzzy [Kim, 1995].

[Connor, 1994] demonstra que redes cujos parâmetros são calculados segundo o critério dos mínimos quadrados são sensíveis a valores espúrios de medição, bem como, a medidas atípicas de carga. Para contornar esse problema ele propõe um filtro para suavizar essa sensibilidade. Tal preocupação é válida se a previsão é feita sobre uma base de dados em que não há nenhum tratamento de erros. O efeito de cargas atípicas devido a feriados pode ser eliminado manualmente.

Capítulo 3

Características da Curva de Carga Nodal

3.1 Análise dos Dados de Carga Nodal Ativa e Reativa

3.1.1 Tratamento de Erros Grosseiros

As medições de potência ativa e reativa utilizadas neste trabalho foram obtidas junto a CPFL e referem-se a um conjunto de subestações situadas na região de Campinas. Os dados coletados correspondem a um período de um ano (5 de junho de 1994 a 31 de junho de 1995) e originam-se diretamente do sistema de aquisição de dados do centro de operação da companhia. Portanto, os dados precedem ao tratamento dos possíveis erros existentes nas medições, uma vez que isso só é feito quando da estimação de estado do sistema. Sendo assim, antes de se utilizar os dados para a obtenção de um modelo de rede neural para as curvas de carga, foi necessário um tratamento dos mesmos, a fim de eliminar os erros grosseiros e omissões.

Os erros grosseiros observados à primeira vista estavam relacionados a falhas nos equipamentos de medição, no envio de dados ou na sua aquisição, acarretando na ausência do valor das medidas correspondentes aos instantes de tempo em que houve uma das falhas acima. Dessa maneira, tornou-se necessário acrescentar pseudo-valores na base de dados, os quais foram calculados a partir dos outros valores presentes.

Pelo fato do crescimento da carga em uma subestação durante o período de um ano ser pequeno, o método utilizado para se obter as pseudo-medidas foi calcular a curva de carga média para um período de quatro semanas e posterior substituição das medidas ausentes nos arquivos de dados pelos valores da curva de carga calculados. Assim, supondo-se, por exemplo,

que não existam os dados de medições correspondentes à primeira semana de agosto, toma-se da curva de carga média calculada os valores pertencentes à primeira semana. Estes valores é que substituíram os valores perdidos. Se fossem os valores da segunda semana, então eles seriam substituídos pelos valores da segunda semana da curva de carga média e assim por diante. Esta técnica é válida porque a falta de algumas medições não é fundamental quando se usam redes neurais, pois elas são robustas em relação à este fato.

Uma outra alternativa para eliminar o erro grosseiro seria substituir o valor perdido por um valor correspondente ao mesmo horário e dia da semana anterior. Contudo, preferiu-se o método da curva de carga média, pelo fato deste evitar a propagação de algum erro grosseiro no medidor, ou seja, o equipamento mediu um valor que não é o verdadeiro.

O ideal, no entanto, é que as medidas a serem inseridas no banco de dados, sobre o qual será feita a identificação do modelo da carga, sejam provenientes de uma base de dados reais depurados, ou de algum processo de filtragem [Connor, 1994], conforme descrito no capítulo anterior. Dessa forma, minimizam-se os efeitos dos erros de medição e outros ruídos inconvenientes na obtenção dos parâmetros do modelo. Neste trabalho, não se tratou do problema de filtragem supondo-se que os dados de carga foram devidamente tratados.

3.1.2 Caracterização da Carga das Subestações

Os testes realizados neste estudo são baseados na demanda medida de potência ativa e reativa de subestações cujas características são apresentadas nas tabelas e figuras a seguir. Pode-se, assim, verificar as especificidades das subestações que foram estudadas, tais como o tipo de consumidor preponderante (número de consumidores atendidos pela subestação em porcentagem) e a carga preponderante percentual (relaciona-se ao tipo de consumidor que requer mais energia da subestação). Esses dados são importantes porque influem na forma da curva de carga. Cabe notar que, neste trabalho, as variáveis de entrada utilizadas foram apenas os valores de carga já medidos, bem como valores correspondentes à hora e ao dia da semana, quando o seu uso foi necessário. É bastante provável que para consumidores com predominância comercial e/ou residencial variáveis relacionadas ao clima e à temperatura [Chen, 1992] possam melhorar a eficiência da rede neural em modelar o comportamento da carga. De fato, para barramentos em que um grande consumidor público/residencial ou comercial (shopping center, por exemplo) estão conectados, nota-se que variáveis explicativas de natureza climática têm importância [Ohlsson, 1994] na descrição do comportamento da carga elétrica, muito embora a forma pela qual essas variáveis são utilizadas dependa do tipo de implementação a ser adotada, conforme descrito no capítulo anterior [Kreider, 1994, El-Sharkawi, 1996]. O mesmo talvez não se pode dizer sobre consumidores industriais. Nesse caso, a carga pode depender também das atividades que a indústria realiza durante a semana, tal como o uso de determinada linha de produção em apenas um dia da semana, tornando-se assim uma carga mais difícil de ser modelada, como será visto no próximo capítulo.

Em relação à potência reativa, observa-se nas figuras que ela possui variações mais rápidas ou mais bruscas. Diferentemente da potência ativa, em que a carga relaciona-se com fatores externos mais previsíveis, a potência reativa está mais relacionada ao comportamento da rede elétrica, tais como chaveamento de banco de capacitores, linhas de transmissão e outros fatores de natureza intrínseca ao controle e operação do sistema e, portanto, possui um caráter mais aleatório [Falcão, 1986]. Nota-se, porém, a presença de uma sazonalidade semanal e diária, como na potência ativa. Esse fato pode ser observado nos gráficos das funções de autocorrelação.

Embora o arquivo de dados obtido da CPFL contenha inúmeras subestações, escolheu-se para estudo aquelas que melhor representam determinado tipo de consumidor de forma a melhor caracterizar alguns tipos de carga frequentemente encontrados em um sistema elétrico. Assim, as tabelas 3.1 a 3.11 e respectivas figuras ilustram os padrões de carga segundo a característica de consumo das subestações escolhidas, apenas com o intuito de mostrar a curva de carga típica semanal apresentada pelas mesmas. Os ciclos diários da curva de carga correspondem aos seguintes dias: domingo, segunda, terça, quarta, quinta, sexta e sábado. Notar que a amplitude do ciclo de domingo e de sábado é menor do que a dos outros dias da semana.

N ^o Subestação	Característica Consumo					
	Residencial	Industrial	Comercial	Rural	Outros	
40	93.8	1.3	4.4	.2	.2	Consumidores (%)
Andorinha	75.4	5.0	17.3	.7	1.5	Consumo (%)

Tabela 3.1: Subestação com carga tipicamente residencial

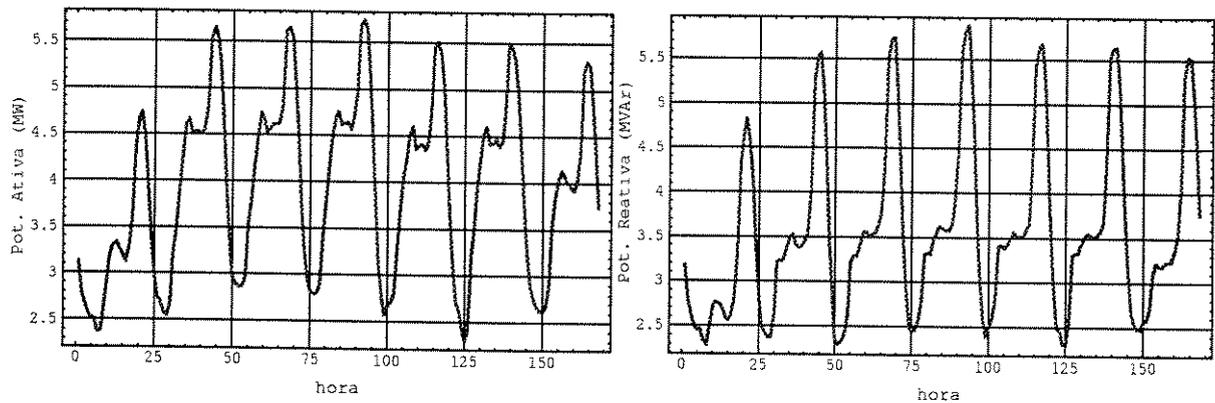


Figura 3.1: Potência ativa e reativa na subestação 40

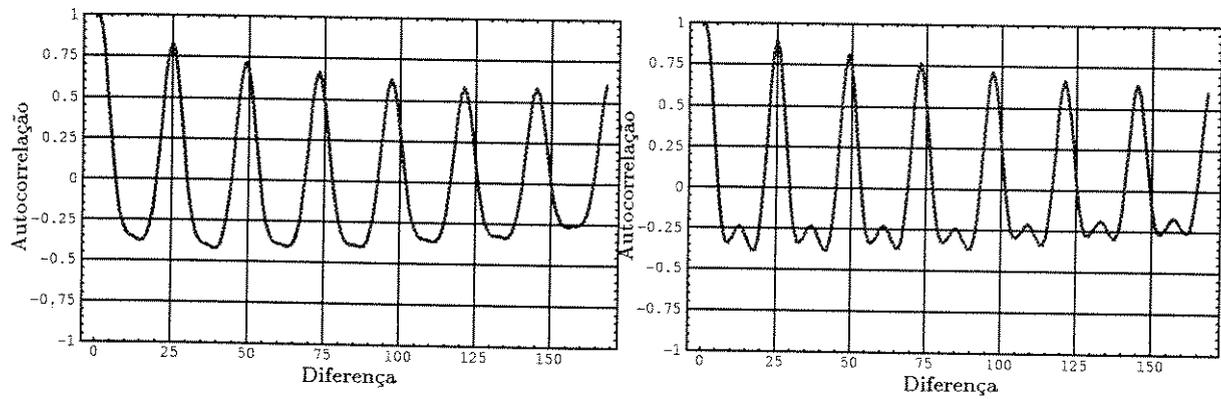


Figura 3.2: Funções de autocorrelação (potência ativa e reativa) na subestação 40

N ^o Subestação	Característica Consumo					
	Residencial	Industrial	Comercial	Rural	Outros	
80	89.8	3.0	4.1	2.8	0.3	Consumidores (%)
Barão Geraldo	26.7	18.3	30.0	4.1	20.6	Consumo (%)

Tabela 3.2: Subestação com grande consumidor público (repartição)

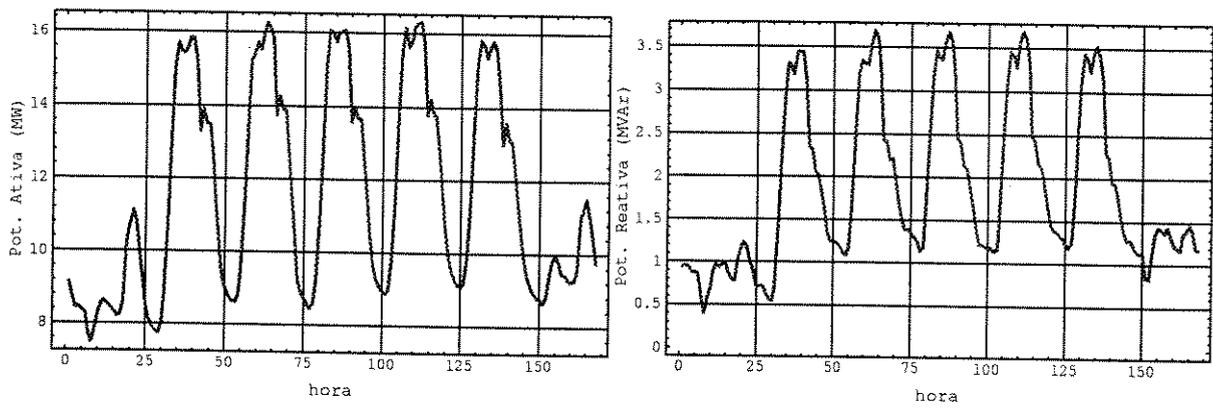


Figura 3.3: Potência ativa e reativa na subestação 80

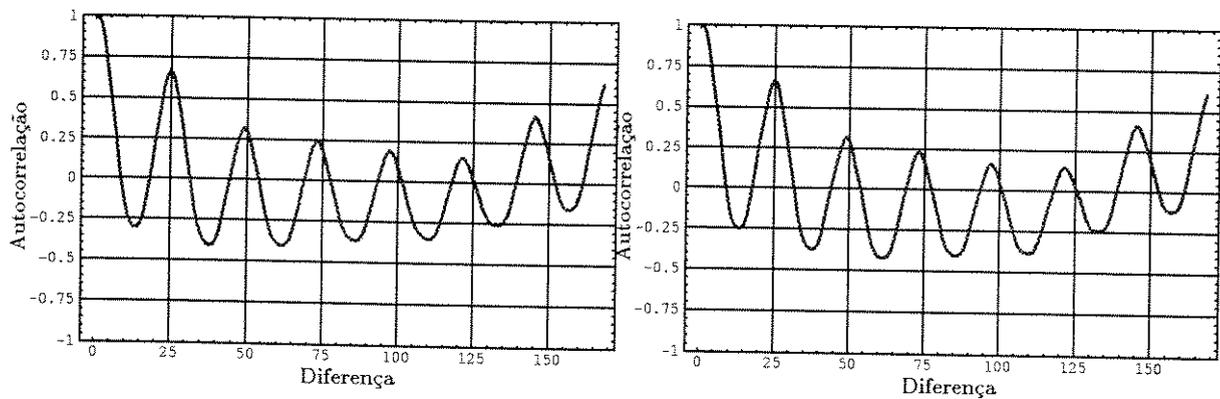


Figura 3.4: Funções de autocorrelação (potência ativa e reativa) na subestação 80

N ^o Subestação	Característica Consumo					
	Residencial	Industrial	Comercial	Rural	Outros	
221	81.5	2.4	15.2	0.1	0.8	Consumidores (%)
Campo Verde	31.4	27.2	18.0	0.1	23.4	Consumo (%)

Tabela 3.3: Subestação com carga equitativa entre consumidores

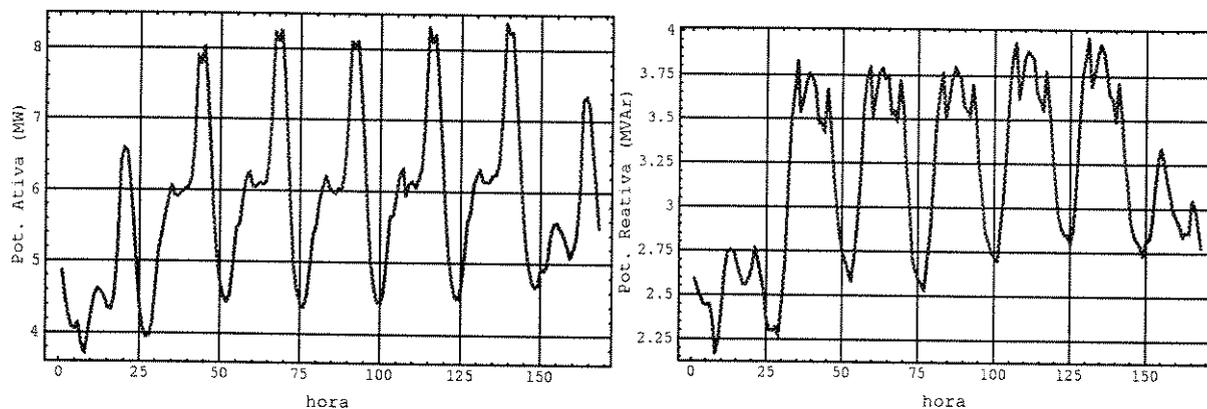


Figura 3.5: Potência ativa e reativa na subestação 221

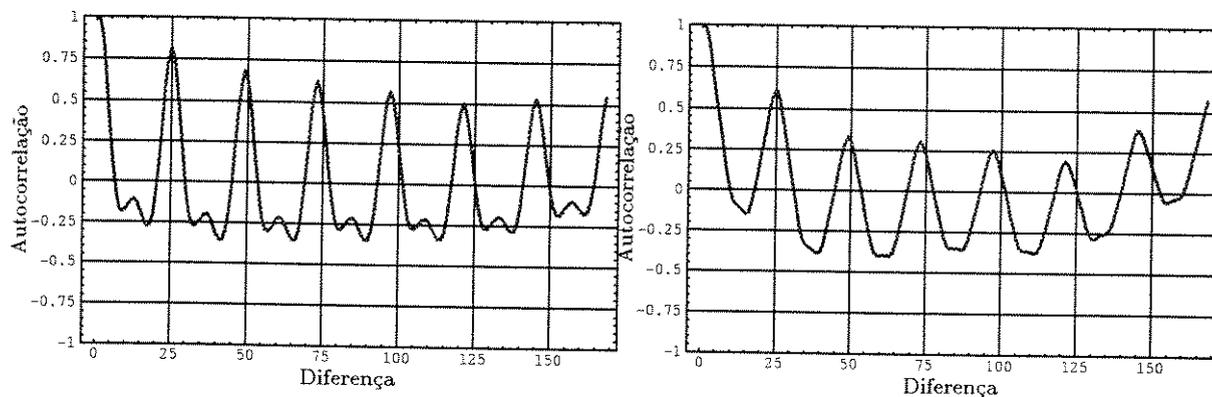


Figura 3.6: Funções de autocorrelação (potência ativa e reativa) na subestação 221

N ^o Subestação	Característica Consumo					
	Residencial	Industrial	Comercial	Rural	Outros	
232	79.7	1.1	18.6	0.0	0.5	Consumidores (%)
Campinas Centro	45.3	7.4	43.1	0.0	4.1	Consumo (%)

Tabela 3.4: Subestação com carga tipicamente residencial e comercial

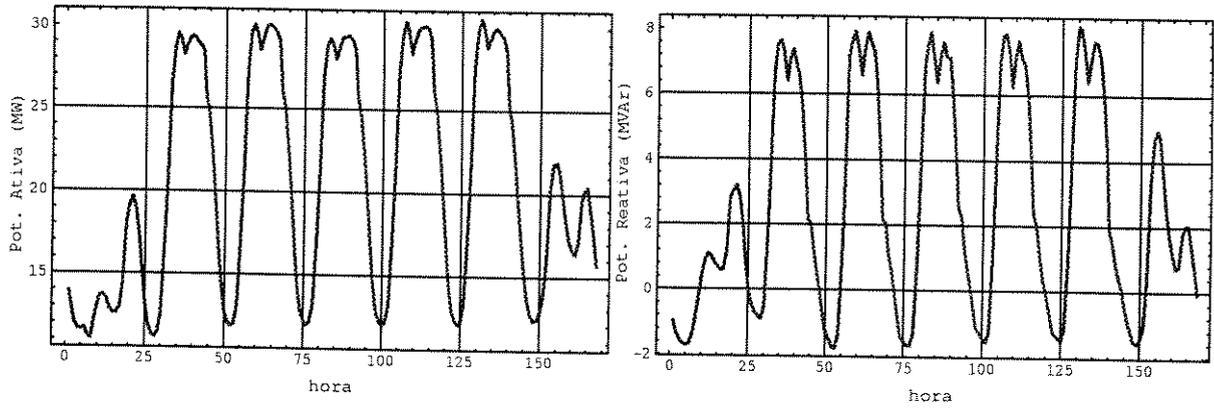


Figura 3.7: Potência ativa e reativa na subestação 232

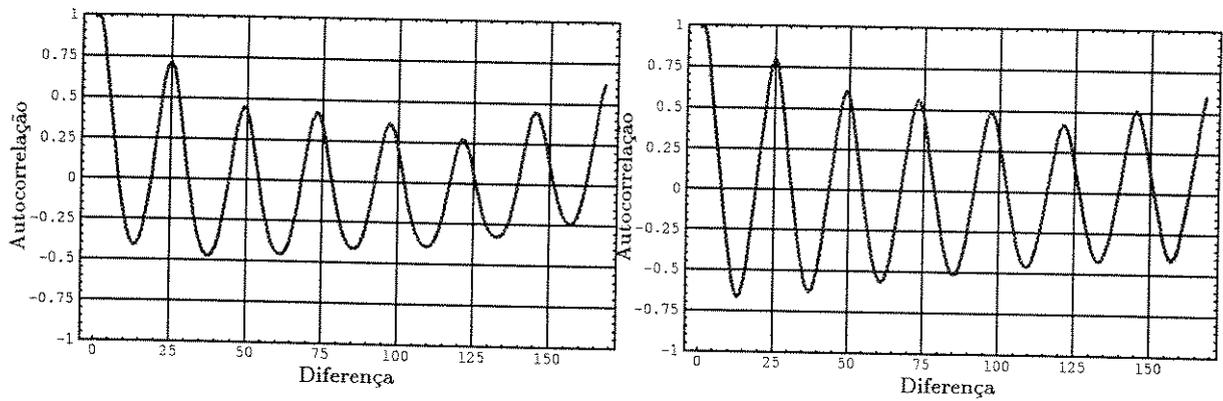


Figura 3.8: Funções de autocorrelação (potência ativa e reativa) na subestação 232

N ^o Subestação	Característica Consumo					
	Residencial	Industrial	Comercial	Rural	Outros	
379	87.7	2.5	2.9	6.7	0.3	Consumidores (%)
DIC	4.7	91.4	1.4	2.2	0.4	Consumo (%)

Tabela 3.5: Subestação com carga tipicamente industrial

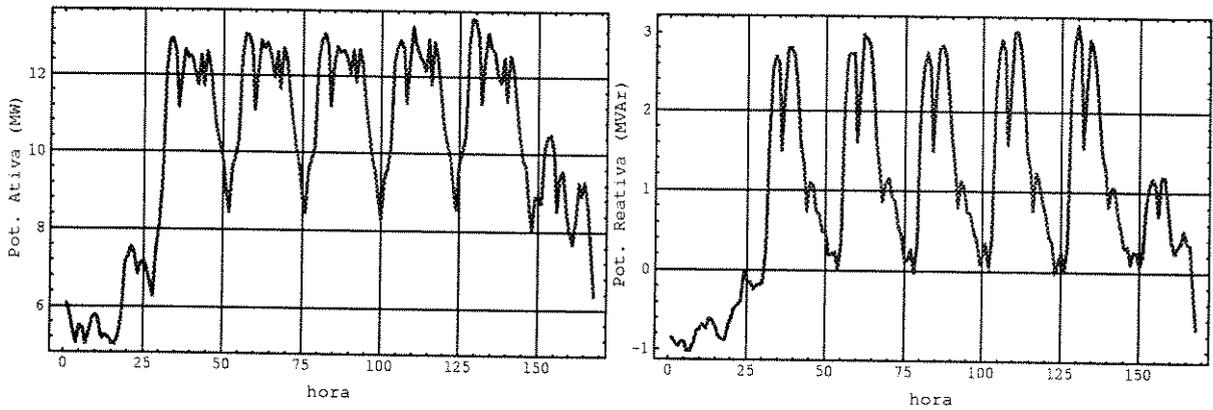


Figura 3.9: Potência ativa e reativa na subestação 40

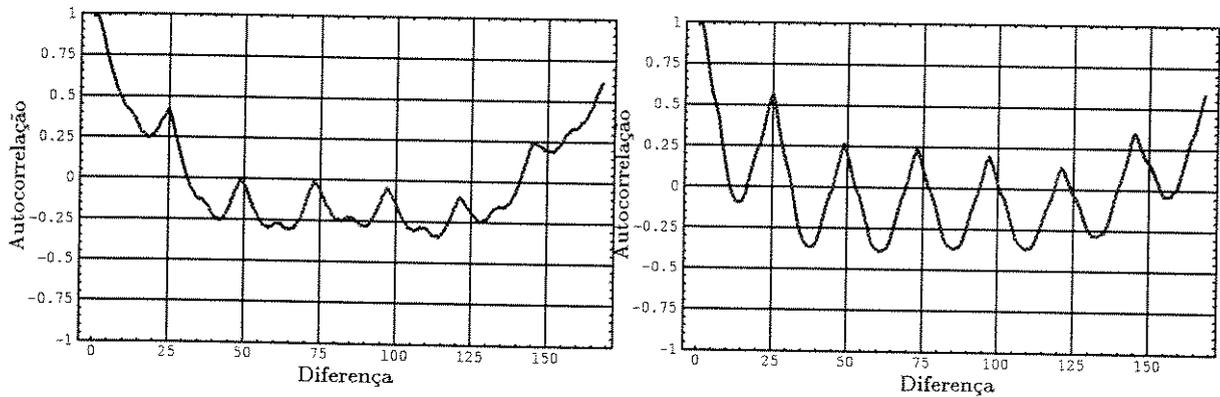


Figura 3.10: Funções de autocorrelação (potência ativa e reativa) na subestação 379

N ^o Subestação	Característica Consumo					
	Residencial	Industrial	Comercial	Rural	Outros	
463	89.8	2.5	7.2	0.1	0.4	Consumidores (%)
Jardim	44.4	41.8	11.6	0.2	1.9	Consumo (%)

Tabela 3.6: Subestação com carga tipicamente residencial e industrial (ind. leves)

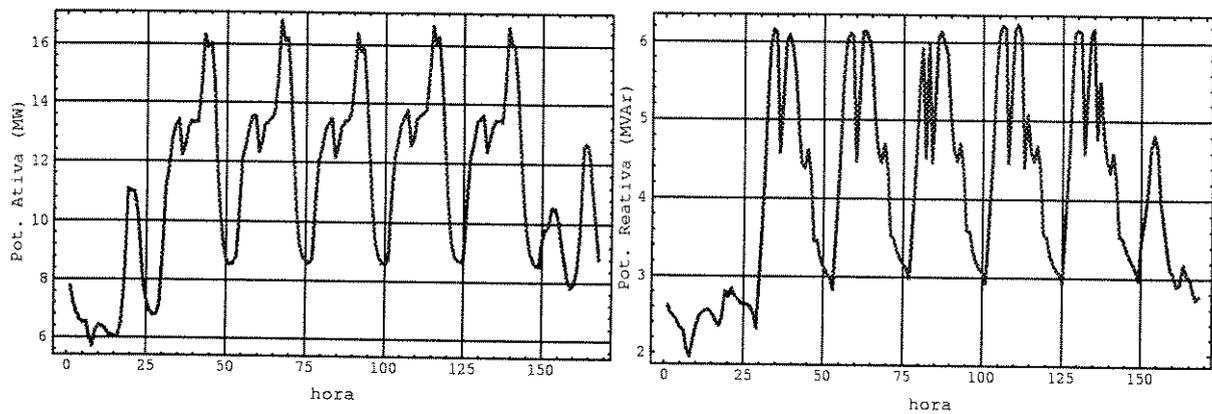


Figura 3.11: Potência ativa e reativa na subestação 463

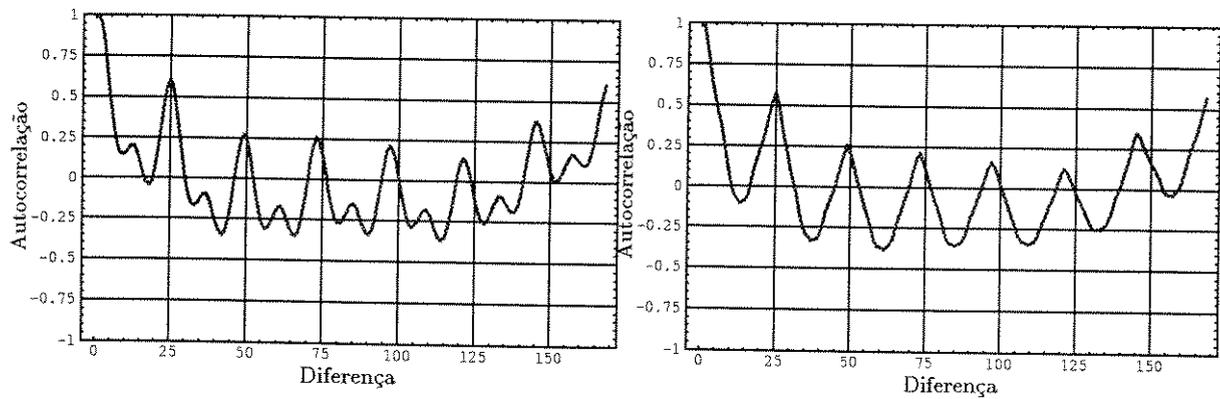


Figura 3.12: Funções de autocorrelação (potência ativa e reativa) na subestação 463

N ^o Subestação	Característica Consumo					
	Residencial	Industrial	Comercial	Rural	Outros	
572	92.9	1.7	4.7	0.3	0.4	Consumidores (%)
Morro Azul	44.4	42.7	5.6	2.7	4.5	Consumo (%)

Tabela 3.7: Subestação com carga tipicamente residencial e industrial (ind. pesadas)

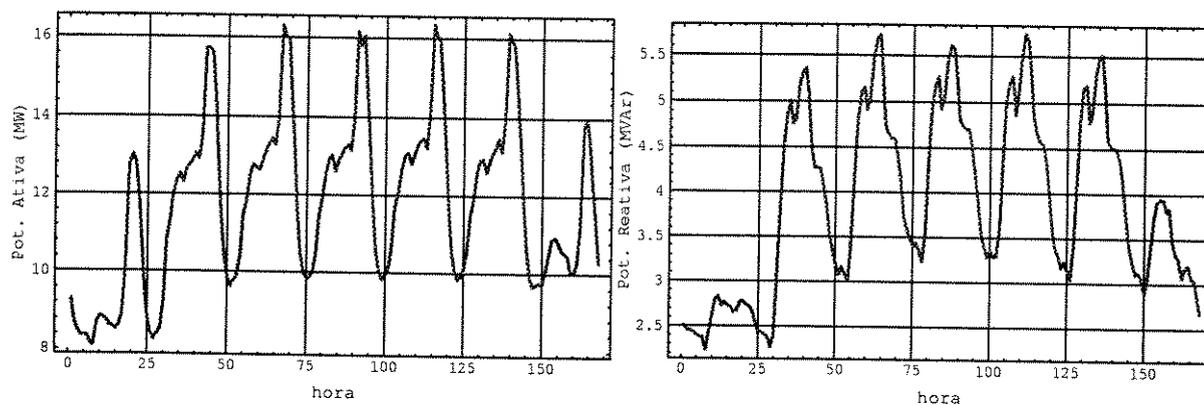


Figura 3.13: Potência ativa e reativa na subestação 572

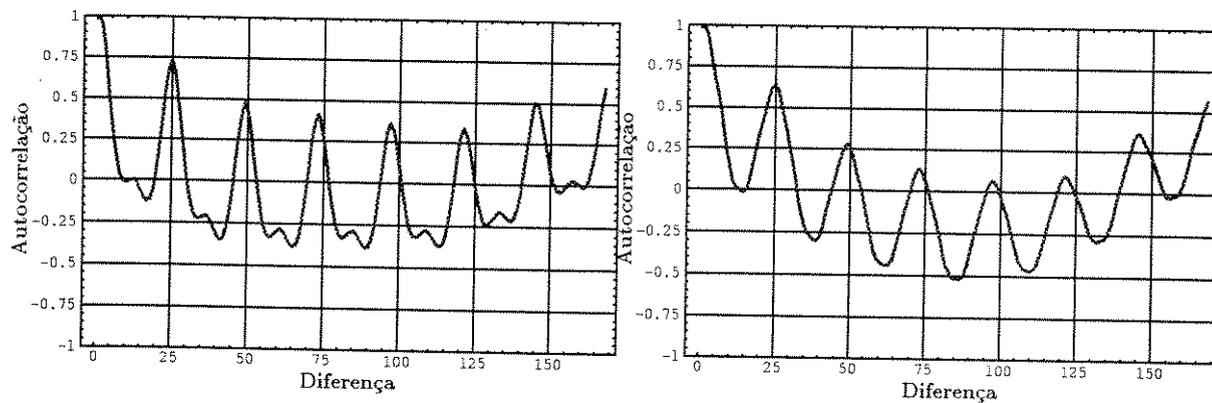


Figura 3.14: Funções de autocorrelação (potência ativa e reativa) na subestação 572

N ^o Subestação	Característica Consumo					
	Residencial	Industrial	Comercial	Rural	Outros	
623	78.6	15.8	3.2	2.0	0.3	Consumidores (%)
NDA	29.1	3.1	45.7	5.9	16.3	Consumo (%)

Tabela 3.8: Subestação com Grande Consumidor Comercial

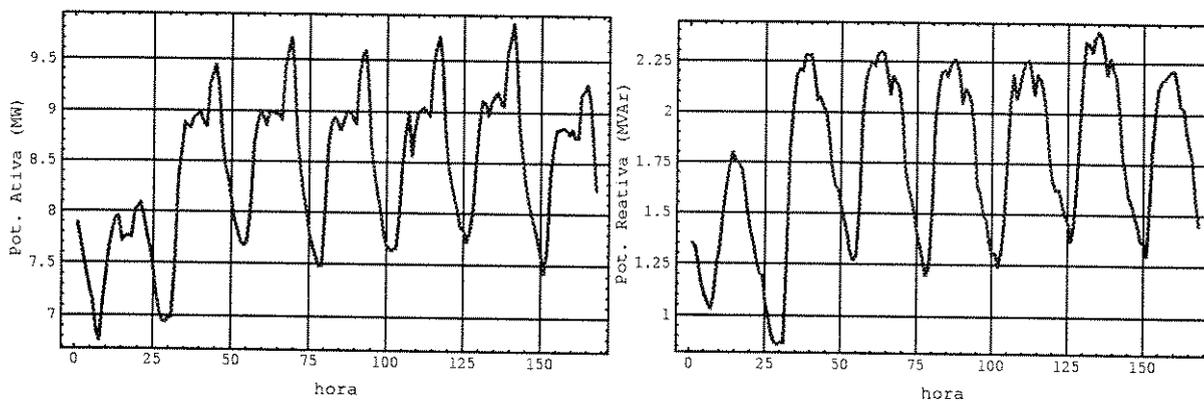


Figura 3.15: Potência ativa e reativa na subestação 623

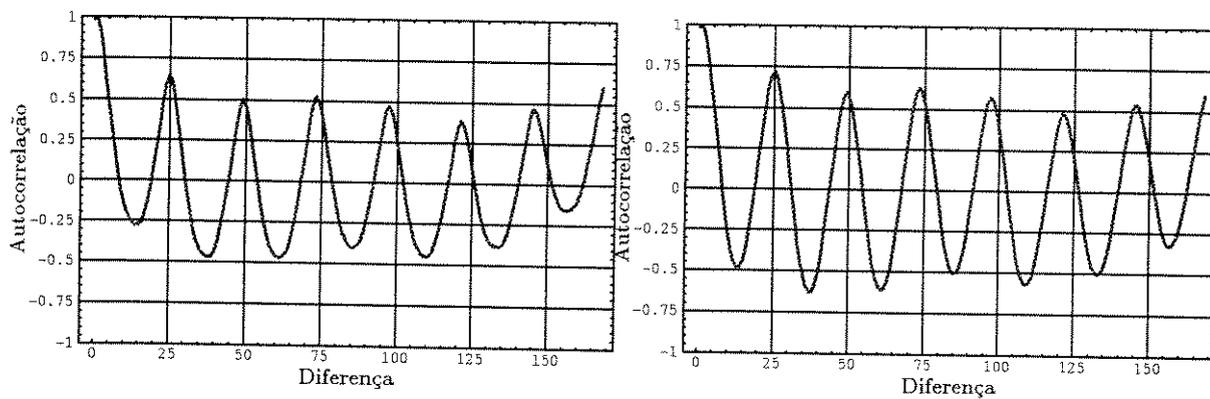


Figura 3.16: Funções de autocorrelação (potência ativa e reativa) na subestação 623

N° Subestação	Característica Consumo					
	Residencial	Industrial	Comercial	Rural	Outros	
668	80.9	2.8	7.5	7.8	1.1	Consumidores (%)
Pinhal	43.6	13.9	13.3	22.1	7.1	Consumo (%)

Tabela 3.9: Subestação com predominância residencial e rural

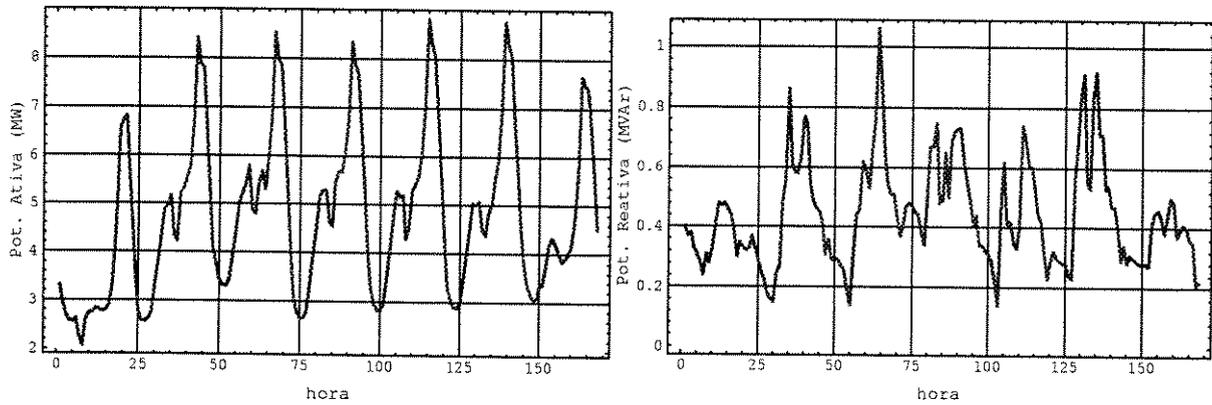


Figura 3.17: Potência ativa e reativa na subestação 668

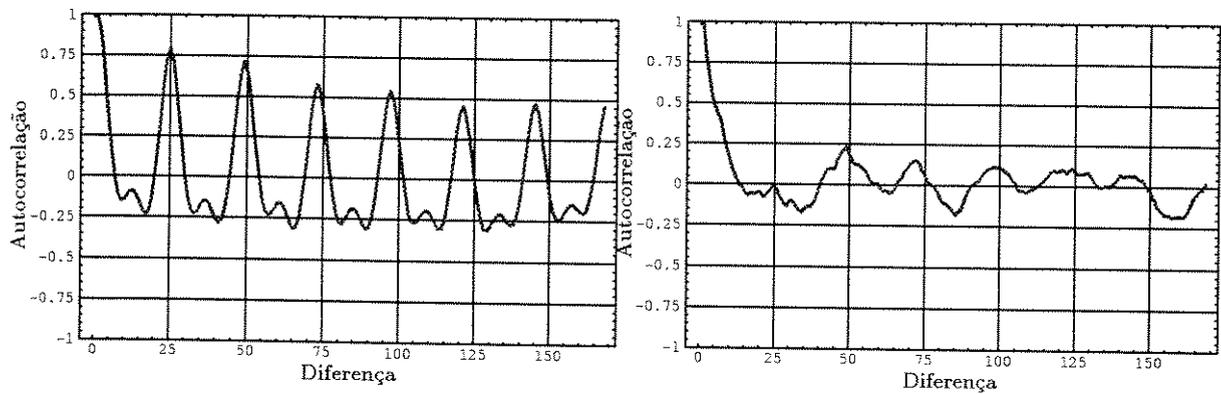


Figura 3.18: Funções de autocorrelação (potência ativa e reativa) na subestação 668

N° Subestação	Característica Consumo					
	Residencial	Industrial	Comercial	Rural	Outros	
762	37.4	10.1	6.5	40.3	5.7	Consumidores (%)
Saltinho	3.2	43.0	29.3	18.7	5.8	Consumo (%)

Tabela 3.10: Subestação com predominância comercial e industrial

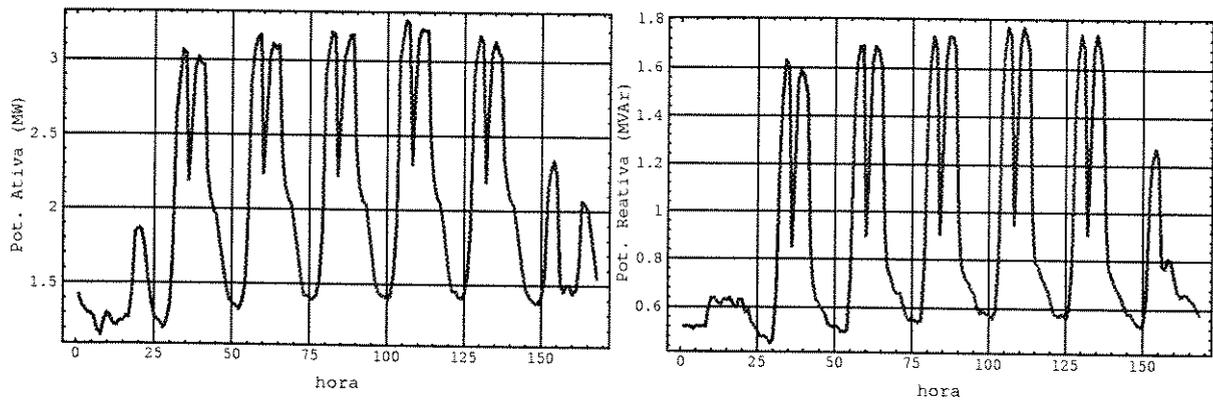


Figura 3.19: Potência ativa e reativa na subestação 762

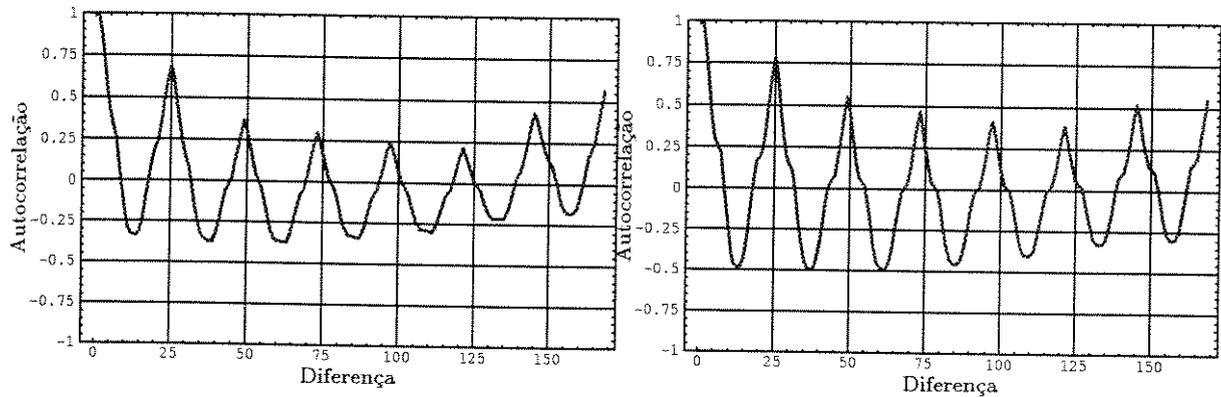


Figura 3.20: Funções de autocorrelação (potência ativa e reativa) na subestação 762

N ^o Subestação	Característica Consumo					
	Residencial	Industrial	Comercial	Rural	Outros	
840	89.9	2.4	5.9	0.1	1.7	Consumidores (%)
Taquaral	35.0	33.3	23.7	0.4	7.5	Consumo (%)

Tabela 3.11: Subestação com carga mista e grandes consumidores comerciais e industriais

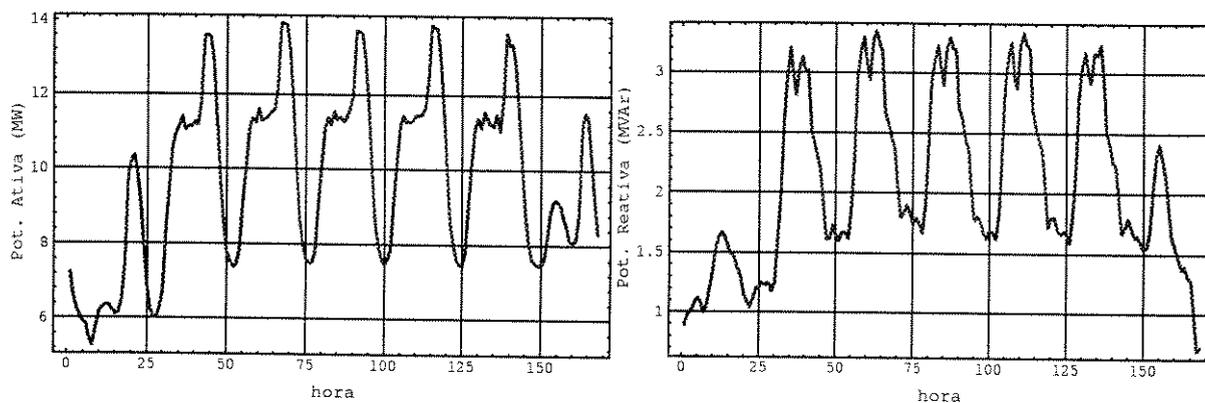


Figura 3.21: Potência ativa e reativa na subestação 840

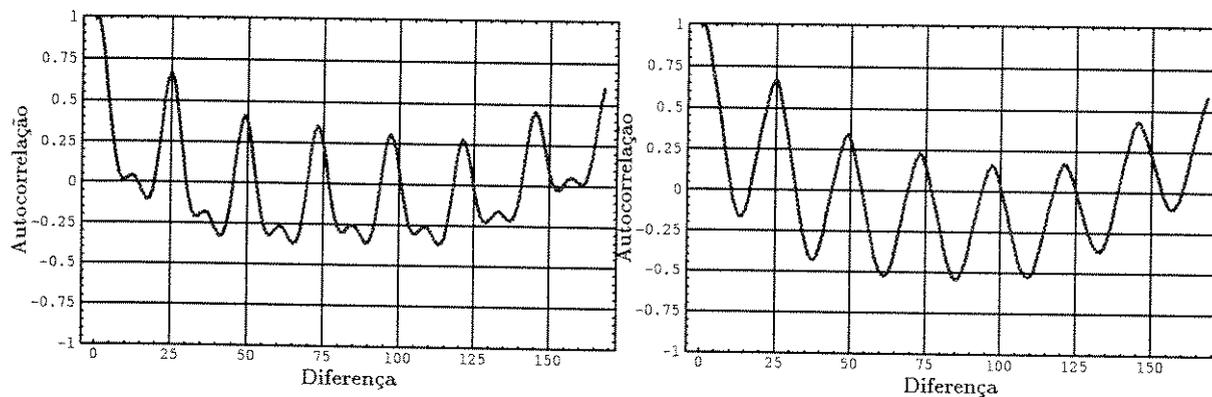


Figura 3.22: Funções de autocorrelação (potência ativa e reativa) na subestação 840

Pode-se notar que o valor da potência reativa está sempre abaixo do valor da potência ativa, pois o valor do fator de potência em geral é alto, e que para algumas subestações a injeção de potência reativa muda de sinal¹. Este fato acrescenta uma dificuldade extra à rede neural, já que ela terá que prever essa mudança.

De acordo com as características expostas sobre o comportamento da potência ativa e reativa, pode-se observar algumas categorias mais marcantes, a saber:

1. Potência ativa e reativa com grau de variação semelhante e lento. Comportamento verificado na subestação 40.
2. Potência ativa com uma variação mais lenta em relação à potência reativa, sendo que esta muda de sinal. Comportamento verificado na subestação 232.
3. Potência ativa e reativa com variação rápida e potência reativa mudando de sinal. Comportamento verificado na subestação 379.
4. As subestações 668 e 80 também apresentam este comportamento de mudança de sinal da potência reativa, mas de uma forma menos periódica que a da subestação 232, não sendo, portanto, visível nos gráficos das curvas de carga típicas dessas subestações.

Para melhor caracterizar o comportamento da carga, as curvas das subestações que representam as categorias acima serão mostradas em maiores detalhes e comparadas com a curva de carga global. Isso é mostrado através das figuras 3.23 a 3.30 que apresentam as curvas de carga ativa e reativa em um intervalo de 4 semanas (20/11/94 a 17/12/94) para as subestações 40, 232 e 379, além da curva de carga global da CPFL. Note que a carga do dia da semana foi separada da carga correspondente ao final de semana, permitindo visualizar melhor os gráficos e mostrar a significativa diferença dos finais de semana, onde se observa uma variação em amplitude menor do que a do dia de semana e uma curva de carga característica possuindo um pico de carga acentuado. Observa-se, ainda que a subestação 379 não apresenta tal característica².

Nas Figuras 3.33 e 3.36 percebe-se uma frequência de oscilação no valor de carga maior do que a apresentada pelas outras subestações ou pela carga global. Ressalta-se, também, a forma suave da curva de carga global e um padrão relativamente constante em relação às cargas das outras subestações. Apesar dessa afirmação ser feita comparando-se apenas a carga de três subestações, verifica-se que a carga de outras subestações apresentam características semelhantes. Nos gráficos de curva de carga para o dia da semana, tem-se que os dias estão ordenados na seguinte seqüência: segunda, terça, quarta, quinta e sexta. Já para o final de semana essa ordenação é: domingo e sábado.

¹ Este efeito está relacionado ao uso de capacitores para fazer a compensação de tensão e ocorre, na maioria das vezes de madrugada em horários não fixados.

² É importante ressaltar que, embora as categorias mostradas representem um comportamento comum da carga em subestações, outros comportamentos diferentes também existem.

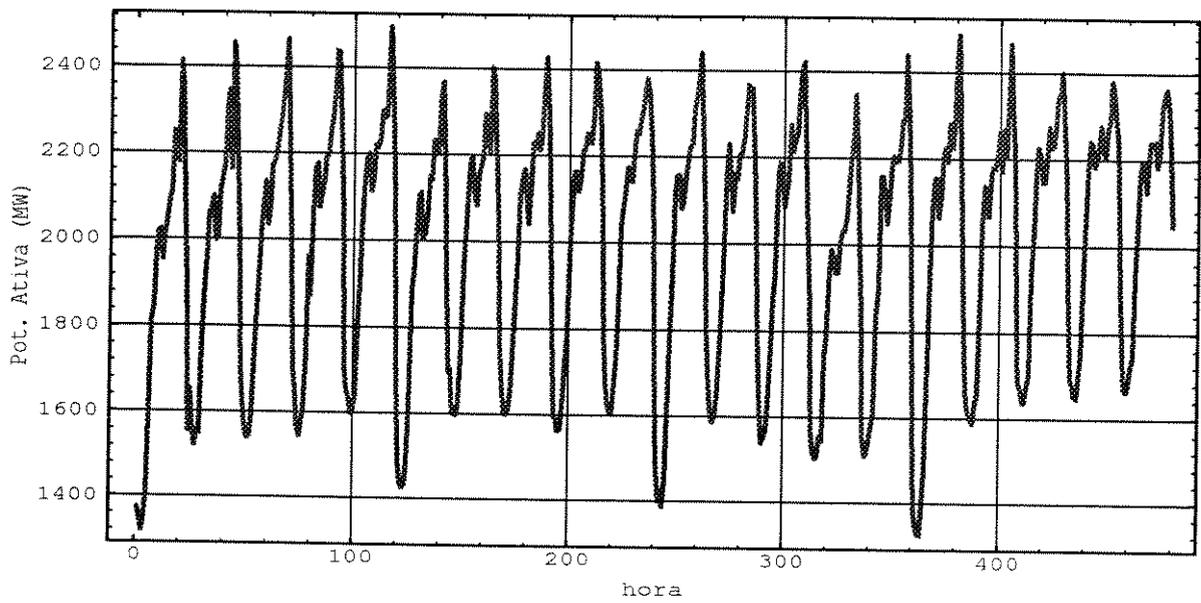


Figura 3.23: Potência ativa global. Dia de Semana

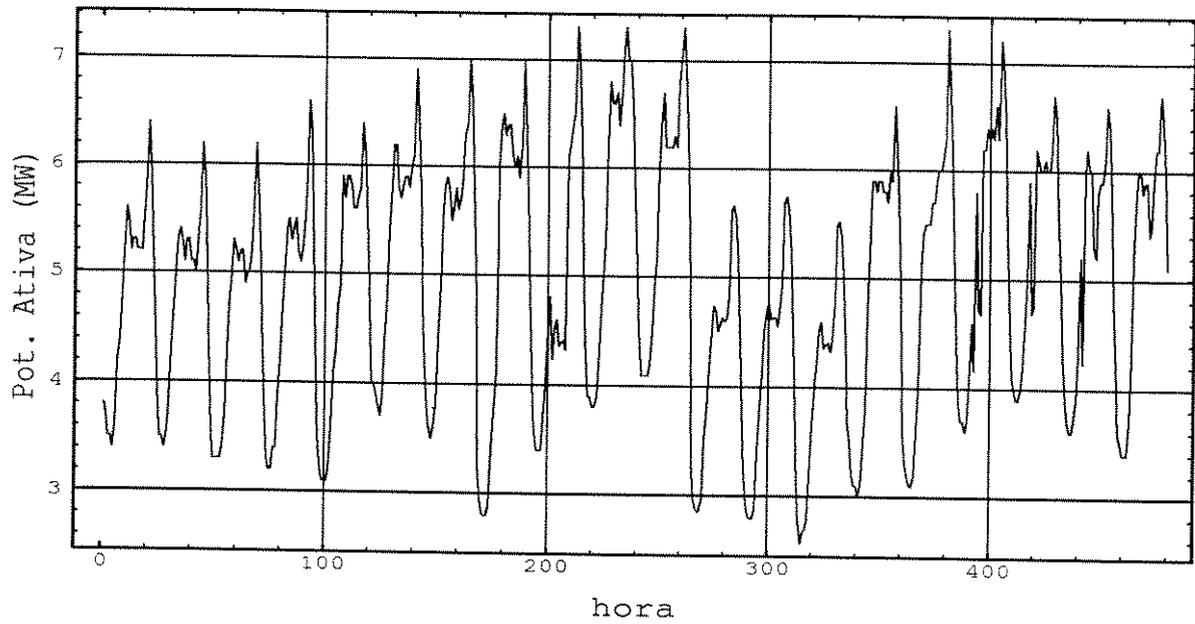


Figura 3.24: Potência ativa subestação 40. Dia de Semana

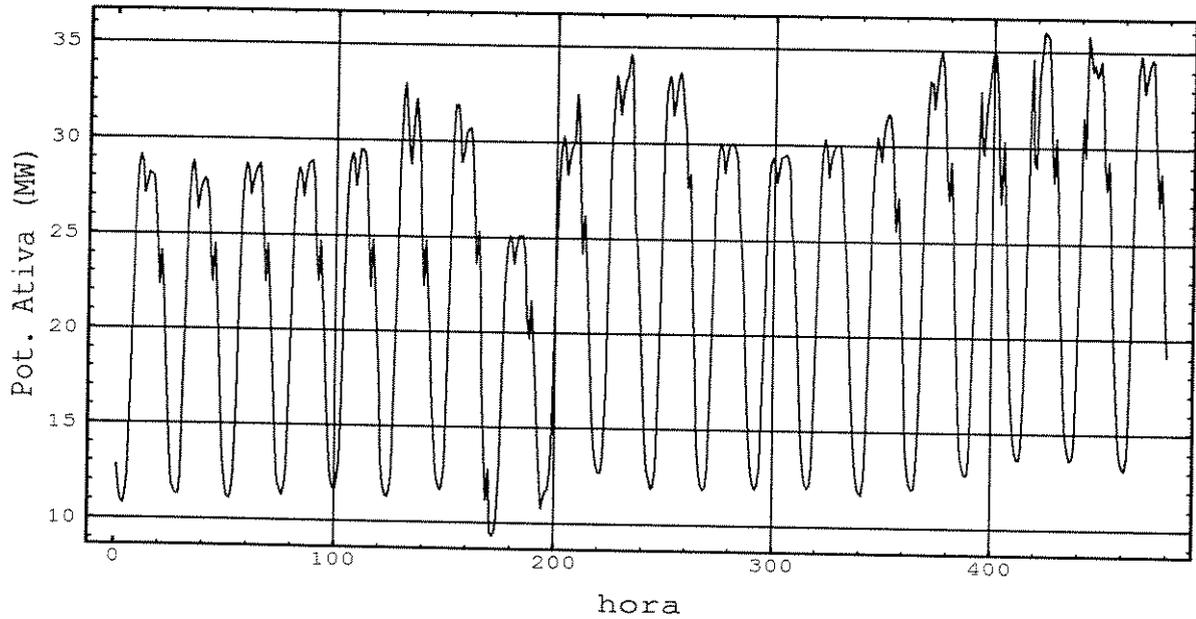


Figura 3.25: Potência ativa subestação 232. Dia de Semana

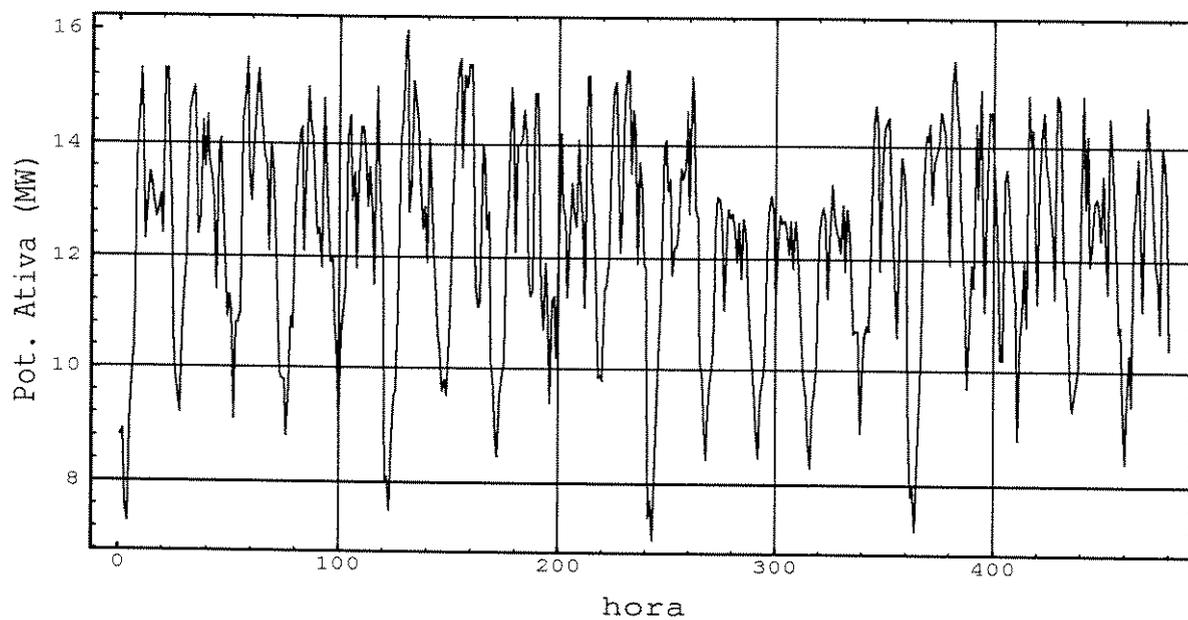


Figura 3.26: Potência ativa subestação 379. Dia de Semana

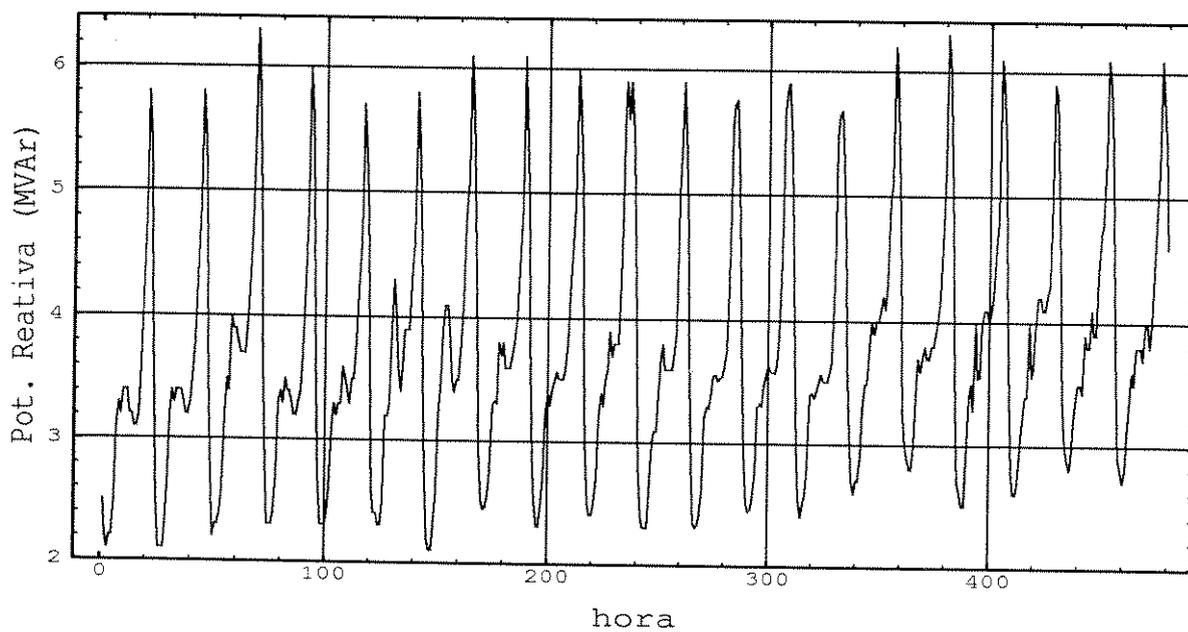


Figura 3.27: Potência reativa subestação 40. Dia de Semana

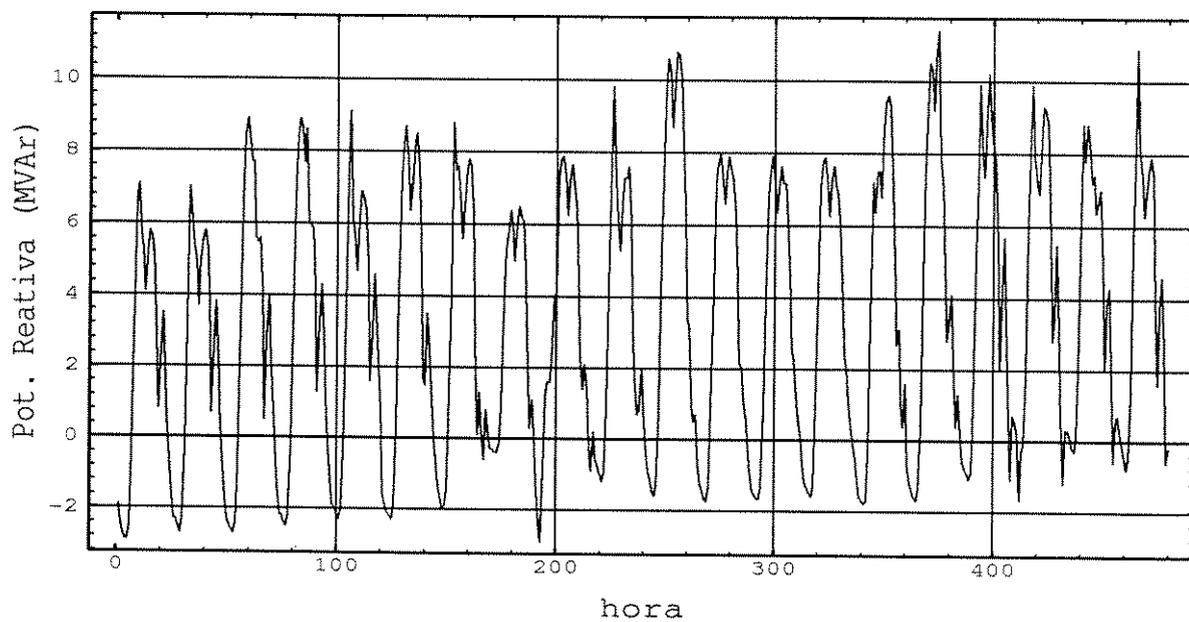


Figura 3.28: Potência reativa subestação 232. Dia de Semana

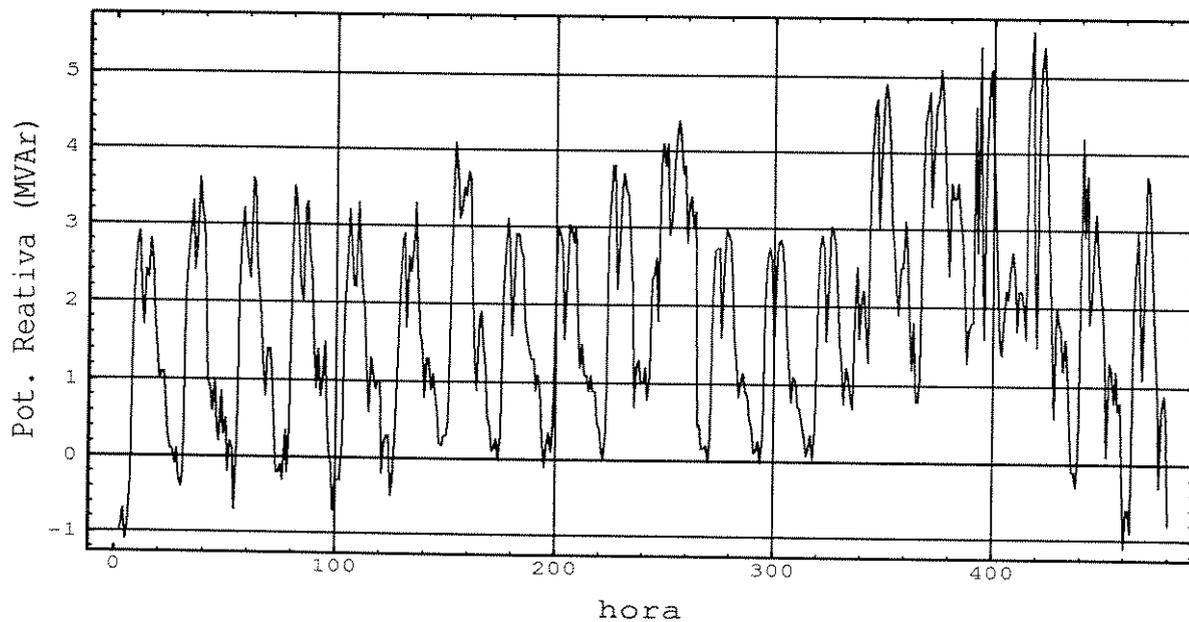


Figura 3.29: Potência reativa subestação 379. Dia de Semana

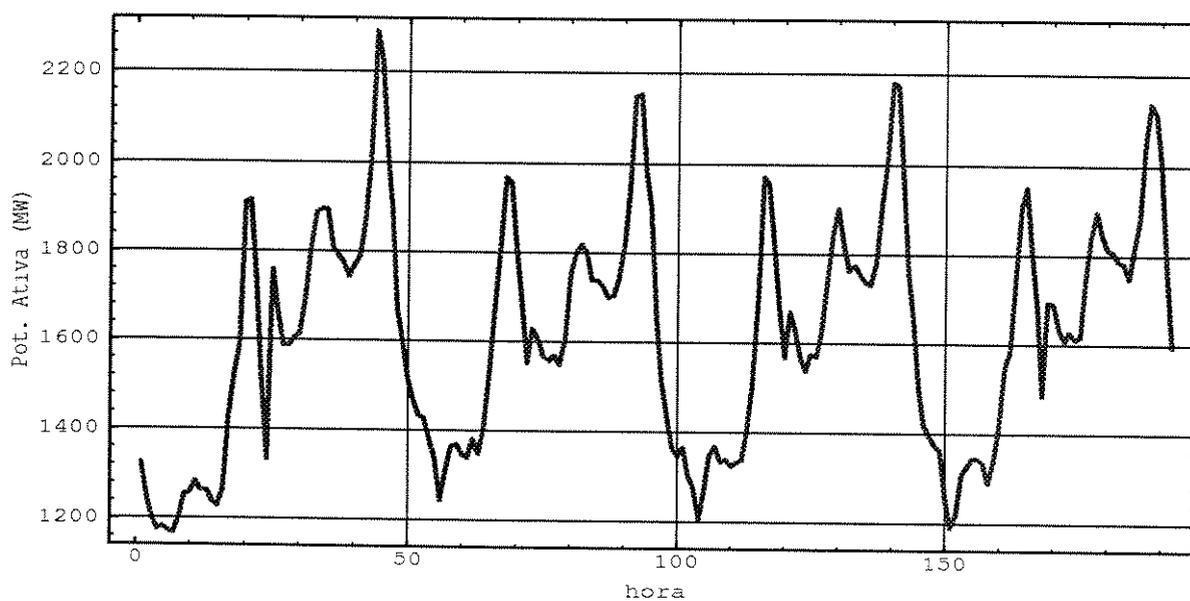


Figura 3.30: Potência ativa global. Fim de Semana

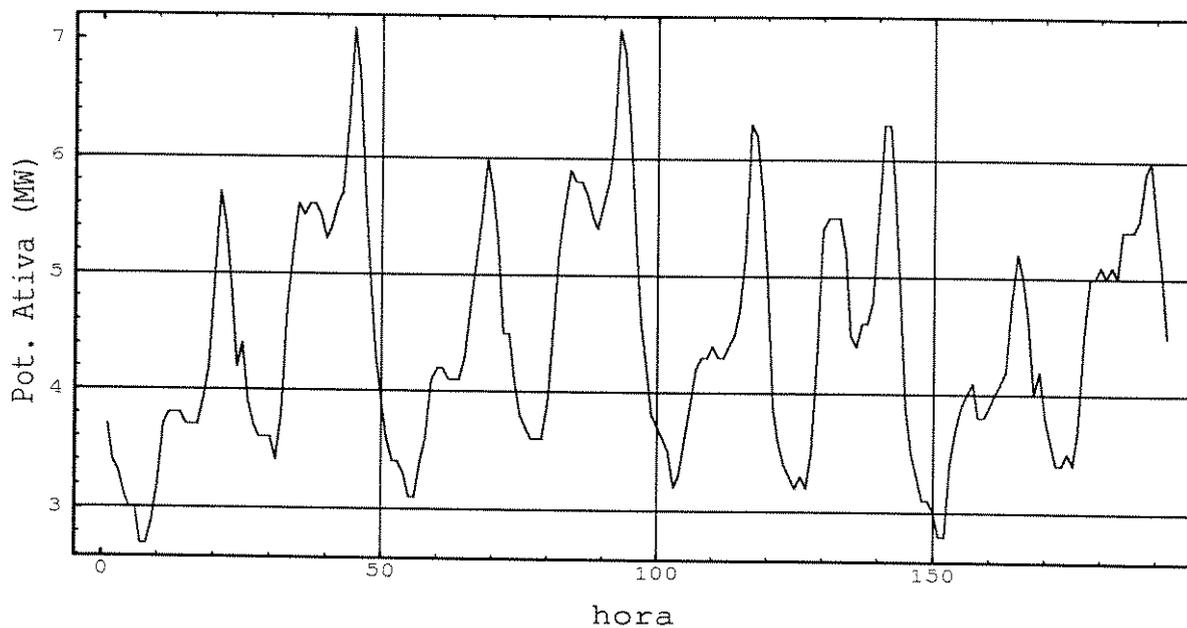


Figura 3.31: Potência ativa subestação 40. Fim de Semana

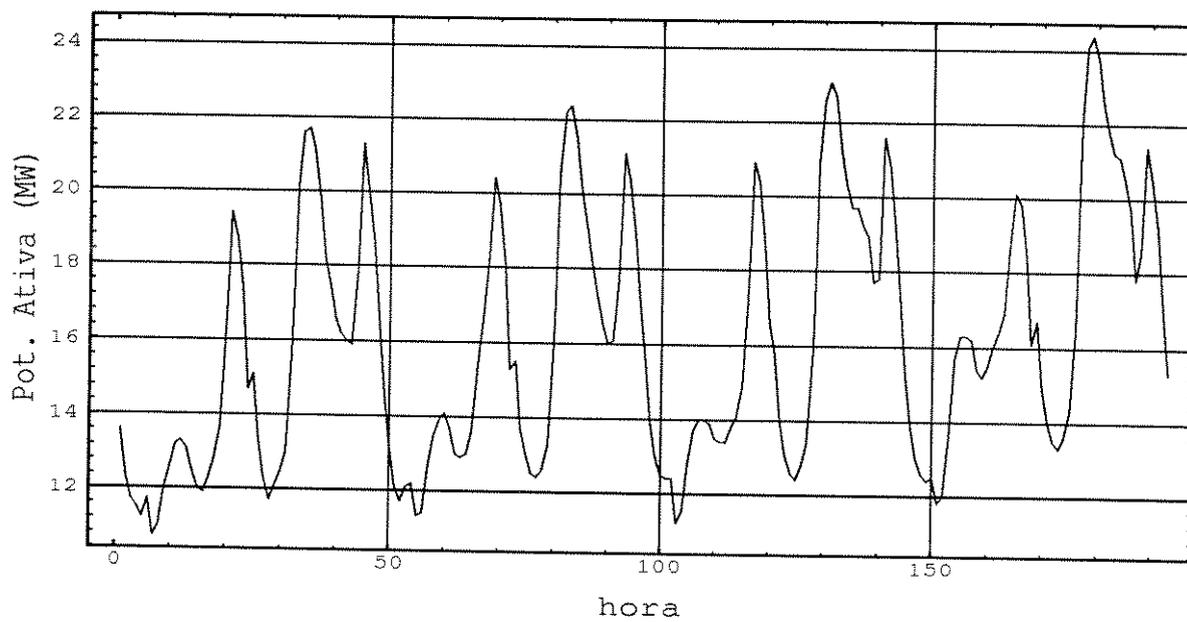


Figura 3.32: Potência ativa subestação 232. Fim de Semana

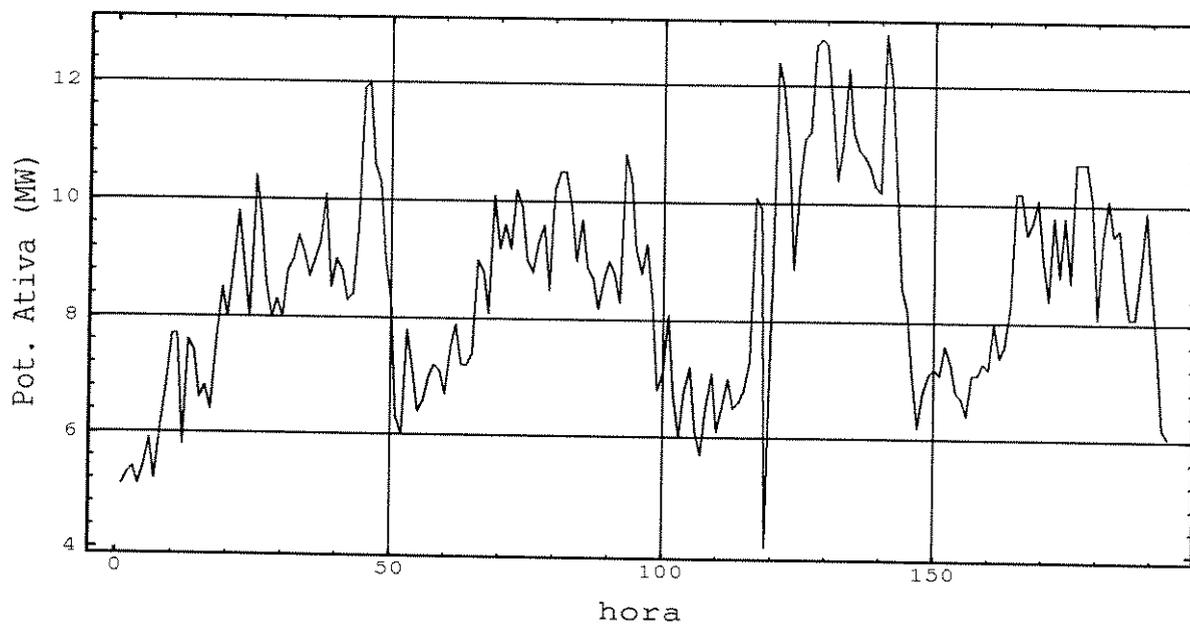


Figura 3.33: Potência ativa subestação 379. Fim de Semana

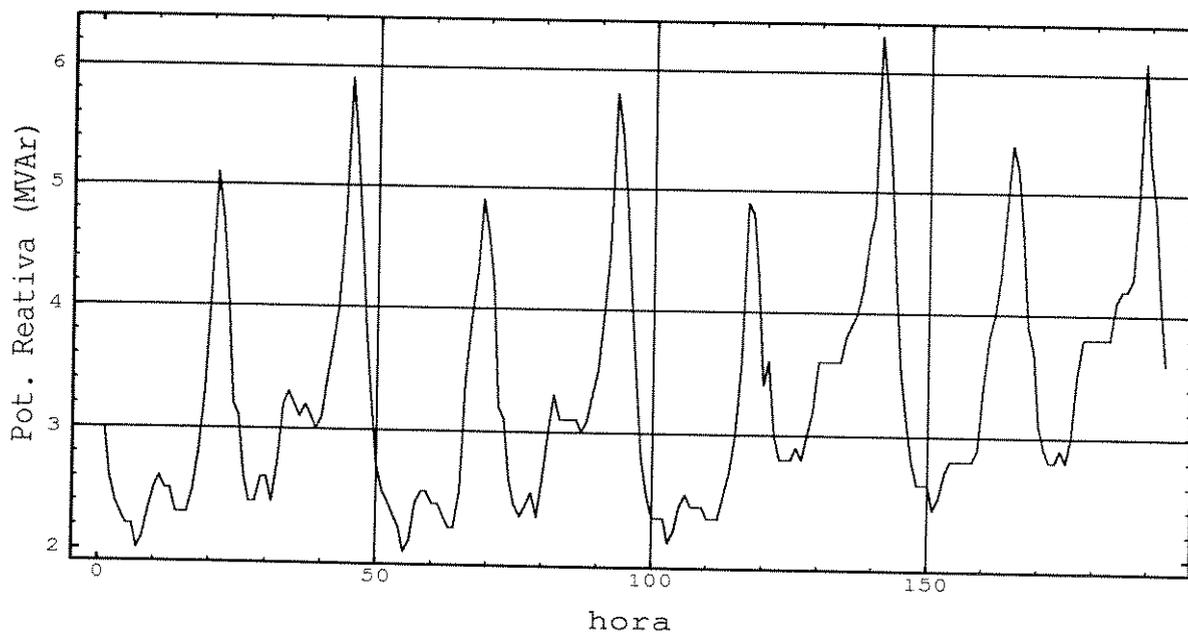


Figura 3.34: Potência reativa subestação 40. Fim de Semana

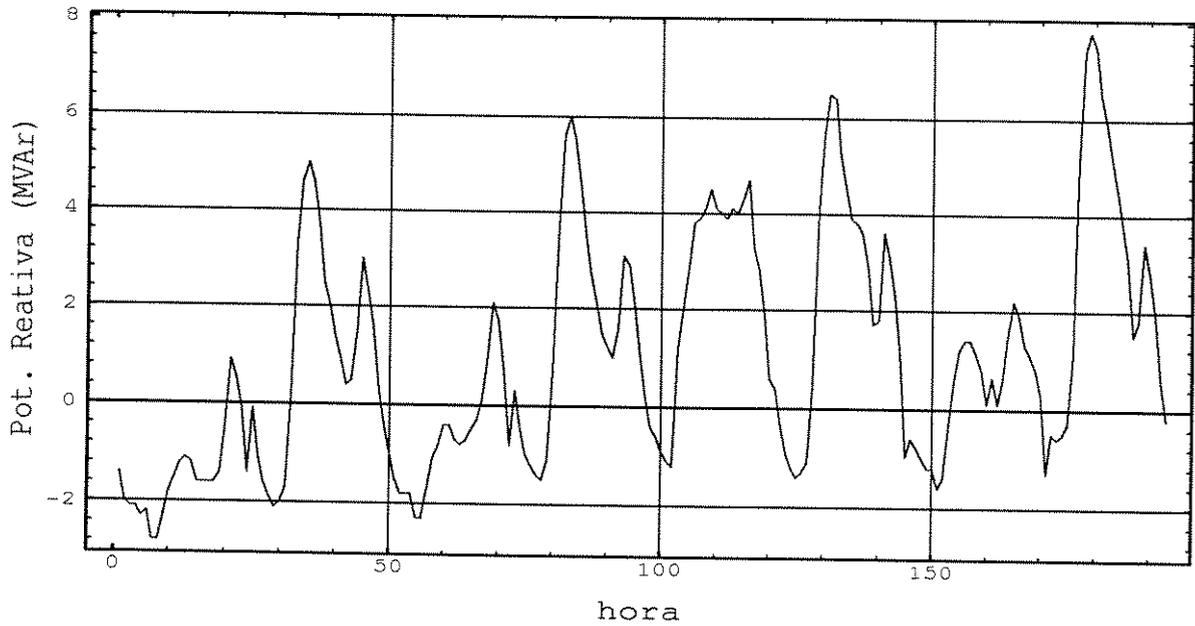


Figura 3.35: Potência reativa subestação 232. Fim de Semana

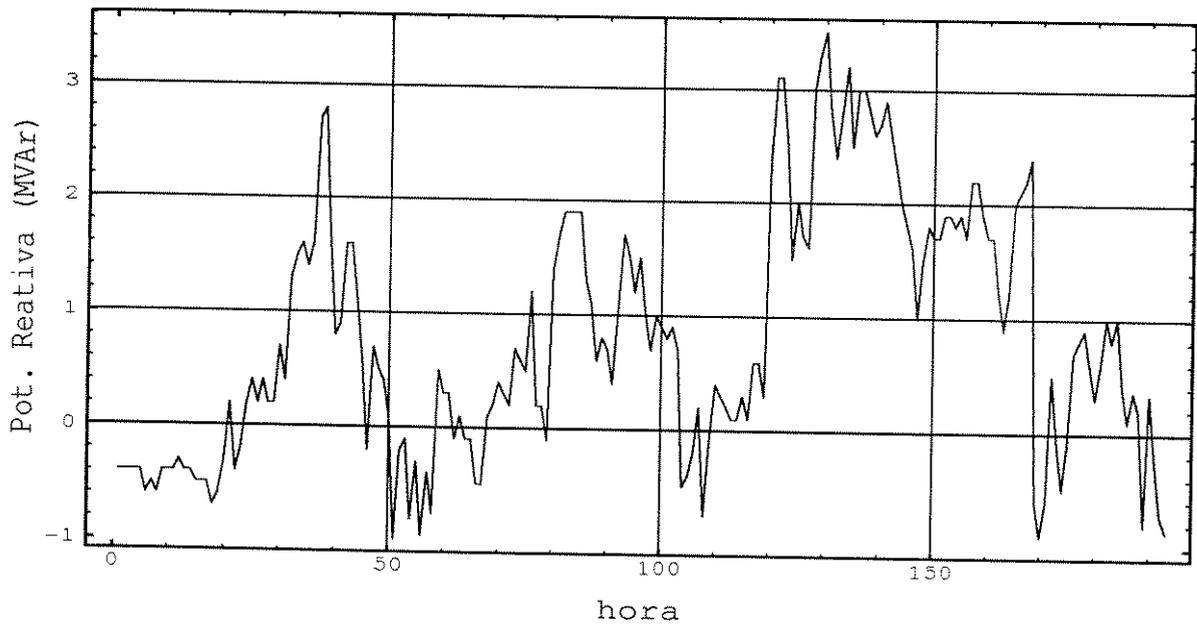


Figura 3.36: Potência reativa subestação 379. Fim de Semana

Capítulo 4

Testes e Resultados

4.1 Considerações Iniciais

Neste capítulo são apresentados resultados de simulação de três modelos de previsão de carga por barramento, todos eles do tipo NARX, sendo que os dois primeiros são implementados através de redes MLP. Ao terceiro modelo é incorporada a capacidade de representar o tempo (memória) e portanto as propriedades dinâmicas de um sinal, sendo esse modelo implementado por uma rede FIR multicamada. Esses modelos foram escolhidos pelo bom desempenho verificado em outras aplicações e pelo nível de adequação ao problema aqui abordado.

O primeiro modelo, que receberá a designação de M, foi baseado no algoritmo proposto por [Chen, 1992] e [Kiartzis, 1994], o segundo modelo, designado por L, é uma adaptação do algoritmo proposto por [Lee, 1992] e o terceiro modelo, designado por F, é uma aplicação de redes FIR [Wan, 1993b] ao problema da previsão de carga. As redes FIR têm características promissoras para a previsão de carga, principalmente da potência reativa (variação mais aleatória), tendo em vista seu extraordinário desempenho na predição não linear de séries temporais caóticas (pulsção de laser) [Wan, 1993a]. Para os três modelos foram estudadas as vantagens de se utilizar particionamento segundo um dos padrões apresentados a seguir:

1. Nenhum particionamento. Uma única rede é utilizada para representar a carga durante a semana.
2. Particionamento em dois grupos: fim de semana ou dia de semana. Uma rede é utilizada para representar a carga do final de semana e outra para a carga de dia de semana.
3. Particionamento em sete grupos, cada um representando um dia da semana. A carga de cada dia da semana é então modelada por uma rede neural.

4. Particionamento com uma rede para determinado grupo de dias da semana. Nesse caso tem-se os seguintes grupos: domingo (grupo1), sábado (grupo 2), segunda (grupo3) e demais dias da semana (terça, quarta, quinta e sexta), designados por dias comerciais no restante deste trabalho. Esse particionamento é uma combinação dos particionamentos 2 e 3, já que os dias comerciais apresentam padrões semelhantes. A segunda-feira não está incluída, por causa do efeito “pós final de semana”. Uma rede é então alocada para cada grupo.

Considerando-se os particionamentos adotados, o universo dos submodelos estudados é o seguinte: M1, M2, M3, L3, L4, F1, F2, F3 e F4. Por associação entre os tipos de modelo e as formas de particionamento tem-se que M1 equivale ao modelo M com particionamento dado pelo item 1, M2 é o modelo M com particionamento dado pelo item 2 e assim por diante. Não foi utilizado o particionamento 4 para o modelo M, nem os particionamentos 1 e 2 para o modelo L, por não serem muito adequados aos modelos. Também não foi considerada a modelagem especial de feriados, devido à dificuldade extra para fazê-la e pela falta de uma base de dados maior para se poder estudar o comportamento da carga durante esses dias. Para cada submodelo são mostrados resultados para diferentes topologias de redes neurais.

Para todos os submodelos a estratégia de aprendizado utilizada foi a de retreinamento (janela de tempo) com um horizonte de previsão de uma hora a frente. Conforme descrito na seção (2.6.3) essa estratégia consiste em adaptar os pesos da rede continuamente, utilizando-se para isso um conjunto de dias ou semanas imediatamente anteriores aos dias de previsão. Os tamanhos das janelas de tempo utilizados neste trabalho são, respectivamente: seis semanas, três semanas e uma semana anterior ao dia de previsão, sendo o retreinamento realizado a cada 24 horas. Para cada submodelo e uma dada janela de tempo são apresentados resultados para três topologias distintas.

4.2 Modelo M

[Chen, 1992] elaborou um modelo para previsão de carga de uma área residencial constituído por um conjunto de rede neurais especializadas em armazenar informações concernentes a hora do dia, dia-da-semana, temperatura, valores passados de carga, ou a um conjunto destes valores. Essas redes neurais são então aglutinadas para formar uma rede neural maior dita rede neural não totalmente conectada, uma vez que os neurônios da camada de entrada não se conectam a todos os neurônios da camada intermediária. As informações sobre a hora do dia e dia da semana são codificadas de forma binária e, em conjunto com medidas de valores horários da temperatura e da carga, são utilizadas para compor o vetor de regressão. Essa rede é então utilizada para fazer a previsão horária de carga.

[Kartzis, 1994] elaborou, por sua vez, um modelo que considerava as mesmas informações (temperatura, dia de semana, hora do dia, valores passados de carga) que [Chen, 1992] para

montar o vetor de entrada da rede. Só que nesse caso o autor modela a carga com uma única rede neural (rede totalmente conectada), ao invés de utilizar um conjunto de estruturas especializadas, além de prever os vinte quatro valores de carga horária de um dia de uma só vez.

Ambos autores apresentam bons resultados utilizando as suas respectivas metodologias. Assim, baseando-se nos fatores mais desejáveis de cada um dos modelos acima, elaborou-se um outro modelo com as seguintes características:

1. Topologia de rede totalmente conectada, por permitir uma implementação e treinamento mais simples .
2. Dia da semana e hora do dia codificados de forma binária.
3. Valores passados de carga horária tomados segundo um modelo sazonal ARI da forma mostrada pela equação 4.1. Essa é a representação adotada por [Chen, 1992]. [Kiarziz, 1994] utiliza um número grande de variáveis representativas dos valores passados da carga, tornando a dimensão do vetor de regressão elevada.

$$\nabla\nabla_{24}\nabla_{168}y(k) = a(k) \quad (4.1)$$

onde $a(k)$ e $y(k)$ representam, respectivamente, ruído branco e carga elétrica. O modelo sazonal acima representa a tendência semanal, diária e horária da carga. Os valores 1 (primeiro ∇), 24 e 168 são escolhidos após análise da função de auto-correlação da carga, mostrando que valores de carga, ocorridos na hora anterior, na mesma hora do dia anterior e há uma semana antes do dia de previsão se relacionam.

As variáveis de entrada descritas nos itens 2 e 3 acima foram adotadas para elaborar o modelo M1. A tabela 4.1 mostra como os elementos da camada de entrada da rede estão organizados. Note que os neurônios de entrada de 13 a 20 representam a equação 4.1. O leitor irá notar que alguns dos elementos ($y(t - 192)$, $y(t - 193)$ e $y(t - 194)$) da equação 4.1 não estão expressos nas variáveis da tabela 4.1, notadamente, os valores $l(i - 8, t)$, $l(i - 8, t - 1)$ e $l(i - 8, t - 2)$. Eles não foram considerados para reduzir a dimensão do vetor, a fim de diminuir o tempo de treinamento.

O modelo M2 é semelhante ao modelo M1 no que se refere às variáveis indicativas de hora e dia da semana. Já com relação aos valores passados de carga, devido à separação entre fim de semana e dias de semana, o modelo sazonal ARI utilizado corresponde à:

$$\nabla^2\nabla_{24}\nabla_{48}y(k) = a(k) \quad (4.2)$$

Nesse caso também abandonou-se os três últimos elementos da equação 4.2. As variáveis de entrada 18, 19 e 20 passam a assumir os valores $l(i - 2, t)$, $l(i - 2, t - 1)$ e $l(i - 2, t - 2)$.

O modelo M3 não contém as variáveis dia-da-semana, pois essa representação passa a não ter sentido, já que está sendo utilizada uma rede para cada dia da semana. Assim, eliminam-se os neurônios de entrada numerados de 6 a 12 na tabela 4.1. O modelo sazonal ARI utilizado também é modificado, sendo representado pela equação 4.3.

$$\nabla^2 \nabla_{168} y(k) = a(k) \quad (4.3)$$

Portanto, eliminam-se também os neurônios de entrada de 15 a 17.

Cabe salientar que separar os padrões em fim-de-semana e dia-de-semana corresponde a dividir a curva semanal da carga em duas. Uma equivalente à curva de carga só para o final de semana e outra equivalente à curva de carga só para dias comerciais. Dessa forma, passa-se a ter dois modelos de previsão diferentes, sendo um para o fim de semana e outro para os dias comerciais da semana e para isso utilizam-se redes neurais diferentes, cada uma modelando um padrão da carga. Esse mesmo raciocínio pode ser aplicado ao considerar cada dia como sendo um padrão diferente. Nesse caso divide-se a curva semanal em sete curvas, tratando-se cada uma com um modelo diferente. Com isso, espera-se que a precisão da previsão melhore, pois cada rede ficaria mais especializada no comportamento característico de cada dia.

4.2.1 Algoritmo de Previsão

A idéia básica do algoritmo apresentado nesta seção é treinar a rede neural com N dias anteriores a um determinado dia de previsão e então utilizá-la, com os pesos estimados segundo o algoritmo de retropropagação (BP), para prever a carga horária de um determinado dia da semana, aqui denominado genericamente de **Dp**. Essa previsão pode ser feita de forma recorrente¹, ou não. No caso da previsão ser feita de modo recorrente, está-se interessado em conhecer o valor da carga m instantes de tempo a frente do instante atual k .

No modelo M a carga $l(i, t)$ do dia de previsão é função de $l(i, t - h)$, $l(i - 1, t - h)$ e $l(i - 7, t - h)$ com $h = 0, 1, 2$ para o caso do dia i ser diferente do dia de previsão e $h = 1, 2$ se o dia de i for o dia de previsão, então para se obter $l(i, k + m)$ é necessário estimar os valores de $l(i, k + m - 1), \dots, l(i, k + 1)$, o que é feito iterando-se o processo de previsão da rede m vezes, ou seja, dado um vetor de regressão \underline{X} , constituído das variáveis mostradas na tabela 4.1, calcula-se um valor estimado $\hat{l}(i, t)$. Esse valor é, então, realimentado de modo a formar um novo vetor $\underline{\hat{X}}$, cujas variáveis 13 e 14 assumem os valores previstos da carga já que o objetivo é obter o valor da carga m instantes de tempo à frente do instante t atual.

¹É necessário esclarecer que utilizar uma rede com recorrência na previsão é diferente de utilizar uma rede recorrente durante a fase de treinamento.

Nó	Descrição						
	Hora						
	1	2	3	...	23	24	
1	0	0	0	...	1	0	
2	0	0	0	...	0	0	
3	0	0	0	...	1	0	
4	0	1	1	...	1	0	
5	1	0	1	...	1	0	
	Dia da Semana						
	D	S	T	Q	Qt	St	Sb
6	1	1	1	1	1	1	1
7	0	1	1	1	1	1	1
8	0	0	1	1	1	1	1
9	0	0	0	1	1	1	1
10	0	0	0	0	1	1	1
11	0	0	0	0	0	1	1
12	0	0	0	0	0	0	1
	Carga						
13	$l(i,t-1)$						
14	$l(i,t-2)$						
15	$l(i-1,t)$						
16	$l(i-1,t-1)$						
17	$l(i-1,t-2)$						
18	$l(i-7,t)$						
19	$l(i-7,t-1)$						
20	$l(i-7,t-2)$						
	i é o dia da previsão t é a hora da previsão						
21	valor de polarização da rede						

Tabela 4.1: Variáveis de entrada utilizadas no modelo M1 e M2

O processo todo é ilustrado na figura 4.1.

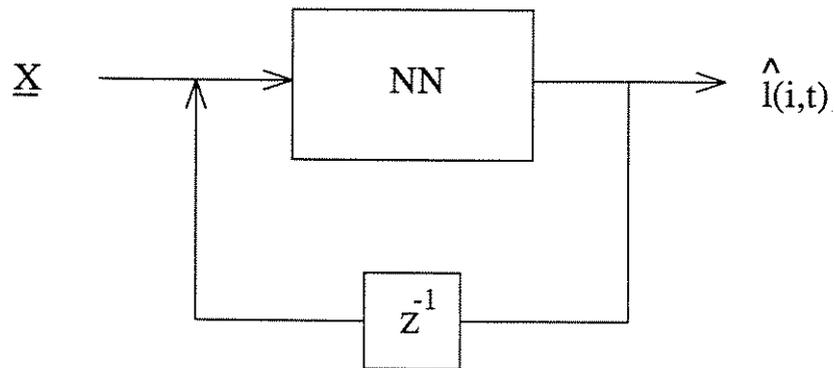


Figura 4.1: Rede neural recorrente

Contudo, se a previsão de carga for feita para apenas uma hora à frente, como neste trabalho, então os valores de carga dos neurônios de entrada de 13 a 20 serão fornecidos pelo sistema gerenciador de dados do centro de controle.

A forma pela qual a previsão é realizada segue, assim, o seguinte procedimento: tendo a carga do dia D_p sido prevista e assumindo-se que as vinte e quatro medições de valores horários de carga ativa e reativa para este dia D_p já estão disponíveis, então a base de dados para treinamento é modificada incorporando-se a ela as novas medições de carga do dia D_p , ao mesmo tempo em que os valores de carga do dia D_p-N são descartados, avançando pois a janela de tempo. Esse procedimento está ilustrado pela figura 4.2.

Conforme descrito no Capítulo 2, acredita-se que essa estratégia de treinamento seja capaz de captar o comportamento não estacionário da carga, já que as medições mais recentes da série temporal são incorporadas à base de dados.

Nesse sentido foram feitos testes a fim de verificar se o processo de treinamento apresentado é viável para a previsão da carga por barramento. Assim elaborou-se o algoritmo 4.1 para implementar o processo acima descrito, o qual é apresentado a seguir:

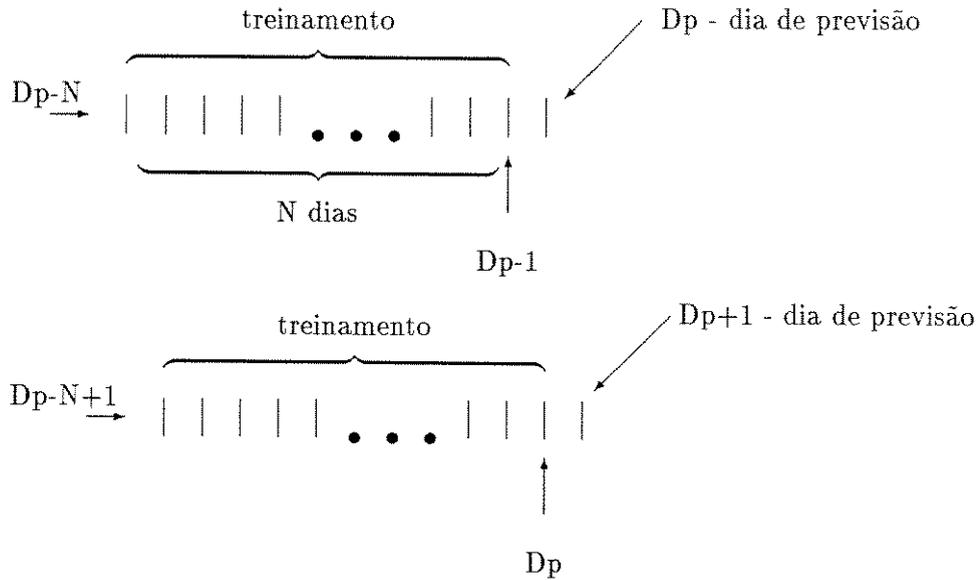


Figura 4.2: Deslocamento da janela de tempo.

Algoritmo 4.1: treinamento e previsão para carga nodal com modelo M1.

1. leitura de parâmetros do algoritmo e da rede neural.
2. cálculo de dia inicial da previsão.
3. montagem da janela de tempo.
4. escolha de uma barra para previsão.
5. leitura de dados para treinamento.
6. treinamento.
7. leitura de dados para previsão.
8. previsão
9. se houve o treinamento para todas as barras

então passe para o item seguinte,
senão volte para o item 4.

10. se todos os dias foram previstos,
então fim.
senão volte para o item 2.

Cabe notar que o algoritmo 4.1 pode ser paralelizado, já que a previsão para uma barra não depende da previsão para outra barra.

O algoritmo 4.1 propõe um processo de treinamento e previsão que, embora simples, tem características globais no sentido de que ele pode ser utilizado independentemente do tipo de particionamento adotado. Na forma em que o algoritmo 4.1 está esboçado, ele é utilizado para representar o modelo M1. Se for utilizado o modelo M2 ou M3, então o algoritmo 4.1 sofre uma modificação, gerando o algoritmo 4.2, o qual difere de 4.1 nos passos 2, 6 e 8. Esse algoritmo é apresentado abaixo:

Algoritmo 4.2: Treinamento e previsão para carga nodal com os modelos M2 e M3.

1. leitura de parâmetros do algoritmo e da rede neural.
2. cálculo de dia inicial da previsão.
3. montagem da janela de tempo, com separação dos grupos segundo particionamento 2 ou 3.
4. escolha de uma barra para previsão.
5. leitura de dados para treinamento.
6. treinamento para o grupo escolhido.
7. leitura de dados para previsão.
8. previsão

9. se houve o treinamento para todas as barras
então passe para o item seguinte,
senão volte para o item 4.

10. se todos os dias foram previstos,
então fim.
senão volte para o item 2.

Resumo dos modelos M1, M2 e M3

Modelo M1: utiliza-se apenas uma rede para todos os dias da semana. Utiliza-se variáveis binárias para representar o dia e a hora. Os valores passados da carga são escolhidos de acordo com a equação 4.1 e o processo de treinamento e previsão é dado pelo algoritmo 4.1.

Modelo M2: utiliza-se uma rede neural para o fim-de-semana e outra para dia-da-semana. As variáveis de entrada são as mesmas do modelo M1, exceto que os valores passados da carga são escolhidos de acordo com a equação 4.2. O treinamento e a previsão são dados pelo algoritmo 4.2.

Modelo M3: utiliza-se uma rede neural para cada dia da semana. Não há variável representativa do dia de semana, apenas da hora e os valores passados da carga são escolhidos segundo a equação 4.3. O treinamento e a previsão são dados pelo algoritmo 4.2.

4.3 Modelo L

Uma vez que os dados disponíveis de medição incluem apenas os valores da carga elétrica, data e hora, procurou-se um outro algoritmo na literatura, usando redes MLP, que pudesse fazer a identificação do modelo de carga elétrica utilizando apenas observações passadas sem desconsiderar os efeitos das variáveis climáticas. Um algoritmo desse tipo é descrito por [Lee, 1992], tendo sido parcialmente utilizado por [Kim, 1995]. Neste trabalho foi implementada uma variação do algoritmo descrito por [Lee, 1992]. Cabe notar que além do algoritmo descrito por Lee possuir um tempo de treinamento reduzido, ele também foi proposto para a previsão de carga global.

Para utilizar o algoritmo acima é necessário separar a carga em padrões correspondentes a determinados grupos de dias da semana. A separação sugerida por [Lee, 1992] é uma combinação das categorias de particionamento 2.a e 2.b apresentadas na seção 2.6.1. Os grupos de dias por ele propostos são os seguintes:

- P1** Carga de domingo, sábado e segunda, tratada como padrões diferentes, ou seja, um modelo de previsão para cada um destes dias.
- P2** Carga de terça, quarta, quinta e sexta tratada como um único padrão, ou seja, um único modelo de previsão para o conjunto de dias acima.

O agrupamento acima corresponde a dizer que a carga para um dia D que se deseja prever é semelhante à carga de outros dias que pertencem ao mesmo padrão. Por exemplo, caso se queira a previsão para uma segunda-feira, então, conforme a classificação apresentada, supõe-se que a carga a ser prevista é semelhante à carga verificada em outras segunda-feiras anteriores. Esse fato é verdadeiro considerando-se a hipótese da carga dentro de um grupo de dias da semana ter um comportamento semelhante que se repete ao longo do tempo, podendo haver pequenas variações devidas a algum distúrbio aleatório. A função de autocorrelação mostra isso, pois observa-se que medidas de cargas ocorridas em intervalos de tempo de 168 horas se correlacionam.

Considerando-se o caso da carga por subestação percebe-se que a classificação adotada acima também é aceitável, principalmente em se tratando de subestações com consumo tipicamente residencial e/ou comercial. Para subestações de consumo industrial verifica-se que há uma variação considerável entre as cargas de dois ou mais dias consecutivos pertencentes a um mesmo padrão, sendo este comentário válido também para a carga reativa.

Uma vez que a carga tenha sido subdividida em padrões, o processo de identificação procede da seguinte maneira:

Seja D_p o dia para o qual se deseja fazer a previsão. Então, para um padrão $P2$ tal que, $D_p \in P2$, toma-se para treinamento os dias $D_p - 1$, $D_p - 2$ e $D_p - 3$, sendo que esses dias devem pertencer todos ao mesmo padrão. Se D_p for um padrão pertencente a $P1$, então toma-se para treinamento os seguintes dias $D_p - 7$, $D_p - 14$ e $D_p - 21$. Em ambas situações, o modelo sazonal ARI correspondente é dado pela equação 4.4:

$$\nabla \nabla_{24} \nabla_{48} y(k) = a(k) \quad (4.4)$$

O algoritmo apresentado por [Lee, 1992] faz uma escolha de padrões um pouco diferente daquela aqui apresentada. Para o caso de D_p pertencer a $P2$, o termo ∇_{48} é desconsiderado. Isso não foi feito neste trabalho por considerar-se essa informação importante, já que nenhuma outra variável que identifique dia e hora está sendo utilizada.

O equacionamento do processo de identificação apresentado é implementado pelas equações 4.5 e 4.6.

$$D_p \in P1 \quad D_p - 7 = F(\hat{W}, D_p - 14, D_p - 21) \quad (4.5)$$

$$D_p \in P2 \quad D_p - 1 = F(\hat{W}, D_p - 2, D_p - 3) \quad (4.6)$$

onde F representa a rede neural e W o vetor de pesos.

Nas equações verifica-se que os valores de carga correspondentes aos dias $D_p - 2$, $D_p - 3$ com $D_p \in P2$ e $D_p - 8$, $D_p - 14$ com $D_p \in P1$ são utilizados para estimar o vetor W , tendo como valores alvo as cargas medidas nos dias $D_p - 1$ com $D_p \in P2$ e $D_p - 7$ com $D_p \in P1$. Dessa forma, é possível propor o modelo de carga L4 apresentado pela equação 4.7 abaixo.

$$\begin{aligned} l(i, t) = & F(W(i, t), l(i, t-1), l(i, t-2), \dots, l(i, t-m), \\ & l(i-1, t-1), l(i-1, t-2), \dots, l(i-1, t-m), \\ & \vdots \\ & l(i-n, t-1), l(i-n, t-2), \dots, l(i-n, t-m)) \end{aligned} \quad (4.7)$$

onde $l(i, t)$ indica valor da carga para o dia i na hora t , n indica índice de deslocamento do dia e m indica índice de deslocamento da hora.

Nesse trabalho assumiu-se os valores de $n = m = 2$ independentemente do padrão. Contudo, deve ficar claro que os valores de l diferem dependendo do padrão utilizado, conforme mostram as equações 4.5 e 4.6.

Cabe notar que o vetor $W(i, t)$ para a equação 4.7 é calculado para cada hora do dia. Esse fato representa uma desvantagem desse algoritmo, pois é necessário armazenar um vetor de pesos para cada hora o que pode representar um problema em termos de espaço em disco. Por outro lado, esse método possui um tempo de treinamento muito reduzido, visto que só um padrão é fornecido à rede a cada hora, o que é uma vantagem considerando-se uma utilização "on-line". Para esse método tem-se um algoritmo de treinamento e previsão semelhante ao algoritmo 4.1, porém modificando-se o passo 3. Cabe notar que a implementação feita neste trabalho difere da proposta por [Lee, 1992] no aspecto de que a previsão de carga aqui adotada é feita sequencialmente usando a carga medida nos instantes anteriores e não a carga prevista pela ANN, como faz Lee. Para finalidade de simulação esse novo algoritmo será identificado como sendo o algoritmo 4.3, mostrado abaixo:

Algoritmo 4.3: Treinamento e previsão para carga nodal com os modelos L3 e L4.

1. leitura de parâmetros do algoritmo e da rede neural.
2. cálculo de dia inicial da previsão.
3. montagem da janela de tempo, com separação dos grupos segundo particionamento 4.
4. escolha de uma barra para previsão.
5. leitura de dados para treinamento.
6. realiza treinamento para o grupo escolhido.
7. leitura de dados para previsão.
8. previsão
9. se houve o treinamento para todas as barras
então passe para o item seguinte,
senão volte para o item 4.
10. se todos os dias foram previstos,
então fim.
senão volte para o item 2.

Ao invés de se escolher os dias utilizados para a obtenção do vetor de pesos W , segundo os padrões estabelecidos no algoritmo 4.3, pode-se obter os dias de treinamento segundo outra classificação, na qual os dias não são mais agrupados de acordo com o período da semana ao qual pertencem, mas sim de acordo com o dia da semana, ou seja, de acordo com o particionamento 3. Logo se D_p é o dia da previsão, então os dias escolhidos para treinamento serão os seguintes: $D_p - 7$, $D_p - 14$ e $D_p - 21$. Esse é o modelo L3, cujo processo de treinamento é descrito somente pela equação 4.5.

Resumo dos modelos L3 e L4

Modelo L3: os dias utilizados para treinamento são agrupados segundo os dias da semana. O modelo é descrito pela equação 4.7 ($n = m = 2$) com processo de treinamento dado pela equação 4.5.

Modelo L4: há quatro grupos: sábado, domingo, segunda e demais dias da semana. O modelo também é descrito pela equação 4.7 ($n = m = 2$) com o treinamento realizado segundo as equações 4.5 ou 4.6 conforme o grupo ao qual o dia de treinamento pertença.

4.4 Modelo F

Ambos os modelos M e L utilizam redes MLP para identificar o comportamento da carga, as quais são redes bastante poderosas, além de serem aproximadores universais para funções contínuas em um dado intervalo conforme descrito no Capítulo 2. Porém cabe lembrar que o mapeamento realizado por redes tipo MLP é estático, ou seja, a rede tem uma estrutura de processamento estática. Essa característica permite uma boa representação de padrões que possuem variação espacial. Elas também podem ser utilizadas para realizar um preditor não linear de uma série temporal. Entretanto, a utilização de um tipo de rede que acrescente o fator tempo em sua representação interna é uma característica desejável, muito embora isso não implique necessariamente em melhor eficiência de previsão em relação às redes MLP. Como foi visto no Capítulo 2, as redes capazes de representar a dinâmica temporal de uma série apresentam recorrência ou filtros temporais em suas sinapses. Neste estudo optou-se por testar redes FIR no problema de previsão de carga por barramento.

Nesta seção é apresentada a implementação do modelo F, o qual é constituído por uma rede tipo FIR. Essa rede foi escolhida em virtude de sua implementação computacional não ser demasiado complexa e por apresentar bons resultados na previsão de séries temporais caóticas, conforme se verifica nos anais do “The Santa Fe Institute Time Series Competition” [Wan, 1993a], onde essa rede obteve os melhores resultados na previsão de 100 passos à frente de uma série temporal representando a intensidade de pulsação de um laser NH_3 . Ela também apresenta a vantagem de, escolhida uma topologia adequada de rede para a modelagem de uma determinada série e uma ordem para os filtros FIR nas sinapses, não ser necessário fazer mais qualquer consideração acerca das relações entre a carga prevista e valores passados de carga. Cabe notar que redes FIR implementam modelos NAR.

A forma pela qual se determina a topologia da rede e a ordem das sinapses é obtida através do método de validação cruzada, sendo que é possível utilizar a função de autocorrelação para

determinar a ordem do filtro nas sinapses que conectam a camada de entrada à primeira camada intermediária [Wan, 1993b].

Para esse modelo os quatro tipos de particionamento apresentados na seção 4.1 foram implementados. Utilizou-se neste estudo uma rede de três camadas. A ordem dos filtros nas sinapses da primeira camada intermediária foi fixada em 20 e a ordem dos filtros nas sinapses da camada de saída foi fixada em 4 para todos os quatro tipos de particionamento. Esses valores foram escolhidos heurísticamente de forma a captar o ciclo diário da carga. Ao observar a função de auto-correlação verifica-se também que valores de carga espaçados de vinte e quatro horas se correlacionam. Cabe salientar que para esse tipo de rede não se constrói um vetor de regressão como é feito para as redes MLP, já que a estrutura dessa rede, conforme definida por [Wan, 1993b], possui um único neurônio de entrada, cuja variável de entrada é um vetor composto por elementos representado p valores de carga ($l(i, t - 1), \dots, l(i, t - p)$), como mostram as equações 2.37, 2.38, 2.39 e a figura 4.3. Na figura as posições de 1 a 20 são ocupadas pelos valores passados da carga. A posição 21 existe para armazenar o valor $l(i, t - p)$ quando se acrescenta o valor $l(i, t)$ no início do vetor (veja a seção 2.8.4). O modelo de carga representado por uma rede FIR é melhor representado pela equação abaixo:

$$l(i, t + 1) = F(l(i, t), \dots, l(i, t - p)) \quad (4.8)$$

onde F é representada pela rede FIR e p é a ordem do modelo NAR (nesse trabalho seu valor é 20).

Os algoritmos de treinamento e previsão são os mesmos que para o modelo M e L, ou seja, são os algoritmos 4.1, 4.2 e 4.3. As janelas de tempo consideradas neste modelo são as mesmas consideradas para o modelo M e o procedimento de treinamento e previsão descrito para o modelo M e mostrado na figura 4.2 permanece válido.

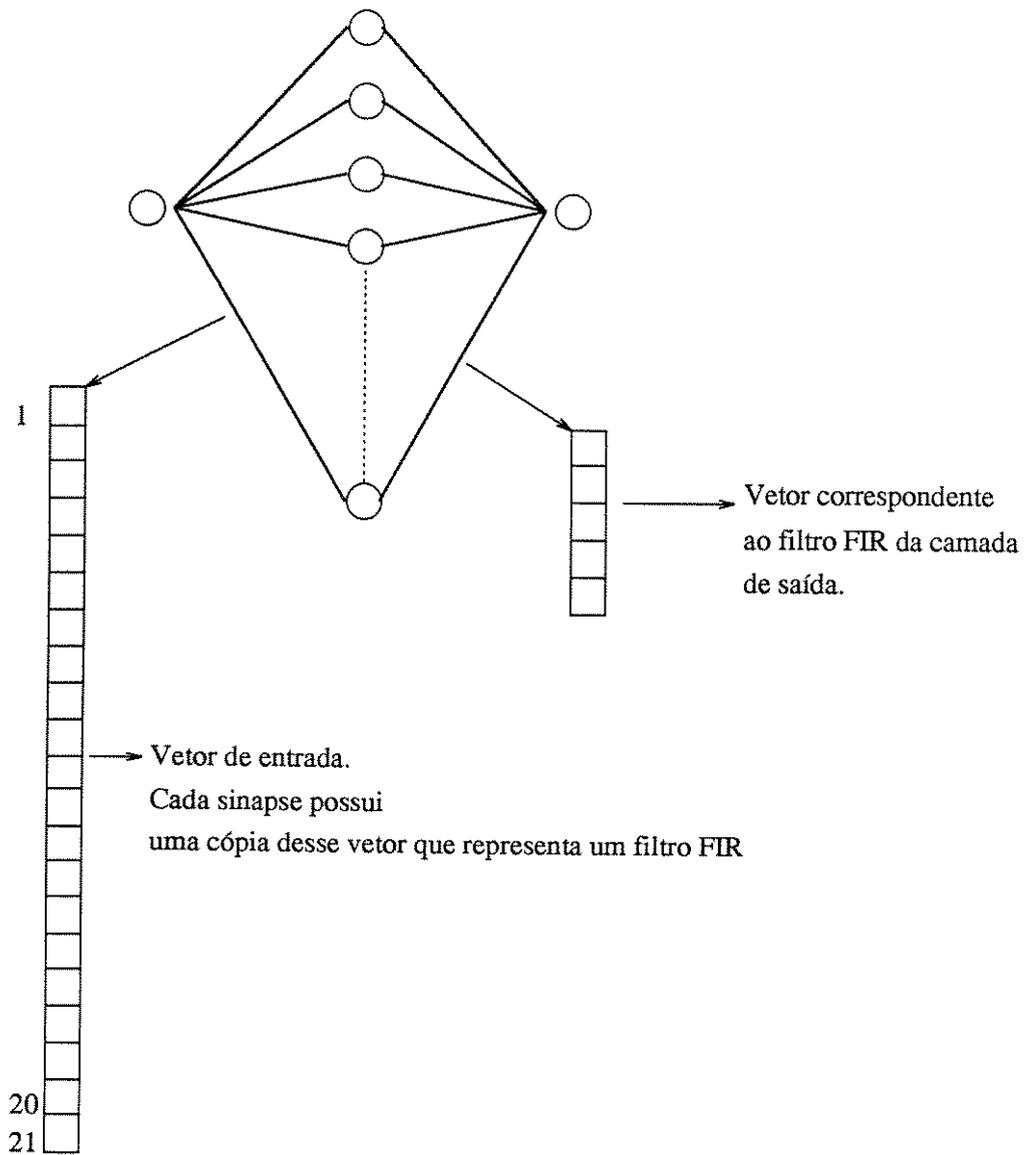


Figura 4.3: Diagrama esquematizado da rede FIR utilizada para previsão de carga

Resumo dos modelos F1, F2, F3 e F4

Modelo F1: utiliza uma só rede para qualquer dia da semana.

Modelo F2: utiliza uma rede para o fim-de-semana e outra para o dia-de-semana.

Modelo F3: utiliza uma rede para cada dia da semana.

Modelo F4: utiliza uma rede para sábado, domingo, e a segunda e uma outra para os demais dias da semana.

OBS: devido as características dessa rede o vetor de entrada consiste apenas de p valores passados de carga, não havendo a necessidade de se representar explicitamente os dias da semana.

4.5 Descrição dos Resultados

Devido à grande quantidade de dados e de simulações o resultado comparativo do desempenho das redes neurais na previsão de carga nodal ativa e reativa obtidos com vários particionamentos e topologias serão apresentados em tabelas, com sua eficiência medida através dos seguintes valores estatísticos:

1. Coeficiente de Variação (CV) (sugerida em [Kreider, 1994]): esta medida representa o desvio padrão dividido pela média do valor de carga real observado no intervalo de teste. Indica o espalhamento do erro. Assim, quanto mais próximo de zero melhor o resultado.

$$CV = \frac{1}{\bar{y}} \left[\frac{1}{N} \sum_{t=1}^N (\hat{y}(t) - y(t))^2 \right]^{1/2} \quad (4.9)$$

2. Erro Máximo Percentual (E_{max}(pu)): Indica qual o máximo erro de previsão obtido. Esta medida será apresentada como percentual da carga base correspondente.

$$E_{max} = \frac{100}{y_{base}} \text{Max} \left(\sum_{t=1}^N |\hat{y}(t) - y(t)| \right) \quad (4.10)$$

3. Erro Médio Percentual (E_{med}(pu)): Valor médio do erro. Indica o valor mais provável do erro. Essa medida também será apresentada como percentual da carga base correspondente.

$$E_{med} = \frac{100}{y_{base}} \frac{1}{N} \sum_{t=1}^N |\hat{y}(t) - y(t)| \quad (4.11)$$

4. Erro absoluto (%): Valor absoluto de erro em percentagem do valor base da carga.

$$E = \frac{100}{y_{base}} |\hat{y}(t) - y(t)| \quad (4.12)$$

onde $y(t)$ é o valor real da série, $\hat{y}(t)$ é o valor previsto da série, N é o número total de horas (descontados feriados e valores espúrios) para os quais foram feitas simulações e $\bar{y}(t)$ é a média do valor medido.

No intuito de normalizar os valores de carga os valores de y_{base} correspondem ao maior valor inteiro da carga máxima verificada em todo conjunto de medidas disponíveis para cada subestação. As potências ativa e reativa de cada subestação possuem valores base diferentes, dados pela tabela 4.2.

Valor Base	Subestação										
	40	80	221	232	379	463	572	623	668	762	840
Pot. Ativa (MW)	9	21	13	38	18	23	24	13	30	5	18
Pot. Reativa (MVAR)	8	7	5	12	6	9	10	4	13	3	7

Tabela 4.2: Valores base para potência ativa e reativa.

As medidas de desempenho acima foram utilizadas para analisar a performance da previsão, para todas as subestações apresentadas no capítulo anterior. Por uma limitação de espaço físico de máquina e pelo elevado tempo de simulação, as simulações para as subestações 80, 221, 463, 572, 623, 668, 762 e 840 foram realizadas para o período de 31/07/94 a 04/02/95, correspondendo a 4536 valores horários. Para as subestações 40, 233 e 379 foram realizadas simulações para o período de 31/07/94 à 24/06/95, correspondendo a 7896 valores horários.

Nas próximas seções serão apresentadas tabelas com os resultados de simulação para os modelos testados em termos das medidas estatísticas apresentadas acima, além de alguns comentários pertinentes. Nas tabelas os melhores resultados estão enfatizados em asterisco. Para todos os modelos, três tamanhos de conjunto de treinamento foram utilizados: S_1 , S_2 e S_3 , como são mostrado abaixo. Os dias correspondem àqueles imediatamente anteriores ao dia de previsão.

- $S_1 = 7$ dias
- $S_2 = 21$ dias

- $S_3 = 42 \text{ dias}$

Além disso, testou-se três diferentes topologias de rede com números variáveis de neurônios na camada intermediária, denotadas por H_1 , H_2 e H_3 . Assim, por exemplo, a expressão $H_1 = 21 : 20 : 1$ utilizada nas próximas seções representa uma rede com 21 neurônios na camada de entrada, 20 neurônios na camada intermediária e 1 neurônio na camada de saída. O neurônio correspondente ao termo de polarização foi omitido na descrição das topologias.

Nas tabelas as informações estão apresentadas na seguinte ordem. Cada coluna representa uma subestação. Para cada subestação são mostradas as estatísticas de erro (CV, Emed e Emax) de previsão que correspondem a uma determinada topologia H , que, por sua vez, corresponde a um tamanho de conjunto de treinamento S . Ao lado do número da subestação se encontram as letras I (industrial), R (residencial) e C (comercial) que indicam o consumo de maior predominância naquela subestação.

Ao final das seções sobre o resultado dos modelos é apresentada uma tabela com os tempos médios de treinamento e previsão. Os valores de tempo estão em segundos. Todos eles normalizados para uma referência comum de modo a facilitar a comparação, uma vez que se utilizou diferentes tipos de computadores para fazer as simulações. Para se fazer a normalização, executou-se em cada um dos computadores utilizados o mesmo programa de previsão, considerando-se apenas um dia de previsão. O tempo de execução foi então medido e o desempenho relativo dos computadores para o problema em questão verificado, permitindo adotar os critérios de desempenho a seguir:

- Desempenho da *SUN SPARCstation Ultra* em relação à *SUN SPARCstation 20*: aprox 1.8 vezes mais rápida.
- Desempenho da *SUN SPARCstation 20* em relação à *SUN SPARCstation 4*: aprox 1.8 vezes mais rápida.
- Desempenho da *SUN SPARCstation 10* foi semelhante ao da *SUN SPARCstation 4*.

4.5.1 Resultados com Modelo M

As tabelas 4.3 e 4.4 mostram os resultados de simulação para previsão de potência ativa e reativa obtidos com o modelo M1. As topologias das redes utilizadas para os testes desse modelo são mostradas abaixo. O resumo dos modelos M1, M2 e M3 se encontra na página 77.

- $H_1 = 21 : 9 : 1$
- $H_2 = 21 : 20 : 1$

- $H_3 = 21 : 30 : 1$

O critério de parada utilizado foi um erro de 0.025 ou 1000 iterações. Valor de η igual a 0.0085 e α igual a 0.85 em todos os casos.

Os resultados para o modelo M2 são apresentados na tabelas 4.5 e 4.6, utilizando as mesmas topologias do modelo M1, tanto para o particionamento de final de semana quanto para o particionamento para o dia da semana. O critério de parada utilizado foi o mesmo. No modelo M3 também foram utilizadas as mesmas topologias de M1 com um critério de parada igual a 2000 iterações ou um erro de 0.005. O erro é menor em relação aos outros modelos pelo fato de a base de dados para M3 conter menos medidas do que a base dos outros dois modelos. Neste caso é desejável fazer com que hajam mais iterações para permitir uma melhor estimação dos pesos.

No final desta seção são apresentados comentários gerais sobre os resultados, bem como são mostrados os tempos médios de computação por subestação para realizar um ciclo de treinamento seguido de previsão, tomados durante o período de simulação descrito na seção 4.5.

$S_1 = 7 \text{ dias } S_2 = 21 \text{ dias } S_3 = 42 \text{ dias.}$
 $H_1 = 21 : 9 : 1 \ H_2 = 21 : 20 : 1 \ H_3 = 21 : 30 : 1.$
 I (industrial), R (residencial) e C (comercial).

treinamento	topologia	estatística	Subestação										
			40 R	80 C	221 R	232 R	379 I	463 R	572 R	623 C	668 R	762 I	840 R
S_1	H_1	CV	0.00220	0.00299	0.00223	0.00063	0.00172	0.00141	0.00126	0.00137	0.00631	0.00264	0.00364
		Emed(pu)	4.72	5.36	4.35	1.95	6.28	3.09	2.65	3.52	4.66	4.27	4.71
		Emax(pu)	40.44	53.86	28.54	13.53	36.61	20.35	18.50	26.31	37.50	33.20	61.67
	H_2	CV	0.00186	0.00245	0.00168	0.00069	0.00164	0.00149	0.00130	0.00119	0.00427	0.00234	0.00357
		Emed(pu)	4.11	4.73	3.66	2.14	6.06	3.19	2.80	3.37	3.30	4.02	4.65
		Emax(pu)	32.39	43.10	21.31	14.89	34.78	21.52	18.71	21.62	24.87	28.80	64.17
H_3	CV	0.00153	0.00202	0.00135	0.00068	0.00157	0.00123	0.00105	0.00093	0.00397	0.00189	0.00326	
	Emed(pu)	3.79	4.50	2.85	2.11	5.87	2.77	2.35	2.70	3.19	3.30	3.69	
	Emax(pu)	27.00	34.62	17.69	14.66	33.44	17.26	14.88	17.15	22.87	22.60	51.39	
S_2	H_1	CV	0.00123	0.00227	0.00145	0.00049	0.00117	0.00110	0.00114	0.00069	0.00483	0.00184	0.00315
		Emed(pu)	2.82	4.14	3.02	1.60	4.47	2.33	2.36	2.23	3.36	3.35	3.39
		Emax(pu)	22.33	41.81	18.92	10.21	24.33	15.87	16.75	11.00	29.03	21.80	38.33
	H_2	CV	0.00125	0.00190	0.00137	0.00049	0.00139	0.00104	0.00088	0.00093	*0.00408	0.00166	0.00316
		Emed(pu)	3.18	3.93	2.43	1.61	5.41	2.36	1.92	2.40	*3.05	2.96	3.13
		Emax(pu)	21.67	33.81	19.46	10.45	28.22	14.48	12.58	17.85	*23.78	20.00	42.39
H_3	CV	0.00117	0.00226	0.00187	0.00050	0.00176	0.00114	0.00112	0.00155	0.00517	0.00208	0.00326	
	Emed(pu)	3.01	4.69	3.87	1.61	6.42	2.49	2.42	4.13	4.08	3.63	4.06	
	Emax(pu)	19.78	40.00	24.92	10.63	37.83	16.26	16.12	30.08	29.23	24.80	47.83	
S_3	H_1	CV	0.00087	*0.00177	*0.00108	0.00049	*0.00086	0.00091	*0.00082	*0.00084	0.00441	*0.00143	0.00316
		Emed(pu)	2.19	*3.12	*2.32	1.54	*3.55	2.04	*1.77	*2.03	3.14	*2.57	2.89
		Emax(pu)	15.11	*32.43	*13.62	10.58	*20.44	12.87	*11.06	*11.23	26.50	*17.20	53.50
	H_2	CV	*0.00086	0.00160	0.00119	*0.00049	0.00099	*0.00090	0.00086	0.00082	0.00472	0.00143	*0.00406
		Emed(pu)	*2.15	3.47	2.58	*1.54	3.75	*2.02	1.89	2.60	3.64	2.59	*2.80
		Emax(pu)	*14.89	28.10	15.31	*10.50	20.56	*12.70	12.17	14.23	27.67	17.00	*42.67
H_3	CV	0.00099	0.00193	0.00121	0.00053	0.00134	0.00097	0.00090	0.00079	0.00467	0.00160	0.00317	
	Emed(pu)	2.55	4.03	2.67	1.68	4.91	2.16	1.95	2.47	3.17	2.83	2.98	
	Emax(pu)	17.11	34.10	15.15	11.34	28.56	13.91	13.08	13.92	28.40	19.40	45.11	

Tabela 4.3: Tabela comparativa com simulações para o modelo M1. Potência Ativa.

$S_1 = 7 \text{ dias } S_2 = 21 \text{ dias } S_3 = 42 \text{ dias.}$
 $H_1 = 21 : 9 : 1 \ H_2 = 21 : 20 : 1 \ H_3 = 21 : 30 : 1.$
 I (industrial), R (residencial) e C (comercial).

treinamento	topologia	estatística	Subestação											
			40 R	80 C	221 R	232 R	379 I	463 R	572 R	623 C	668 R	762 I	840 R	
S_1	H_1	CV	0.00181	0.00511	0.00218	0.00681	0.01027	0.01079	0.00419	0.00343	0.01081	0.00388	0.00419	
		Emed(pu)	3.45	7.46	7.07	9.44	11.34	9.64	6.96	6.83	5.22	5.25	6.64	
		Emax(pu)	33.63	48.86	40.80	57.83	63.50	100.00	48.10	40.00	33.85	36.33	38.43	
	H_2	CV	0.00162	0.00398	0.00208	0.00573	0.01031	0.01017	0.00263	0.00410	0.01199	0.00320	0.00355	
		Emed(pu)	3.27	5.96	6.83	8.11	11.71	7.89	4.73	8.05	5.06	4.48	5.50	
		Emax(pu)	30.12	37.14	38.20	47.75	62.33	100.00	28.80	49.00	39.77	29.33	32.86	
H_3	CV	0.00171	0.00327	0.00168	0.00562	0.01006	0.00924	0.00220	0.00294	0.00943	0.00265	0.00308		
	Emed(pu)	3.39	4.95	5.57	8.05	11.34	7.37	3.98	6.01	4.21	3.78	4.54		
	Emax(pu)	32.12	31.29	30.80	46.58	60.83	100.00	23.60	34.25	29.92	23.67	29.43		
S_2	H_1	CV	0.00093	0.00248	0.00169	0.00492	0.00682	0.01041	*0.00164	0.00230	0.00748	0.00228	0.00223	
		Emed(pu)	2.35	3.92	5.10	6.99	7.79	6.82	*2.85	4.77	3.38	3.24	3.39	
		Emax(pu)	16.12	22.57	33.20	41.08	40.00	77.78	*18.00	26.00	24.00	20.67	21.00	
	H_2	CV	0.00084	0.00275	0.00174	0.00385	0.00719	0.00925	0.00178	0.00303	0.00835	0.00228	0.00226	
		Emed(pu)	2.16	4.32	5.54	5.49	8.21	6.97	3.19	6.15	3.80	3.26	3.36	
		Emax(pu)	14.38	25.71	32.80	31.92	41.50	100.00	18.80	35.50	26.54	20.67	21.43	
H_3	CV	0.00106	0.00253	0.00171	0.00538	0.01062	0.00944	0.00196	0.00290	0.00934	0.00229	0.00250		
	Emed(pu)	2.58	3.93	5.64	7.75	12.01	7.64	3.41	5.87	4.21	3.29	3.67		
	Emax(pu)	18.50	22.86	31.60	44.42	64.33	100.00	21.10	34.00	29.69	20.33	24.00		
S_3	H_1	CV	*0.00077	*0.00229	*0.00127	0.00369	*0.00627	*0.01158	0.00191	*0.00220	*0.00698	*0.00202	*0.00205	
		Emed(pu)	*1.89	*3.51	*4.12	5.15	*7.11	*6.73	3.24	*2.50	*3.20	*2.92	*3.11	
		Emax(pu)	*13.13	*21.86	*23.20	31.00	*37.88	*100.00	21.60	*26.50	*21.77	*18.38	*19.14	
	H_2	CV	0.00078	0.00247	0.00156	*0.00858	0.00740	0.01276	0.00173	0.00250	0.00865	0.00206	0.00222	
		Emed(pu)	2.02	3.67	5.05	*5.05	8.37	9.13	2.97	5.29	4.21	2.97	3.19	
		Emax(pu)	13.25	23.43	28.60	*30.08	44.50	100.00	19.40	28.00	26.69	18.67	21.14	
H_3	CV	0.00078	0.00284	0.00177	0.00365	0.00711	0.01085	0.00188	0.00291	0.00932	0.00206	0.00223		
	Emed(pu)	2.04	4.43	5.92	5.09	8.03	7.82	3.33	5.94	4.32	3.04	3.32		
	Emax(pu)	13.25	26.86	32.20	30.75	43.17	100.00	20.80	34.00	29.54	18.33	20.86		

Tabela 4.4: Tabela comparativa com simulações para o modelo M1. Potência Reativa.

$S_1 = 7$ dias $S_2 = 21$ dias $S_3 = 42$ dias.
 $H_1 = 21 : 9 : 1$ $H_2 = 21 : 20 : 1$ $H_3 = 21 : 30 : 1$.
 I (industrial), R (residencial) e C (comercial).

treinamento	topologia	esta-tística	Subestação											
			40 R	80 C	221 R	232 R	379 I	463 R	572 R	623 C	668 R	762 I	840 R	
S ₁	H ₁	CV	0.00236	0.00312	0.00228	0.00070	0.00205	0.00166	0.00144	0.00158	0.00577	0.00233	0.00360	
		Emed(pu)	4.64	5.33	4.43	2.15	7.27	3.53	3.14	3.77	4.01	4.14	4.01	4.89
		Emax(pu)	44.11	57.62	30.62	15.26	44.78	24.09	20.63	31.15	34.60	34.60	28.60	58.11
	H ₂	CV	0.00190	0.00249	0.00158	0.00069	0.00165	0.00146	0.00125	0.00111	0.00494	0.00222	0.00336	
		Emed(pu)	4.20	4.54	3.51	2.19	6.21	3.17	2.79	3.07	3.64	3.80	4.25	
		Emax(pu)	34.78	45.62	19.69	14.61	34.94	20.78	17.83	20.23	29.40	27.40	53.17	
H ₃	CV	0.00183	0.00235	0.00162	0.00078	0.00165	0.00141	0.00121	0.00103	0.00487	0.00225	0.00346		
	Emed(pu)	4.03	4.60	3.55	2.50	6.19	3.08	2.68	2.96	3.60	3.78	4.32		
	Emax(pu)	33.67	42.24	20.31	16.37	34.94	20.48	17.29	18.69	28.77	27.80	45.61		
S ₂	H ₁	CV	0.00134	0.00212	0.00121	0.00061	0.00137	0.00111	0.00102	0.00100	0.00475	0.00175	0.00328	
		Emed(pu)	3.06	4.23	2.56	1.88	4.96	2.49	2.28	2.49	3.49	3.11	3.39	
		Emax(pu)	23.89	37.57	15.62	13.37	29.56	15.83	14.42	18.08	28.20	21.20	58.56	
	H ₂	CV	0.00138	0.00212	0.00121	0.00058	0.00134	0.00111	0.00102	0.00100	0.00475	0.00175	0.00328	
		Emed(pu)	3.02	4.23	2.56	1.81	4.80	2.49	2.28	2.49	3.49	3.11	3.39	
		Emax(pu)	24.67	37.57	15.62	12.76	28.72	15.83	14.42	18.08	28.20	21.20	58.56	
H ₃	CV	0.00138	0.00224	*0.00118	0.00062	0.00136	0.00118	0.00102	0.00110	0.00426	0.00173	0.00317		
	Emed(pu)	3.04	4.46	*2.46	1.90	4.99	2.56	2.28	2.77	3.26	3.05	3.30		
	Emax(pu)	25.22	40.10	*15.62	13.45	29.22	17.00	14.62	21.77	24.87	21.00	43.00		
S ₃	H ₁	CV	0.00139	*0.00190	0.00116	0.00056	*0.00122	*0.00101	*0.00093	*0.00068	*0.00390	*0.00154	*0.00310	
		Emed(pu)	2.78	*3.84	2.53	1.75	*4.49	*2.23	*1.60	*2.18	*2.93	*2.77	*2.96	
		Emax(pu)	26.11	*30.67	14.92	11.97	*25.89	*14.80	*13.46	*11.85	*23.60	*18.00	*43.80	
	H ₂	CV	*0.00126	0.00190	0.00119	*0.00064	0.00133	0.00101	0.00097	0.00090	0.00481	0.00153	0.00316	
		Emed(pu)	*2.78	3.65	2.49	*1.68	4.97	2.27	2.13	2.62	3.49	2.78	3.00	
		Emax(pu)	*21.78	34.43	15.69	*11.68	28.17	14.35	13.92	16.85	28.37	18.00	56.28	
H ₃	CV	0.00145	0.00206	0.00124	0.00054	0.00131	0.00101	0.00097	0.00095	0.00469	0.00157	0.00320		
	Emed(pu)	2.99	4.18	2.52	1.70	4.93	2.27	2.11	2.66	3.49	2.82	3.09		
	Emax(pu)	27.22	36.86	16.15	11.82	27.33	14.22	13.92	17.69	27.73	18.80	46.56		

Tabela 4.5: Tabela comparativa com simulações para o modelo M2. Potência Ativa.

$S_1 = 7$ dias $S_2 = 21$ dias $S_3 = 42$ dias.
 $H_1 = 21 : 9 : 1$ $H_2 = 21 : 20 : 1$ $H_3 = 21 : 30 : 1$.
 I (industrial), R (residencial) e C (comercial).

treinamento	topologia	estática	Subestação										
			40 R	80 C	221 R	232 R	379 I	463 R	572 R	623 C	688 R	762 I	840 R
S_1	H_1	CV	0.00164	0.00443	0.00227	0.00653	0.01083	0.01011	0.00312	0.00395	0.01107	0.00313	0.00373
		Emed(pu)	3.27	6.32	6.92	8.83	11.87	8.32	5.43	7.76	4.89	4.39	5.71
		Emax(pu)	30.63	43.14	44.00	56.08	63.67	100.00	34.90	47.00	36.08	28.67	33.14
	H_2	CV	0.00158	0.00410	0.00207	0.00572	0.00979	0.00969	0.00301	0.00342	0.00928	0.00286	0.00347
		Emed(pu)	3.21	5.80	6.56	8.16	11.06	7.82	5.32	6.78	4.27	4.14	5.27
		Emax(pu)	29.37	40.86	38.20	47.42	59.50	100.00	32.90	40.75	29.62	25.67	32.43
H_3	CV	0.00175	0.00383	0.00202	0.00515	0.00952	0.01087	0.00293	0.00346	0.01006	0.00303	0.00351	
	Emed(pu)	3.36	5.53	6.46	7.47	10.68	8.10	5.31	7.00	4.49	4.37	5.42	
	Emax(pu)	32.88	37.14	37.80	42.25	57.83	100.00	31.70	39.75	31.23	26.67	32.43	
S_2	H_1	CV	0.00142	0.00375	0.00177	0.00401	0.00745	*0.00988	0.00230	0.00292	0.01026	0.00252	0.00260
		Emed(pu)	2.71	5.20	5.53	5.68	8.30	*6.88	4.01	5.92	4.17	3.51	3.93
		Emax(pu)	26.38	37.57	34.00	33.42	45.67	*100.00	25.50	34.00	33.77	23.33	24.57
	H_2	CV	0.00127	0.00337	0.00178	0.00407	0.00779	0.00940	0.00219	0.00291	0.00893	0.00242	0.00270
		Emed(pu)	2.74	4.91	5.80	5.82	8.74	6.75	3.78	5.74	4.05	3.46	4.01
		Emax(pu)	23.00	33.14	33.00	33.83	47.50	100.00	24.50	33.50	28.46	22.00	24.14
H_3	CV	0.00124	0.00340	0.00175	0.00406	0.00810	0.00995	0.00213	0.00296	0.00927	0.00244	0.00282	
	Emed(pu)	2.77	4.71	5.52	5.81	8.96	7.57	3.82	6.07	4.09	3.53	4.03	
	Emax(pu)	22.50	34.00	32.80	33.58	49.17	100.00	22.90	34.25	30.08	22.00	27.14	
S_3	H_1	CV	*0.00127	*0.00283	*0.00108	0.00387	*0.00681	0.01265	*0.00198	*0.00282	*0.00778	0.00225	*0.00222
		Emed(pu)	*2.88	*4.14	*6.18	5.44	*7.60	8.48	*3.58	*5.18	*3.56	3.22	*3.39
		Emax(pu)	*23.25	*27.14	*32.60	32.42	*41.93	100.00	*21.70	*29.50	*25.00	20.33	*20.71
	H_2	CV	0.00119	0.00303	0.00169	*0.00377	0.00771	0.01228	0.00224	0.00275	0.00870	*0.00221	0.00245
		Emed(pu)	2.56	4.42	5.46	*5.88	8.72	8.50	3.77	5.56	4.06	*3.18	3.71
		Emax(pu)	21.50	28.86	31.80	*30.75	46.50	100.00	24.80	32.25	26.92	*19.00	22.86
H_3	CV	0.00140	0.00450	0.00202	0.00524	0.00917	0.01258	0.00290	0.00348	0.01048	0.00325	0.00323	
	Emed(pu)	2.71	6.73	6.80	7.16	10.37	17.93	5.30	7.37	5.07	4.74	5.17	
	Emax(pu)	25.62	43.57	36.40	44.92	55.50	100.00	31.60	39.25	32.77	28.33	29.14	

Tabela 4.6: Tabela comparativa com simulações para o modelo M2. Potência Reativa.

$S_1 = 7$ dias $S_2 = 21$ dias $S_3 = 42$ dias.
 $H_1 = 21 : 9 : 1$ $H_2 = 21 : 20 : 1$ $H_3 = 21 : 30 : 1$.
 I (industrial), R (residencial) e C (comercial).

treinamento	topologia	estatística	Subestação											
			40 R	80 C	221 R	232 R	379 I	463 R	572 R	623 C	668 R	762 I	840 R	
S_1	H_1	CV	0.00261	0.00317	0.00229	0.00136	0.00207	0.00204	0.00176	0.00110	0.00727	0.00274	0.00351	
		Emed(pu)	6.31	5.93	4.97	3.95	7.70	4.41	3.90	3.37	5.36	4.57	5.09	
		Emax(pu)	46.22	57.00	28.69	29.58	43.06	29.43	25.17	19.77	43.27	33.80	59.28	
	H_2	CV	0.00272	0.00334	0.00216	0.00136	0.00202	0.00193	0.00179	0.00109	0.00720	0.00274	0.00349	
		Emed(pu)	6.48	5.82	4.76	3.91	7.48	4.18	3.96	3.33	5.28	4.55	5.07	
		Emax(pu)	48.89	58.57	27.46	30.03	42.94	27.87	25.08	19.77	42.60	33.60	54.06	
H_3	CV	0.00256	0.00371	0.00221	0.00127	0.00182	0.00196	0.00176	0.00112	0.00768	0.00270	0.00355		
	Emed(pu)	5.84	6.15	4.85	3.53	6.73	4.30	3.89	3.33	5.52	4.49	5.25		
	Emax(pu)	45.11	67.43	28.08	28.68	38.83	27.78	24.83	20.00	45.80	33.00	62.50		
S_2	H_1	CV	0.00221	0.00244	0.00165	0.00114	0.00166	0.00157	0.00134	0.00091	0.00546	0.00240	0.00331	
		Emed(pu)	5.22	4.46	3.35	2.89	6.22	2.87	2.73	2.65	3.82	3.66	3.62	
		Emax(pu)	39.67	44.43	22.00	26.37	34.89	23.57	19.79	17.00	32.90	30.20	48.22	
	H_2	CV	0.00251	0.00250	0.00180	0.00115	0.00226	0.00160	0.00145	0.00109	0.00646	0.00242	0.00334	
		Emed(pu)	5.47	4.94	3.67	2.90	8.23	3.00	2.97	3.06	4.41	3.75	3.89	
		Emax(pu)	46.44	44.76	23.54	26.39	48.61	24.35	21.54	20.23	38.97	30.60	43.50	
H_3	CV	0.00177	0.00244	0.00167	0.00102	0.00168	0.00157	0.00143	0.00103	0.00654	0.00242	0.00337		
	Emed(pu)	3.60	4.72	3.47	2.15	5.89	2.89	2.85	2.96	4.43	3.74	3.82		
	Emax(pu)	33.00	43.67	22.08	24.42	37.39	24.22	21.46	19.23	39.53	28.20	47.78		
S_3	H_1	CV	0.00202	0.00217	0.00147	0.00110	0.00157	0.00148	*0.00124	0.00094	0.00559	*0.00228	0.00323	
		Emed(pu)	4.74	3.94	2.93	2.70	5.86	2.68	*2.46	2.36	3.80	*3.32	3.24	
		Emax(pu)	36.11	38.81	19.46	25.63	33.28	22.78	*18.29	18.69	33.77	*28.20	50.83	
	H_2	CV	0.00184	0.00212	0.00151	0.00101	*0.00150	0.00148	0.00126	0.00089	0.00632	0.00234	0.00326	
		Emed(pu)	3.52	4.11	3.10	2.08	*6.25	2.73	2.56	2.61	4.30	3.55	3.35	
		Emax(pu)	35.00	38.19	19.69	24.26	*38.17	22.65	18.67	16.38	37.87	28.40	49.33	
H_3	CV	*0.00172	*0.00200	*0.00142	*0.00100	0.00183	*0.00148	0.00122	*0.00083	*0.00586	0.00229	*0.00320		
	Emed(pu)	*3.52	*3.85	*2.85	*2.07	6.21	*2.59	2.48	*2.33	*9.05	3.40	*4.18		
	Emax(pu)	*32.22	*36.10	*19.08	*24.18	41.39	*22.18	17.92	*15.02	*32.33	28.80	*46.00		

Tabela 4.7: Tabela comparativa com simulações para o modelo M3. Potência Ativa.

$S_1 = 7$ dias $S_2 = 21$ dias $S_3 = 42$ dias.
 $H_1 = 21 : 9 : 1$ $H_2 = 21 : 20 : 1$ $H_3 = 21 : 30 : 1$.
 I (industrial), R (residencial) e C (comercial).

treinamento	topologia	estatística	Subestação											
			40 R	80 C	221 R	232 R	379 I	463 R	572 R	623 C	668 R	762 I	840 R	
S_1	H_1	CV	0.00240	0.00583	0.00251	0.00794	0.01248	0.00927	0.00405	0.00448	0.01316	0.00372	0.00415	
		Emed(pu)	4.88	8.85	7.86	11.16	13.82	8.71	7.13	8.63	5.94	5.21	6.74	
	H_2	Emed(pu)	44.50	55.29	46.80	66.67	76.83	100.00	43.10	52.25	41.46	34.33	37.00	
		CV	0.00252	0.00575	0.00257	0.00733	0.01207	0.00972	0.00399	0.00426	0.01348	0.00370	0.00417	
	S_2	H_1	Emed(pu)	5.06	8.63	7.78	10.41	13.35	9.02	7.13	8.55	5.87	5.17	6.66
			Emed(pu)	46.00	55.43	50.00	60.92	74.50	100.00	42.10	49.25	44.00	34.00	37.00
H_2		CV	0.00223	0.00573	0.00248	0.00669	0.01064	0.00860	0.00393	0.00417	0.01337	0.00374	0.00426	
		Emed(pu)	4.49	8.60	7.59	9.46	11.93	8.76	7.09	8.51	5.83	5.22	6.72	
H_3		Emed(pu)	41.25	54.29	46.80	55.58	64.00	100.00	43.00	48.75	43.46	34.00	38.57	
		CV	0.00201	0.00443	0.00187	0.00653	0.00988	0.00988	0.00309	0.00326	0.01291	0.00317	0.00293	
S_3	H_1	Emed(pu)	4.24	6.05	5.69	9.11	11.05	*0.00962	5.13	6.23	5.23	4.17	4.30	
		Emed(pu)	37.12	44.29	36.60	53.67	60.00	*100.00	35.40	37.25	42.77	30.00	26.00	
	H_2	CV	0.00253	0.00465	0.00199	0.00748	0.01267	0.00982	0.00325	0.00357	0.01373	0.00325	0.00338	
		Emed(pu)	5.05	6.48	6.23	10.35	14.02	7.44	5.49	7.14	5.64	4.33	5.09	
	H_3	Emed(pu)	47.25	46.43	36.60	63.17	75.50	100.00	37.10	42.25	45.54	30.67	31.71	
		CV	0.00159	0.00453	0.00202	0.00515	0.00956	0.00856	0.00316	0.00346	0.01244	0.00329	0.00337	
S_3	H_1	Emed(pu)	3.14	6.37	6.22	6.61	10.35	6.74	5.41	6.98	5.28	4.37	4.99	
		Emed(pu)	28.75	45.14	38.80	44.67	60.00	100.00	35.70	40.25	41.15	31.00	31.57	
	H_2	CV	0.00157	*0.00871	*0.00173	*0.00496	*0.00721	0.01171	*0.00281	*0.00274	*0.00973	0.00306	*0.00268	
		Emed(pu)	2.92	*5.09	*5.01	*5.64	*7.75	6.44	*4.27	*5.40	*4.12	3.96	*3.76	
	H_3	Emed(pu)	29.50	*37.29	*34.40	*42.17	*45.50	99.00	*83.40	*32.25	*22.23	29.00	*28.29	
		CV	0.00160	0.00381	0.00182	0.00514	0.00857	0.01188	0.00277	0.00296	0.01159	0.00294	0.00266	
S_3	H_1	Emed(pu)	3.07	5.33	5.67	6.29	9.35	6.46	4.47	5.90	4.83	3.91	3.89	
		Emed(pu)	29.87	37.43	35.00	46.50	52.67	100.00	32.20	34.75	36.77	27.67	25.57	
	H_2	CV	*0.00141	0.00372	0.00175	0.00508	0.00955	0.01142	0.00271	0.00286	0.01089	*0.00290	0.00275	
		Emed(pu)	*2.90	5.31	5.53	6.27	10.31	6.56	4.45	5.64	4.53	*3.91	3.86	
	H_3	Emed(pu)	*28.00	36.57	33.60	45.42	59.83	100.00	31.40	34.25	36.08	*27.67	26.43	

Tabela 4.8: Tabela comparativa com simulações para o modelo M3. Potência Reativa.

4.5.2 Comentários sobre os resultados do modelo M

Os tempos de CPU apresentados para o modelo M1 a partir da referência S_2H_1 foram tomados em uma estação *SUN SPARCstation Ultra*. Os tempos do modelos M3 em uma *SUN SPARCstation 10* e do modelo M2 a partir da referência S_2H_1 em uma *SUN SPARCstation 20*. Demais tempos foram tomados em uma *SUN SPARCstation 4*. Os tempos da tabela estão referenciados para uma *SUN SPARCstation 20*

modelo	referência	tempo
		médio
M1	S_1H_1	42.57
	S_1H_2	12.77
	S_1H_3	4.53
	S_2H_1	254.46
	S_2H_2	84.18
	S_2H_3	187.66
	S_3H_1	158.25
	S_3H_2	188.94
	S_3H_3	287.17
M2	S_1H_1	13.44
	S_1H_2	3.86
	S_1H_3	5.10
	S_2H_1	33.34
	S_2H_2	26.33
	S_2H_3	10.29
	S_3H_1	123.39
	S_3H_2	115.75
	S_3H_3	146.84
M3	S_1H_1	3.26
	S_1H_2	5.17
	S_1H_3	10.50
	S_2H_1	13.75
	S_2H_2	24.50
	S_2H_3	39.36
	S_3H_1	19.44
	S_3H_2	7.13
	S_3H_3	12.47

Tabela 4.9: Tempos de execução em segundos para as diversas variações no modelo M.

As características gerais de desempenho do modelo M são as seguintes:

- Para o modelo M1 a variação do erro entre o conjunto de treinamento S_1 e suas respectivas topologias para o erro no conjunto de treinamento S_3 é maior para a potência reativa do que para a potência ativa, ou seja, o número de neurônios na rede teve uma influência maior no erro de previsão da potência reativa do que na potência ativa.
- No caso do modelo M2 a variação acima diminui, ou seja, o efeito do número de neurônios na rede no erro de previsão é reduzido.
- Muito embora o erro de previsão com o modelo M2 seja maior do que com o modelo M1, percebe-se que os dois são bem próximos. Assim, o modelo M2 acaba sendo uma alternativa ao modelo M1, considerando-se que o tempo de treinamento deste é menor, conforme se observa na tabela 4.9.
- Pode-se observar uma tendência de melhor desempenho de previsão quando se aumenta o tamanho do conjunto de treinamento e se diminui o número de neurônios. Por outro lado, para um conjunto de treinamento pequeno a topologia que apresenta melhor resultado é aquela que possui mais neurônios. No caso do modelo M3 esse fato é observado para os três tamanhos de conjunto de treinamento utilizados, à exceção do conjunto S_3 na tabela 4.8, para o qual o melhor desempenho se encontra no conjunto que utilizou a menor topologia. Isto aconteceu porque o número de iterações utilizado para treinamento foi pequeno, favorecendo a topologia que possui mais neurônios. Já que, neste caso, ela possui maior flexibilidade em relação à topologia com menor número de neurônios. De outra forma, o acréscimo de informação no conjunto de treinamento faz com que a rede com menor número possível de neurônios seja a que apresente melhores resultados.
- O modelo M3 apresenta uma performance inferior à dos modelos M1 e M2. Cabe notar que para ele os melhores erros de previsão se apresentam nas topologias que possuem maior número de neurônios.

4.5.3 Resultados com Modelo L

Uma vez que o número de dias utilizados para se compor o vetor de treinamento no caso do modelo L é previamente estabelecido e fixo, conforme apresentado na seção 4.3, então as únicas variações testadas foram as de topologia, as quais são apresentadas abaixo, sendo idênticas para os modelos L3 e L4, cujo resumo se encontra na página 81.

- $H_1 = 9 : 4 : 1$
- $H_2 = 9 : 8 : 1$
- $H_3 = 9 : 18 : 1$

O critério de parada utilizado foi um erro de 0.0005 ou 1000 iterações. Valor de η igual a 0.1. O valor do erro é pequeno porque o número de elementos apresentados para treinamento da rede é muito reduzido. Assim, é necessário avançar no processo de adaptação dos pesos a fim de estimá-los convenientemente.

As tabelas 4.10 e 4.11 mostram os resultados de simulação obtidos com os modelos L3 e L4, respectivamente, para potência ativa e reativa.

S_1 Conjunto de treinamento fixo. Ver descrição na seção .
 $H_1 = 9 : 4 : 1$ $H_2 = 9 : 8 : 1$ $H_3 = 9 : 18 : 1$.
 I (industrial), R (residencial) e C (comercial).

conjunto estimação	topo- logia	esta- tística	Subestação											
			40 R	80 C	221 R	232 R	379 I	463 R	572 R	623 C	668 R	762 I	840 R	
S_1	H_1	CV	*0.00211	*0.00228	*0.00186	*0.00117	*0.00196	*0.00104	*0.00148	*0.00098	0.00653	*0.00200	*0.00817	
		Emed(pu)	*4.96	*5.00	*4.11	*3.14	*6.74	*3.54	*3.13	*2.96	5.08	*3.48	*3.98	
		Emax(pu)	*88.00	*40.10	*23.69	*26.63	*36.67	*23.48	*21.75	*17.84	33.83	*24.40	*54.22	
	H_2	CV	0.00218	0.00229	0.00191	0.00118	0.00164	0.00164	0.00151	0.00100	*0.00680	0.00202	0.00318	
		Emed(pu)	5.18	5.03	4.18	3.21	5.79	3.55	3.18	3.00	*4.77	3.51	4.03	
		Emax(pu)	39.11	38.95	23.85	26.87	36.00	23.61	22.17	18.15	*35.53	24.20	55.39	
	H_3	CV	0.00217	0.00232	0.00200	0.00117	0.00169	0.00171	0.00152	0.00101	0.00672	0.00202	0.00321	
		Emed(pu)	5.21	5.16	4.33	3.16	5.94	3.69	3.22	3.05	5.19	3.56	4.14	
		Emax(pu)	39.00	40.19	25.77	26.84	37.00	24.70	22.33	18.46	36.30	23.80	57.44	
	S_1	H_1	CV	*0.00185	*0.00486	*0.00183	*0.00491	*0.00945	*0.00696	*0.00930	*0.00885	0.01171	*0.00291	*0.00854
			Emed(pu)	*3.77	*7.65	*6.06	*9.01	*10.15	*6.48	*6.11	*7.67	5.14	*4.20	*5.83
			Emax(pu)	*84.13	*45.87	*33.20	*52.67	*59.38	*78.00	*36.20	*38.75	38.31	*25.67	*31.43
H_2		CV	0.00191	0.00489	0.00185	0.00646	0.00952	0.00707	0.00350	0.00377	*0.01144	0.00295	0.00360	
		Emed(pu)	3.93	7.78	6.10	9.24	10.37	6.60	6.45	8.01	*6.05	4.26	5.93	
		Emax(pu)	35.38	44.86	34.20	53.75	59.50	73.00	35.90	41.50	*37.23	26.67	32.00	
H_3		CV	0.00191	0.00504	0.00189	0.00641	0.00976	0.00704	0.00353	0.00384	0.01201	0.00294	0.00368	
		Emed(pu)	3.97	8.03	6.31	9.13	10.50	6.76	6.56	8.18	5.36	4.30	6.06	
		Emax(pu)	35.00	45.71	34.60	53.08	60.67	62.78	37.70	41.00	38.85	26.33	32.43	

Tabela 4.10: Tabela comparativa com simulações para o modelo L3

S_1 . Conjunto de treinamento fixo. Ver descrição na seção
 $H_1 = 9 : 4 : 1$ $H_2 = 9 : 8 : 1$ $H_3 = 9 : 18 : 1$.
 I (industrial), R (residencial) e C (comercial).

conjunto estimação	topo- logia	esta- tística	Subestação											
			40 R	80 C	221 R	232 R	379 I	463 R	572 R	623 C	668 R	762 I	840 R	
S_1	H_1	CV	*0.00198	*0.00280	*0.00164	*0.00128	*0.00162	*0.00160	*0.00189	*0.00092	0.00554	*0.00213	*0.00315	
		Emed(pu)	*4.56	*4.56	*3.52	*2.95	*5.48	*3.19	*2.82	*3.40	*3.99	*3.40	*3.76	
		Emax(pu)	*34.78	*41.58	*20.69	*27.97	*36.61	*23.52	*20.71	*16.62	*26.60	*26.60	*46.61	
		CV	0.00203	0.00230	0.00166	0.00126	0.00163	0.00160	0.00141	0.00094	*0.00642	0.00219	0.00316	
		Emed(pu)	4.79	4.61	3.57	3.02	5.52	3.22	2.88	2.87	*3.88	3.49	3.79	
		Emax(pu)	36.56	41.43	21.69	29.42	36.67	24.04	20.96	17.00	*32.43	27.60	47.50	
	H_2	CV	0.00202	0.00236	0.00169	0.00125	0.00166	0.00166	0.00143	0.00095	0.00571	0.00216	0.00317	
		Emed(pu)	4.82	4.77	3.63	3.02	5.63	3.30	2.91	2.90	4.08	3.51	3.89	
		Emax(pu)	36.33	42.43	21.62	29.32	37.39	24.87	20.83	16.85	33.90	27.20	47.61	
		H_3	CV	*0.00170	*0.00449	*0.00182	*0.00322	*0.00854	*0.00648	*0.00295	*0.00310	0.00978	*0.00308	*0.00823
			Emed(pu)	*3.53	*6.57	*5.68	*8.49	*9.29	*5.92	*5.20	*6.83	4.24	*4.15	*5.11
			Emax(pu)	*31.98	*42.86	*34.80	*52.75	*52.60	*96.56	*31.80	*36.60	32.00	*28.33	*28.00
CV	0.00176		0.00455	0.00183	0.00638	0.00875	0.00648	0.00311	0.00328	*0.00966	0.00312	0.00331		
Emed(pu)	3.70		6.70	5.74	8.71	9.56	5.99	5.47	6.84	*4.22	4.25	5.28		
Emax(pu)	32.25		43.43	35.00	54.17	54.00	94.56	34.20	36.50	*31.62	28.67	30.00		
S_1	H_3	CV	0.00175	0.00470	0.00185	0.00634	0.00885	0.00653	0.00314	0.00333	0.01003	0.00310	0.00336	
		Emed(pu)	3.71	6.92	5.86	8.68	9.64	6.17	5.58	6.94	4.40	4.30	5.35	
		Emax(pu)	32.25	45.14	35.20	54.17	55.00	92.56	34.80	37.00	32.85	28.67	30.29	

Tabela 4.11: Tabela comparativa com simulações para o modelo L4

4.5.4 Comentários sobre os resultados do modelo L

Os tempos de CPU apresentados para o modelo L foram tomados em uma estação *SUN SPARCstation 4*.

modelo	referência	tempo
		médio
L3	S_1H_1	0.57
	S_1H_2	0.78
	S_1H_3	1.24
L4	S_1H_1	0.59
	S_1H_2	0.80
	S_1H_3	1.24

Tabela 4.12: Tempos de execução em segundos para as diversas variações no modelo L.

Da mesma forma que para o modelo M, no caso do modelo L a performance de L3 foi inferior a de L4, sendo que o maior número de neurônios na topologia da rede não acarretou a degradação de seu desempenho. Cabe notar o reduzido tempo de treinamento desse modelo, embora cada hora do dia tenha o seu conjunto de pesos que precisa ser ajustado.

4.5.5 Resultados com Modelo F

As tabelas 4.13 e 4.14 mostram os resultados de simulação obtidos com o modelo F1 para potência ativa e reativa. As topologias das redes utilizadas para os testes desse modelo são as seguintes.

- $H_1 = 1 \times 0 : 5 \times 20 : 1 \times 4$
- $H_2 = 1 \times 0 : 10 \times 20 : 1 \times 4$
- $H_3 = 1 \times 0 : 15 \times 20 : 1 \times 4$

O critério de parada utilizado foi um erro de de 0.025 ou 1000 iterações. Valor de η igual a 0.005. Não se utilizou um termo de momento para esta rede.

Os resultados para o modelo F2 são apresentados na tabelas 4.15 e 4.16, utilizando as mesmas topologias do modelo F1, no que se refere a modelagem para o dia de semana. Para os finais de semana as topologias utilizadas são as seguintes:

- $H_1 = 1 \times 0 : 3 \times 20 : 1 \times 4$
- $H_2 = 1 \times 0 : 6 \times 20 : 1 \times 4$
- $H_3 = 1 \times 0 : 9 \times 20 : 1 \times 4$

Esta mesma topologia é utilizada no modelo F4, cujos resultados são apresentados nas tabelas 4.19 e 4.20. Já no modelo F3 utiliza-se uma topologia idêntica à do modelo F1, com o critério de parada igual a 1000 iterações ou erro igual a 0.005 e resultados apresentados nas tabelas 4.17 e 4.18. Para os casos dos modelos F2 e F4 os critérios de parada são 1000 iterações ou erro igual a 0.01.

A notação $H_1 : 1 \times 0 : 3 \times 20 : 1 \times 4$ de topologia acima indica uma rede FIR $1 : 3 : 1$ constituída de um neurônio na camada de entrada, três neurônios na camada intermediária e 1 neurônio na camada de saída, cuja ordem dos filtros nas sinapses $0 \times 20 \times 4$ indica um filtro de ordem zero na camada de entrada (essa camada não possui sinapses, nem filtros), um filtro de ordem vinte nas sinapses da camada intermediária e um filtro de ordem quatro nas sinapses da camada de saída. O resumo deste modelo se encontra na página 84.

A ordem dos filtros $0 \times 20 \times 4$ foi escolhida de forma a captar o ciclo diário da carga, uma vez que a soma das ordens dos filtros é igual a 24, correspondendo à 24 medições horárias de carga.

$S_1 = 7 \text{ dias } S_2 = 21 \text{ dias } S_3 = 42 \text{ dias.}$
 $H_1 = 1 \times 0 : 5 \times 20 : 1 \times 4 H_2 = 1 \times 0 : 10 \times 20 : 1 \times 4 H_3 = 1 \times 0 : 15 \times 20 : 1 \times 4.$
 I (industrial), R (residencial) e C (comercial).

treinamento	topologia	esta- física	Subestação											
			40 R	80 C	221 R	232 R	379 I	463 R	572 R	623 C	668 R	762 I	840 R	
S ₁	H ₁	CV	0.00184	0.00259	0.00198	0.00125	0.00198	0.00194	0.00172	0.00113	0.00454	0.00359	0.00339	
		Emed(pu)	4.64	6.15	4.63	3.64	7.73	4.38	4.07	3.71	3.79	6.94	4.88	
		Emax(pu)	32.22	44.19	23.77	27.97	40.50	27.48	23.42	18.85	25.77	41.40	45.72	
	H ₂	CV	0.00235	0.00308	0.00245	0.00115	0.00240	0.00175	0.00190	0.00174	0.00606	0.00377	0.00358	
		Emed(pu)	5.54	7.36	5.42	3.23	9.15	3.88	4.40	5.36	4.96	7.11	5.64	
		Emax(pu)	42.33	51.33	31.23	25.92	49.39	25.13	26.42	30.85	34.37	44.40	55.67	
H ₃	CV	0.00238	0.00358	0.00223	0.00114	0.00230	0.00166	0.00159	0.00186	0.00611	0.00347	0.00349		
	Emed(pu)	5.67	8.18	4.89	3.22	8.71	3.59	3.68	5.37	4.95	6.56	5.19		
	Emax(pu)	42.67	61.52	28.69	23.55	48.22	23.96	22.08	34.69	34.93	40.60	59.89		
S ₂	H ₁	CV	*0.00119	0.00208	0.00146	0.00118	0.00161	0.00156	0.00139	0.00085	*0.00398	0.00294	0.00319	
		Emed(pu)	*2.99	4.88	3.23	3.63	6.20	3.39	3.24	2.79	*8.18	5.61	3.54	
		Emax(pu)	*20.89	35.71	18.46	25.11	33.39	22.17	19.04	14.46	*23.18	34.00	45.56	
	H ₂	CV	0.00174	0.00235	0.00164	0.00104	0.00188	0.00141	0.00116	0.00115	0.00457	0.00268	0.00329	
		Emed(pu)	4.46	5.58	3.70	2.64	7.39	3.00	2.60	3.80	3.94	4.83	3.84	
		Emax(pu)	30.22	39.86	20.62	23.53	38.00	19.70	16.46	19.08	25.40	29.60	53.78	
H ₃	CV	0.00129	0.00249	0.00184	0.00111	0.00202	0.00144	0.00115	0.00147	0.00547	0.00258	0.00333		
	Emed(pu)	3.31	5.68	4.06	3.09	7.74	3.04	2.64	4.77	4.57	4.56	4.12		
	Emax(pu)	22.22	43.05	23.62	25.03	41.72	21.13	16.25	25.46	30.67	29.60	49.61		
S ₃	H ₁	CV	0.00142	0.00198	*0.00142	0.00106	*0.00142	0.00140	0.00119	*0.00076	0.00472	0.00254	*0.00316	
		Emed(pu)	3.32	4.43	*3.18	2.83	*6.34	2.84	2.66	*2.42	3.58	4.54	*3.34	
		Emax(pu)	25.44	34.57	*17.99	24.13	*29.67	20.96	16.71	*13.23	27.67	29.80	*47.28	
	H ₂	CV	0.00165	*0.00195	0.00156	*0.00099	0.00167	*0.00138	*0.00108	0.00086	0.00570	0.00256	0.00319	
		Emed(pu)	4.18	*4.88	3.50	*2.43	6.47	*2.78	*2.40	2.80	4.62	4.62	3.54	
		Emax(pu)	28.22	*33.52	18.92	*21.80	34.39	*20.57	*15.54	14.54	32.60	30.40	44.17	
H ₃	CV	0.00209	0.00234	0.00171	0.00101	0.00192	0.00143	0.00114	0.00114	0.00564	*0.00262	0.00328		
	Emed(pu)	4.74	5.57	3.90	2.60	7.41	2.94	2.55	3.75	4.67	*4.51	4.14		
	Emax(pu)	37.78	39.76	21.00	21.39	39.61	20.83	16.12	19.15	32.13	*29.00	56.61		

Tabela 4.13: Tabela comparativa com simulações para o modelo F1. Potência Ativa.

$S_1 = 7 \text{ dias } S_2 = 21 \text{ dias } S_3 = 42 \text{ dias.}$
 $H_1 = 1 \times 0 : 5 \times 20 : 1 \times 4 \ H_2 = 1 \times 0 : 10 \times 20 : 1 \times 4 \ H_3 = 1 \times 0 : 15 \times 20 : 1 \times 4.$
 I (industrial), R (residencial) e C (comercial).

treina- mento	topo- logia	esta- tística	Subestação											
			40 R	80 C	221 R	232 R	379 I	463 R	572 R	623 C	668 R	762 I	840 R	
S ₁	H ₁	CV	0.00180	0.00483	0.00237	0.00841	0.01100	0.00825	0.00315	0.00356	0.01011	0.00488	0.00354	
		Emed(pu)	4.10	7.45	8.10	12.52	12.71	10.07	5.72	7.64	5.21	7.45	5.81	
		Emax(pu)	32.38	44.86	42.20	67.67	63.67	100.00	34.30	39.50	30.31	42.00	31.29	
	H ₂	CV	0.00192	0.00464	0.00266	0.00951	0.01581	0.01567	0.00336	0.00481	0.01186	0.00523	0.00420	
		Emed(pu)	4.05	7.13	8.98	13.57	17.62	15.54	6.32	9.62	6.15	7.95	6.58	
		Emax(pu)	35.25	43.00	47.80	78.17	91.67	100.00	35.50	56.75	35.08	45.67	38.86	
H ₃	CV	0.00194	0.00552	0.00266	0.00891	0.01655	0.01704	0.00356	0.00458	0.01259	0.00486	0.00487		
	Emed(pu)	4.06	8.64	9.05	12.78	18.30	17.10	6.56	9.42	6.13	7.26	7.56		
	Emax(pu)	34.75	50.71	48.00	73.67	100.00	100.00	37.60	52.75	39.00	42.00	43.71		
S ₂	H ₁	CV	0.00122	0.00347	0.00194	0.00577	0.00939	0.00858	0.00211	0.00292	*0.00789	*0.00826	0.00275	
		Emed(pu)	2.97	5.33	6.59	8.29	10.87	9.41	3.75	6.22	*3.68	*4.69	4.28	
		Emax(pu)	21.38	33.00	34.80	47.92	55.67	100.00	23.30	32.75	*25.08	*29.33	23.71	
	H ₂	CV	0.00116	0.00402	0.00245	0.00577	0.01098	0.01073	0.00242	0.00340	0.01092	0.00353	0.00313	
		Emed(pu)	2.81	6.28	8.39	8.29	12.65	13.23	4.48	7.33	5.33	5.19	4.87	
		Emax(pu)	20.50	37.86	44.00	47.58	65.00	100.00	26.20	37.75	32.92	31.33	28.71	
H ₃	CV	0.00127	0.00392	0.00267	0.00636	0.01489	0.01375	0.00261	0.00446	0.01232	0.00347	0.00339		
	Emed(pu)	3.08	5.97	9.12	8.96	16.51	15.55	4.85	9.32	6.18	5.14	5.34		
	Emax(pu)	22.38	37.14	48.00	53.50	90.83	100.00	28.00	50.75	37.54	31.00	31.00		
S ₃	H ₁	CV	0.00133	*0.00813	*0.00174	0.00577	*0.00708	*0.00767	*0.00195	*0.00245	0.00855	0.00331	*0.00258	
		Emed(pu)	3.06	*4.66	*5.81	8.29	*8.67	*8.60	*3.87	*4.98	4.09	4.80	*3.92	
		Emax(pu)	23.88	*30.00	*31.40	47.75	*46.17	*106.00	*22.00	*27.50	26.92	29.00	*24.43	
	H ₂	CV	*0.00110	0.00338	0.00202	*0.00567	0.00854	0.01011	0.00204	0.00282	0.00923	0.00328	0.00285	
		Emed(pu)	*2.80	5.06	6.68	*7.88	9.73	12.30	3.58	5.88	4.68	4.87	4.49	
		Emax(pu)	*19.00	32.57	35.20	*46.42	51.00	100.00	22.70	32.00	28.00	29.00	25.71	
H ₃	CV	0.00170	0.00369	0.00205	0.00644	0.01007	0.00922	0.00241	0.00321	0.01059	0.00328	0.00269		
	Emed(pu)	4.04	5.63	6.85	9.52	11.54	11.06	4.28	6.73	5.38	4.71	4.17		
	Emax(pu)	30.12	33.71	36.40	50.50	60.33	100.00	26.10	36.25	31.38	29.67	24.00		

Tabela 4.14: Tabela comparativa com simulações para o modelo F1. Potência Reativa.

$S_1 = 7 \text{ dias } S_2 = 21 \text{ dias } S_3 = 42 \text{ dias.}$
 $H_1 = 1 \times 0 : 5 \times 20 : 1 \times 4 H_2 = 1 \times 0 : 10 \times 20 : 1 \times 4 H_3 = 1 \times 0 : 15 \times 20 : 1 \times 4.$
 Dia de Semana
 $H_1 = 1 \times 0 : 3 \times 20 : 1 \times 4 H_2 = 1 \times 0 : 6 \times 20 : 1 \times 4 H_3 = 1 \times 0 : 9 \times 20 : 1 \times 4.$
 Outro grupo.
 I (industrial), R (residencial) e C (comercial).

treinamento	topologia	estatística	Subestação											
			40 R	80 C	221 R	232 R	379 I	463 R	572 R	623 C	668 R	762 I	840 R	
S ₁	H ₁	CV	0.00175	0.00427	0.00202	0.00638	0.00936	*0.00791	0.00316	0.00374	0.00999	0.00350	0.00334	
		Emed(pu)	3.82	6.64	6.64	8.74	10.65	*8.43	5.52	7.54	4.73	5.20	5.21	
		Emax(pu)	31.62	41.00	37.40	54.00	55.00	*100.00	35.40	43.50	31.62	31.00	30.57	
	H ₂	CV	0.00176	0.00411	0.00233	0.00607	0.01024	0.00998	0.00312	0.00439	0.01018	0.00371	0.00363	
		Emed(pu)	3.46	6.05	7.91	8.35	11.51	10.96	5.77	9.06	5.00	5.39	5.70	
		Emax(pu)	32.88	39.43	42.20	51.75	62.33	100.00	33.40	49.25	31.08	33.00	33.29	
H ₃	CV	0.00174	0.00438	0.00229	0.00620	0.00982	0.01136	0.00347	0.00463	0.01049	0.00351	0.00361		
	Emed(pu)	3.41	6.32	7.55	8.57	11.16	11.45	6.24	9.17	5.11	5.05	5.38		
	Emax(pu)	32.75	43.14	42.40	52.00	58.33	100.00	38.20	54.25	32.62	32.00	34.29		
S ₂	H ₁	CV	0.00142	0.00319	*0.00175	0.00515	0.00762	0.00938	0.00221	0.00278	0.00875	0.00314	0.00262	
		Emed(pu)	3.05	4.86	*5.69	7.01	8.56	10.70	3.81	5.60	4.04	4.55	3.92	
		Emax(pu)	26.12	31.43	*32.60	44.17	46.00	100.00	25.00	31.50	27.85	28.33	24.00	
	H ₂	CV	0.00175	0.00367	0.00197	0.00605	0.01109	0.01019	0.00258	0.00402	0.01076	0.00322	0.00304	
		Emed(pu)	3.63	5.27	6.53	8.44	12.24	11.37	4.57	8.01	4.95	4.74	4.65	
		Emax(pu)	32.12	36.14	36.40	50.50	68.50	100.00	28.40	46.50	33.62	28.67	28.43	
H ₃	CV	0.00156	0.00380	0.00248	0.00581	0.01146	0.01164	0.00284	0.00442	0.01090	0.00328	0.00353		
	Emed(pu)	3.15	5.61	8.23	7.82	12.83	12.77	5.13	9.13	5.19	4.83	5.43		
	Emax(pu)	29.00	36.29	45.00	50.25	68.50	100.00	31.00	50.75	34.15	28.67	31.86		
S ₃	H ₁	CV	0.00132	*0.00288	0.00178	*0.00477	*0.00748	0.01051	*0.00207	*0.00244	*0.00804	*0.00300	*0.00245	
		Emed(pu)	2.86	*3.97	5.76	*6.88	*8.28	8.86	*3.47	*4.82	*3.77	*4.17	*3.58	
		Emax(pu)	24.12	*29.00	33.20	*40.33	*46.17	100.00	*28.60	*29.25	*28.31	*27.67	*23.00	
	H ₂	CV	0.00128	0.00327	0.00180	0.00499	0.00838	0.01147	0.00213	0.00289	0.00904	0.00293	0.00267	
		Emed(pu)	2.73	4.64	5.89	6.70	9.31	11.18	3.76	5.86	4.18	4.19	3.98	
		Emax(pu)	23.62	32.00	33.40	42.92	51.33	100.00	23.10	33.75	28.00	26.67	25.14	
H ₃	CV	*0.00104	0.00335	0.00197	0.00524	0.00936	0.01014	0.00243	0.00326	0.00940	0.00311	0.00296		
	Emed(pu)	*2.54	4.88	6.56	6.97	10.50	10.51	4.34	6.63	4.40	4.48	4.55		
	Emax(pu)	*18.25	32.37	35.20	45.58	56.67	100.00	26.50	37.50	29.85	27.67	27.14		

Tabela 4.16: Tabela comparativa com simulações para o modelo F2. Potência Reativa.

$S_1 = 7 \text{ dias } S_2 = 21 \text{ dias } S_3 = 42 \text{ dias.}$
 $H_1 = 1 \times 0 : 5 \times 20 : 1 \times 4 \ H_2 = 1 \times 0 : 10 \times 20 : 1 \times 4 \ H_3 = 1 \times 0 : 15 \times 20 : 1 \times 4.$
 I (industrial), R (residencial) e C (comercial).

treinamento	topologia	estatística	Subestação											
			40 R	80 C	221 R	232 R	379 I	463 R	572 R	623 C	668 R	762 I	840 R	
S ₁	H ₁	CV	0.00259	0.00291	0.00281	0.00138	0.00232	0.00237	0.00215	0.00156	0.00788	0.00317	0.00371	
		Emed(pu)	6.29	6.43	6.05	3.93	8.50	5.06	4.57	4.71	5.69	5.57	5.89	
		Emax(pu)	46.22	50.90	36.00	30.76	49.06	34.09	31.37	28.00	47.13	38.60	62.78	
	H ₂	CV	0.00253	0.00298	0.00284	0.00130	0.00213	0.00202	0.00191	0.00151	0.00735	0.00288	0.00366	
		Emed(pu)	6.09	6.42	6.05	3.69	7.94	4.22	4.14	4.57	5.50	5.02	5.62	
		Emax(pu)	45.00	52.57	37.00	29.05	45.39	29.70	27.04	27.31	43.17	35.20	52.11	
H ₃	CV	0.00242	0.00322	0.00264	0.00127	0.00209	0.00181	0.00184	0.00151	0.00755	0.00272	0.00348		
	Emed(pu)	5.82	6.78	5.66	3.55	7.64	3.93	3.98	4.49	5.50	4.80	5.13		
	Emax(pu)	43.22	56.76	34.23	28.71	45.06	26.17	26.58	27.54	45.07	33.20	56.67		
S ₂	H ₁	CV	0.00177	0.00261	0.00185	0.00105	0.00169	0.00179	0.00142	0.00101	0.00593	0.00259	0.00330	
		Emed(pu)	4.11	5.69	4.02	2.58	6.33	3.76	3.16	3.23	4.45	4.44	4.21	
		Emax(pu)	32.00	44.95	23.85	24.50	35.67	26.26	20.33	17.31	34.07	31.20	56.39	
	H ₂	CV	0.00199	0.00253	0.00196	0.00104	0.00210	0.00166	0.00171	0.00133	0.00584	0.00252	0.00342	
		Emed(pu)	4.48	5.42	4.28	2.59	7.80	3.43	3.62	4.09	4.48	4.28	4.39	
		Emax(pu)	35.11	44.43	24.77	23.66	44.22	24.35	25.08	23.92	34.27	31.00	57.94	
H ₃	CV	0.00199	0.00257	0.00213	0.00106	0.00198	0.00162	0.00152	0.00127	0.00632	0.00248	0.00340		
	Emed(pu)	4.42	5.58	4.53	2.69	7.34	3.35	3.35	4.04	4.71	4.21	4.58		
	Emax(pu)	36.44	45.29	27.46	24.00	42.17	23.65	21.96	22.23	37.33	30.40	55.89		
S ₃	H ₁	CV	*0.00152	*0.00198	*0.00154	0.00100	*0.00157	0.00164	*0.00140	*0.00091	*0.00584	*0.00245	*0.00327	
		Emed(pu)	*3.47	*4.04	*3.33	2.35	*3.66	3.45	*2.91	*2.81	*3.90	*3.96	*3.86	
		Emax(pu)	*27.33	*35.62	*19.85	23.34	*33.67	23.91	*20.54	*16.08	*31.90	*30.20	*42.78	
	H ₂	CV	0.00178	0.00237	0.00181	0.00102	0.00176	0.00158	0.00155	0.00113	0.00580	0.00250	0.00330	
		Emed(pu)	4.06	4.99	3.93	2.40	6.54	3.24	3.33	3.42	4.38	4.22	4.06	
		Emax(pu)	32.22	42.00	23.38	23.92	37.44	23.48	22.46	20.38	33.63	29.60	35.33	
H ₃	CV	0.00174	0.00231	0.00198	*0.00098	0.00192	*0.00154	0.00138	0.00117	0.00627	0.00251	0.00346		
	Emed(pu)	4.07	4.82	4.26	*2.26	7.19	*3.16	3.05	3.74	4.59	4.25	4.60		
	Emax(pu)	31.44	40.67	25.77	*22.71	40.72	*22.74	19.79	19.69	37.37	31.00	42.17		

Tabela 4.17: Tabela comparativa com simulações para o modelo F3. Potência Ativa.

$S_1 = 7$ dias $S_2 = 21$ dias $S_3 = 42$ dias.
 $H_1 = 1 \times 0 : 5 \times 20 : 1 \times 4$ $H_2 = 1 \times 0 : 10 \times 20 : 1 \times 4$ $H_3 = 1 \times 0 : 15 \times 20 : 1 \times 4$.
 I (industrial), R (residencial) e C (comercial).

treinamento	topologia	estática	Subestação											
			40 R	80 C	221 R	232 R	379 I	463 R	572 R	623 C	668 R	762 I	840 R	
S ₁	H ₁	CV	0.00251	0.00571	0.00288	0.00761	0.01306	0.00906	0.00418	0.00466	0.01304	0.00409	0.00482	
		Emed(pu)	5.72	8.33	9.00	10.78	14.42	11.43	7.59	9.65	6.19	6.09	7.53	
		Emax(pu)	45.00	55.14	55.60	63.67	80.83	100.00	45.20	53.75	41.15	36.33	44.14	
	H ₂	CV	0.00221	0.00612	0.00284	0.00745	0.01211	0.01139	0.00402	0.00463	0.01298	0.00393	0.00492	
		Emed(pu)	4.85	8.59	9.14	10.38	13.70	12.72	7.26	9.60	6.10	5.83	7.50	
		Emax(pu)	40.25	60.86	51.60	62.42	71.67	100.00	43.90	52.00	40.85	35.00	46.00	
H ₃	CV	0.00224	0.00579	0.00249	0.00730	0.01091	0.01344	0.00401	0.00534	0.01282	0.00393	0.00467		
	Emed(pu)	4.82	8.25	8.10	10.15	12.07	13.13	7.21	10.63	6.03	5.75	7.29		
	Emax(pu)	41.00	57.29	46.20	61.75	65.83	100.00	44.10	63.25	39.69	35.00	43.14		
S ₂	H ₁	CV	0.00172	0.00395	0.00208	0.00563	0.00982	*0.00829	0.00265	0.00331	0.01118	0.00328	0.00319	
		Emed(pu)	3.85	5.54	6.86	7.78	10.97	*9.87	4.73	6.80	4.90	4.67	4.88	
		Emax(pu)	31.13	39.43	38.40	47.33	59.83	*100.00	29.00	38.00	36.23	29.33	29.71	
	H ₂	CV	0.00159	0.00443	0.00226	0.00565	0.01139	0.00907	0.00320	0.00403	0.01224	0.00358	0.00363	
		Emed(pu)	28.87	43.71	41.60	48.25	68.33	100.00	35.00	46.00	39.69	32.67	33.43	
		Emax(pu)	3.69	6.26	7.69	7.52	12.32	0.00968	0.00321	0.00421	0.01333	0.00340	0.00381	
H ₃	CV	32.25	41.14	41.20	46.17	66.83	100.00	34.00	48.75	42.00	31.33	35.00		
	Emed(pu)	*0.00148	*0.00367	*0.00176	*0.00802	*0.00884	0.00828	*0.00226	*0.00291	*0.00962	*0.00810	*0.00266		
	Emax(pu)	*3.29	*5.26	*5.63	*6.73	*9.62	9.49	*3.83	*5.82	*4.39	*4.26	*4.06		
S ₃	H ₁	CV	*26.68	*35.71	*32.80	*43.25	*55.33	100.00	*25.80	*84.25	*30.92	*28.33	*25.00	
		Emed(pu)	0.00163	0.00404	0.00206	0.00531	0.00971	0.00893	0.00260	0.00328	0.01064	0.00324	0.00320	
		Emax(pu)	3.52	5.94	6.76	7.23	10.88	10.81	4.45	6.82	4.81	4.65	4.92	
	H ₂	CV	30.00	39.43	38.00	45.25	59.33	100.00	29.40	37.75	34.31	29.33	29.71	
		Emed(pu)	0.00160	0.00402	0.00215	0.00559	0.01125	0.00987	0.00297	0.00386	0.01104	0.00318	0.00330	
		Emax(pu)	3.49	5.79	7.06	7.61	12.62	11.54	5.18	7.83	5.20	4.55	5.20	
H ₃	CV	29.00	39.43	39.00	47.50	67.50	100.00	33.30	45.25	35.00	29.00	30.00		
	Emed(pu)													

Tabela 4.18: Tabela comparativa com simulações para o modelo F3. Potência Reativa.

$S_1 = 7 \text{ dias } S_2 = 21 \text{ dias } S_3 = 42 \text{ dias.}$
 $H_1 = 1 \times 0 : 5 \times 20 : 1 \times 4 H_2 = 1 \times 0 : 10 \times 20 : 1 \times 4 H_3 = 1 \times 0 : 15 \times 20 : 1 \times 4.$
 Dia de Semana
 $H_1 = 1 \times 0 : 3 \times 20 : 1 \times 4 H_2 = 1 \times 0 : 6 \times 20 : 1 \times 4 H_3 = 1 \times 0 : 9 \times 20 : 1 \times 4.$
 Outro grupo.
 I (industrial), R (residencial) e C (comercial).

treinamento	topologia	estatística	Subestação											
			40 R	80 C	221 R	232 R	379 I	463 R	572 R	623 C	668 R	762 I	840 R	
S ₁	H ₁	CV	0.00350	0.00458	0.00413	0.00212	0.00296	0.00295	0.00255	0.00214	0.01061	0.00427	0.00430	
		Emed(pu)	8.49	10.03	8.67	6.28	10.63	6.43	5.76	5.86	8.46	7.36	7.70	
	Emed(pu)	62.44	80.62	53.54	46.97	64.00	42.35	35.83	40.46	59.77	51.60	73.89		
	CV	0.00356	0.00451	0.00351	0.00191	0.00261	0.00278	0.00245	0.00236	0.01061	0.00421	0.00434		
	Emed(pu)	8.55	9.63	7.79	5.61	9.65	6.02	5.47	5.79	8.38	7.16	7.54		
	Emed(pu)	63.67	79.86	44.38	42.47	54.61	40.26	34.92	46.23	60.80	51.80	75.33		
S ₂	H ₁	CV	0.00358	0.00479	0.00347	0.00174	0.00257	0.00256	0.00221	0.01004	0.00451	0.00421		
		Emed(pu)	8.43	9.62	7.63	5.33	9.37	5.66	5.47	5.56	7.83	7.67	6.98	
	Emed(pu)	64.56	86.19	44.23	38.03	55.06	36.35	35.71	43.31	58.30	55.40	69.39		
	CV	0.00278	0.00348	0.00278	0.00122	0.00223	0.00204	0.00208	0.00155	0.00875	0.00304	0.00395		
	Emed(pu)	6.34	7.37	6.03	3.85	8.33	4.57	4.61	4.26	6.94	5.56	6.03		
	Emed(pu)	50.78	59.62	35.31	26.24	45.50	28.87	29.79	28.38	50.43	36.40	66.83		
S ₃	H ₁	CV	0.00231	0.00371	0.00268	0.00112	0.00174	0.00197	0.00181	0.00171	0.00756	0.00277	0.00370	
		Emed(pu)	5.31	7.92	5.72	3.51	6.51	4.33	4.08	4.29	5.74	5.15	5.17	
	Emed(pu)	41.44	65.43	35.08	24.11	36.72	28.04	25.38	33.38	44.37	32.40	60.67		
	CV	0.00232	0.00378	0.00272	0.00109	0.00174	0.00176	0.00184	0.00159	0.00758	0.00277	0.00363		
	Emed(pu)	5.39	8.09	5.78	3.45	6.39	3.96	4.09	3.93	5.91	5.10	5.10		
	Emed(pu)	42.11	66.67	35.23	23.24	37.33	24.96	26.25	30.23	44.03	32.40	52.06		
S ₃	H ₁	CV	*0.00208	0.00314	0.00220	0.00115	0.00209	0.00176	0.00183	0.00138	0.00798	0.00259	0.00353	
		Emed(pu)	*4.86	6.26	4.90	3.43	8.00	3.96	4.15	4.27	6.26	4.97	4.82	
	Emed(pu)	*37.87	55.43	27.69	25.47	43.67	24.00	25.54	24.31	46.37	30.20	61.56		
	CV	0.00223	*0.00276	0.00218	*0.00097	0.00169	0.00161	0.00158	0.00132	0.00746	*0.00230	0.00344		
	Emed(pu)	5.00	*5.98	4.76	*3.08	6.41	3.58	3.61	3.85	5.75	4.87	4.57		
	Emed(pu)	40.78	*48.98	27.46	*20.79	35.28	22.13	22.12	23.62	41.87	*26.80	57.33		
H ₃	CV	0.00209	0.00337	*0.00210	0.00101	*0.00168	*0.00157	*0.00149	*0.00145	*0.00764	0.00235	*0.00340		
	Emed(pu)	4.86	6.71	*4.88	3.15	*6.16	*3.63	*3.44	*3.85	*5.73	4.46	*4.86		
	Emed(pu)	37.78	59.76	*26.40	21.42	*34.17	*22.26	*20.75	*26.92	*44.87	27.40	*60.28		

Tabela 4.19: Tabela comparativa com simulações para o modelo F4. Potência Ativa.

$S_1 = 7 \text{ dias } S_2 = 21 \text{ dias } S_3 = 42 \text{ dias.}$
 $H_1 = 1 \times 0 : 5 \times 20 : 1 \times 4 H_2 = 1 \times 0 : 10 \times 20 : 1 \times 4 H_3 = 1 \times 0 : 15 \times 20 : 1 \times 4.$
 Dia de Semana
 $H_1 = 1 \times 0 : 3 \times 20 : 1 \times 4 H_2 = 1 \times 0 : 6 \times 20 : 1 \times 4 H_3 = 1 \times 0 : 9 \times 20 : 1 \times 4.$
 Outro grupo.
 I (industrial), R (residencial) e C (comercial).

treinamento	topologia	estafísica	Subestação											
			40 R	80 C	221 R	232 R	379 I	463 R	572 R	623 C	668 R	762 I	840 R	
S ₁	H ₁	CV	0.00292	0.00787	0.00334	0.01080	0.01579	0.02687	0.00561	0.00714	0.01719	0.00564	0.00676	
		Emed(pu)	6.56	17.50	11.31	15.17	17.62	19.74	10.38	14.32	8.17	8.20	10.90	
	Emax(pu)	53.00	72.71	60.60	90.58	96.00	100.00	60.20	84.00	53.85	51.00	60.29		
	CV	0.00297	0.00792	0.00322	0.01017	0.01386	0.03369	0.00589	0.00618	0.01661	0.00529	0.00618		
	Emed(pu)	6.46	12.38	10.80	14.02	15.69	21.53	10.95	12.60	7.89	7.71	9.90		
	Emax(pu)	53.88	74.29	56.60	86.67	83.83	100.00	63.40	71.00	52.38	47.33	55.71		
S ₂	H ₁	CV	0.00298	0.00814	0.00324	0.00962	0.01391	0.03933	0.00598	0.00613	0.01656	0.00546	0.00617	
		Emed(pu)	6.36	12.73	10.99	13.51	15.65	23.04	10.91	12.94	7.72	8.05	9.75	
	Emax(pu)	54.88	74.86	58.00	80.67	84.17	100.00	64.30	69.25	52.54	49.00	55.86		
	CV	0.00224	0.00599	0.00268	0.00792	0.01297	0.01901	0.00468	0.00486	0.01477	0.00396	0.00421		
	Emed(pu)	5.08	9.29	8.95	10.93	14.19	15.63	8.47	10.07	6.82	6.05	6.81		
	Emax(pu)	40.38	56.14	48.00	67.50	78.67	100.00	50.80	54.75	46.00	34.67	38.00		
S ₃	H ₁	CV	0.00230	0.00638	0.00247	0.00712	0.01121	0.01872	0.00428	0.00472	0.01392	0.00350	0.00420	
		Emed(pu)	4.99	10.16	8.23	10.17	12.37	15.85	7.43	9.82	6.25	5.38	6.82	
	Emax(pu)	42.12	57.29	44.40	56.75	69.33	100.00	48.10	52.50	44.92	30.33	36.14		
	CV	0.00227	0.00635	0.00239	0.00700	0.01099	0.02075	0.00423	0.00467	0.01424	0.00334	0.00416		
	Emed(pu)	4.89	10.04	7.93	9.94	12.34	16.28	7.34	9.74	6.35	5.19	6.56		
	Emax(pu)	41.37	58.00	43.80	57.75	66.17	100.00	46.80	51.75	45.85	28.33	37.43		
S ₃	H ₁	CV	0.00196	*0.00513	0.00225	0.00655	0.01202	*0.01102	0.00348	0.00445	0.01607	0.00367	0.00383	
		Emed(pu)	4.40	*7.92	7.62	9.24	13.47	*10.79	6.48	9.35	6.79	5.59	6.29	
	Emax(pu)	35.50	*47.86	40.00	55.00	71.17	*100.00	36.60	50.25	53.00	31.33	34.00		
	CV	*0.00181	0.00527	0.00231	0.00609	0.00954	0.01412	*0.00359	0.00411	0.01413	0.00309	*0.00341		
	Emed(pu)	*4.13	8.23	7.43	8.62	10.89	12.81	*8.84	8.59	6.29	4.81	*5.48		
	Emax(pu)	*32.89	49.29	43.20	51.00	56.50	100.00	*39.80	47.00	45.38	26.33	*30.86		
H ₃	CV	0.00188	0.00518	*0.00218	*0.00604	*0.00664	0.01738	0.00386	*0.00409	*0.01418	*0.00303	0.00360		
	Emed(pu)	4.22	8.06	*7.16	*8.62	*10.88	14.58	6.84	*8.55	*6.17	*4.74	5.71		
		Emax(pu)	34.00	46.29	*40.40	*50.33	*55.67	100.00	42.40	*46.75	*46.46	32.14		

Tabela 4.20: Tabela comparativa com simulações para o modelo F4. Potência Reativa.

4.5.6 Comentários sobre os resultados do modelo F

Os tempos de CPU apresentados para o modelo F1 a partir da referência S_2H_1 foram tomados em uma estação *SUN SPARCstation Ultra*, bem como os tempos do modelo F2 a partir da referência S_3H_1 . Outros tempos dos modelos F2 e demais modelos foram tomados em uma *SUN SPARCstation 4*. Para facilitar a análise as medidas de tempo apresentadas estão referenciadas para uma *SUN SPARCstation 20*.

modelo	referência	tempo médio	modelo	referência	tempo médio
F1	S_1H_1	42.17		S_1H_1	2.82
	S_1H_2	51.06		S_1H_2	3.08
	S_1H_3	21.38		S_1H_3	3.99
	S_2H_1	100.60		S_2H_1	3.06
	S_2H_2	192.76	F3	S_2H_2	2.63
	S_2H_3	167.81		S_2H_3	3.40
	S_3H_1	430.90		S_3H_1	19.42
	S_3H_2	504.43		S_3H_2	7.10
	S_3H_3	729.37		S_3H_3	7.62
F2	S_1H_1	25.23		S_1H_1	14.07
	S_1H_2	35.12		S_1H_2	14.96
	S_1H_3	30.67		S_1H_3	12.55
	S_2H_1	35.02		S_2H_1	38.01
	S_2H_2	61.87	F4	S_2H_2	91.63
	S_2H_3	72.36		S_2H_3	95.47
	S_3H_1	42.26		S_3H_1	107.21
	S_3H_2	85.74		S_3H_2	205.92
	S_3H_3	111.12		S_3H_3	282.00

Tabela 4.21: Tempos de execução em segundos para as diversas variações no modelo F.

A características gerais de desempenho para esse modelo são as seguintes:

- Nota-se que o modelo F2 apresentou o melhor desempenho.
- Da mesma forma que no modelo M o aumento do conjunto de treinamento acarretou uma melhoria no desempenho da previsão. Contudo, para o modelo F o aumento do número de neurônios, de forma geral, não resultou em um erro de previsão melhor, ou seja, independentemente do tamanho do conjunto de treinamento a topologia que apresentou melhor resultado é a que possui menor número de neurônios. Para o modelo F4, porém, se percebe que o conjunto de treinamento S_1 para a previsão de potência reativa apresenta melhor desempenho quando se utiliza mais neurônios em sua topologia.
- O modelo F3 novamente apresentou a pior performance.

4.5.7 Comparações de Resultados

$S_1 = 7$ dias $S_2 = 21$ dias $S_3 = 42$ dias. Modelo M e F. $S_1 =$. Valor fixo para o modelo L.
 $H_1 = 21 : 9 : 1$ $H_2 = 21 : 20 : 1$ $H_3 = 21 : 30 : 1$. (M1, M2 e M3)
 $H_1 = 9 : 4 : 1$ $H_2 = 9 : 8 : 1$ $H_3 = 9 : 18 : 1$. (L3 e L4)
 $H_1 = 1 \times 0 : 5 \times 20 : 1 \times 4$ $H_2 = 1 \times 0 : 10 \times 20 : 1 \times 4$ $H_3 = 1 \times 0 : 15 \times 20 : 1 \times 4$. (F1 e F3). Dia de Semana. (F2 e F4)
 $H_1 = 1 \times 0 : 3 \times 20 : 1 \times 4$ $H_2 = 1 \times 0 : 6 \times 20 : 1 \times 4$ $H_3 = 1 \times 0 : 9 \times 20 : 1 \times 4$. Outro grupo (F2 e F4).
 I (industrial), R (residencial) e C (comercial).

modelo	esta- tística	Subestação											
		40 R	80 C	221 R	232 R	379 I	463 R	572 R	623 C	668 R	762 I	840 R	
M1	ref	* $S_3 H_2$	* $S_3 H_1$	* $S_3 H_1$	* $S_3 H_2$	* $S_3 H_1$	* $S_3 H_2$	* $S_3 H_1$	* $S_3 H_1$	* $S_3 H_2$	* $S_3 H_1$	* $S_3 H_2$	
	CV	*0.00086	*0.00177	*0.00106	*0.00049	*0.00098	*0.00080	*0.00082	*0.00064	0.00403	*0.00148	*0.00308	
	Emed	*2.15	*3.12	*2.32	*1.54	*3.55	*2.02	*1.77	*2.03	3.05	*2.57	*2.80	
	Emax	*14.89	*32.43	*13.62	*10.50	*20.44	*12.70	*11.98	*11.23	23.73	*17.20	*42.67	
F1	ref	$S_2 H_1$	$S_3 H_2$	$S_3 H_1$	$S_3 H_2$	$S_3 H_1$	$S_3 H_2$	$S_3 H_1$	$S_3 H_1$	$S_2 H_1$	$S_3 H_3$	$S_3 H_1$	
	CV	0.00119	0.00195	0.00142	0.00099	0.00142	0.00138	0.00108	0.00076	0.00398	0.00252	0.00316	
	Emed	2.99	4.38	3.13	2.43	5.34	2.78	2.40	2.42	3.13	4.51	3.34	
	Emax	20.89	33.52	17.69	21.50	29.67	20.57	15.54	13.23	23.13	29.00	47.28	
M2	ref	$S_3 H_2$	$S_3 H_1$	$S_2 H_3$	$S_3 H_2$	$S_3 H_1$							
	CV	0.00125	0.00199	0.00118	0.00054	0.00122	0.00101	0.00093	0.00068	0.00399	0.00154	0.00310	
	Emed	2.78	3.54	2.45	1.68	4.49	2.23	1.99	2.13	2.93	2.77	2.96	
	Emax	21.78	36.67	15.62	11.68	25.89	14.30	13.46	11.85	23.60	18.00	43.50	
F2	ref	$S_3 H_3$	$S_3 H_1$	$S_3 H_1$	$S_3 H_3$	$S_3 H_1$	$S_3 H_1$	$S_3 H_1$	$S_3 H_1$	* $S_3 H_1$	$S_3 H_1$	$S_3 H_1$	
	CV	0.00125	0.00179	0.00132	0.00096	0.00129	0.00134	0.00110	0.00080	*0.00409	0.00216	0.00318	
	Emed	3.01	3.47	2.86	2.15	4.66	2.67	2.39	2.50	*2.87	3.54	3.38	
	Emax	22.11	31.57	16.46	21.53	28.11	19.57	16.00	13.92	*24.67	26.40	41.89	
M3	ref	$S_3 H_3$	$S_3 H_3$	$S_3 H_3$	$S_3 H_3$	$S_3 H_2$	$S_3 H_3$	$S_3 H_1$	$S_3 H_3$	$S_3 H_3$	$S_3 H_1$	$S_3 H_3$	
	CV	0.00172	0.00200	0.00142	0.00100	0.00150	0.00143	0.00124	0.00083	0.00535	0.00228	0.00320	
	Emed	3.52	3.85	2.85	2.07	5.25	2.59	2.46	2.33	3.65	3.32	3.18	
	Emax	32.22	36.19	19.08	24.18	33.17	23.13	18.29	15.62	32.33	28.20	46.00	
F3	ref	$S_3 H_1$	$S_3 H_2$	$S_3 H_3$	$S_3 H_2$	$S_3 H_3$	$S_3 H_2$	$S_3 H_3$					
	CV	0.00208	0.00276	0.00210	0.00097	0.00163	0.00157	0.00149	0.00145	0.00754	0.00230	0.00340	
	Emed	4.85	5.98	4.68	3.08	6.16	3.53	3.44	3.85	5.73	4.37	4.36	
	Emax	37.67	48.38	26.46	20.79	34.17	22.26	20.75	26.92	44.37	26.80	60.28	
L3	ref	$S_1 H_1$	$S_1 H_2$	$S_1 H_1$	$S_1 H_1$								
	CV	0.00211	0.00228	0.00186	0.00117	0.00165	0.00164	0.00148	0.00098	0.00630	0.00200	0.00317	
	Emed	4.96	5.00	4.11	3.14	5.74	3.54	3.13	2.95	4.77	3.48	3.98	
	Emax	38.00	40.10	23.69	26.63	36.67	23.48	21.75	17.54	35.53	24.40	54.22	
L4	ref	$S_1 H_1$	$S_1 H_2$	$S_1 H_1$	$S_1 H_1$								
	CV	0.00196	0.00230	0.00164	0.00123	0.00162	0.00160	0.00139	0.00092	0.00542	0.00213	0.00315	
	Emed	4.56	4.56	3.52	2.95	5.46	3.19	2.82	2.82	3.83	3.40	3.76	
	Emax	34.78	41.38	20.69	27.97	36.61	23.52	20.71	16.62	32.43	26.60	46.61	
F4	ref	$S_3 H_1$	$S_3 H_1$	$S_3 H_1$	$S_3 H_3$	$S_3 H_1$	$S_3 H_3$	$S_3 H_1$					
	CV	0.00152	0.00198	0.00154	0.00098	0.00157	0.00154	0.00140	0.00091	0.00534	0.00245	0.00327	
	Emed	3.47	4.04	3.33	2.26	5.66	3.16	2.91	2.81	3.90	3.95	3.86	
	Emax	27.33	35.62	19.85	22.71	33.67	22.74	20.54	16.08	31.90	30.20	42.78	

Tabela 4.22: Comparação entre os modelos que apresentaram melhor desempenho. Potência Ativa.

$S_1 = 7$ dias $S_2 = 21$ dias $S_3 = 42$ dias. Modelo M e F. Valor fixo para o modelo L.
 $H_1 = 21 : 9 : 1$ $H_2 = 21 : 20 : 1$ $H_3 = 21 : 30 : 1$. (M1, M2 e M3)
 $H_1 = 9 : 4 : 1$ $H_2 = 9 : 8 : 1$ $H_3 = 9 : 18 : 1$. (L3 e L4)
 $H_1 = 1 \times 0 : 5 \times 20 : 1 \times 4$ $H_2 = 1 \times 0 : 10 \times 20 : 1 \times 4$ $H_3 = 1 \times 0 : 15 \times 20 : 1 \times 4$. (F1 e F3). Dia de Semana. (F2 e F4)
 $H_1 = 1 \times 0 : 3 \times 20 : 1 \times 4$ $H_2 = 1 \times 0 : 6 \times 20 : 1 \times 4$ $H_3 = 1 \times 0 : 9 \times 20 : 1 \times 4$. Outro grupo. (F2 e F4)
 I (industrial), R (residencial) e C (comercial).

modelo	Subestação													
	40 R	80 C	221 R	232 R	379 I	463 R	572 R	623 C	668 R	762 I	840 R			
M1	ref	* $S_3 H_1$	* $S_3 H_1$	* $S_3 H_2$	* $S_3 H_1$	$S_3 H_1$	* $S_2 H_1$	* $S_3 H_1$	* $S_3 H_1$	* $S_3 H_1$	* $S_3 H_1$			
	CV	*0.00077	*0.00229	*0.00127	*0.00358	*0.00627	*0.00164	*0.00220	*0.00688	*0.00202	*0.00206			
	Emed	*1.99	*3.51	*4.12	*5.05	*7.11	6.73	*2.85	*4.45	*3.20	*2.92			
	Emax	*13.13	*21.86	*23.20	*30.08	*37.83	0.00	*18.00	*25.50	*21.77	*18.33			
F1	ref	$S_3 H_2$	$S_3 H_1$	$S_3 H_1$	$S_3 H_2$	$S_3 H_1$	$S_3 H_1$	$S_3 H_1$	$S_3 H_1$	$S_2 H_1$	$S_3 H_1$			
	CV	0.00110	0.00313	0.00174	0.00557	0.00768	0.00767	0.00195	0.00245	0.00789	0.00325			
	Emed	2.80	4.65	5.81	7.88	8.67	8.60	3.37	4.98	3.68	4.69			
	Emax	19.00	30.00	31.40	46.42	46.17	0.00	22.00	27.50	25.08	29.33			
M2	ref	$S_3 H_1$	$S_3 H_1$	$S_3 H_1$	$S_3 H_2$	$S_3 H_1$	$S_3 H_1$	$S_3 H_1$	$S_3 H_1$	$S_3 H_2$	$S_3 H_1$			
	CV	0.00127	0.00283	0.00168	0.00377	0.00681	0.00988	0.00198	0.00252	0.00778	0.00221			
	Emed	2.38	4.14	5.18	5.36	7.60	6.38	3.58	5.13	3.56	3.15			
	Emax	23.25	27.14	32.60	30.75	41.33	0.00	21.70	29.50	25.00	19.00			
F2	ref	$S_3 H_3$	$S_3 H_1$	$S_3 H_1$	$S_3 H_1$	$S_3 H_1$	$S_3 H_1$	$S_3 H_1$	$S_3 H_1$	$S_3 H_1$	$S_3 H_1$			
	CV	0.00104	0.00289	0.00175	0.00477	0.00746	0.00791	0.00207	0.00244	0.00864	0.00300			
	Emed	2.54	3.97	5.69	6.38	8.28	8.43	3.47	4.82	3.77	4.17			
	Emax	18.25	29.00	32.60	40.33	45.17	0.00	23.60	29.25	28.31	27.67			
M3	ref	$S_3 H_3$	$S_3 H_1$	$S_3 H_1$	$S_3 H_1$	$S_3 H_1$	$S_3 H_1$	$S_3 H_1$	$S_3 H_1$	$S_3 H_3$	$S_3 H_1$			
	CV	0.00141	0.00371	0.00173	0.00466	0.00721	0.00852	0.00281	0.00274	0.00973	0.00268			
	Emed	2.90	5.09	5.01	5.64	7.75	6.15	4.27	5.40	4.12	3.91			
	Emax	26.00	37.29	34.40	42.17	45.50	0.00	33.40	32.25	32.23	27.67			
F3	ref	$S_3 H_2$	$S_3 H_1$	$S_3 H_3$	$S_3 H_3$	$S_3 H_3$	$S_3 H_1$	$S_3 H_2$	$S_3 H_3$	$S_3 H_3$	$S_3 H_2$			
	CV	0.00181	0.00513	0.00218	0.00604	0.00954	0.01102	0.00359	0.00409	0.01418	0.00303			
	Emed	4.13	7.92	7.16	8.62	10.83	10.79	6.34	8.55	6.17	4.74			
	Emax	32.38	47.86	40.40	50.33	55.67	0.00	39.80	46.75	46.46	25.67			
L3	ref	$S_1 H_1$	$S_1 H_1$	$S_1 H_1$	$S_1 H_1$	$S_1 H_1$	$S_1 H_1$	$S_1 H_1$	$S_1 H_2$	$S_1 H_1$	$S_1 H_1$			
	CV	0.00185	0.00486	0.00183	0.00631	0.00945	0.00696	0.00330	0.00365	0.01144	0.00291			
	Emed	3.77	7.65	6.06	9.01	10.15	6.48	6.11	7.67	5.05	4.20			
	Emax	34.13	45.57	33.20	52.67	59.33	78.00	35.20	38.75	37.23	25.67			
L4	ref	$S_1 H_1$	$S_1 H_1$	$S_1 H_1$	$S_1 H_1$	$S_1 H_1$	* $S_1 H_1$	$S_1 H_1$	$S_1 H_2$	$S_1 H_1$	$S_1 H_1$			
	CV	0.00170	0.00449	0.00182	0.00622	0.00854	*0.00643	0.00295	0.00319	0.00966	0.00303			
	Emed	3.53	6.57	5.68	8.49	9.29	*5.92	5.20	6.63	4.22	4.15			
	Emax	31.38	42.86	34.80	52.75	52.50	*95.56	31.80	36.50	31.62	28.33			
F4	ref	$S_3 H_1$	$S_3 H_1$	$S_3 H_1$	$S_3 H_1$	$S_3 H_1$	$S_2 H_1$	$S_3 H_1$	$S_3 H_1$	$S_3 H_1$	$S_3 H_1$			
	CV	0.00148	0.00367	0.00175	0.00502	0.00884	0.00829	0.00226	0.00291	0.00962	0.00310			
	Emed	3.29	5.26	5.63	6.73	9.62	9.37	3.83	5.82	4.39	4.26			
	Emax	26.63	35.71	32.80	43.25	55.33	0.00	25.80	34.25	30.92	28.33			

Tabela 4.23: Comparação entre os modelos que apresentaram melhor desempenho. Potência Reativa.

4.5.8 Gráficos dos Resultados

Gráficos comparando os valores medidos e previstos de carga e outras relações de erro serão apresentados apenas para os modelos cujos índices *Emed* apresentam os menores valores, evitando um excesso de informação. As curvas de carga nos gráficos estão dispostas na seguinte ordem: o primeiro gráfico na página contém as curvas de carga de terça a sexta, nessa ordem; o segundo gráfico indica o erro de previsão correspondente; o terceiro gráfico contém as curvas de carga de sábado a segunda, nessa ordem e o quarto gráfico o erro de previsão correspondente ao gráfico anterior. **Nos gráficos, a carga prevista é a curva pontilhada e a carga medida a curva cheia.** Os gráficos apresentados correspondem a 4 semanas de simulação.

Os gráficos de percentil indicam o percentual do número total de medidas de erro de previsão que estão abaixo de um determinado valor de erro médio absoluto. Assim, por exemplo, ao se tomar o gráfico de percentil das figuras 4.5 e 4.6 vê-se que 80% dos erros estão abaixo do valor 4%.

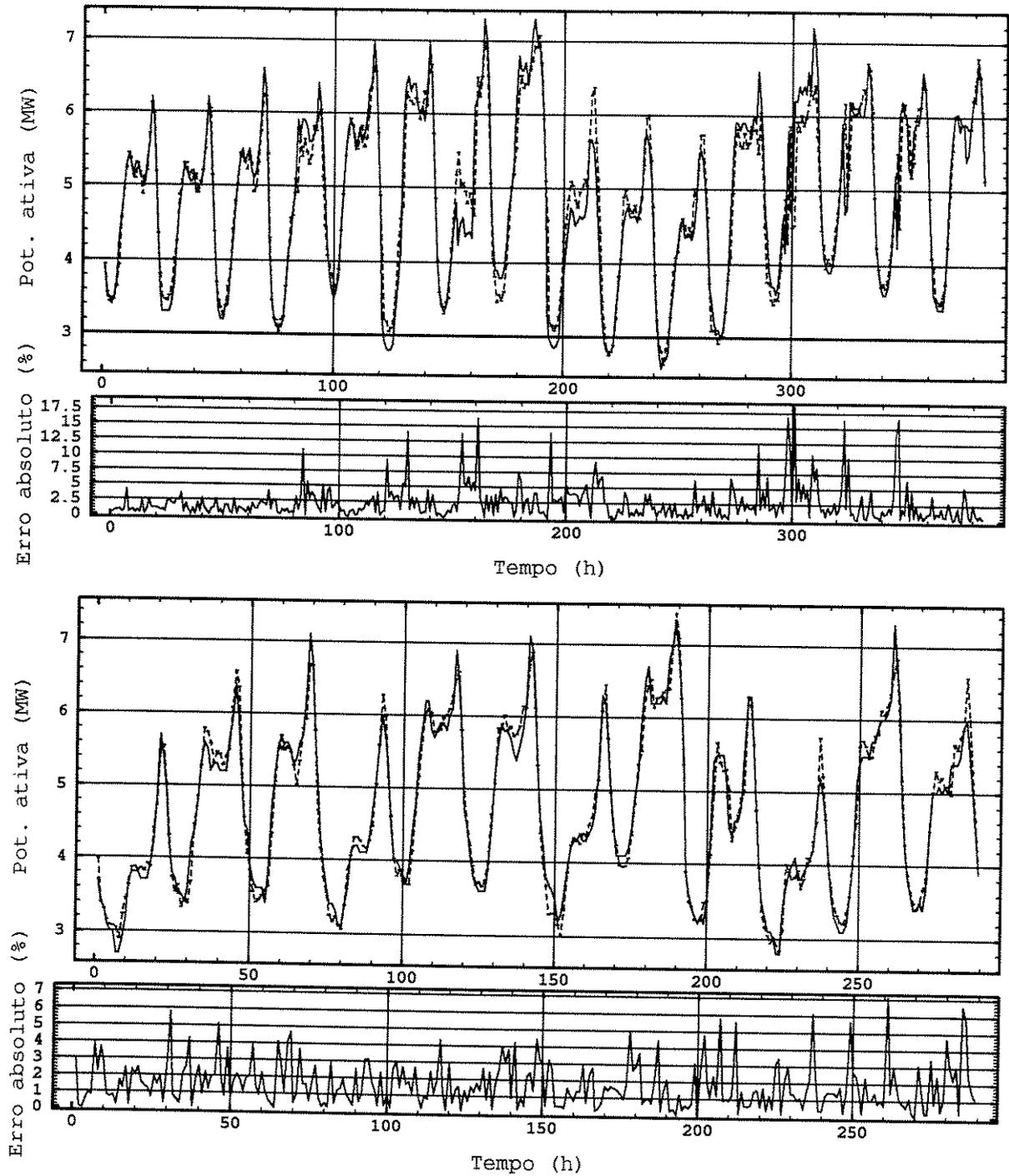


Figura 4.4: Potência ativa prevista, real e erro. Subestação 40. Modelo M1.

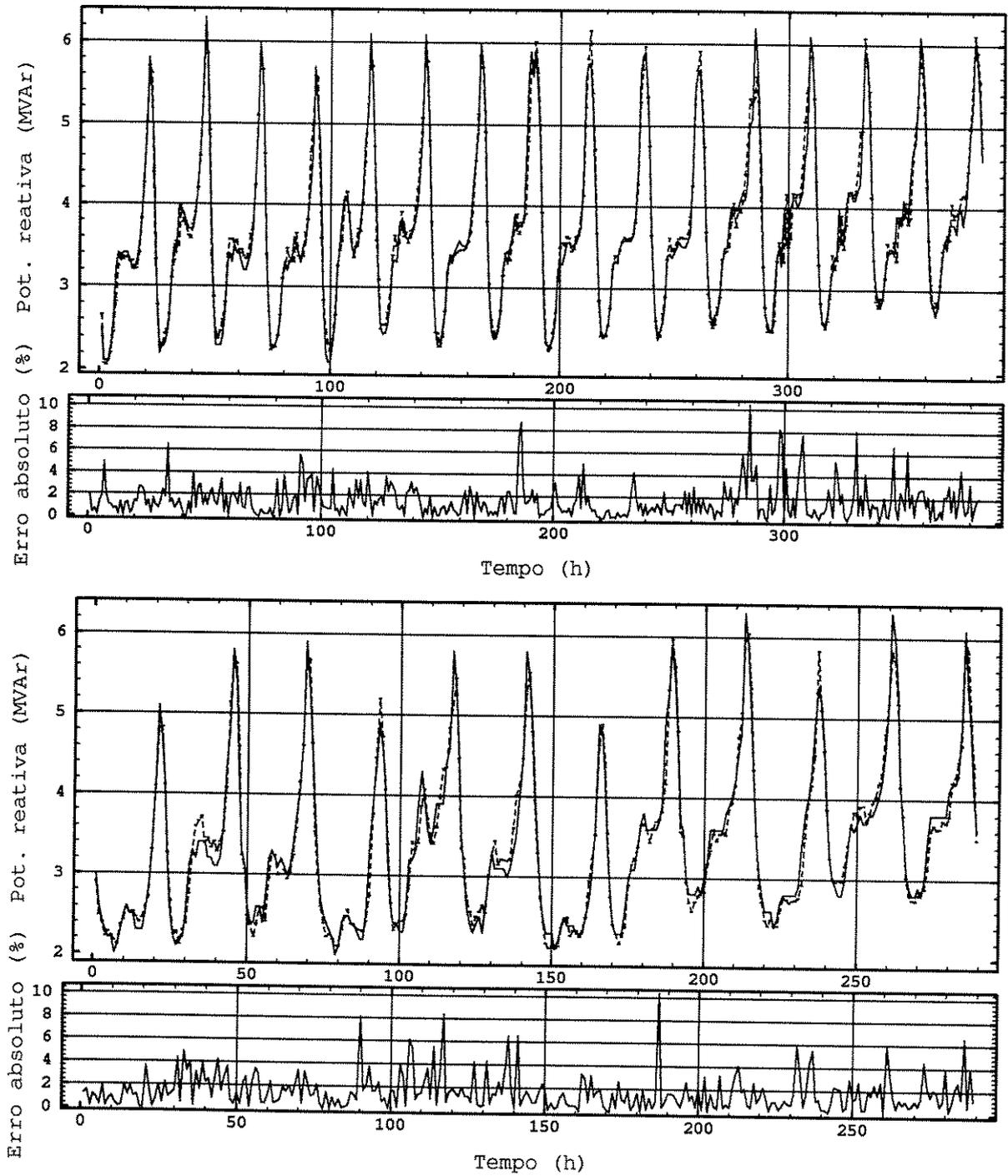


Figura 4.5: Potência reativa prevista, real e erro. Subestação 40. Modelo M1.

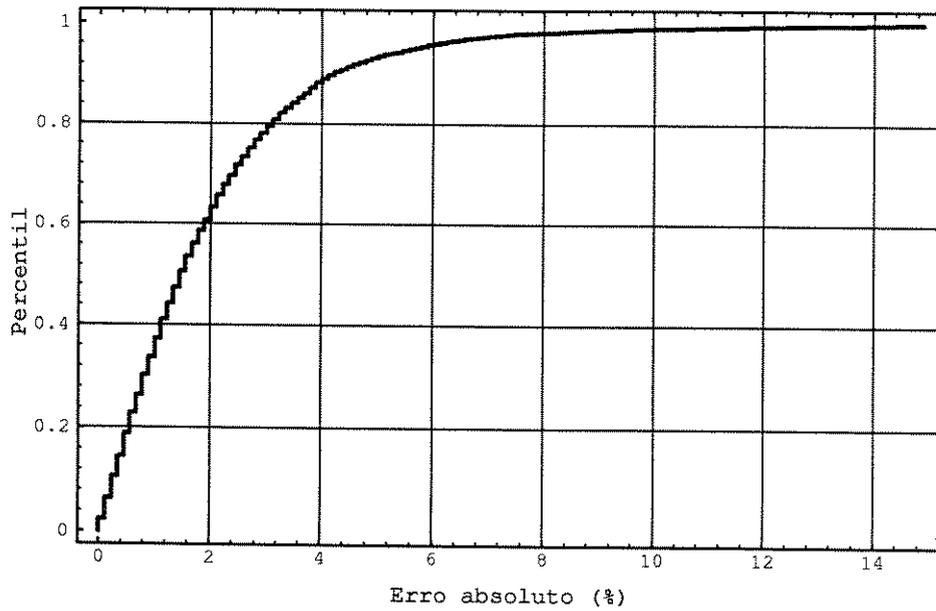


Figura 4.6: Gráfico da porcentagem de medidas em função do erro absoluto. Barra 40, potência ativa. Modelo M1

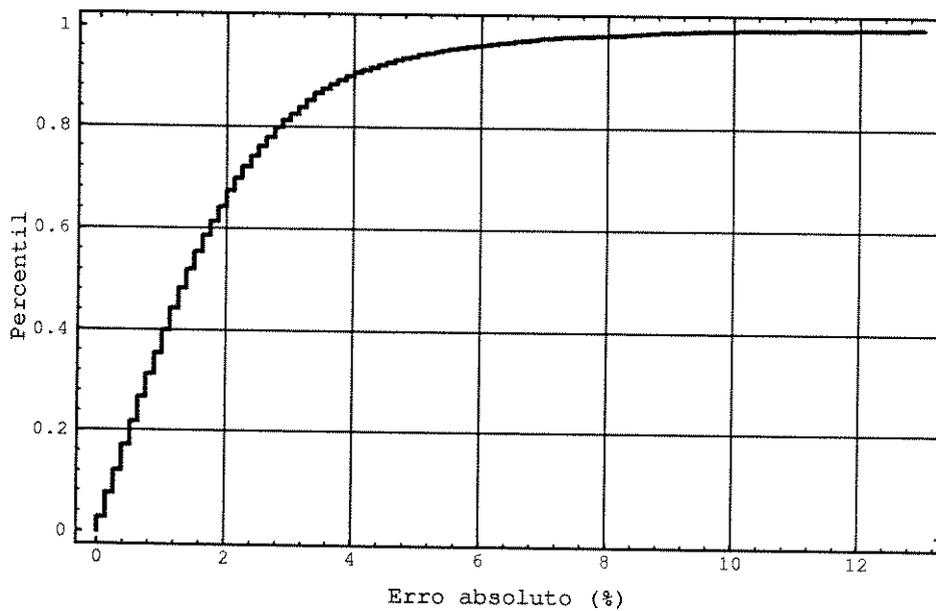


Figura 4.7: Gráfico da porcentagem de medidas em função do erro absoluto. Barra 40, potência reativa. Modelo M1

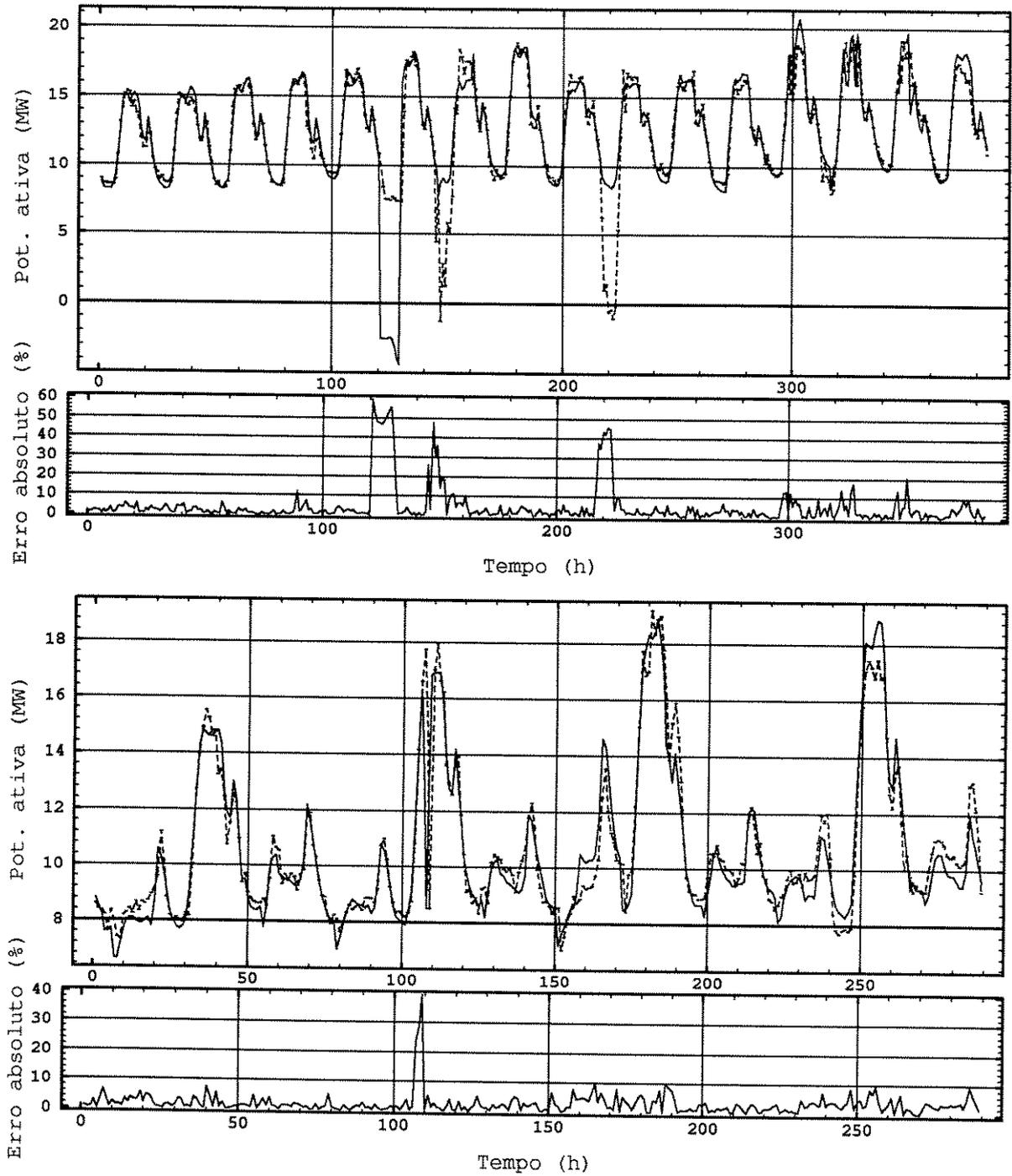


Figura 4.8: Potência ativa prevista, real e erro. Subestação 80. Modelo M1.

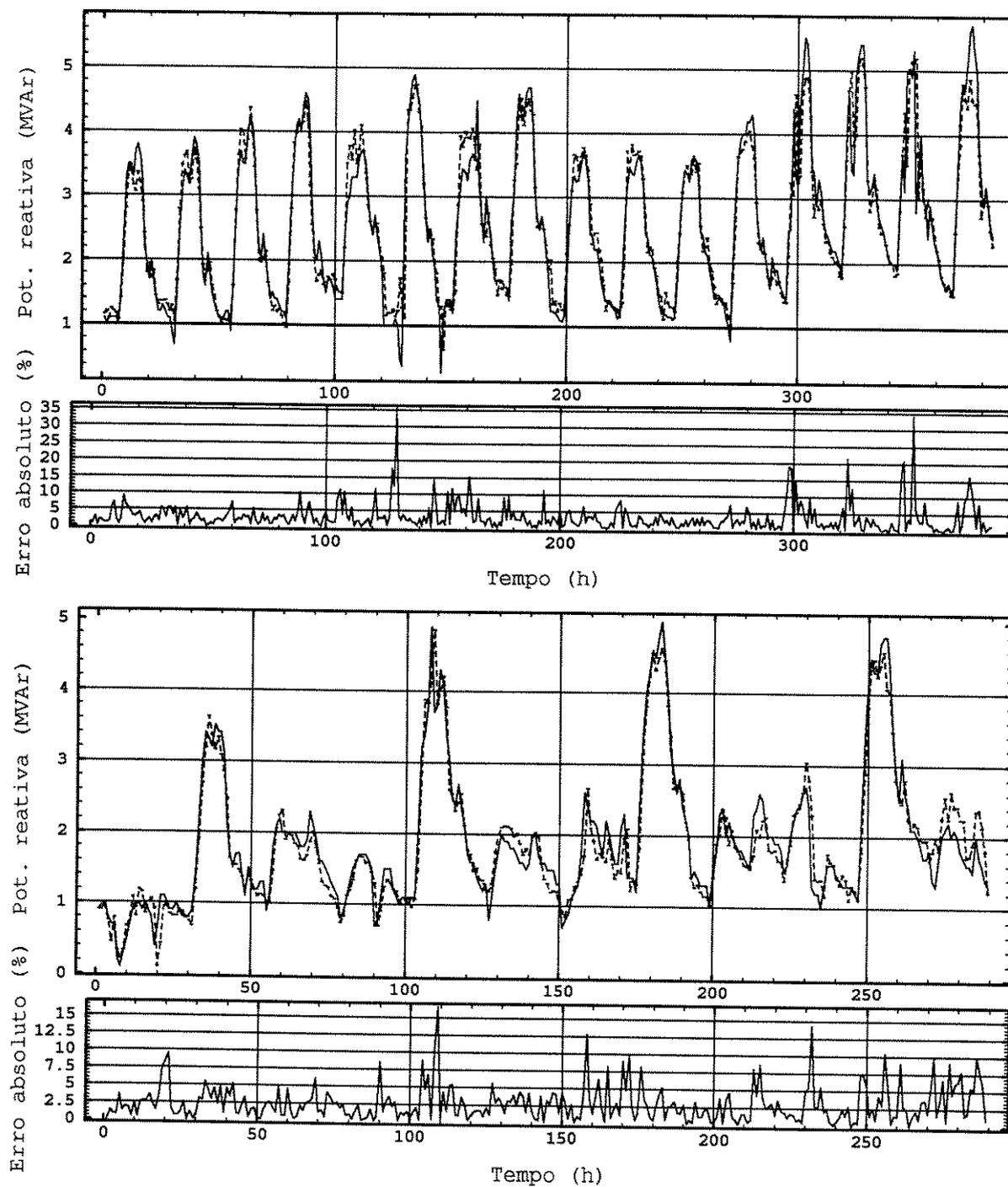


Figura 4.9: Potência reativa prevista, real e erro. Subestação 80. Modelo M1.

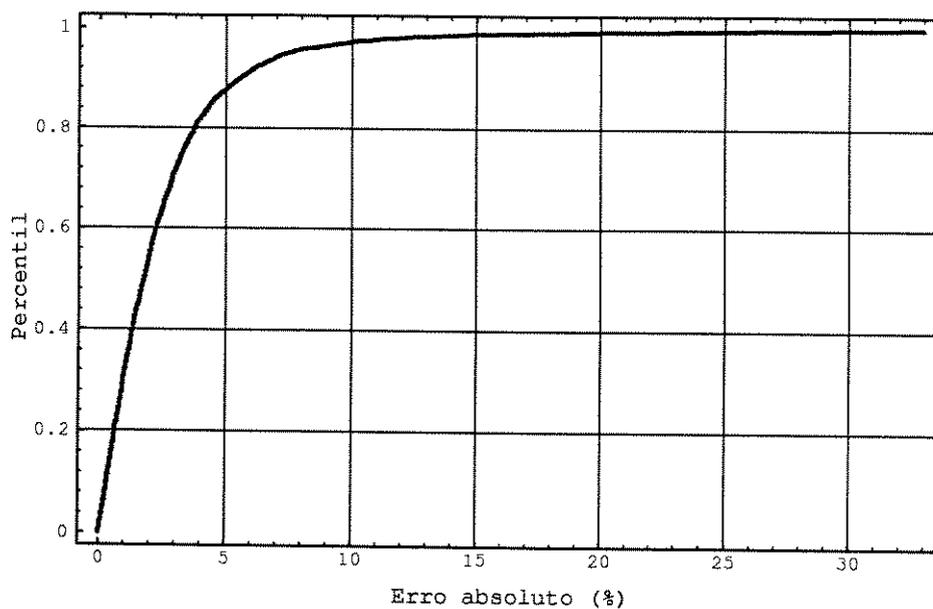


Figura 4.10: Gráfico da porcentagem de medidas em função do erro absoluto. Barra 80, potência ativa. Modelo M1

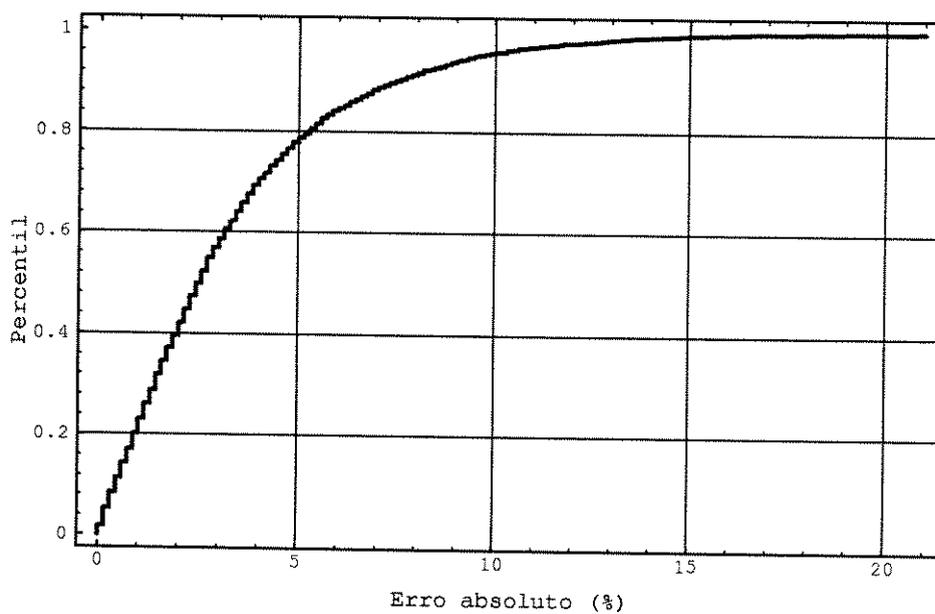


Figura 4.11: Gráfico da porcentagem de medidas em função do erro absoluto. Barra 80, potência reativa. Modelo M1

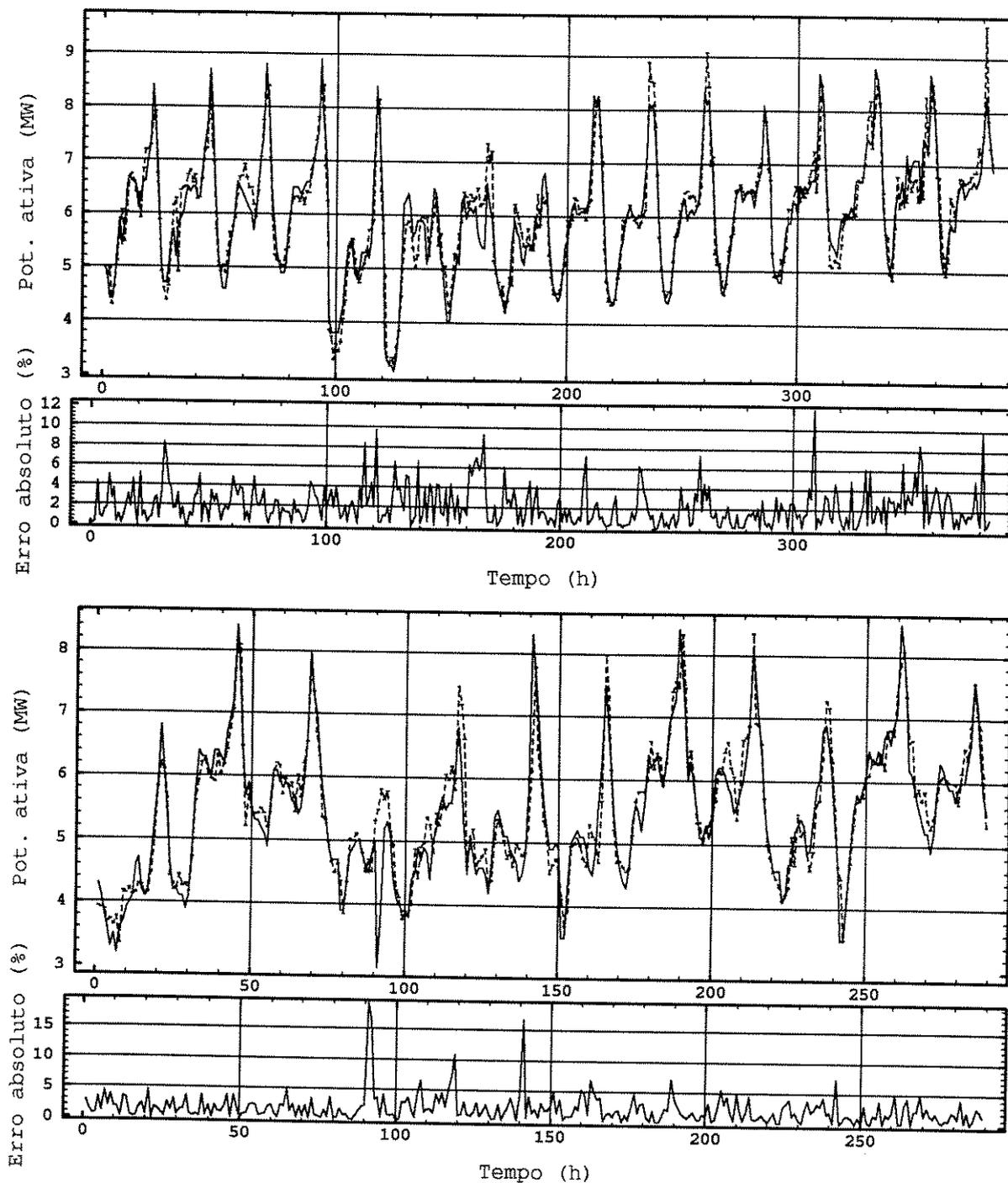


Figura 4.12: Potência ativa prevista, real e erro. Subestação 221. Modelo M1.

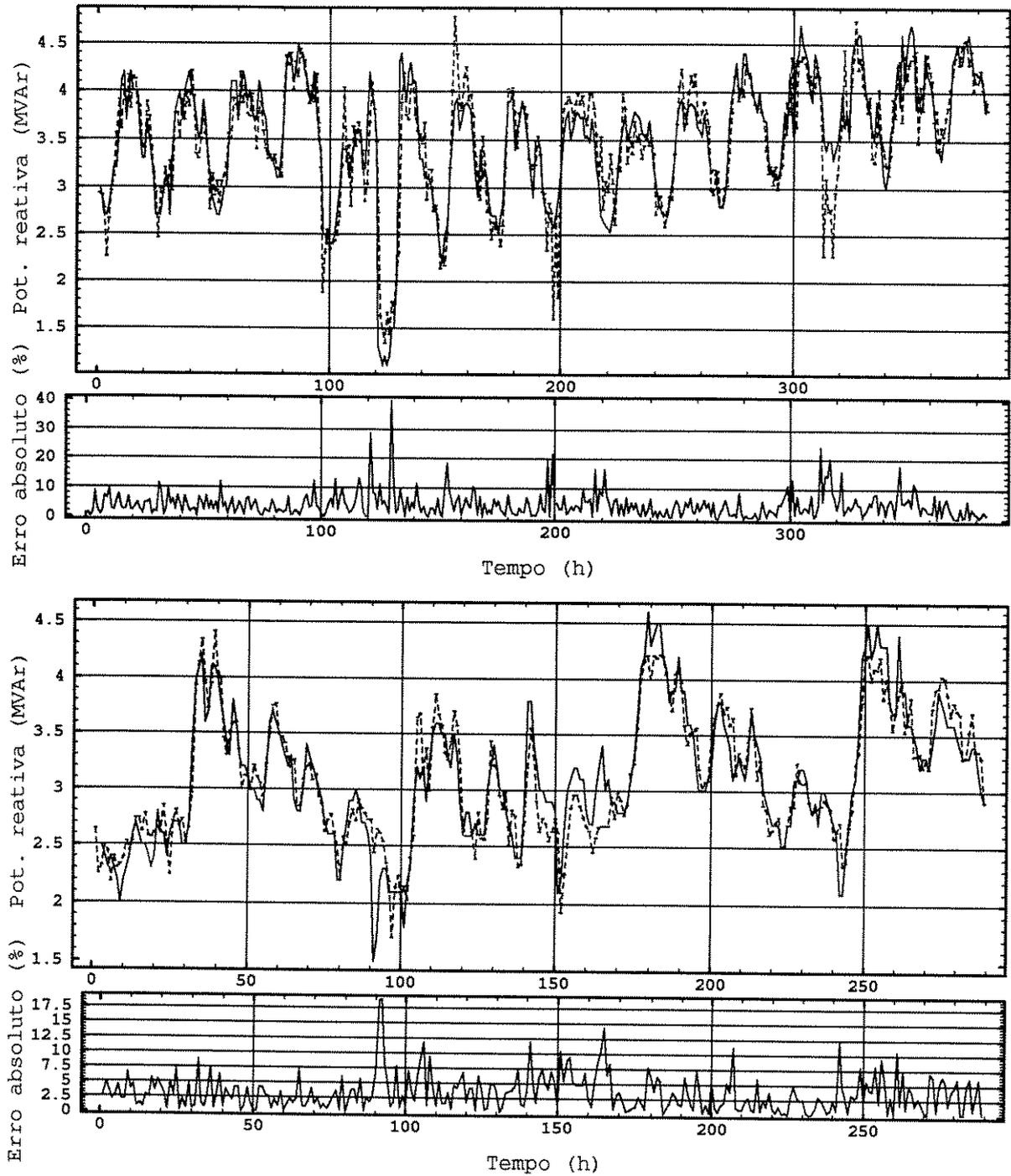


Figura 4.13: Potência reativa prevista, real e erro. Subestação 221. Modelo M1.

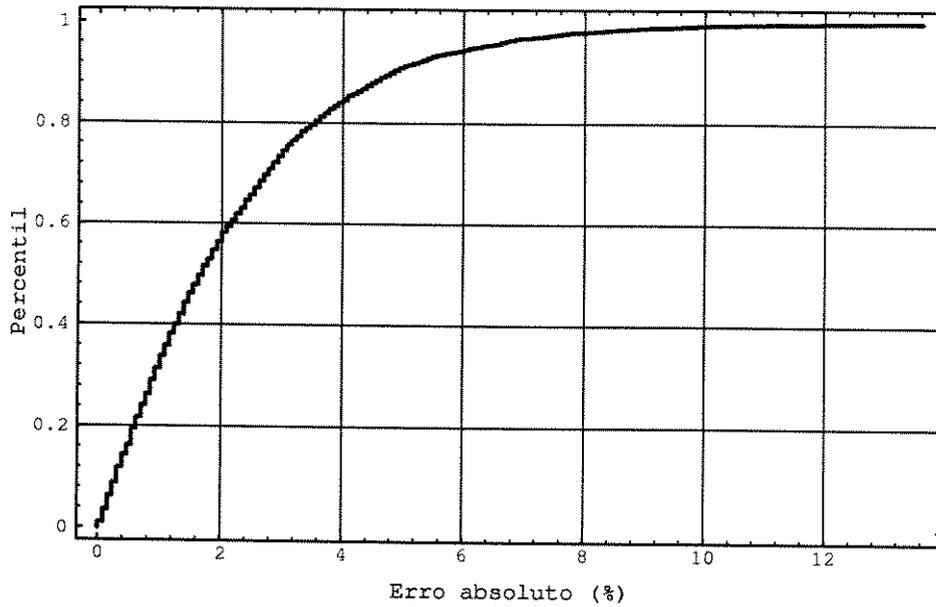


Figura 4.14: Gráfico da percentagem de medidas em função do erro absoluto. Barra 221, potência ativa. Modelo M1

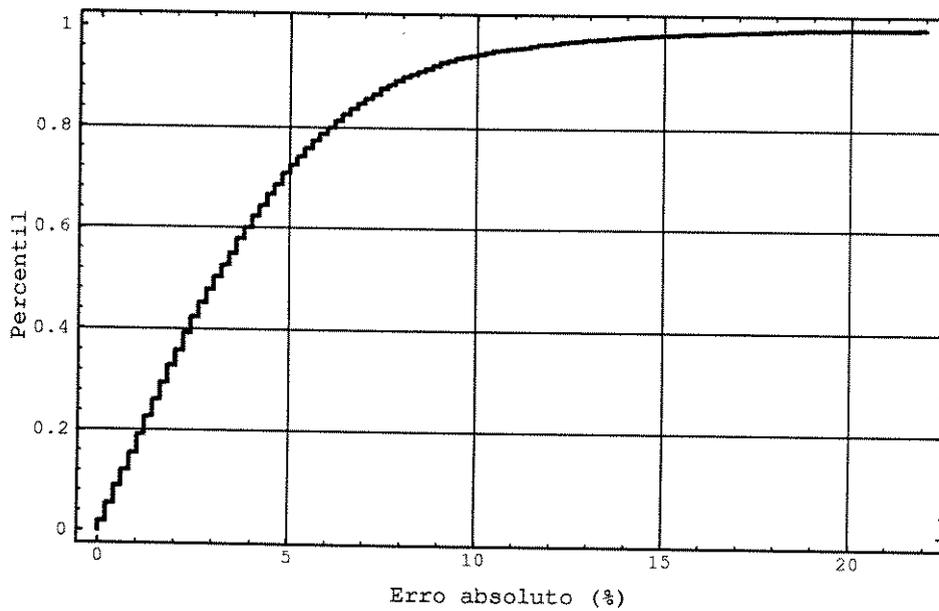


Figura 4.15: Gráfico da percentagem de medidas em função do erro absoluto. Barra 221, potência reativa. Modelo M1

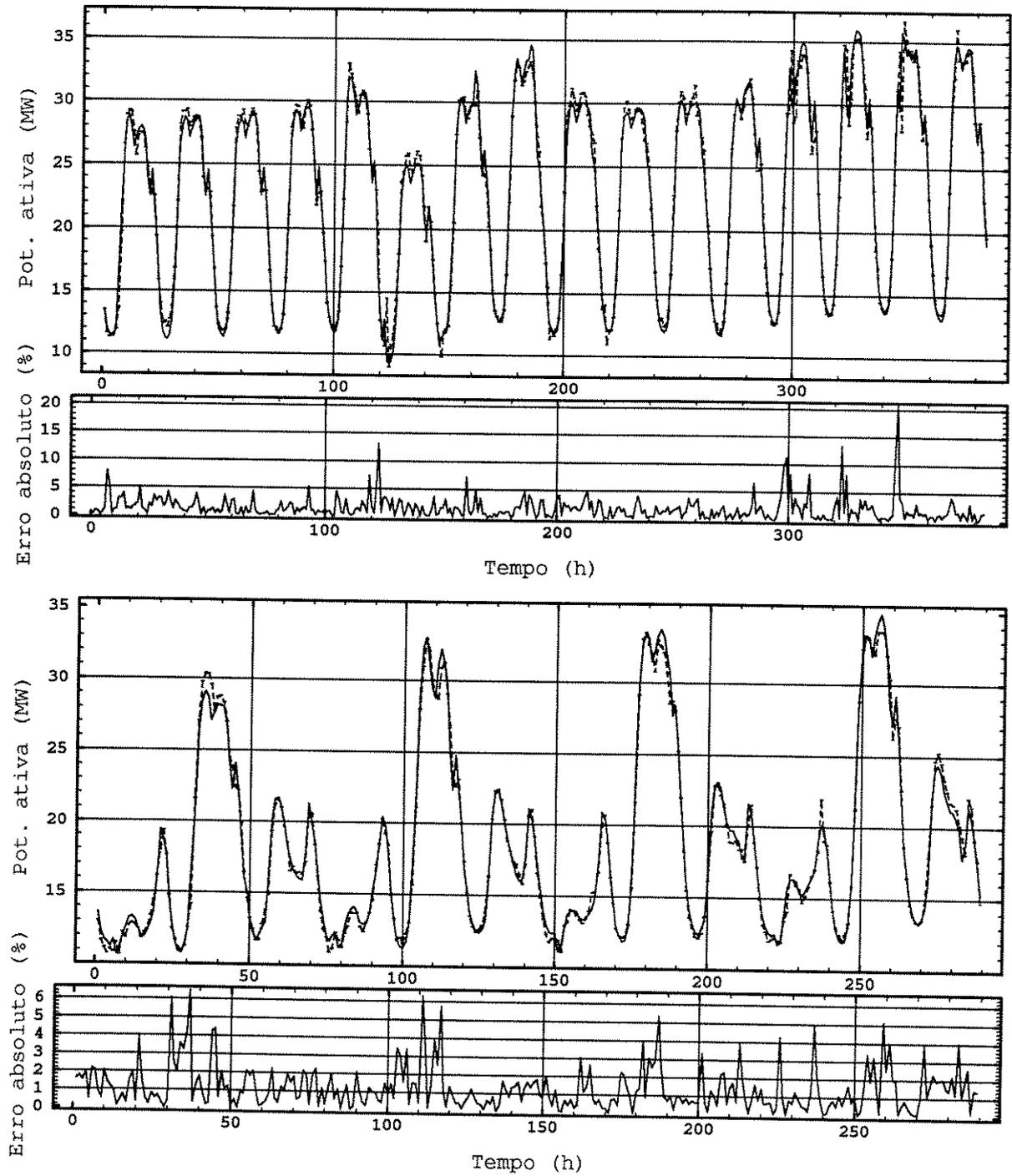


Figura 4.16: Potência ativa prevista, real e erro. Subestação 232. Modelo M1.

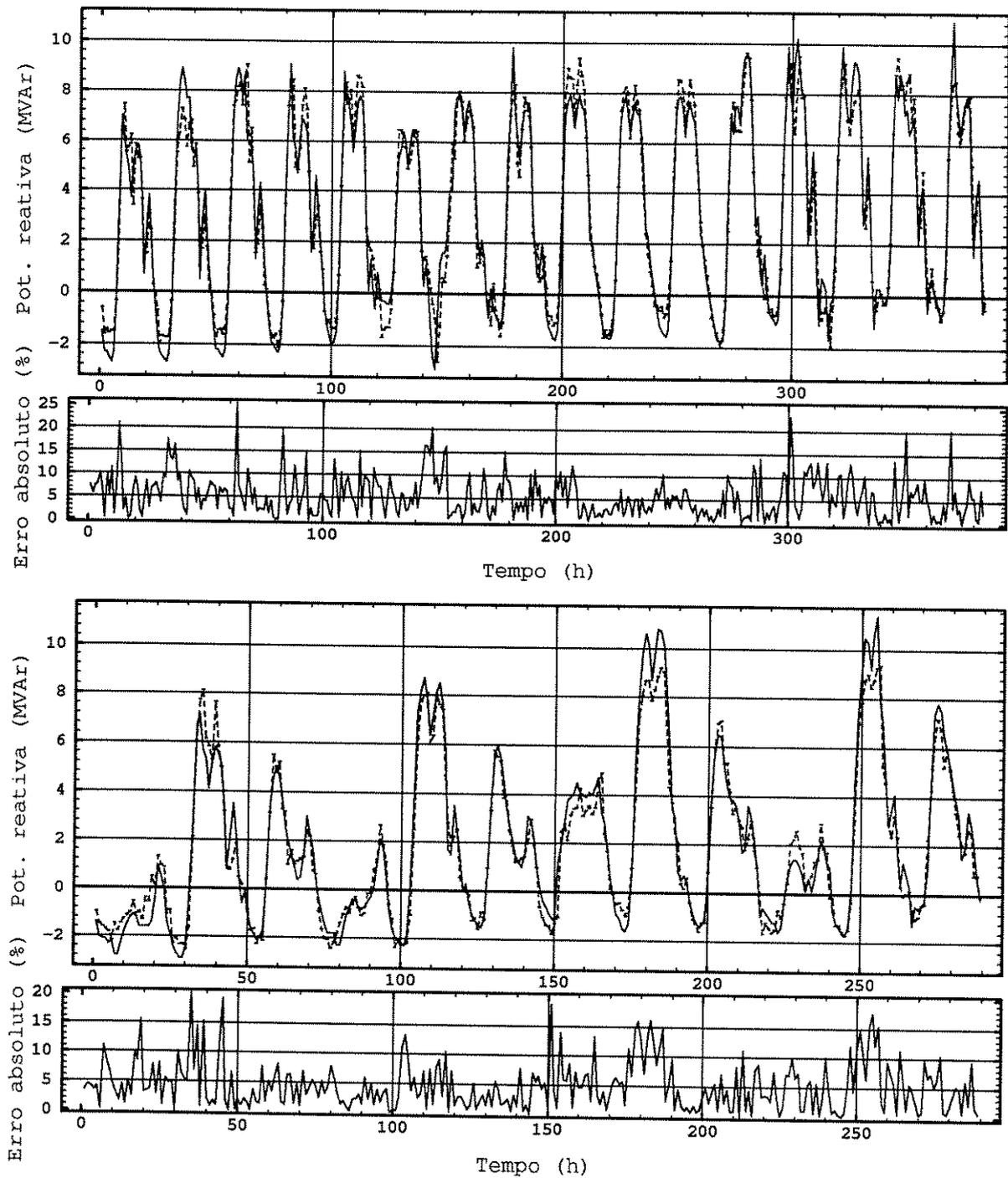


Figura 4.17: Potência reativa prevista, real e erro. Subestação 232. Modelo M1.

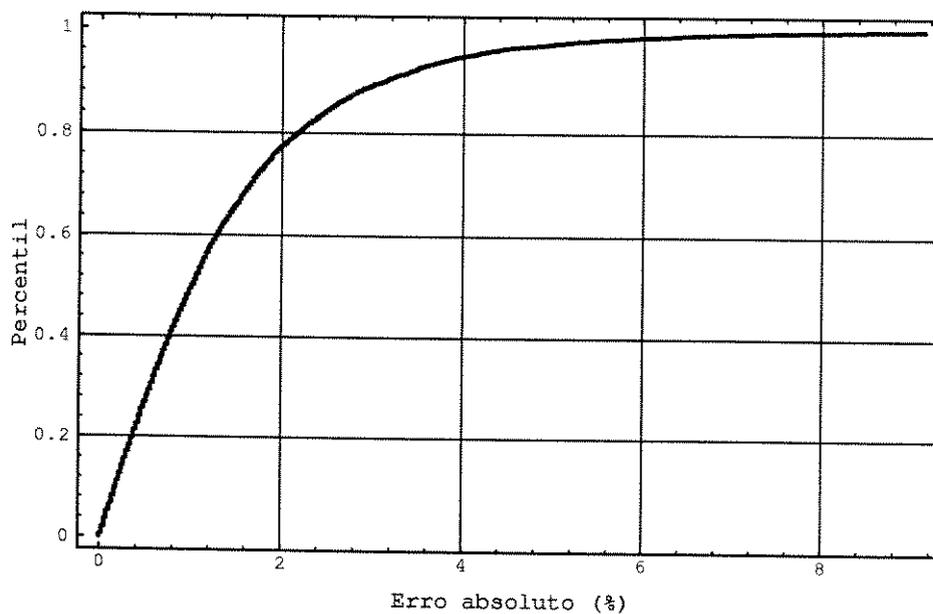


Figura 4.18: Gráfico da porcentagem de medidas em função do erro absoluto. Barra 232, potência ativa. Modelo M1

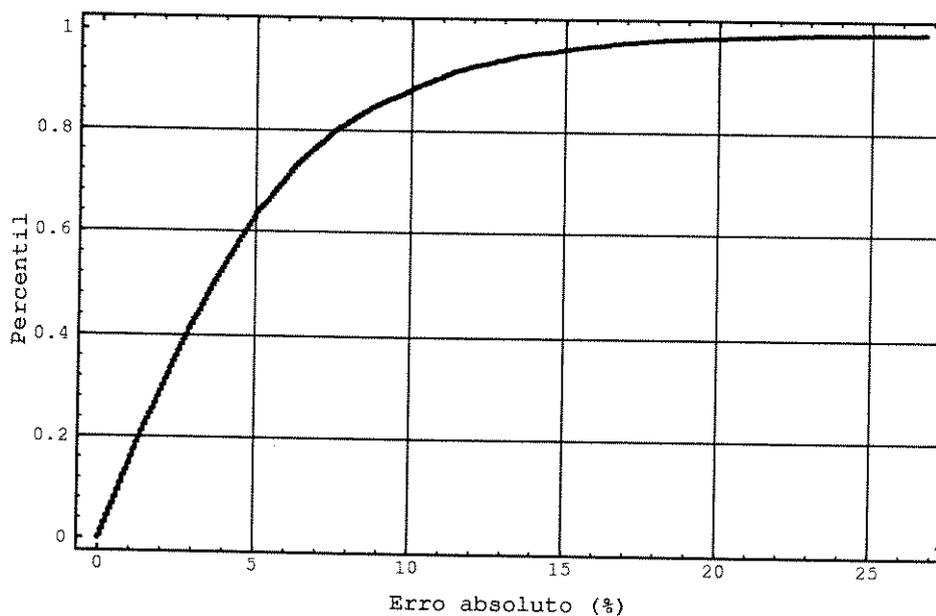


Figura 4.19: Gráfico da porcentagem de medidas em função do erro absoluto. Barra 232, potência reativa. Modelo M1

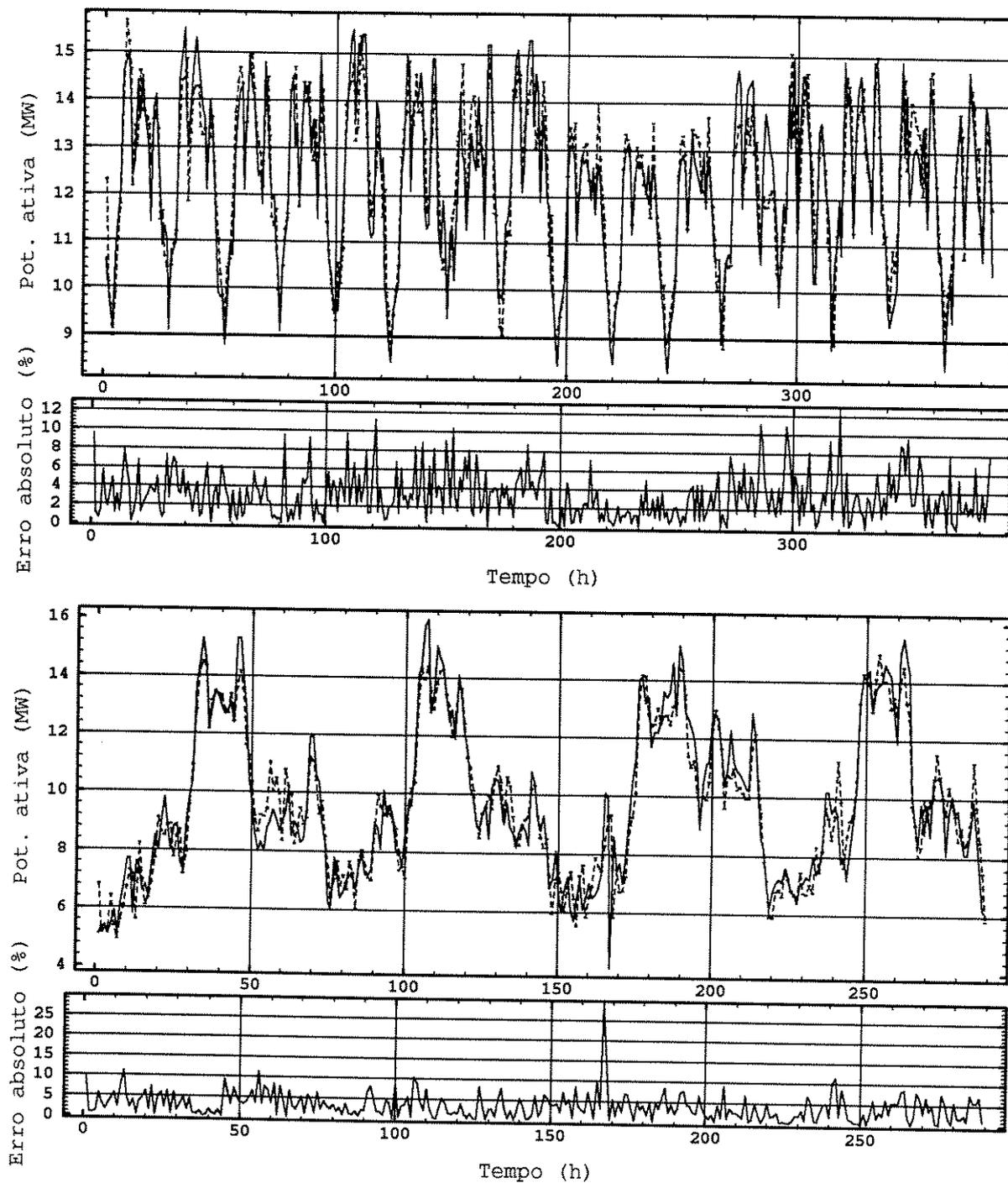


Figura 4.20: Potência ativa prevista, real e erro. Subestação 379. Modelo M1.

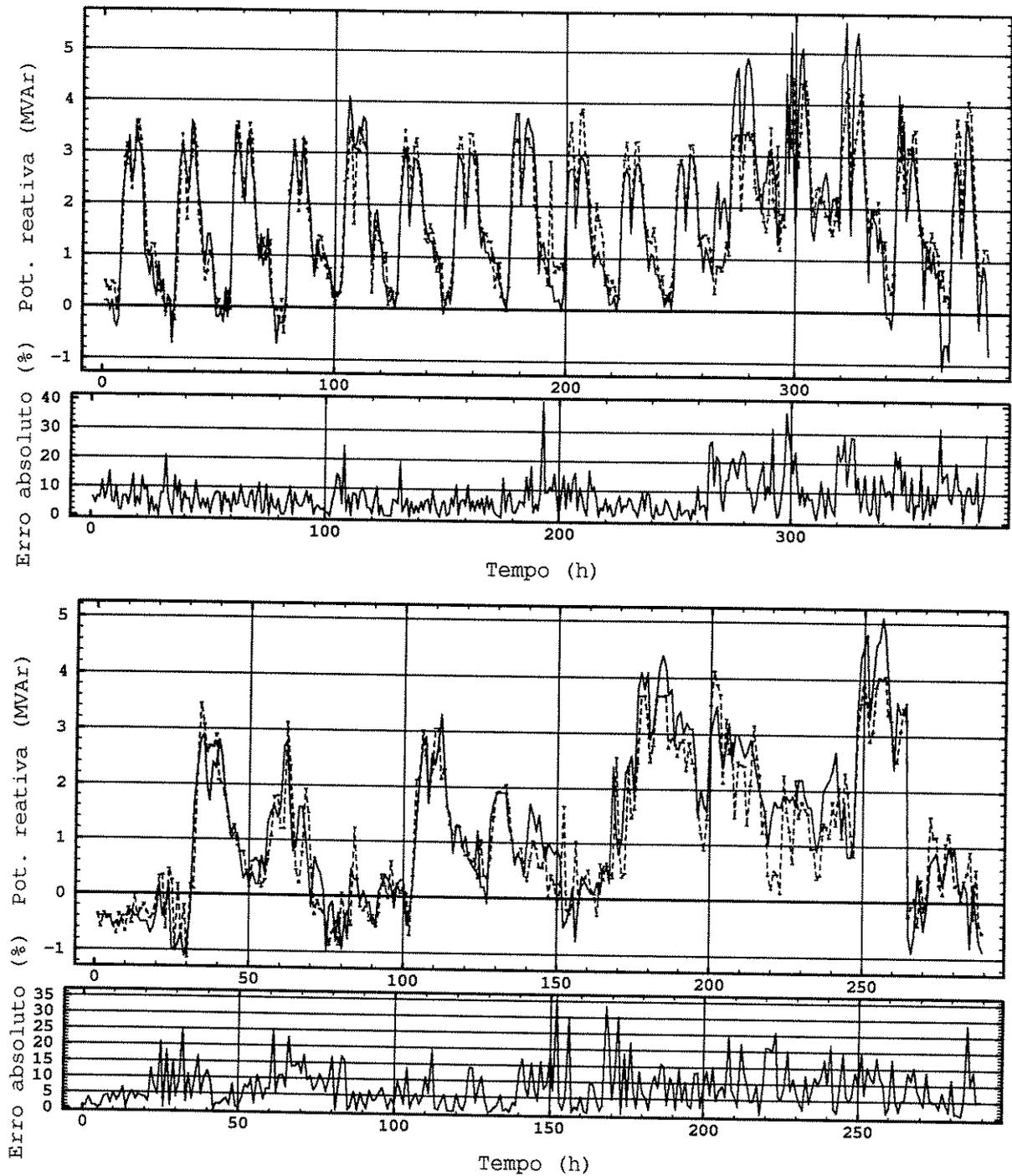


Figura 4.21: Potência reativa prevista, real e erro. Subestação 379. Modelo M1.

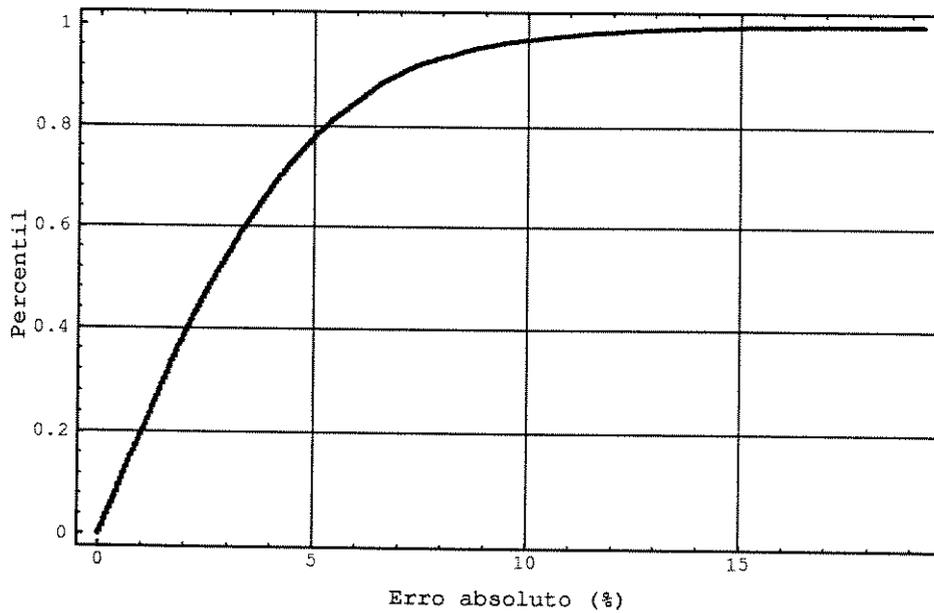


Figura 4.22: Gráfico da porcentagem de medidas em função do erro absoluto. Barra 379, potência ativa. Modelo M1

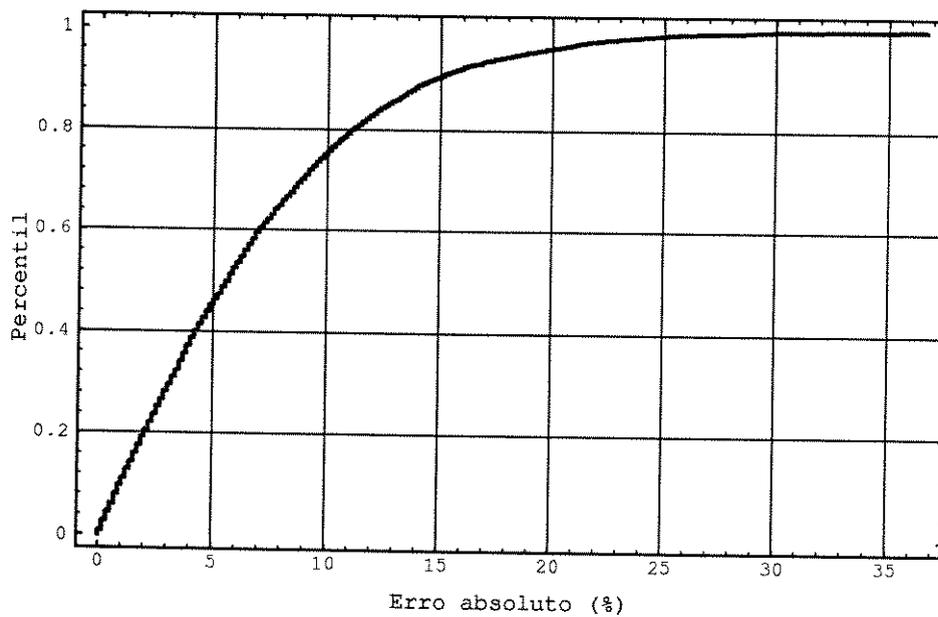


Figura 4.23: Gráfico da porcentagem de medidas em função do erro absoluto. Barra 379, potência reativa. Modelo M1

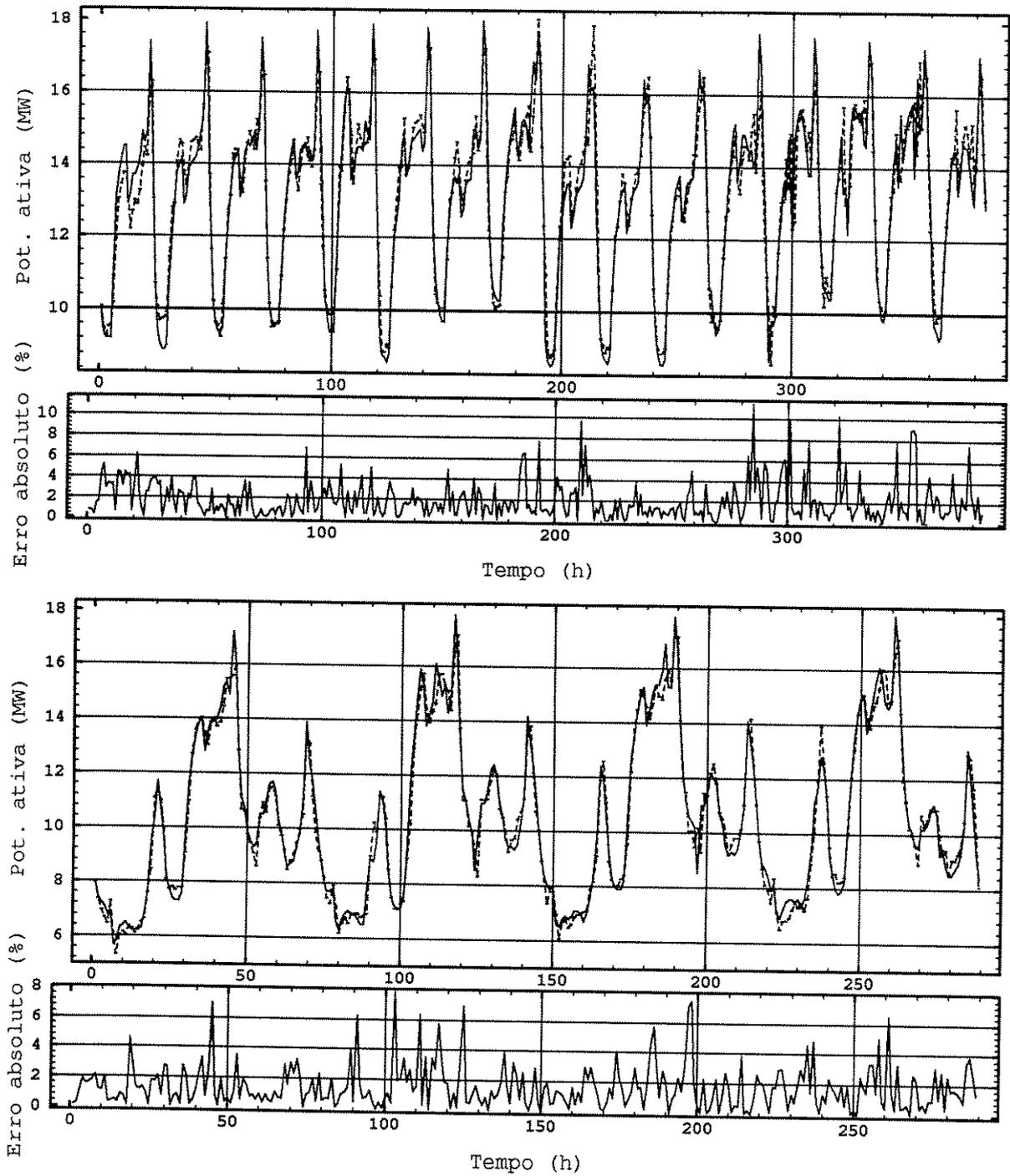


Figura 4.24: Potência ativa prevista, real e erro. Subestação 463. Modelo M1.

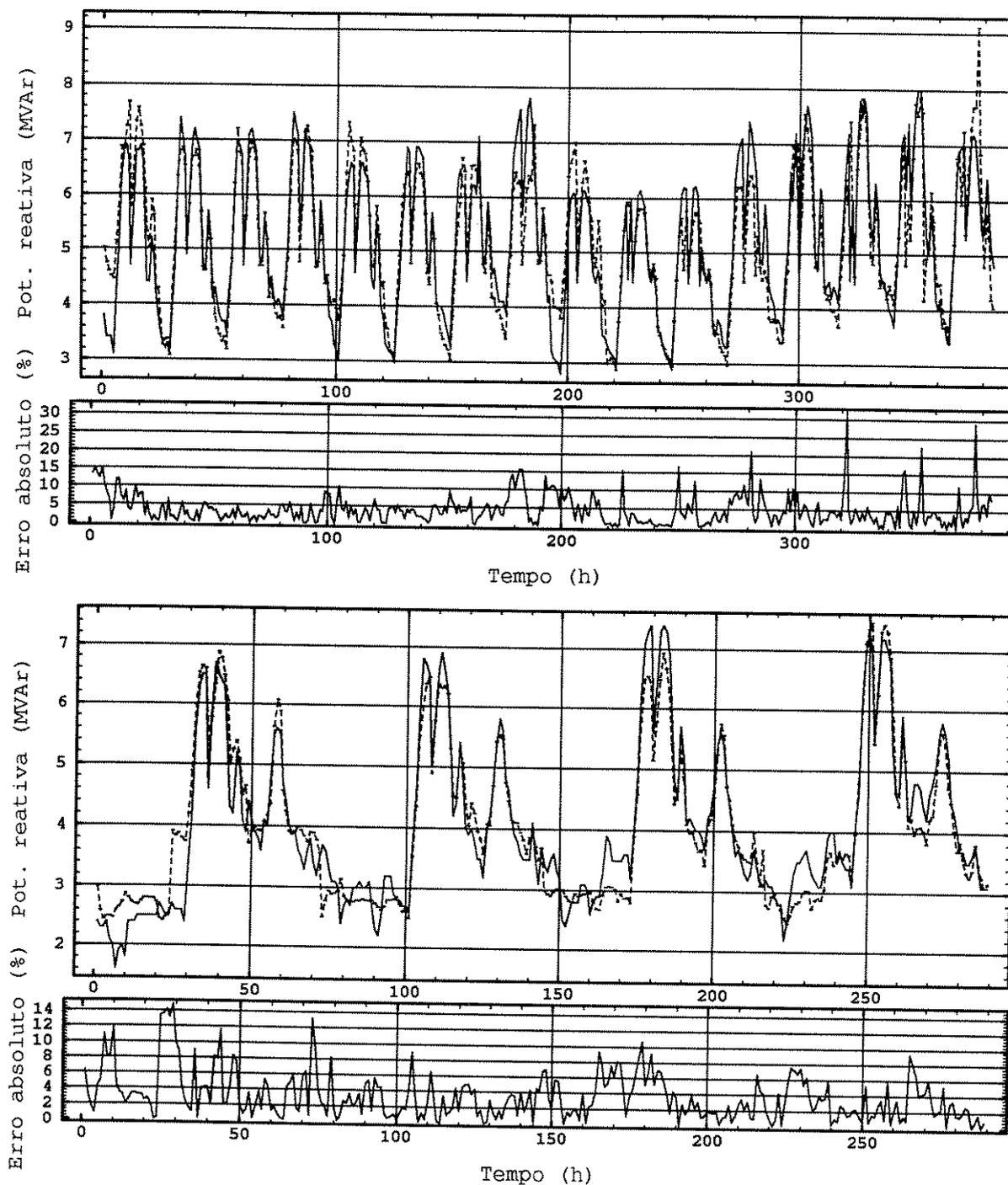


Figura 4.25: Potência reativa prevista, real e erro. Subestação 463. Modelo L4.

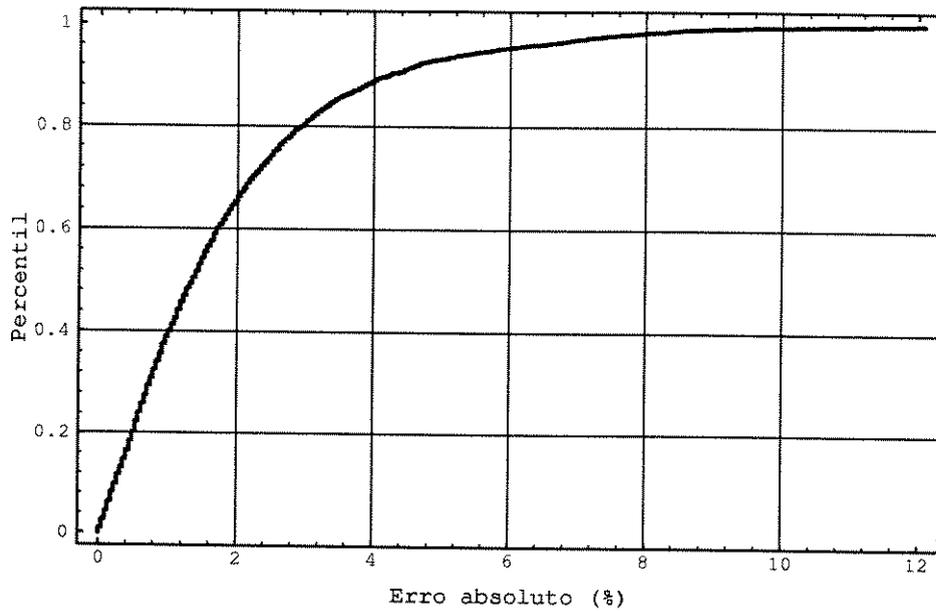


Figura 4.26: Gráfico da porcentagem de medidas em função do erro absoluto. Barra 463, potência ativa. Modelo M1

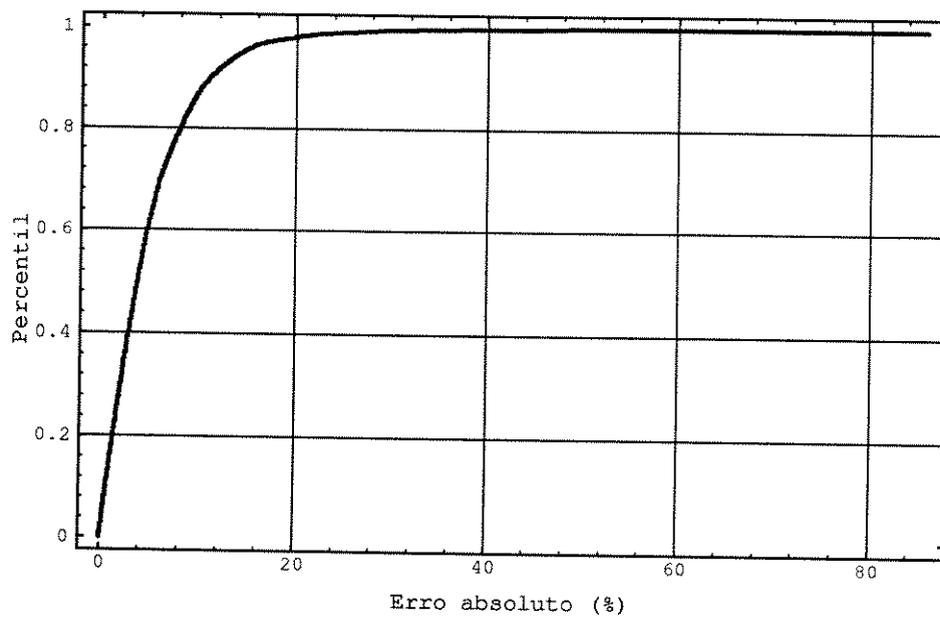


Figura 4.27: Gráfico da porcentagem de medidas em função do erro absoluto. Barra 463, potência reativa. Modelo L4

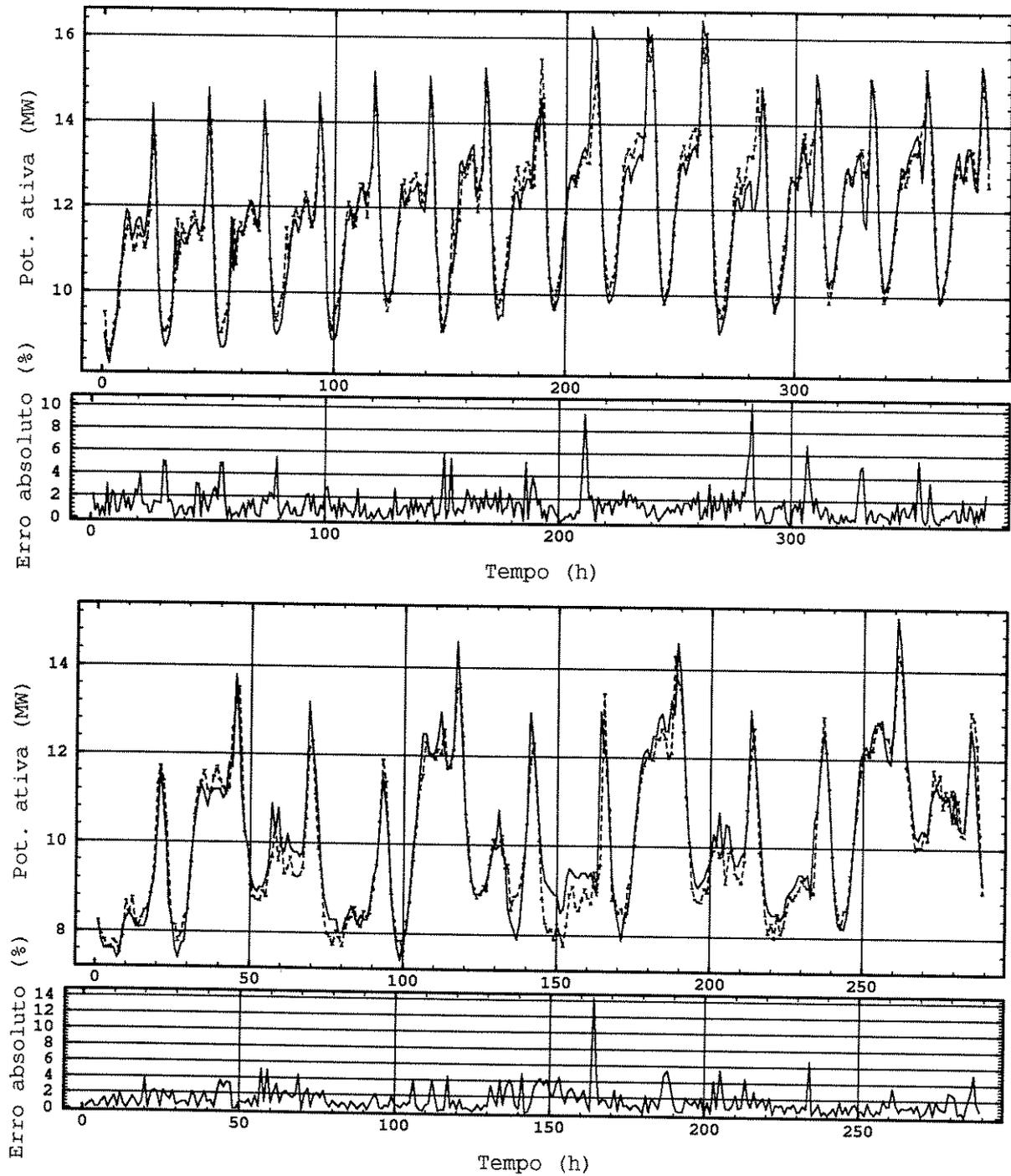


Figura 4.28: Potência ativa prevista, real e erro. Subestação 572. Modelo M1.

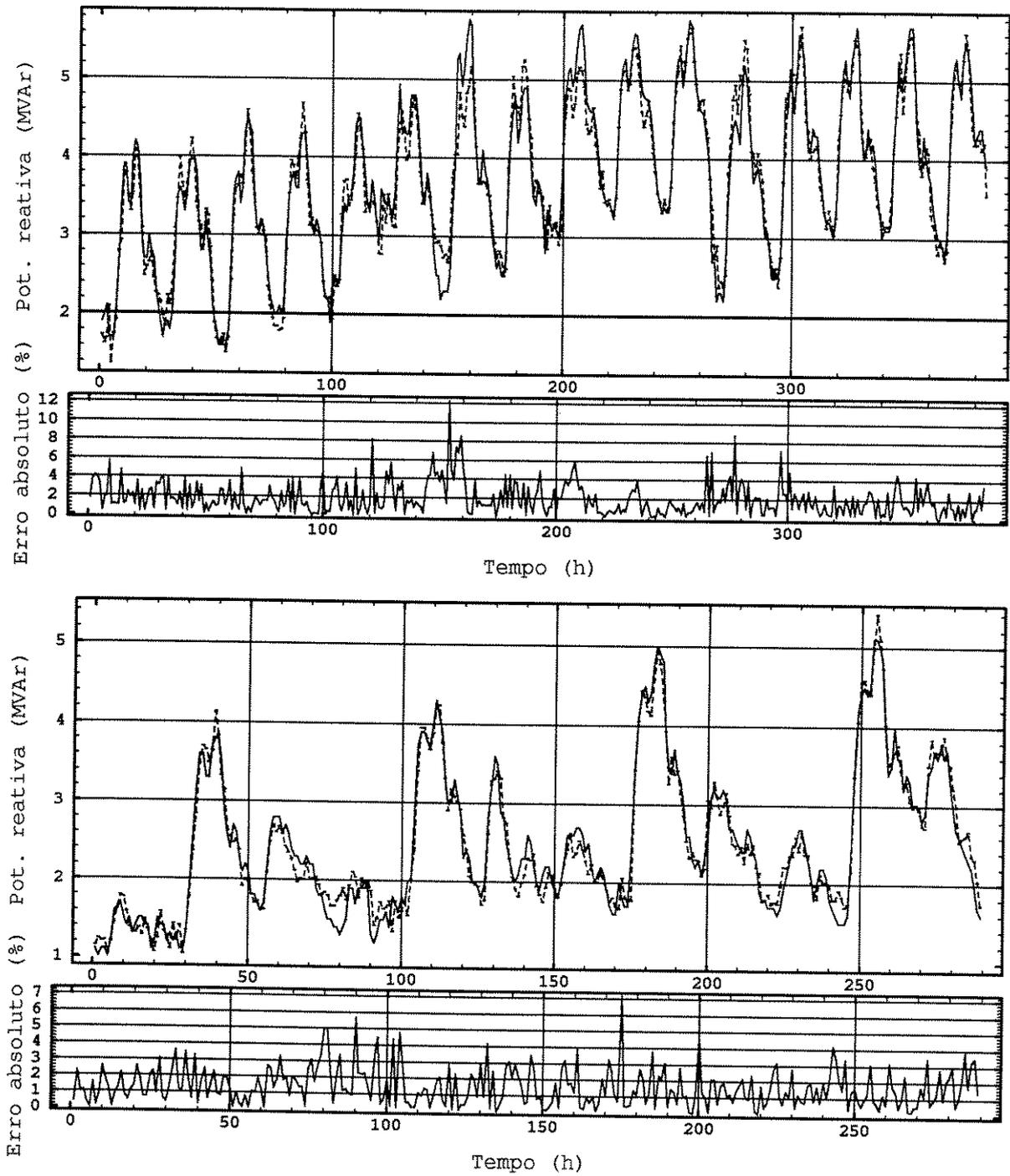


Figura 4.29: Potência reativa prevista, real e erro. Subestação 572. Modelo M1.

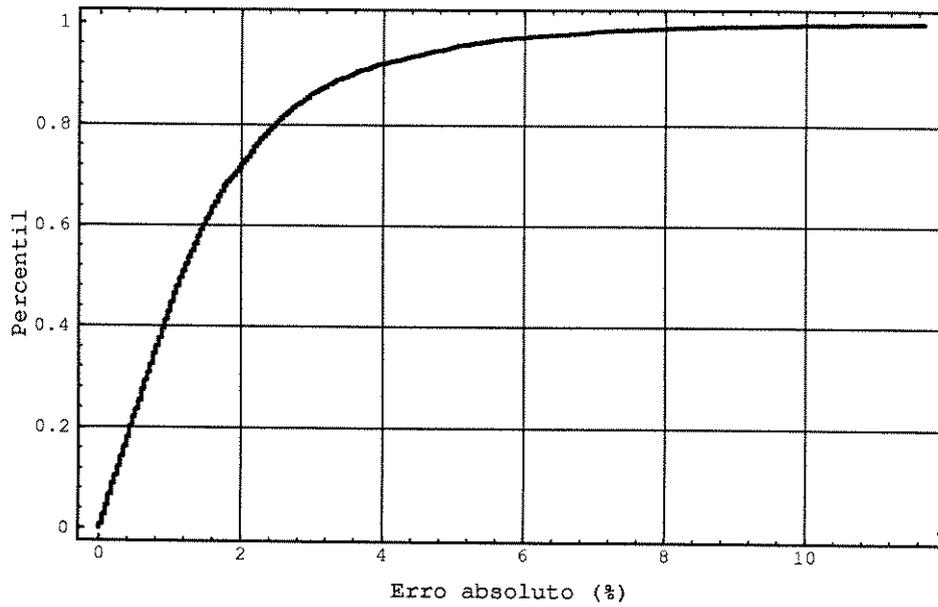


Figura 4.30: Gráfico da porcentagem de medidas em função do erro absoluto. Barra 572, potência ativa. Modelo M1

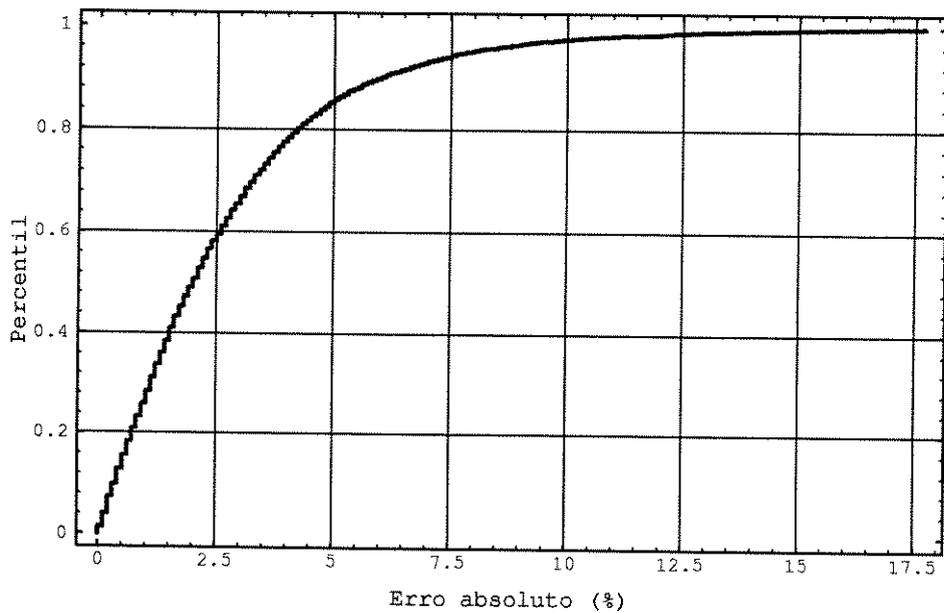


Figura 4.31: Gráfico da porcentagem de medidas em função do erro absoluto. Barra 572, potência reativa. Modelo M1

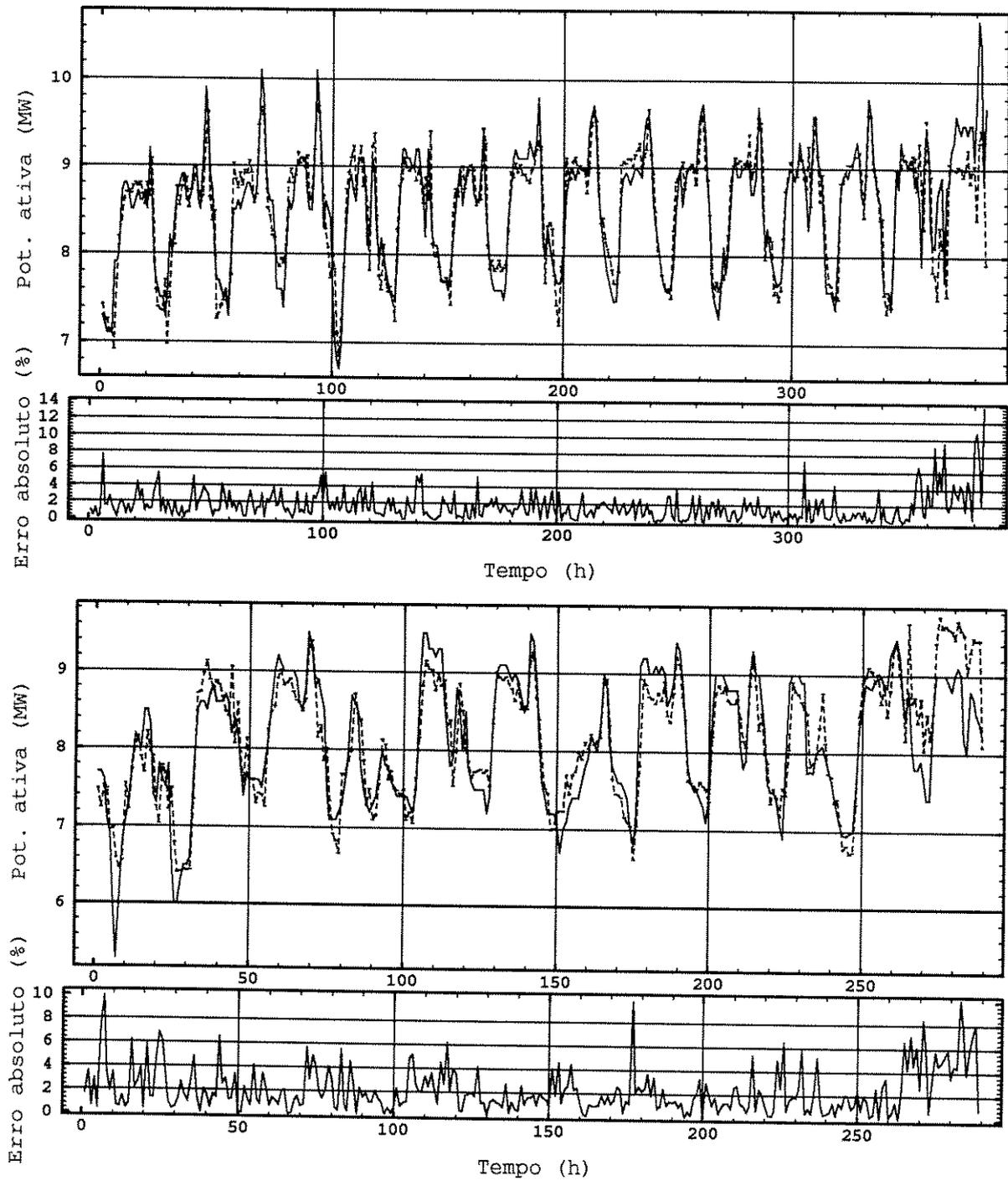


Figura 4.32: Potência ativa prevista, real e erro. Subestação 623. Modelo M1.

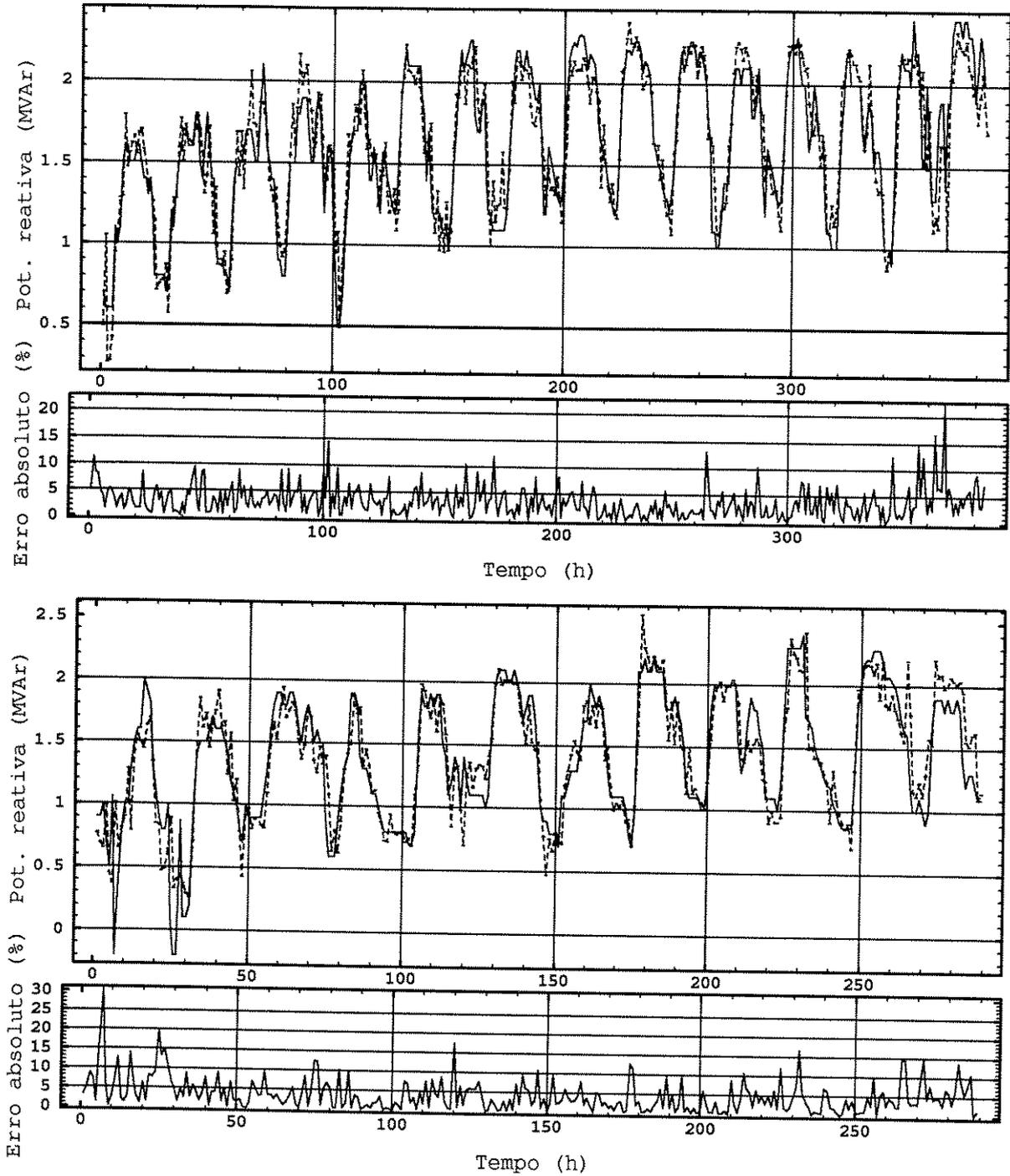


Figura 4.33: Potência reativa prevista, real e erro. Subestação 623. Modelo M1.

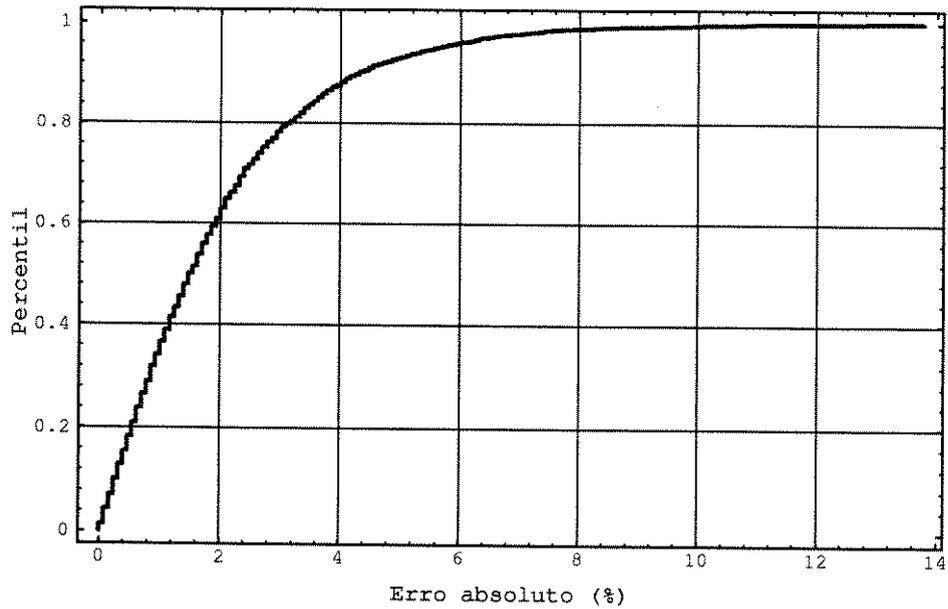


Figura 4.34: Gráfico da porcentagem de medidas em função do erro absoluto. Barra 623, potência ativa. Modelo M1

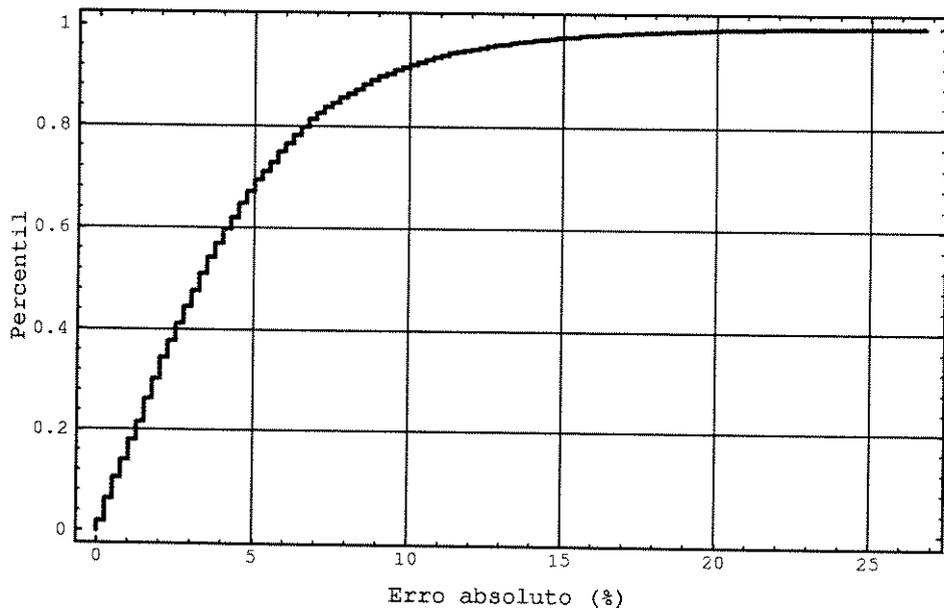


Figura 4.35: Gráfico da porcentagem de medidas em função do erro absoluto. Barra 623, potência reativa. Modelo M1

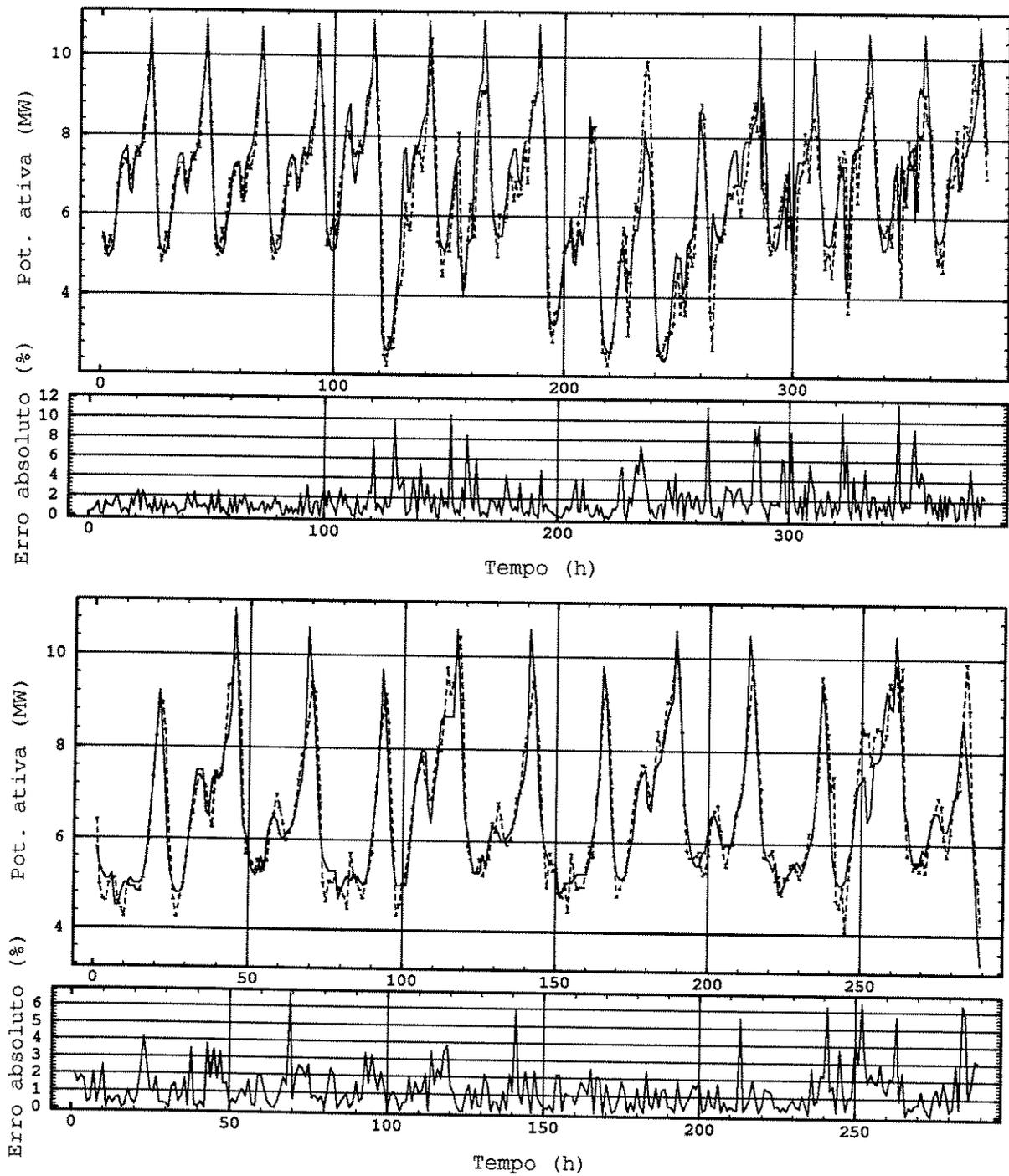


Figura 4.36: Potência ativa prevista, real e erro. Subestação 668. Modelo F2.

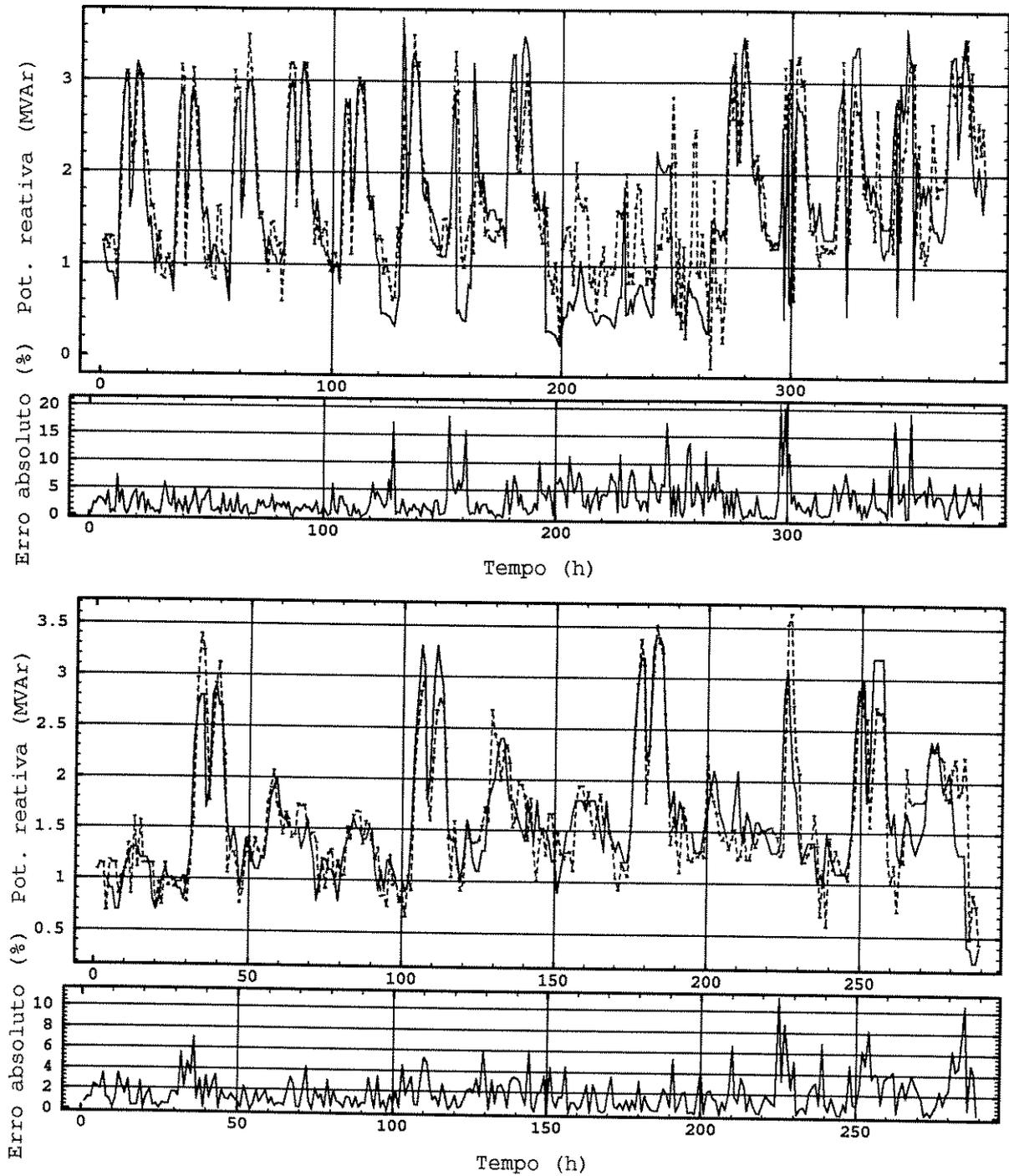


Figura 4.37: Potência reativa prevista, real e erro. Subestação 668. Modelo M1.

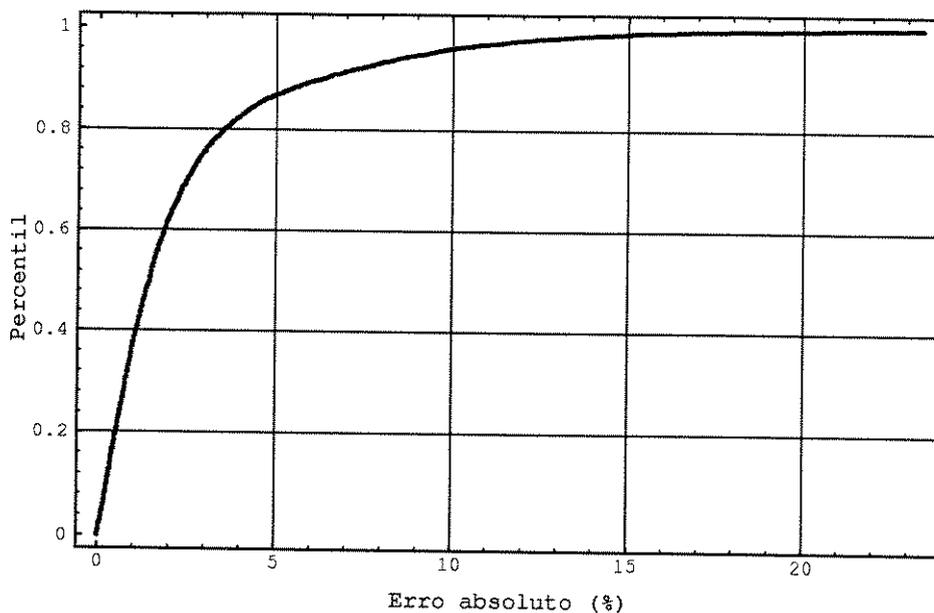


Figura 4.38: Gráfico da porcentagem de medidas em função do erro absoluto. Barra 668, potência ativa. Modelo F2

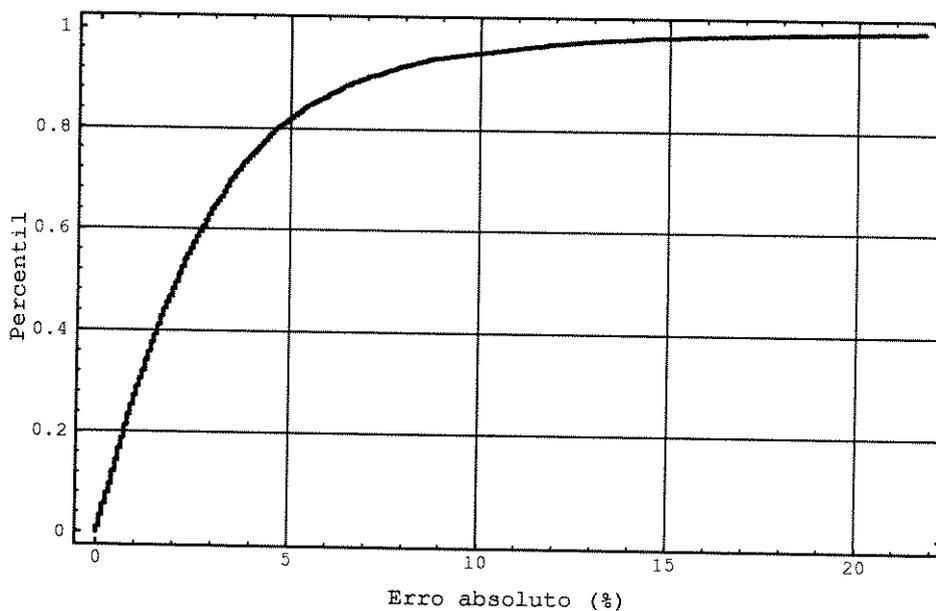


Figura 4.39: Gráfico da porcentagem de medidas em função do erro absoluto. Barra 668, potência reativa. Modelo M1

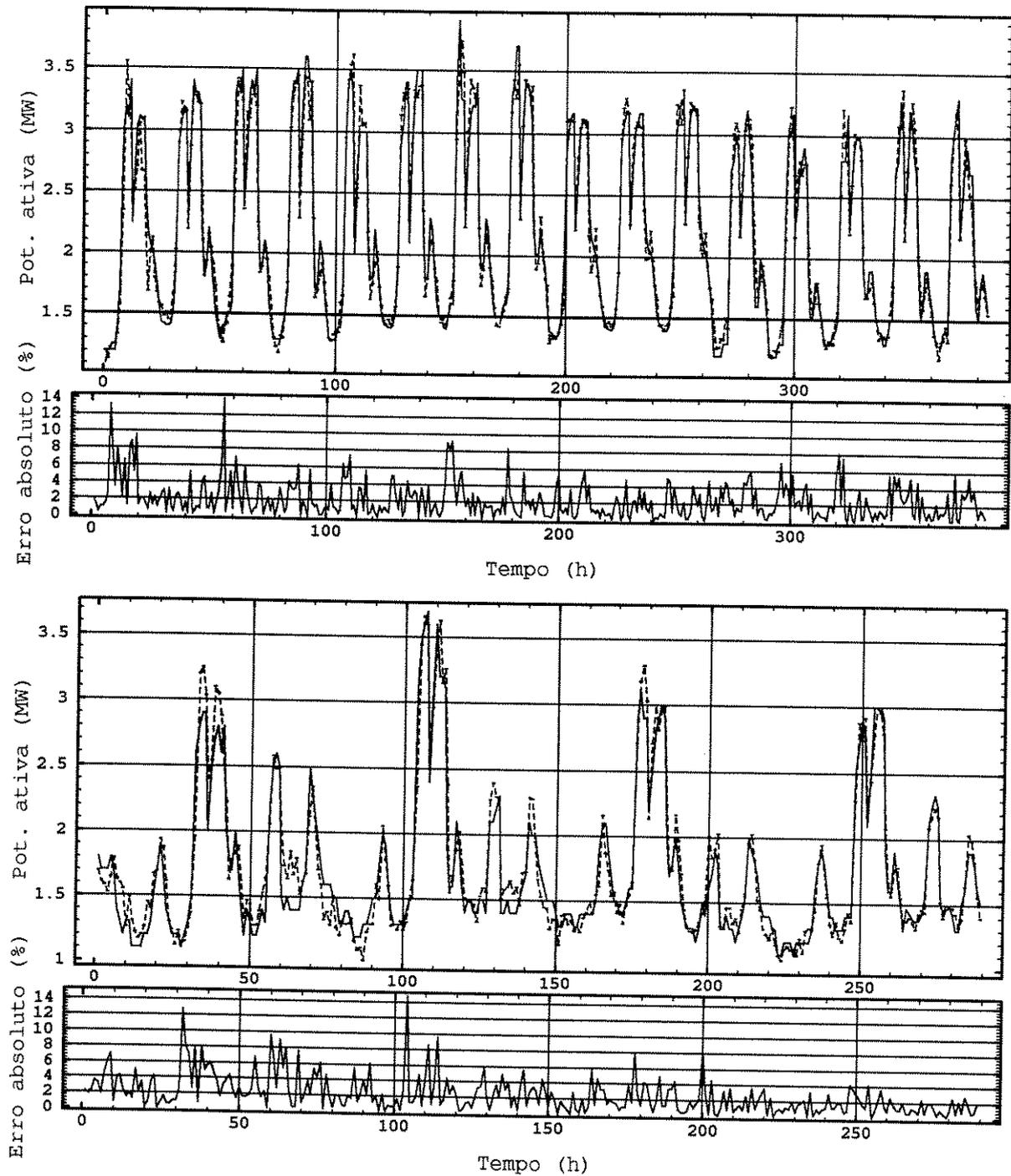


Figura 4.40: Potência ativa prevista, real e erro. Subestação 762. Modelo M1.

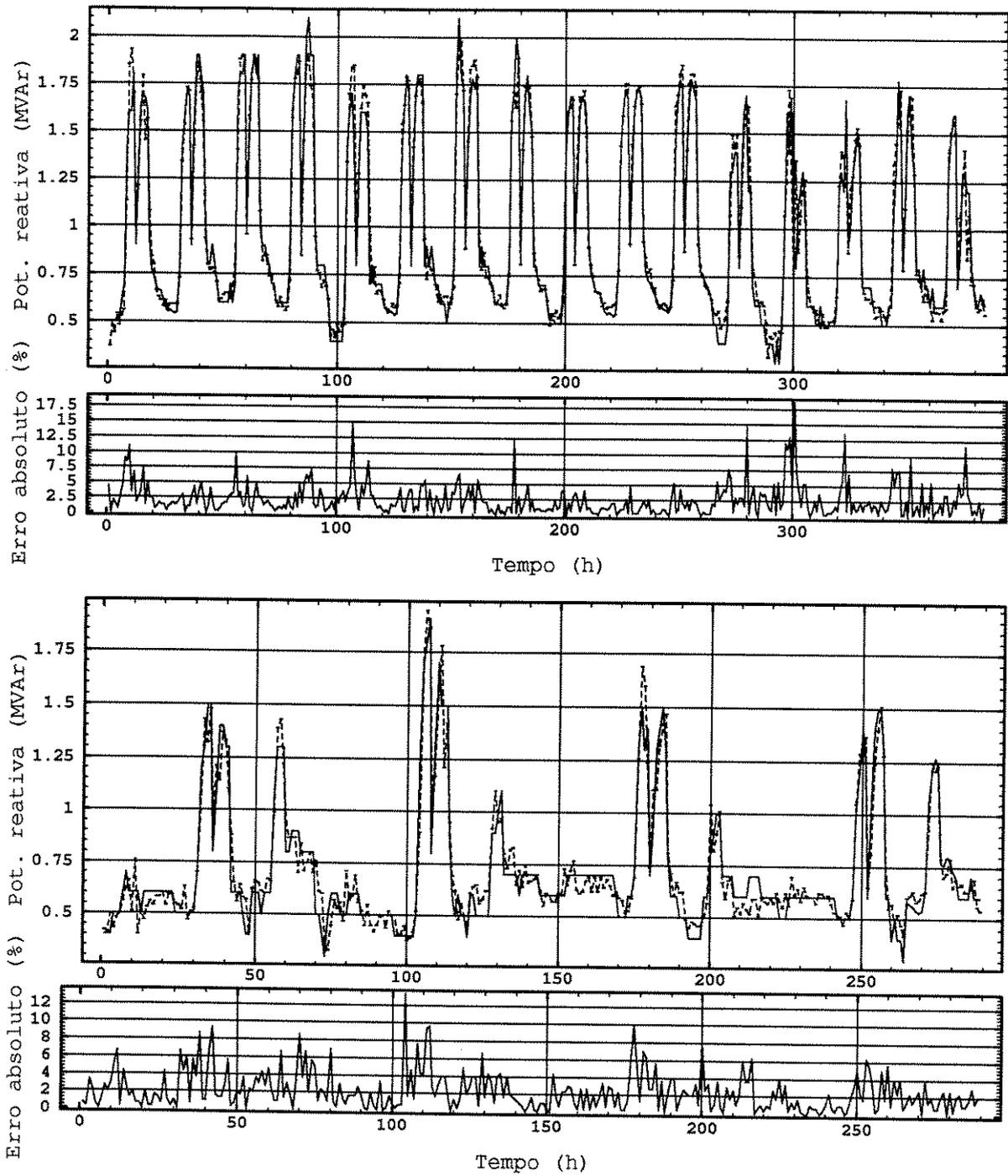


Figura 4.41: Potência reativa prevista, real e erro. Subestação 762. Modelo M1.

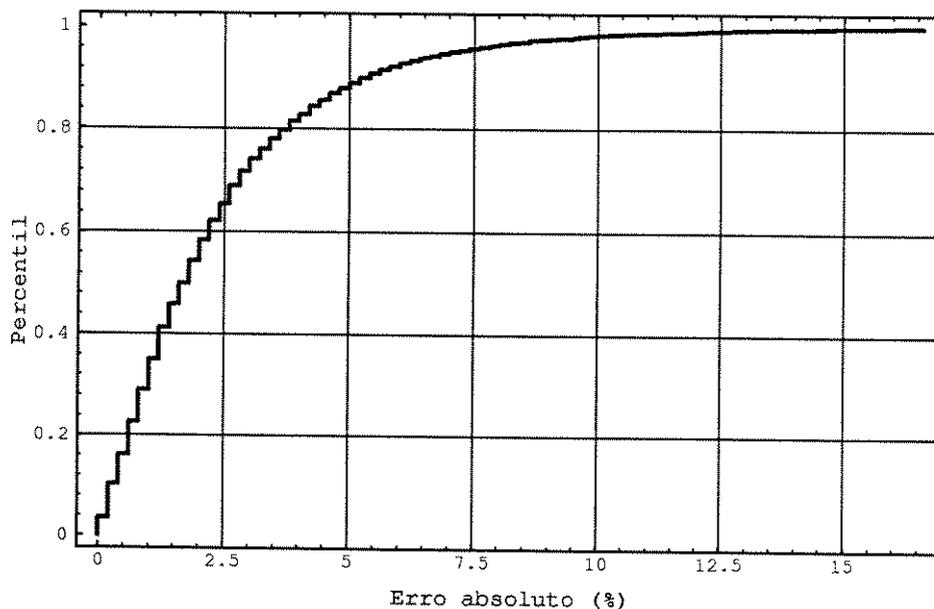


Figura 4.42: Gráfico da porcentagem de medidas em função do erro absoluto. Barra 762, potência ativa. Modelo M1

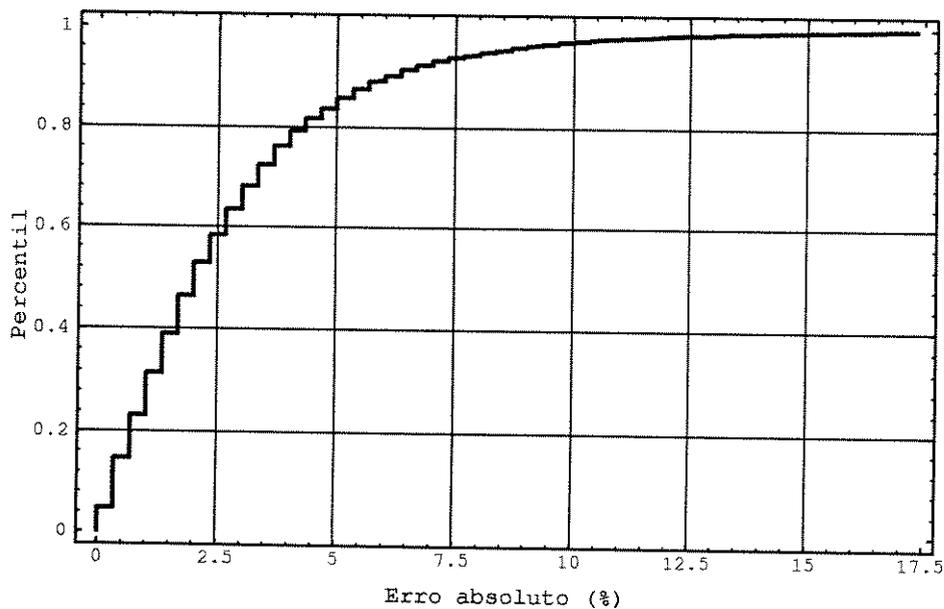


Figura 4.43: Gráfico da porcentagem de medidas em função do erro absoluto. Barra 762, potência reativa. Modelo M1

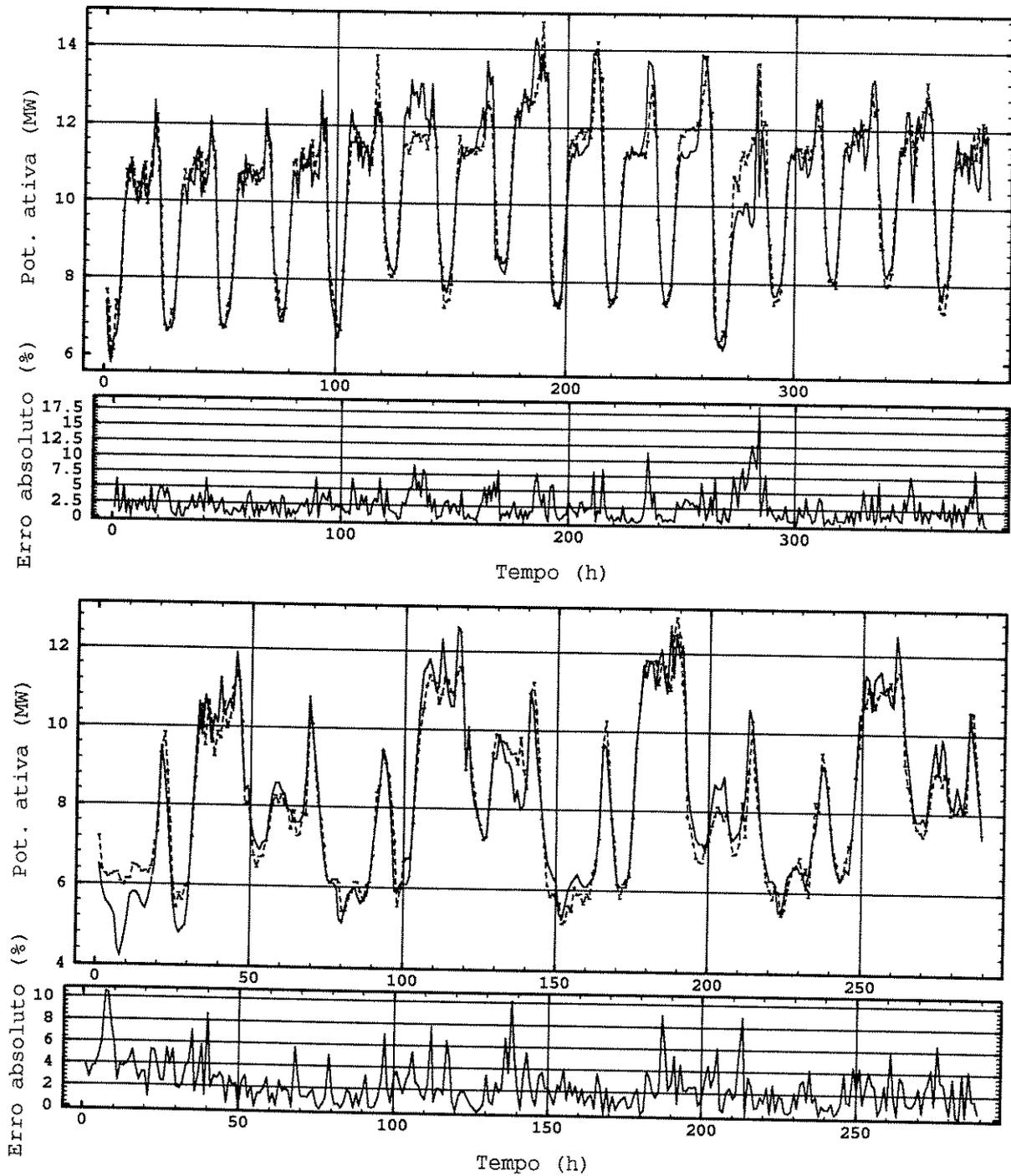


Figura 4.44: Potência ativa prevista, real e erro. Subestação 840. Modelo M1.

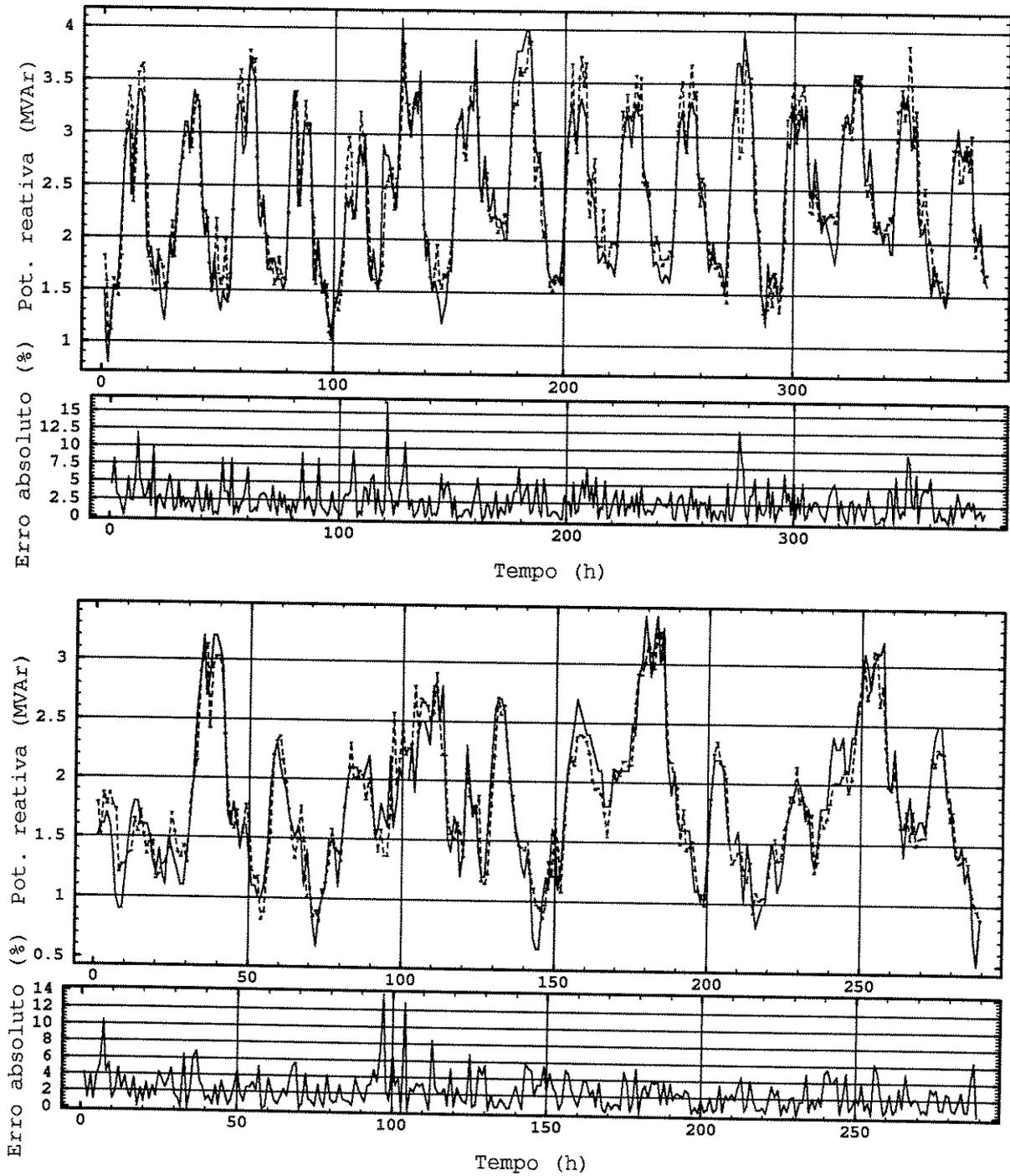


Figura 4.45: Potência reativa prevista, real e erro. Subestação 840. Modelo M1.

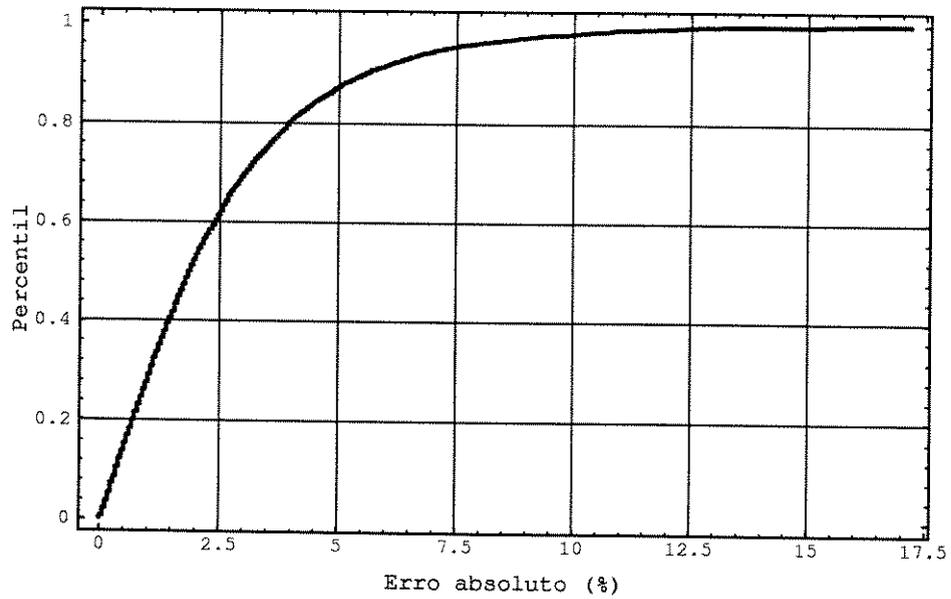


Figura 4.46: Gráfico da porcentagem de medidas em função do erro absoluto. Barra 840, potência ativa. Modelo M1

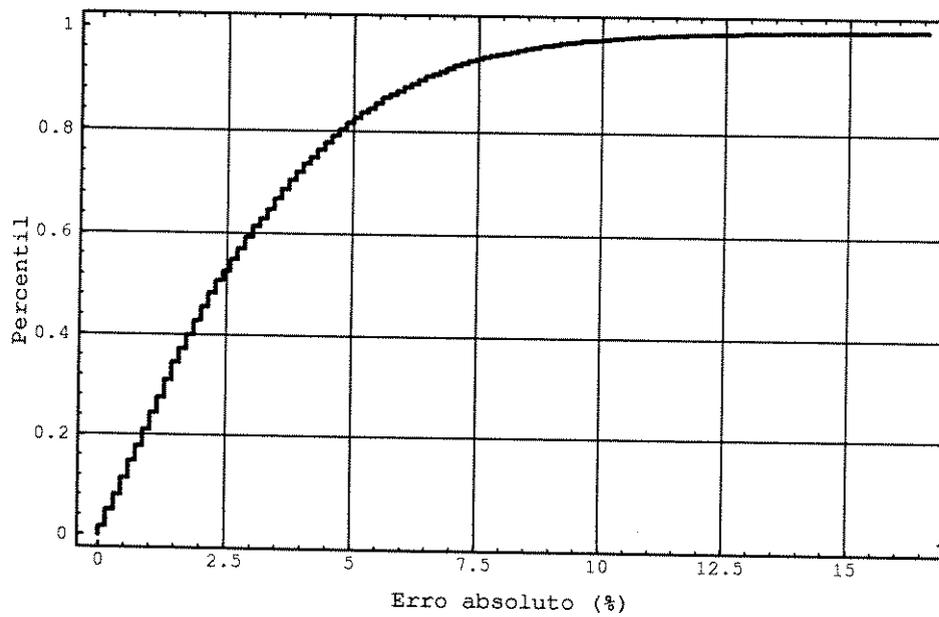


Figura 4.47: Gráfico da porcentagem de medidas em função do erro absoluto. Barra 840, potência reativa. Modelo M1

A análise das tabelas 4.22 e 4.23 e dos gráficos permite tirar as seguintes conclusões:

Das tabelas 4.22 e 4.23 percebe-se que os modelos que apresentam melhor desempenho por ordem crescente de erro médio são:

1. Potência Ativa: M1, M2 , F2, M3, F1, L4, L3 e F3.
2. Potência Reativa: M1, M2, M3, F2, F1, F4, L4, L3 e F3.

Dentro do contexto de cada modelo, o particionamento 3 (uma rede para cada dia da semana) foi o que apresentou pior resultado, indicando que a excessiva especialização pode ser prejudicial no processo de previsão, tendo em vista as variáveis de entrada utilizadas.

Nos gráficos é possível notar que a potência reativa apresenta maior nível de ruído do que a potência ativa, conforme visto no capítulo 3, sendo que a curva de carga para o final de semana, tanto para a potência ativa, quanto para a potência reativa é a que apresenta maior quantidade de ruído.

Nota-se nos gráficos alguns valores espúrios ou anormais, cuja presença não afeta a capacidade de previsão da rede exceto localmente. Em especial, no primeiro gráfico da figura 4.8, observa-se uma medida espúria de potência na curva de carga da quarta-feira, cuja incorporação ao conjunto de treinamento acarretou um erro de previsão elevado para a curva de carga da quarta-feira seguinte.

Capítulo 5

Previsão do Estado Futuro

5.1 Considerações Iniciais

A operação de um sistema de energia elétrica precisa ser feita de forma confiável e econômica, a fim de que toda a carga e intercâmbio de energia sejam atendidos a um custo que seja mínimo para a empresa e que nenhuma restrição de operação dos equipamentos que compõem o sistema de potência (geradores, transformadores e linhas de transmissão) seja violada. Isso significa que a demanda de potência deve ser atendida de forma que os limites operativos de cada componente do sistema sejam respeitados. Para isso é necessário que o operador no centro de controle de energia faça estudos visando garantir a confiabilidade do sistema, ou seja, ele verifica se o sistema opera de forma segura. Esse estudo é denominado de análise de segurança, a qual é constituída basicamente de quatro funções [Wood, 1996]:

1. Monitoração do Sistema.
2. Análise Estática de Contingência.
3. Análise de Segurança Dinâmica.
4. Fluxo de Carga Ótimo com Restrição de Segurança.

cujas definições se encontram no Capítulo 1.

Uma vez que a função 1 relaciona-se com a aquisição de medidas e que a função 4 é uma extensão da função 2, então restringir-se-á as considerações seguintes para as funções 2 e 3.

A análise estática de contingência verifica se a saída de operação de um transformador, linha de transmissão, gerador ou carga, acarretará alguma sobrecarga de fluxo em outras linhas,

sobrecarga de geração ou sobretensão em alguma barra do sistema. Já a análise de segurança dinâmica verifica se os geradores manterão o sincronismo quando da ocorrência de um defeito. Em ambos os casos é necessário obter o estado do sistema através do uso de algoritmo de fluxo de carga AC, o qual permite obter os valores das variáveis (V, θ) , a partir dos valores das injeções independentes de potência P e Q , através da solução das equações 5.1 e 5.2, que é obtida de forma iterativa utilizando o método de Newton-Raphson ou suas versões desacopladas [Monticelli, 1983]. A figura 5.1 ilustra o processo acima.

$$P_k = V_k \sum_{m \in K} V_m (G_{km} \cos \theta_{km} + B_{km} \sin \theta_{km}) \quad (5.1)$$

$$Q_k = V_k \sum_{m \in K} V_m (G_{km} \sin \theta_{km} + B_{km} \cos \theta_{km}) \quad (5.2)$$

onde K é o conjunto de m barras adjacentes à barra k , incluindo a própria barra k , G é a condutância e B a susceptância.

Uma vez que não tem sentido fazer análise de segurança para o ponto de operação atual, pois nada poderá se fazer, então é preciso que essa análise seja feita considerando o estado futuro do sistema que depende do perfil de carga futuro, já que o estado é função da carga ativa e reativa como mostram as equações 5.1 e 5.2.

É possível obter diretamente uma previsão do estado, considerando as tensões nodais (V, θ) como variáveis de estado, através de estimadores dinâmicos de estado. Contudo há benefícios em utilizar a previsão de carga nodal e depois o estado os quais são considerados abaixo [Couto, 1994]:

- Injeções nodais de potência ativa e reativa são medidas mais apropriadas para representar a dinâmica de interesse do sistema.
- São valores que guardam maior independência entre si.
- A modelagem pode ser feita através de técnicas estatísticas já conhecidas.

Tendo em vista a abordagem exposta na figura 5.1, procurou-se neste capítulo avaliar o desempenho da previsão de estado por meio de fluxos de carga em relação ao desempenho da previsão de carga nodal ativa e reativa. Para isso foi montado um exemplo considerando inicialmente um sistema de duas barras e outro estudo considerando o sistema IEEE14 modificado da seguinte maneira: manteve-se os dados de linha de transmissão e utilizou-se como injeções de potência em 11 barras PQ as cargas previstas nas subestações da CPFL.

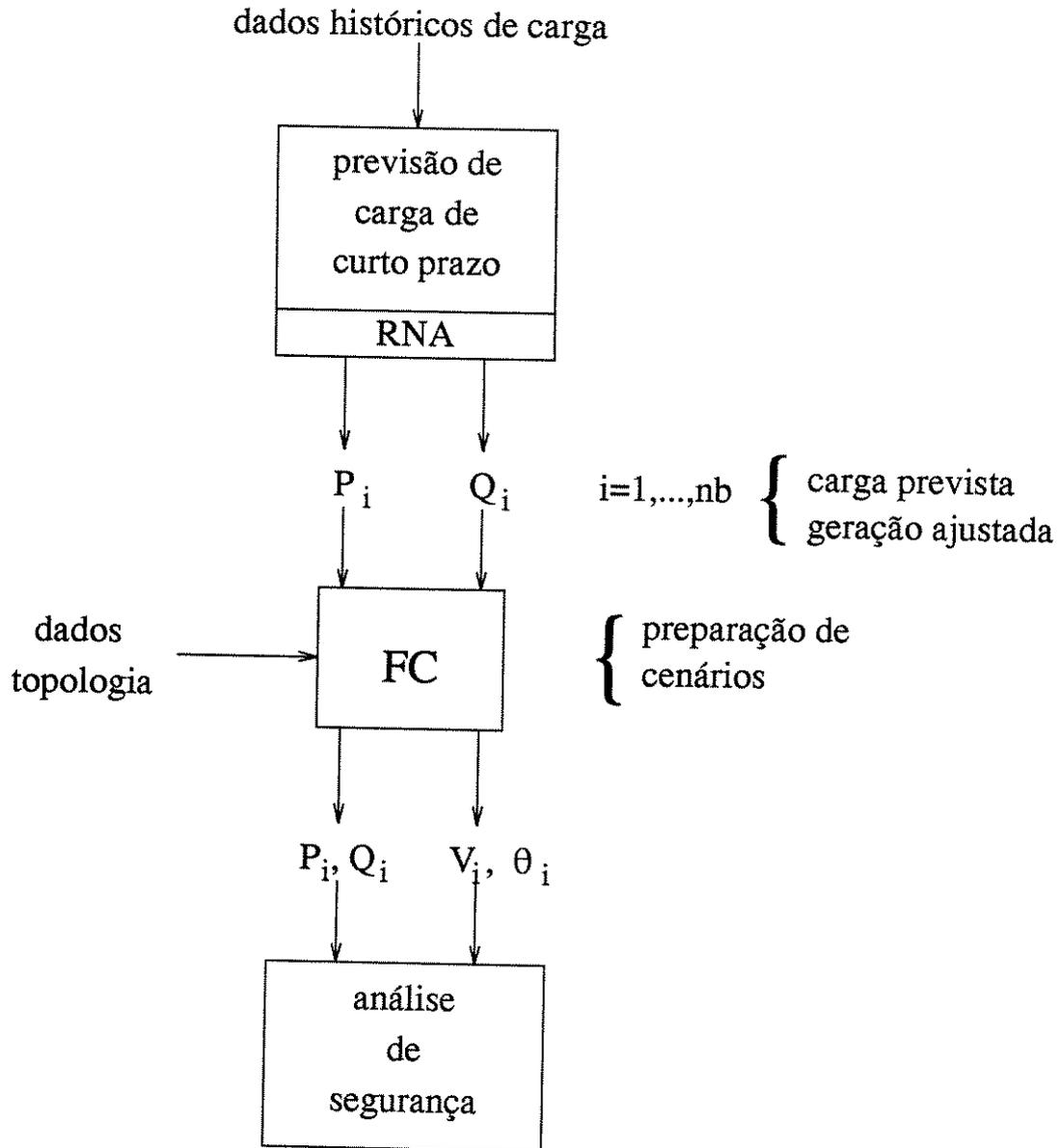


Figura 5.1: Diagrama de blocos para o problema da análise de segurança com previsão de carga.

5.2 Efeito da Previsão de Carga no Estado

Para exemplificar a idéia exposta anteriormente, nesta seção será mostrado como varia o estado do sistema (V e θ) em função da variação em P e Q , ou seja, será mostrada a sensibilidade do erro na previsão do estado do sistema em função do erro na previsão de carga. Para isso será utilizado um exemplo simples com um sistema de duas barras conforme mostra a figura 5.2.

Sistema de duas barras

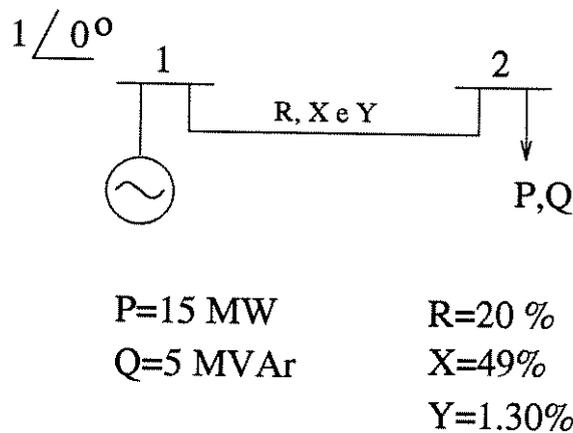


Figura 5.2: Sistema para exemplo de previsão de estado.

Para o sistema acima considerou-se uma carga base de $P = 15 \text{ MW}$ e $Q = 5 \text{ MVAr}$. Em relação a essa carga base será feito um simples estudo para ilustrar a variação no estado do sistema quando se varia a carga ativa e reativa em $\pm 10\%$. São abordados nove casos conforme mostrado abaixo.

- Caso 1: P e Q base.
- Casos 2 e 3: P não varia, Q varia.
- Casos 4 e 5: P varia, Q não varia.
- Casos 6 a 9: P e Q variam.

A tabela 5.1 mostra as injeções de potência na barra 2 e a tabela 5.2 mostra o estado correspondente a essas injeções, bem como o erro em relação ao caso base.

Pelo exemplo é possível observar alguns fatos importantes que se repetirão, quando da análise do estado para o sistema IEEE 14 barras:

- Uma imprecisão na potência reativa tem pequena influência no estado (V, θ) , se comparado com uma imprecisão na potência ativa.
- A sensibilidade da magnitude da tensão em relação à variação nas duas potências é bem menor quando comparada com a sensibilidade do ângulo,
- Em outras palavras: a previsão do estado por essa abordagem exige uma maior precisão da potência ativa que na potência reativa.

caso	barra	P	Q
1	2	-15.00	-5.00
2	2	-15.00	-4.50
3	2	-15.00	-5.50
4	2	-13.50	-5.00
5	2	-16.50	-5.00
6	2	-16.50	-5.50
7	2	-13.50	-5.50
8	2	-16.50	-4.50
9	2	-13.50	-4.50

Tabela 5.1: Valor das injeções de potência na barra 2

caso	barra	tensão	ângulo	erro (%) V	erro (%) θ
1	2	0.9429	-3.9256	-	-
2	2	0.9456	-3.9771	0.28	1.31
3	2	0.9402	-3.8741	0.29	1.31
4	2	0.9465	-3.4713	0.38	11.57
5	2	0.9384	-4.3992	0.48	12.06
6	2	0.9356	-4.3493	0.77	10.80
7	2	0.9437	-3.4201	0.08	12.88
8	2	0.9413	-4.4762	0.17	14.03
9	2	0.9492	-3.5501	0.67	9.57

Tabela 5.2: Estado na barra 2

5.2.1 Análise do Erro de Previsão de Carga no Estado do Sistema IEEE14

Nesta seção é feita uma análise da sensibilidade do estado do sistema IEEE14 barras em relação ao erro de previsão nas injeções de potência obtidas através dos modelos de ANN de melhor desempenho encontrados nos testes.

O Sistema IEEE14 é composto de 14 barras: 1 barra $V\theta$, 4 barras PV e 9 barras PQ. Uma vez que as simulações de previsão foram feitas para 11 barras do tipo PQ, então duas das barras PV foram convertidas em PQ, notadamente as barras 3 e 6. Para as barras PV restantes a geração de potência ativa foi fixada em 25 MW e a tensão em $V = 1pu$. O Valor de V e θ para a barra *slack* foi fixado em: $V = 1pu$ e $\theta = 0$. Os dados do IEEE14 se encontram no Apêndice A. A correspondência entre as subestações e os números das barras é apresentada na tabela abaixo:

Subestação	40 R	80 C	221 R	232 R	379 I	463 R	572 R	623 C	668 R	762 I	840 R
Nº Barra	3	4	5	7	8	9	10	11	12	13	14

Tabela 5.3: Relação entre o número da barra e o número da subestação.

A análise estatística dos dados foi realizada para o período de 31/07/94 a 04/02/95 totalizando 4536 valores horários, sem descontar a presença de feriados, os quais não foram considerados na análise. Assim para cada hora do intervalo acima foi rodado um fluxo de carga para os casos apresentados abaixo, sendo que o resultado estatístico para esses casos é apresentado na tabela 5.5, onde os erros são valores percentuais relativos. Já as tabelas 5.6 e 5.7 apresentam uma estatística do erro nos fluxos das linhas, onde os erros são em relação às bases apresentadas na tabela 5.4 para cada ramo do sistema. Os gráficos mostram o erro de previsão da tensão e do ângulo em relação ao erro de previsão da potência reativa e ativa. Note que a presença de um valor espúrio na previsão de potência é refletida na previsão de estado, como era de se esperar.

- Caso 1: P não varia, Q varia.
- Caso 2: P varia, Q não varia.
- Caso 3: P varia, Q varia.

base (MVA)	ramo									
	1	2	3	4	5	6	7	8	9	10
	55	67	45	45	45	34	89	45	22	55
base (MVA)	ramo									
	11	12	13	14	15	16	17	18	19	20
	27	22	34	67	55	22	34	22	22	22

Tabela 5.4: Tabela com as potências base para o fluxo nos ramos

A observação do comportamento do erro de previsão de estado para os casos acima permite obter as seguintes conclusões:

1. O erro de previsão da potência ativa afeta pouco ($< 0.10\%$) a magnitude da tensão; porém ocasiona um erro médio significativo na fase da tensão ($< 6.5\%$), o que era esperado tendo em vista o acoplamento $P - \theta$ nos sistemas de transmissão.
2. O erro de previsão da potência reativa pouco influencia a fase da tensão ($< 0.5\%$), como se espera devido ao desacoplamento $Q - \theta$; porém, surpreendentemente, afeta muito pouco a previsão da magnitude de tensão ($< 0.25\%$), tendo em vista o acoplamento $Q - V$. Uma razão para isso pode ser encontrada no pequeno valor dos elementos da matriz B'' , que é a matriz de sensibilidade $Q - V$.
3. O erro médio percentual relativo de previsão do ângulo é maior do que o erro médio de previsão da potência ativa.
4. O erro nas medidas de fluxo ativo e reativo é baixo, porque esses valores variam linearmente com o erro tensão que é reduzido.
5. O erro médio absoluto do ângulo é da ordem de 0.375 graus, que é um valor próximo ao obtido por [Silva, 1993].

fluxo	esta- tística	Subestação										
		40	80	221	232	379	463	572	623	668	762	840
		Magnitude										
caso 1	CV	0.00001	0.00001	0.00001	0.00004	0.00005	0.00005	0.00004	0.00002	0.00002	0.00001	0.00003
	Emed(pu)	0.04	0.05	0.04	0.17	0.23	0.16	0.14	0.08	0.08	0.04	0.12
	Emax(pu)	0.24	0.33	0.25	1.12	1.43	1.02	1.08	0.55	0.73	0.31	0.79
caso 2	CV	0.00001	0.00001	0.00001	0.00001	0.00002	0.00001	0.00001	0.00001	0.00001	0.00001	0.00002
	Emed(pu)	0.03	0.06	0.05	0.07	0.07	0.07	0.07	0.04	0.07	0.03	0.08
	Emax(pu)	0.21	0.34	0.28	0.42	0.44	0.40	0.41	0.24	0.51	0.14	0.46
caso 3	CV	0.00001	0.00002	0.00002	0.00005	0.00006	0.00005	0.00005	0.00003	0.00003	0.00001	0.00004
	Emed(pu)	0.06	0.10	0.07	0.21	0.26	0.20	0.18	0.11	0.11	0.05	0.17
	Emax(pu)	0.39	0.59	0.45	1.40	1.65	1.36	1.39	0.76	0.81	0.35	1.07
		Angulo										
caso 1	CV	0.00005	0.00004	0.00004	0.00002	0.00002	0.00003	0.00003	0.00003	0.00011	0.00004	0.00003
	Emed(pu)	0.08	0.12	0.06	0.06	0.05	0.08	0.09	0.09	0.47	0.15	0.12
	Emax(pu)	1.98	1.69	1.42	0.90	0.72	1.03	1.00	1.09	4.31	1.42	1.07
caso 2	CV	0.00101	0.00096	0.00098	0.00081	0.00079	0.00082	0.00083	0.00089	0.00107	0.00102	0.00088
	Emed(pu)	5.97	5.42	5.71	4.27	4.22	4.38	4.43	4.93	6.28	5.99	4.75
	Emax(pu)	52.11	38.57	43.61	26.03	25.58	26.75	26.98	31.57	44.43	43.55	29.22
caso 3	CV	0.00100	0.00095	0.00098	0.00080	0.00080	0.00082	0.00083	0.00090	0.00110	0.00103	0.00087
	Emed(pu)	5.94	5.37	5.69	4.26	4.22	4.35	4.40	4.92	6.39	6.02	4.72
	Emax(pu)	51.57	38.43	43.43	25.73	25.72	26.30	26.70	31.44	44.71	43.63	29.07

Tabela 5.5: Estatísticas sobre o estado do sistema (IEEE14).

ramo	Caso											
	1				2				3			
	CV	Emed(pu)	Emax(pu)		CV	Emed(pu)	Emax(pu)		CV	Emed(pu)	Emax(pu)	
1	0.00007	0.02	0.36		0.00140	1.39	10.36		0.00140	1.39	10.36	
2	0.00004	0.02	0.60		0.00119	2.99	22.24		0.00119	2.99	22.24	
3	0.00009	0.02	0.67		0.00054	0.54	3.78		0.00054	0.54	3.78	
4	0.00002	0.01	0.22		0.00057	0.89	6.67		0.00057	0.89	6.67	
5	0.00007	0.03	0.44		0.00038	0.32	2.44		0.00038	0.32	2.44	
6	0.00020	0.02	0.88		0.00076	0.52	4.12		0.00076	0.52	4.12	
7	0.00009	0.09	1.01		0.00086	1.27	9.10		0.00086	1.27	9.10	
8	0.00006	0.12	1.11		0.00071	2.04	14.44		0.00071	2.04	14.44	
9	0.00011	0.14	1.82		0.00082	1.70	13.18		0.00082	1.70	13.18	
10	0.00027	0.17	2.00		0.00181	1.63	13.09		0.00181	1.63	13.09	
11	0.00009	0.20	1.85		0.00043	1.38	8.52		0.00043	1.38	8.52	
12	0.00021	0.21	1.82		0.00168	2.04	15.00		0.00168	2.04	15.00	
13	0.00013	0.13	1.47		0.00097	1.08	7.94		0.00097	1.08	7.94	
14	0.00000	0.00	0.00		0.00129	0.94	5.52		0.00129	0.94	5.52	
15	0.00093	0.10	0.91		0.00697	0.83	7.27		0.00697	0.83	7.27	
16	0.00087	0.21	2.27		0.00397	1.52	10.45		0.00397	1.52	10.45	
17	0.00125	0.11	1.18		0.01774	0.89	13.82		0.01774	0.89	13.82	
18	0.00016	0.21	2.27		0.00066	1.38	8.64		0.00066	1.38	8.64	
19	0.00229	0.22	1.82		0.01122	1.31	11.36		0.01122	1.31	11.36	
20	0.00017	0.20	2.27		0.00125	1.48	15.45		0.00125	1.48	15.45	

Tabela 5.6: Estatísticas do fluxo de potência ativa (IEEE14).

ramo	Caso											
	1				2				3			
	CV	Emed(pu)	Emax(pu)		CV	Emed(pu)	Emax(pu)		CV	Emed(pu)	Emax(pu)	
1	0.00008	0.00	0.18		0.00061	0.08	1.27		0.00061	0.08	1.27	
2	0.00107	0.24	1.04		0.00068	0.16	2.39		0.00068	0.16	2.39	
3	0.00181	0.39	1.11		0.00137	0.24	2.89		0.00137	0.24	2.89	
4	0.00161	0.65	0.89		0.00138	0.46	5.56		0.00138	0.46	5.56	
5	0.00144	0.45	0.89		0.00193	0.51	2.67		0.00193	0.51	2.67	
6	0.00213	0.41	1.47		0.00177	0.28	3.24		0.00177	0.28	3.24	
7	0.00195	0.47	1.35		0.00074	0.17	3.26		0.00074	0.17	3.26	
8	0.00126	1.21	2.22		0.00038	0.31	3.78		0.00038	0.31	3.78	
9	0.00110	0.80	2.27		0.00034	0.24	2.73		0.00034	0.24	2.73	
10	0.00070	0.25	2.00		0.00081	0.28	3.09		0.00081	0.28	3.09	
11	0.00153	1.31	2.96		0.00034	0.28	2.96		0.00034	0.28	2.96	
12	0.00988	1.09	3.18		0.00268	0.29	4.09		0.00268	0.29	4.09	
13	0.00246	0.82	2.35		0.00124	0.15	1.76		0.00124	0.15	1.76	
14	0.00458	0.55	0.45		0.00049	0.04	0.45		0.00049	0.04	0.45	
15	0.00374	0.66	1.64		0.00064	0.12	5.64		0.00064	0.12	5.64	
16	0.00538	1.36	3.18		0.00182	0.28	2.73		0.00182	0.28	2.73	
17	0.00695	0.53	1.76		0.00401	0.14	1.47		0.00401	0.14	1.47	
18	0.00275	1.33	3.18		0.00109	0.28	3.18		0.00109	0.28	3.18	
19	0.00361	0.90	2.73		0.00118	0.25	3.18		0.00118	0.25	3.18	
20	0.00197	1.08	3.64		0.00089	0.22	2.27		0.00089	0.22	2.27	

Tabela 5.7: Estatísticas do fluxo de potência reativa (IEEE14).

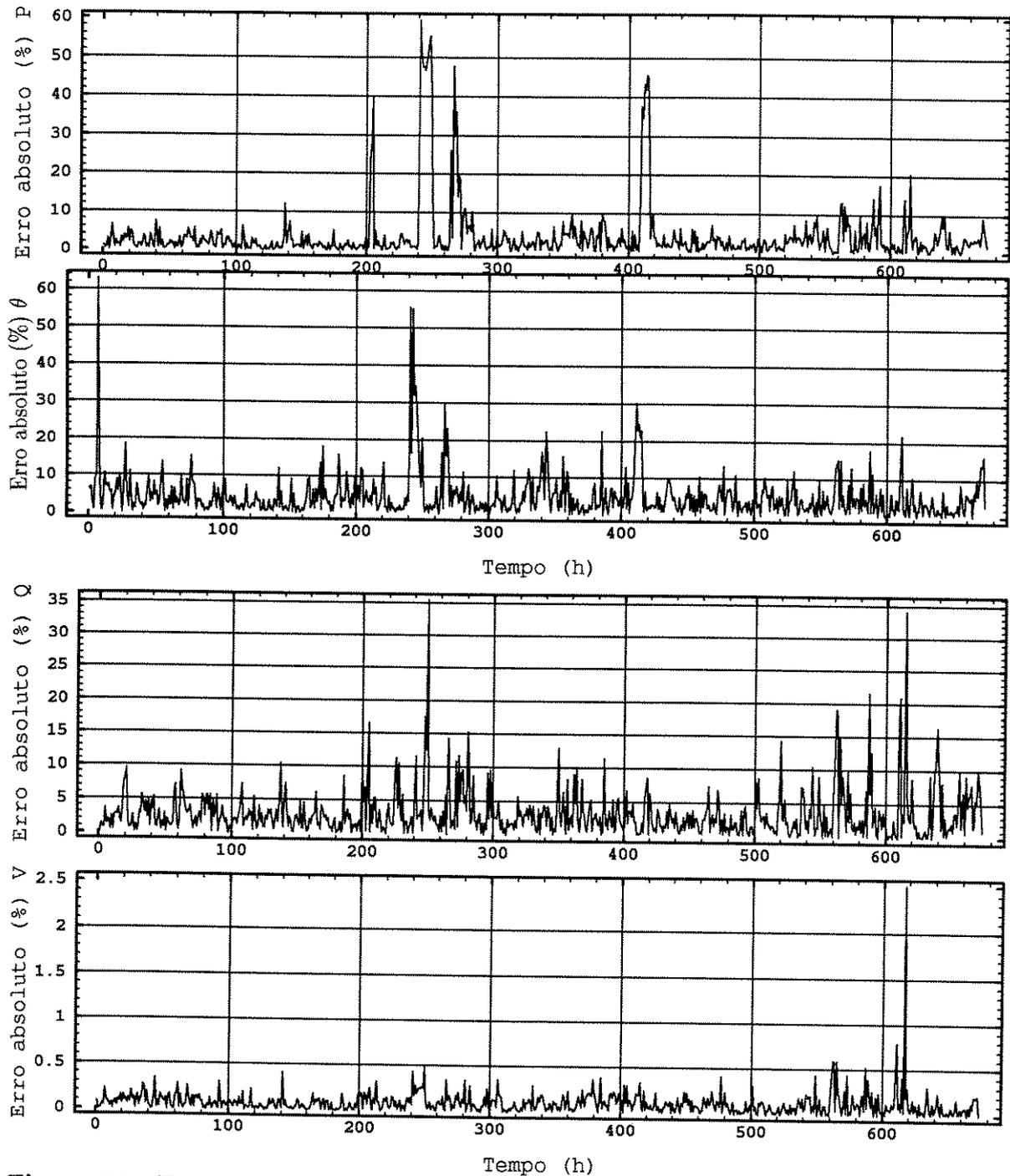


Figura 5.3: Erro na potência ativa e reativa e sua relação com o erro no ângulo e na tensão. Subestação 80

Capítulo 6

Conclusões

Este trabalho realizou uma análise da previsão de carga nodal ativa e reativa, a partir de dados reais de subestações com carga residencial, comercial e industrial, através de redes neurais artificiais.

Para isso foram propostos três modelos: Os dois primeiros utilizando rede MLP estática e o terceiro utilizando rede FIR, os quais foram designados por M que utiliza variáveis binárias para representar dia e hora, L que utiliza uma rede neural para cada hora do dia e F que por representar o tempo, incorpora automaticamente o ciclo diário da carga. Para cada um destes modelos foi designado um tipo de particionamento do conjunto de dados e que são definidos abaixo:

1. nenhum particionamento.
2. particionamento entre dia de semana e final de semana.
3. particionamento em cada dia da semana.
4. particionamento entre os dias: sábado, domingo, segunda e demais dias da semana.

Assim, formou-se o seguinte conjunto de modelos: M1, M2, M3, F1, F2, F3, F4, L3 e L4, cujos desempenhos em termos de erro médio de previsão foram analisados permitindo obter as conclusões que são apresentadas a seguir:

O modelo M1 apresentou os melhores resultados em comparação a todos os outros modelos, inclusive os de processamento temporal, sendo esse resultado válido para as variáveis, topologias e conjunto de treinamentos usados no presente estudo. O erro médio de previsão

com esse modelo é da ordem de 2.45% para a potência ativa e 4.09% para a potência reativa¹.

A forma de particionamento 3 foi a que apresentou o pior desempenho para todos os modelos, indicando que o excesso de especialização das redes neurais pode ser prejudicial à previsão.

O modelo M2, apesar de não ter apresentado o melhor desempenho, é uma alternativa ao modelo M1, tendo em vista que seu erro médio de previsão (2.72% potência ativa e 4.75% potência reativa) é comparável ao deste modelo, sendo que seu tempo de execução é menor.

De uma forma geral, verificou-se que quanto maior o conjunto de treinamento, melhor o desempenho da previsão ao utilizar menos neurônios.

Para todos os modelos observou-se a importância de encontrar uma relação adequada entre conjunto de treinamento X tamanho de rede, a fim de se diminuir o erro de previsão, indicando que é desejável elaborar um algoritmo construtivo ou de poda capaz de selecionar o número de neurônios necessários para obter o menor erro de precisão.

Conforme foi exposto no capítulo 5, o objetivo da previsão de carga nodal neste trabalho é obter uma previsão do estado do sistema de energia, para poder realizar a análise de segurança do mesmo. Dessa forma, apesar da previsão de potência reativa ser pior do que a de potência ativa, verificou-se para um sistema de duas barras e para um sistema de 14 barras (IEEE14) que o erro de previsão na potência ativa têm um efeito maior no erro de previsão do estado do que o erro de previsão na potência reativa.

6.1 Sugestão para Trabalhos Futuros

Os itens abaixo indicam algumas sugestões de estudos futuros que podem ser realizados:

- Verificação do desempenho de previsão de potência ativa e reativa através do uso de modelos NARMAX implementados por meio de redes recorrentes.
- Implementação de um modelo de previsão cujo algoritmo de treinamento seja rápido e eficiente, utilizando-se: métodos de otimização de segunda ordem, evitando, assim, que o usuário tenha que definir η e α , uma rede RBF (“Radial Basis Function”), ou outra metodologia que permita atingir esse objetivo.
- Implementação e estudo de algoritmos construtivos, uma vez que a topologia utilizada afeta o desempenho da previsão.
- Implementação e estudo de algoritmos que façam a seleção de variáveis de entrada.

¹Valores percentuais são obtidos em relação a uma base de potência.

- Desenvolvimento de um algoritmo paralelizado para realizar a previsão de carga nodal.
- Desenvolvimento de um algoritmo que gere índices de confiabilidade para para os valores previstos.
- Estudo da influência da previsão de carga na análise de segurança.

Apêndice A

Descrição do IEEE14 barras

Esse apêndice descreve o sistema IEEE14 de forma simplificada. As tabelas A.1 e A.2 são baseadas no formato CDF descrito pelo IEEE. Os valores de tensão 10000 devem ser compreendidos como sendo 1.0000 pu. Os tipos de barras possíveis são: 2 barra *slack*, 1 barra PV e 0 barra PQ. Muito embora a tabela A.1 mostre valores de potência consumidos nas barras PQ, os valores realmente utilizados nas simulações foram aqueles previstos pelos modelos do Capítulo 4.

Na tabela A.2 a barra origem e a barra destino indicam um ramo conectando essas duas barras, sendo que a numeração dos mesmos neste trabalho segue a ordem em que eles são apresentados na tabela. Os valores de R, X E Y estão em porcentagem na base 100 MVA.

N° barra	tipo	tensão	ângulo	P gerado	Q gerado	P consumido	Q consumido
1	2	1000	0.	0.0	0.0	0.0	0.0
2	1	1000	0.	25.0	0.0	0.0	0.0
3	0	1000	0.	0.0	0.0	15.0	5.0
4	0	1000	0.	0.0	0.0	15.0	5.0
5	0	1000	0.	0.0	0.0	15.0	5.0
6	1	1000	0.	25.0	0.0	0.0	0.0
7	0	1000	0.	0.0	0.0	15.0	5.0
8	0	1000	0.	0.0	0.0	15.0	5.0
9	0	1000	0.	0.0	0.0	15.0	5.0
10	0	1000	0.	0.0	0.0	15.0	5.0
11	0	1000	0.	0.0	0.0	15.0	5.0
12	0	1000	0.	0.0	0.0	15.0	5.0
13	0	1000	0.	0.0	0.0	15.0	5.0
14	0	1000	0.	0.0	0.0	15.0	5.0

Tabela A.1: Dados de barra do IEEE14 modificado.

barra origem hline 1	barra destino 2	R (%) 1.94	X (%) 59.2	Y (%) 5.28
1	5	5.40	22.3	4.92
2	3	4.70	19.8	4.38
2	4	5.81	17.6	3.74
2	5	5.70	17.4	3.40
3	4	6.70	17.1	3.46
4	5	1.34	4.2	1.28
4	7	0.0	20.9	0.00
4	9	0.0	55.6	0.00
5	6	0.0	25.2	0.00
6	11	9.5	19.9	0.00
6	12	12.3	25.6	0.00
6	13	6.61	13.0	0.00
7	8	0.0	17.6	0.00
7	9	0.0	11.0	0.00
9	10	3.18	8.5	0.00
9	14	12.7	27.4	0.00
10	11	8.2	19.2	0.00
12	13	22.1	20.0	0.00
13	14	17.1	34.8	0.00

Tabela A.2: Dados de linha do IEEE14.

Bibliografia

- [Abraham, 1983] Bovas Abraham and Johannes Ledolter. *Statistical Methods for Forecasting*. John Wiley & Sons, 1983.
- [Abu-El-Magdad, 1981] Mohamed A. Abu-El-Magdad and Naresh K. Sinha. Two new algorithms for on-line modelling and forecasting of the load demand of a multinode power system. *IEEE Trans on Power Apparatus and Systems*, 100(7), July 1981.
- [Bakirtzis, 1994] A. G. Bakirtzis, J. B. Theocharis, S. J. Kiartzis, and K.J. Satsios. Short term forecasting using fuzzy neural networks. IEEE/PES Winter Meeting, Jan 30 - Feb 03 1994. New York.
- [Bakirtzis, 1996] A. G. Bakirtzis, V. Petridis, S. J. Klartzis, M. C. Alexiadis, and A. H. Maissis. A neural network short term load forecasting model for the greek power system. *IEEE Transactions on Power Systems*, 11(2):858–869, May 1996.
- [Baumann, 1993] T. Baumann, H. Strasser, and H. Landrichter. Short term load forecasting methods in comparison: Kohonen learning, backpropagation learning, multiple regression analysis and kalman filters. In *Proceedings of the 11th PSCC*, 1993.
- [Box, 1976] G.E.P. Box and G.M. Jenkins. *Time Series Analysis Forecasting and Control*. Holden Day, 1976.
- [Campo, 1987] R. Campo and P. Ruiz. Adaptive weather-sensitive short term load forecasting. *IEEE Trans on Power Apparatus and systems*, 2(3), August 1987.
- [Chen, 1992] S.T. Chen, D.C. Yu, and A.R. Moghaddmjo. Weather sensitive short term load forecasting using nonfully connected neural network. *IEEE trans on Power Systems*, 7(3), August 1992.
- [Christiaanse, 1971] W.R. Christiaanse. Short term load forecasting using general exponential smoothing. *IEEE Trans on Power Apparatus and Systems*, PAS-90(2), March 1971.
- [Connor, 1994] J.T. Connor, R.D. Martin, and L.E. Atlas. Recurrent neural networks and robust time series prediction. *IEEE Trans on Neural Networks*, 5(2), March 1994.

- [Cottrell, 1995] Marie Cottrell, Bernard Girard, Yvone Girard, Morgan Mangeas, and Corine Muller. Neural modelling for time series: A statistical stepwise method for weight elimination. *IEEE Transactions on Neural Networks*, November 1995.
- [Couto, 1994] M. B. do Couto Filho and A. M. Leite da Silva. Estimação e previsão do estado operativo de sistemas de potência. Anais do 10º Congresso Brasileiro de Automática, 1994. Rio de Janeiro, RJ.
- [Czernichow, 1993] T. Czernichow, A. Germond, B. Dorizzi, and P. Caire. Improving recurrent network load forecasting. ANNPS, 1993. Yokohama-Japan.
- [Debs, 1988] Atif S. Debs. *Modern Power Systems Control and Operation*. Kluwer Academic Publishers, 1988.
- [Djukanovic, 1995] M. Djukanovic, S. Ruzic, B. Babic, D. J. Sobajic, and Y-H Pao. A neural net based short term load forecasting using moving window procedure. *Electrical Power & Energy Systems*, 17(6):391–397, 1995.
- [El-Sharkawi, 1996] Mohammed El-Sharkawi and Dagmar Niebur. Artificial neural networks with applications to power systems, 1996. IEEE Power Engineering Society (Tutorial Course). NY.
- [Falcão, 1986] D.M. Falcão and U.H. Bezerra. Short term forecasting of nodal active and reactive load in electric power systems. In *Proceedings 2nd IEE Int. Conf. on Power Systems Monitoring and Control*, 1986. Durhan, UK.
- [Fan, 1994] J.Y. Fan and J.D. McDonald. A real time implementation of short term load forecasting for distribution power systems. *IEEE Trans on Power Systems*, 9(2), May 1994.
- [Feuston, 1994] Bradley P. Feuston and John H. Thurtell. Generalized nonlinear regression with ensemble of neural nets: The great energy predictor shootout. *ASHRAE Transactions*, 100, 1994.
- [Freeman, 1991] J.A. Freeman and D.M. Skapura. *Neural Networks: Algorithms, Applications and Programming Techniques*. Addison-Wesley, 1991.
- [Galiana, 1974] Francisco D. Galiana, Edmund Handschin, and Albert R. Fiechter. Identification of stochastic electric load models from physical data. *IEEE transactions on automatic Control*, 19(6), December 1974.
- [Germond, 1996] A. J. Germond, A. Piras, B. Buchenel, K. Imhof, and Y. Jaccard. Heterogeneous artificial neural network for short term load forecasting. *IEEE Transactions on Power Systems*, 11(5):397–402, February 1996.
- [Gross, 1987] G. Gross and F.D. Galiana. Short term load forecasting. *Proceedings of the IEEE*, 75(12), Dec 1987.
- [Gupta, 1971] P.C. Gupta. A stochastic approach to peak power-demand forecasting in electric utility systems. *IEEE Trans on Power Apparatus and Systems*, 90(2), March/April 1971.

- [Hagan, 1987] M.T. Hagan and S.M. Behr. The time series approach to short term load forecasting. *IEEE Trans on Power Systems*, 2(3), August 1987.
- [Handschin, 1988] E. Handschin and Ch. Dörnemann. Bus load modelling and forecasting. *IEEE Trans on Power Systems*, 3(2), May 1988.
- [Haykin, 1994] Simon Haykin. *Neural Networks. A Comprehensive Foundation*. Macmillan Publishing Company, 1994.
- [Horton, 1987] Jerry S. Horton and David P. Gross Computer Configurations. *Proceedings of the IEEE*, December 1990. Apud Eduardo Nicola Ferraz Zagari. Escalonamento em tempo real das funções avançadas de análise de rede elétrica de um moderno centro de controle. Master's thesis, UNICAMP, 1996.
- [Hsu, 1991a] Yuan-Yih Hsu and Chien chuen Yang. Design of artificial neural networks for short term load forecasting. part i: Self-organising feature maps for day type identification. *IEE Proceedings-C*, 138(5):407–413, 1991.
- [Hsu, 1991b] Yuan-Yih Hsu and Chien chuen Yang. Design of artificial neural networks for short term load forecasting. part ii: Multilayer feedforward networks for peak load and valley load forecasting. *IEE Proceedings-C*, 138(5):414–418, 1991.
- [Jain, 1996] Anil K. Jain and Jianchang Mao. Artificial neural networks: A tutorial. *IEEE Computer Magazine*, March 1996.
- [Keyhani, 1983] A. Keyhani and S. M. Miri. On-line weather-sensitive and industrial group bus load forecasting for microprocessor-based applications. *IEEE Trans on Power Apparatus and Systems*, 102(12), December 1983.
- [Khotanzad, 1995] A. Khotanzad, R. C. Hwang, A. Abaye, and D. Maratukulam. An adaptive modular artificial neural network hourly load forecaster and its implementation at electric utilities. *IEEE Trans on Power Systems*, 10(3), August 1995.
- [Kiartzis, 1994] S.J. Kiartzis, A.G. Bakirtzis, and V. Petridis. Neural networks application to short term load forecasting. ISAP-94, 1994. Montpellier-Fr.
- [Kim, 1995] Kwang-Ho Kim, Jong-Keun Park, Kab-Ju Hwang, and Sung-Hak Kim. Implementation of hybrid short term load forecasting system using artificial neural networks and fuzzy expert systems. *IEEE Trans on Power Systems*, 10(3), August 1995.
- [Kreider, 1994] Jan F. Kreider and Jeff S. Haberl. Predicting hourly building energy use: the great energy predictor shootout-overview and discussion results. *ASHRAE Transactions*, 100, 1994.
- [Lee, 1992] K.Y. Lee, Y.T. Cha, and J.H. Park. Short term load forecasting using an artificial neural network. *IEEE Trans on Power Systems*, 7(1), February 1992.
- [MacKay, 1994] David J. C. MacKay. Bayesian nonlinear modeling for the prediction competition. *ASHRAE Transactions*, 100, 1994.

- [Mohammed, 1995] O. Mohammed, D. Park, R. Merchant, T. Dinh, C. Tong, and A. Azeem. Practical experiences with an adaptive neural network short-term load forecasting system. *IEEE Transactions on Power Systems*, 10(1):254–259, February 1995.
- [Monticelli, 1983] Alcir J. Monticelli. *Fluxo de Carga em Redes de Energia Elétrica*. Editora Edgard Blücher Ltda, 1983.
- [Mori, 1994] H. Mori and H. Kobayzashi. A fuzzy neural net for short term load forecasting. ISAP-94, 1994. Montpellier-Fr.
- [Mori, 1993] Hiroyuki Mori and Toshiji Ogasawara. A recurrent neural network for short term load forecasting. ANNPS, 1993. Yokohama-Japan.
- [Narendra, 1990] Kumpati S. Narendra. Identification and control of dynamical systems using neural networks. *IEEE Transactions on Neural Networks*, March 1990.
- [Narendra, 1996] Kumpati S. Narendra. Neural networks for control. *Proceedings of the IEEE*, October 1996.
- [Nazem, 1988] Sufi M. Nazem. *Applied Time Series Analysis for Business and Economic Forecasting*. Marcel Dekker, Inc., 1988.
- [Nerrand, 1993] O. Nerrand, P. Roussel-Ragot, L. Personnaz, and G. Dreyfuz. Neural networks and nonlinear adaptive filtering: Unifying concepts and new algorithms. *Neural Computation*, 5:165–199, 1993.
- [Nix, 1995] David A. Nix and Andreas S. Weigend. Learning local error bars for nonlinear regression. In G. Tesauro, D. S. Touretzky, and T. K. Leen, editors, *Advances in Neural Information Processing Systems*, pages 489–496. MIT Press, 1995. Cambridge,MA.
- [Ohlsson, 1994] Mattias B.O. Ohlsson, Carsten O. Peterson, Hong Pi, Thorstein S. Rögnvaldsson, and Bo P. W. Söderberg. Predicting system loads with artificial neural networks-methods and results from the great energy predictor shootout. *ASHRAE Transactions*, 100, 1994.
- [Papalexopoulos, 1990] Alex D. Papalexopoulos and Timothy C. Hesterberg. A regression-based approach to short term load forecasting. *IEEE Trans. on Power Systems*, 5(4):1535–1544, Nov 1990.
- [Papalexopoulos, 1994] A.D. Papalexopoulos, S. Hao, and T.M Peng. An implementation of a neural network based load forecasting model for the ems. IEEE/PES Winter Meeting, Jan 30 - Feb 03 1994. New York.
- [Park, 1991] D.C. Park, M.A. El-Sharkawi, R.J. Marks II, L.E Atlas, and M.J. Damborg. Electric load forecasting using an artificial neural network. *IEEE Trans. on Power Systems*, 6:442–449, May 1991.
- [Park, 1996] Young R. Park, Thomas J. Murray, and Chung Chen. Predicting sun spots using a layered perceptron neural networks. *IEEE Transactions on Neural Networks*, March 1996.

- [Peng, 1992] T.M. Peng, N.F. Hubele, and G.G. Karady. Advancements in the application of neural networks for short term load forecasting. *IEEE Trans. on Power Systems*, 7(1), February 1992.
- [Rumelhart, 1986] D.E. Rumelhart, G.E. Hinton, and R.J. Williams. Learning internal representations by error propagation. In D.E. Rumelhart and J.L. McClelland, editors, *Parallel Distributed Processing: Explorations in The Microstructures of Cognition.*, pages 318–362. MIT Press, 1986. Cambridge, MA.
- [Rush, 1993] Dorn H. Rush and Bill G. Horne. Progress in supervised neural networks: What's new since lippmann ? *IEEE Signal Processing Magazine*, January 1993.
- [Sato, 1994] R. Sato, E. Tanaka, and J. Hasegawa. Daily load forecasting using a neural network combined with regression analysis. ISAP-94, 1994. Montpellier-Fr.
- [Shang, 1996] Yi Shang and Benjamin W. Wah. Global optimization for neural network training. *IEEE Computer Magazine*, March 1996.
- [Sharkawi, 1993] M. A. El Sharkawi, R. J. Marks II, S. Oh, and C. M. Brace. Data partitioning for training a layered perceptron to forecast electric load. ANNPS, 1993. Yokohama-Japan.
- [Silva, 1993] A. P. Alves da Silva, A. M. Leite da Silva, J. C. S. de Souza, and M. B. do Couto Filho. State forecasting based on artificial neural networks, Ago 1993.
- [Sjöberg, 1995] Jonas Sjöberg. *Non-Linear system identification with neural networks*. PhD thesis, Linköping University, S-581,83, Linköping, Sweden, 1995. Dept. of Electrical Engineering.
- [Srinivasan, 1995] Dipti Srinivasan, C. S. Chang, and A. C. Liew. Demand forecasting using fuzzy neural computation, with special emphasis on weekend and public holiday forecasting. *IEEE Transactions on Power Systems*, 10(4):1897–1903, November 1995.
- [Strasser, 1990] H. Strasser, N. Friemelt, and G. Schellstede. Short term load forecasting using multiple regression analysis or adaptive regression analysis. In *Proceedings of the 10th PSCC*, Graz., 1990.
- [Vemuri, 1981] S. Vemuri, W.L. Huang, and D.J. Nelson. On line algorithms for forecasting hourly loads of an electric utility. *IEEE Trans on Power Apparatus and Systems*, PAS-100(8), August 1981.
- [Von Zuben, 1996] Fernando José Von Zuben. *Modelos paramétricos e não paramétricos de Redes Neurais Artificiais e Aplicações*. PhD thesis, Universidade Estadual de Campinas, Fevereiro 1996.
- [Wan, 1993a] Eric A. Wan. Finite impulse response neural networks for autoregressive time series prediction. In A. Weigend and N. Gershenfeld, editors, *Proceedings of the NATO Advanced Workshop on Time Series Prediction and Analysis*. Addison-Wesley, 1993. Santa Fe, NM.

-
- [Wan, 1993b] Eric A. Wan. *Finite Impulse Response Neural Networks with Applications in Time Series Prediction*. PhD thesis, Stanford University, November 1993.
- [Widrow, 1990] B. Widrow and M.A. Lehr. 30 years of adaptive neural networks: Perceptrons, madaline and backpropagation. *Proceedings of the IEEE*, September 1990.
- [Williams, 1989] Ronald J. Williams and David Zipser. A learning algorithm for continually running fully recurrent neural networks. *Neural Computation*, 1:270–280, 1989.
- [Wood, 1996] Allen J. Wood and Bruce F. Wollenberg. *Power Generation, Operation and Control*. John Wiley & Sons, Inc, 1996.
- [Zagari, 1996] Eduardo Nicola Ferraz Zagari. Escalonamento em tempo real das funções avançadas de análise de rede elétrica de um moderno centro de controle. Master's thesis, UNICAMP, 1996.