

**SILVIO MICHEL DE ROCCO**

**PREVISÃO DE DEMANDAS PARA  
CONTROLE DOS MONTANTES DE USO  
DO SISTEMA DE TRANSMISSÃO PARA  
DISTRIBUIDORAS DE ENERGIA  
ELÉTRICA**

Dissertação apresentada como requisito parcial para a obtenção do grau de Mestre, no Programa de Pós-Graduação em Engenharia Elétrica, Setor de Tecnologia, Universidade Federal do Paraná.

Orientador: Dr. Alexandre Rasi Aoki.

**Curitiba, 2008.**

## **AGRADECIMENTOS**

Agradeço a Deus, em primeiro lugar, pelas oportunidades oferecidas na minha vida.

Aos meus Pais, Anibal e Marlene, pelo apoio.

A minha esposa, Luciane, e meus filhos, Cristian e Thais, pela paciência.

Ao Dr. Alexandre Rasi Aoki pela dedicação, paciência e empenho e finais de semana comprometidos com a orientação.

Aos Profs. Drs. Marcus Vinicius Lamar (UNB), Thelma S. Piazza Fernandes (UFPR) e Elizete Maria Lourenço (UFPR) pelas contribuições como avaliadores deste trabalho.

Aos Pesquisadores Vitor Hugo Ferreira e Fabio Alessandro Guerra (LACTEC) pelo apoio no desenvolvimento das simulações.

A COPEL Distribuição pela liberação parcial de tempo para execução desta capacitação profissional. Destacando, o Sr. Antonio Justino Spinello pelo apoio e incentivo.

## SUMÁRIO

<b>AGRADECIMENTOS</b> .....	<b>i</b>
<b>SUMÁRIO</b> .....	<b>ii</b>
<b>LISTA DE TABELAS</b> .....	<b>v</b>
<b>LISTA DE FIGURAS</b> .....	<b>vii</b>
<b>LISTA DE SIGLAS</b> .....	<b>x</b>
<b>LISTA DE SÍMBOLOS</b> .....	<b>xi</b>
<b>RESUMO</b> .....	<b>xii</b>
<b>ABSTRACT</b> .....	<b>xiii</b>
<b>1 INTRODUÇÃO</b> .....	<b>1</b>
<b>1.1 INTRODUÇÃO</b> .....	<b>1</b>
1.1.1 Monitoramento .....	2
1.1.2 Contratação.....	3
<b>1.2 DESCRIÇÃO DO PROBLEMA</b> .....	<b>4</b>
<b>1.3 OBJETIVOS</b> .....	<b>5</b>
1.3.1 Objetivo Geral .....	5
1.3.2 Objetivos Específicos .....	5
<b>1.4 ESTADO DA ARTE</b> .....	<b>6</b>
1.4.1 Previsão por Técnicas de Inteligência Artificial.....	8
1.4.2 Previsão de Demanda a Curto Prazo .....	11
1.4.3 Previsão de Demanda a Longo Prazo .....	12
1.4.4 Considerações Finais sobre Previsão.....	12
<b>1.5 ESTRUTURA DA DISSERTAÇÃO</b> .....	<b>13</b>
<b>2 CARACTERIZAÇÃO E ANÁLISE DOS DADOS</b> .....	<b>15</b>
<b>2.1 INTRODUÇÃO</b> .....	<b>15</b>
<b>2.2 LEVANTAMENTO DA BASE DE DADOS</b> .....	<b>15</b>
<b>2.3 ANÁLISE GERAL DA BASE DE DADOS</b> .....	<b>17</b>
<b>2.4 ANÁLISE DA MÉDIA, TENDÊNCIAS E DESVIOS PADRÃO</b> .....	<b>23</b>
<b>2.5 ANÁLISE E INTERPOLAÇÃO DOS DADOS FALTANTES</b> .....	<b>25</b>
<b>2.6 REMOÇÃO DOS <i>OUTLIERS</i></b> .....	<b>27</b>
<b>2.7 AUTOCORRELAÇÃO</b> .....	<b>30</b>

2.8	<b>CONSIDERAÇÕES FINAIS DO CAPÍTULO</b> .....	35
<b>3</b>	<b>METODOLOGIA DE PREVISÃO DE SÉRIES TEMPORAIS</b> .....	<b>37</b>
3.1	<b>INTRODUÇÃO</b> .....	37
3.2	<b>MÁQUINAS DE VETOR SUPORTE PARA REGRESSÃO</b> .....	39
3.2.1	Preparação dos Dados.....	42
3.2.2	Implementação .....	44
3.3	<b>REDES NEURAS ARTIFICIAIS</b> .....	45
3.3.1	Rede Neural de Função de Base Radial .....	48
3.3.2	Teorema de Micchelli.....	52
3.3.3	Aplicações .....	53
3.4	<b>CONSIDERAÇÕES SOBRE O CAPÍTULO</b> .....	53
<b>4</b>	<b>RESULTADOS OBTIDOS</b> .....	<b>55</b>
4.1	<b>INTRODUÇÃO</b> .....	55
4.2	<b>PREVISÃO USANDO MÁQUINAS DE VETOR SUPORTE PARA REGRESSÃO – SVR</b> .....	55
4.2.1	Parâmetros Básicos que Variaram Durante o Estudo.....	57
4.2.2	Período de Interesse do Histórico para Análise.....	57
4.2.3	Número de Simulações de Treinamentos e Testes Utilizados.....	57
4.2.4	Sazonalidade.....	58
4.2.5	Número de Previsões.....	58
4.2.6	Tempo de Processamento .....	58
4.2.7	Alternativas Usadas nos Estudos com SVR .....	59
4.3	<b>RESULTADOS DO SVR</b> .....	60
4.3.1	SE Campo do Assobio.....	61
4.3.2	SE Cascavel.....	72
4.3.3	SE Ibiporã.....	74
4.3.4	SE Foz do Chopin.....	76
4.3.5	SE Londrina.....	78
4.3.6	SE Pato Branco.....	80
4.3.7	SE Umbará.....	82
4.3.8	Resumo dos Resultados do SVR para Março/2007.....	84
4.3.9	Previsões Efetuadas para o Mês de Agosto/2006.....	86
4.3.10	Considerações Sobre o Uso de SVR .....	88
4.4	<b>RESULTADOS DAS RNA-RBF</b> .....	89

4.5	COMPARAÇÃO SVR VERSUS RNA-RBF.....	93
5	CONCLUSÕES E DESENVOLVIMENTOS FUTUROS .....	96
5.1	CONCLUSÕES FINAIS .....	96
5.1.1	Uso da SVR para Previsões do MUST .....	96
5.1.2	Uso de RNA-RBF para Previsões do MUST. ....	98
5.1.3	Análise Comparativa SVR versus RNA-RBF .....	98
5.2	SUGESTÃO PARA TRABALHOS FUTUROS .....	99
	REFERÊNCIAS BIBLIOGRÁFICAS.....	100
A	CARACTERIZAÇÃO E ANÁLISE DOS DADOS .....	108
A.1	ANÁLISE DA MÉDIA, TENDÊNCIAS E DESVIOS PADRÕES .....	108

## LISTA DE TABELAS

Tabela 2.1 – Interpretação de Correlação.....	31
Tabela 2.2 – Análise Estatística dos dados ajustados das SEs. ....	31
Tabela 4.1 – SE CAS – Erro em kW da Alternativa 1. ....	62
Tabela 4.2 – SE CAS – Erro Absoluto em kW da Alternativa 1.....	62
Tabela 4.3 – SE CAS – Erro Percentual da Alternativa 1. ....	63
Tabela 4.4 – SE CAS – Erro Absoluto Percentual da Alternativa 1. ....	63
Tabela 4.5 – SE CAS – Erro em kW da Alternativa 2. ....	64
Tabela 4.6 – SE CAS – Erro Absoluto em kW da Alternativa 2.....	64
Tabela 4.7 – SE CAS – Erro Percentual da Alternativa 2. ....	64
Tabela 4.8 – SE CAS – Erro Absoluto Percentual da Alternativa 2. ....	65
Tabela 4.9 – SE CAS – Erro em kW da Alternativa 3. ....	65
Tabela 4.10 – SE CAS – Erro Absoluto em kW da Alternativa 3.....	65
Tabela 4.11 – SE CAS – Erro Percentual da Alternativa 3. ....	65
Tabela 4.12 – SE CAS – Erro Absoluto Percentual da Alternativa 3. ....	66
Tabela 4.13 – SE CAS – Erro em kW da Alternativa 4. ....	66
Tabela 4.14 – SE CAS – Erro Absoluto em kW da Alternativa 4.....	66
Tabela 4.15 – SE CAS – Erro Percentual da Alternativa 4. ....	66
Tabela 4.16 – SE CAS – Erro Absoluto Percentual da Alternativa 4. ....	67
Tabela 4.17 – SE CAS – Erros Percentuais para as 4 Alternativas.....	67
Tabela 4.18 – SE CAS Erros Percentuais Sem Finais de Semana - Alternativa 1.....	69
Tabela 4.19 – SE CAS Erros Percentuais Sem Finais de Semana - Alternativa 2.....	70
Tabela 4.20 – SE CAS Erros Percentuais Comparativos - Alternativas 1 e 2, desconsiderando-se os finais de semana das estatísticas.....	71
Tabela 4.21 – SE CEL Comparativo entre as 4 Alternativas com Finais de Semana. ....	72
Tabela 4.22 – SE CEL Comparativo entre as Alternativas 1 e 2 sem Finais de Semana.....	73
Tabela 4.23 – SE IBI Comparativo entre as 4 Alternativas com Finais de Semana. ....	74
Tabela 4.24 – SE IBI Comparativo entre as Alternativas 1 e 2 sem Finais de Semana. ....	75
Tabela 4.25 – SE FOZ Comparativo entre as 4 Alternativas com Finais de Semana. ....	76
Tabela 4.26 – SE FOZ Comparativo entre as Alternativas 1 e 2 sem Finais de Semana.....	77
Tabela 4.27 – SE LNA Comparativo entre as 4 Alternativas com Finais de Semana.....	78

Tabela 4.28 – SE FOZ Comparativo entre as Alternativas 1 e 2 sem Finais de Semana.....	79
Tabela 4.29 – SE PTO Comparativo entre as 4 Alternativas com Finais de Semana. ....	80
Tabela 4.30 – SE FOZ Comparativo entre as Alternativas 1 e 2 sem Finais de Semana.....	81
Tabela 4.31 – SE UMB Comparativo entre as 4 Alternativas com Finais de Semana.....	82
Tabela 4.32 – SE UMB Comparativo entre as Alternativas 1 e 2 sem Finais de Semana. ....	83
Tabela 4.33 – Resumo das Previsões Usando SVR para Março/2007 com Finais de Semana.	84
Tabela 4.34 – Resumo das Previsões Usando SVR para Março/2007 sem Finais de Semana.	84
Tabela 4.35 – Resumo das Previsões Usando SVR do dia D+1 com Finais de Semana. ....	85
Tabela 4.36 – Resumo das Previsões Usando SVR do dia D+1 sem Finais de Semana.....	85
Tabela 4.37 – Resumo Comparativo das Previsões Usando SVR do dia D+1 com e sem Finais de Semana.....	86
Tabela 4.38 – Resumo Comparativo de Todas as SEs - dias D+1, ..., D+7.....	87
Tabela 4.39 – Resumo Comparativo de Todas as SEs - dia D+1.....	87
Tabela 4.40 – Número de Neurônios na Camada Escondida – Março/2007.....	90
Tabela 4.41 – Comparativo para D+1 com RNA-RBF – Março/2007.....	91
Tabela 4.42 – Número de Neurônios na Camada Escondida – Agosto/2006.....	92
Tabela 4.43 – Comparativo para D+1 com RNA-RBF – Agosto/2006. ....	92
Tabela 4.44 – Comparativo das Alternativas 1 e 2 com SVR e RNA-RBF – Março/2007. ....	93
Tabela 4.45 – Comparativo das Alternativas 1 e 2 com SVR e RNA-RBF – Agosto/2006. ....	94

## LISTA DE FIGURAS

Figura 1.1 – Ponto de Conexão com a Rede Básica.....	2
Figura 2.1 – Dados originais da SE CAS – Período 2004/2006.....	19
Figura 2.2 – Dados originais da SE CEL – Período 2004/2006.....	20
Figura 2.3 – Dados originais da SE LNA – Período 2004/2006.....	20
Figura 2.4 – Dados originais da SE IBI – Período 2004/2006.....	21
Figura 2.5 – Dados originais da SE FOZ – Período 2004/2006.....	21
Figura 2.6 – Dados originais da SE PTO – Período 2004/2006.....	22
Figura 2.7 – Dados originais da SE UMB – Período 2004/2006.....	22
Figura 2.8 – Dados originais da SE CAS – Período 2004/2006 – Discretização plurianual....	23
Figura 2.9 – Dados originais da SE CAS – Período 2004/2006 – Discretização anual.....	24
Figura 2.10 – Dados originais da SE CAS – Período 2005/2006 – Curva de tendência.....	25
Figura 2.11 – Dados da SE CEL com Interpolação – Período 2004/2006.....	26
Figura 2.12 – Dados da SE LON com Interpolação – Período 2005/2006.....	26
Figura 2.13 – Dados da SE IBI com Interpolação – Período 2005/2006.....	27
Figura 2.14 – Dados da SE CAS de Segunda-feira – Com e Sem <i>Outliers</i> .....	27
Figura 2.15 – Dados da SE CAS sem <i>Outliers</i> – Período 2005/2006.....	28
Figura 2.16 – Dados da SE CEL sem <i>Outliers</i> – Período 2005/2006.....	28
Figura 2.17 – Dados da SE LNA sem <i>Outliers</i> – Período 2005/2006.....	28
Figura 2.18 – Dados da SE IBI sem <i>Outliers</i> – Período 2005/2006.....	29
Figura 2.19 – Dados da SE FOZ sem <i>Outliers</i> – Período 2005/2006.....	29
Figura 2.20 – Dados da SE PTO sem <i>Outliers</i> – Período 2005/2006.....	29
Figura 2.21 – Dados da SE UMB sem <i>Outliers</i> – Período 2005/2006.....	30
Figura 2.22 – Autocorrelação dos dados da SE CAS – Período 2005/2006.....	32
Figura 2.23 – Autocorrelação dos dados da SE CEL – Período 2005/2006.....	32
Figura 2.24 – Autocorrelação dos dados da SE LNA – Período 2005/2006.....	33
Figura 2.25 – Autocorrelação dos dados da SE IBI – Período 2005/2006.....	33
Figura 2.26 – Autocorrelação dos dados da SE FOZ – Período 2005/2006.....	34
Figura 2.27 – Autocorrelação dos dados da SE PTO – Período 2005/2006.....	34
Figura 2.28 – Autocorrelação dos dados da SE UMB – Período 2005/2006.....	35
Figura 3.1 – Representação de Vetores de Suporte.....	44



Figura 3.3 – Estrutura Geral de uma Rede Neural RN-MLP. ....	47
Figura 3.4 – Estrutura Geral de uma Rede Neural RN-RBF. ....	48
Figura 3.5 – MLP x RBF. ....	49
Figura 4.1 – Erro Absoluto Médio Percentual da SE CAS. ....	68
Figura 4.2 – Erro Médio Percentual da SE CAS. ....	68
Figura 4.3 – Comparativo das Alternativas 1 e 2. ....	71
Figura 4.4 – SE CEL Comparativo entre as 4 Alternativas com Finais de Semana. ....	72
Figura 4.5 – SE CEL Comparativo entre as Alternativas 1 e 2 sem Finais de Semana. ....	73
Figura 4.6 – SE IBI Comparativo entre as 4 Alternativas com Finais de Semana. ....	74
Figura 4.7 – SE IBI Comparativo entre as Alternativas 1 e 2 sem Finais de Semana. ....	75
Figura 4.8 – SE FOZ Comparativo entre as 4 Alternativas com Finais de Semana. ....	76
Figura 4.9 – SE FOZ Comparativo entre as Alternativas 1 e 2 sem Finais de Semana. ....	77
Figura 4.10 – SE LNA Comparativo entre as 4 Alternativas com Finais de Semana. ....	78
Figura 4.11 – SE LNA Comparativo entre as Alternativas 1 e 2 sem Finais de Semana. ....	79
Figura 4.12 – SE PTO Comparativo entre as 4 Alternativas com Finais de Semana. ....	80
Figura 4.13 – SE PTO Comparativo entre as Alternativas 1 e 2 sem Finais de Semana. ....	81
Figura 4.14 – SE UMB Comparativo entre as 4 Alternativas com Finais de Semana. ....	82
Figura 4.15 – SE UMB Comparativo entre as Alternativas 1 e 2 sem Finais de Semana. ....	83
Figura A.1 – SE CEL Dados dos Anos de 2004 a 2006. ....	108
Figura A.2 – SE CEL Dados dos Anos de 2004 a 2006 e Curva de Tendência. ....	109
Figura A.3 – SE LNA Dados dos Anos de 2004 a 2006 e Curva de Tendência. ....	110
Figura A.4 – SE LNA Dados dos Anos de 2005 a 2006 e Curva de Tendência. ....	110
Figura A.5 – SE LNA Dados dos Anos de 2004 a 2006 e Curva de Tendência. ....	111
Figura A.6 – SE IBI Dados dos Anos de 2004 a 2006 e Curva de Tendência. ....	112
Figura A.7 – SE IBI Dados dos Anos de 2005 a 2006 e Curva de Tendência. ....	112
Figura A.8 – SE IBI Dados dos Anos de 2004 a 2006 e Curva de Tendência. ....	113
Figura A.9 – SE FOZ Dados dos Anos de 2004 a 2006 e Curva de Tendência. ....	114
Figura A.10 – SE FOZ Dados dos Anos de 2005 a 2006 e Curva de Tendência. ....	114
Figura A.11 – SE FOZ Dados dos Anos de 2004 a 2006 e Curva de Tendência. ....	115
Figura A.12 – SE PTO Dados dos Anos de 2004 a 2006 e Curva de Tendência. ....	116
Figura A.13 – SE PTO Dados dos Anos de 2004 a 2006 e Curva de Tendência. ....	116
Figura A.14 – SE PTO Dados dos Anos de 2005 a 2006 e Curva de Tendência. ....	117
Figura A.15 – SE UMB Dados dos Anos de 2004 a 2006 e Curva de Tendência. ....	118
Figura A.16 – SE UMB Dados dos Anos de 2004 a 2006 e Curva de Tendência. ....	118

Figura A.17 – SE UMB Dados dos Anos de 2004 a 2006 e Curva de Tendência. .... 119

## LISTA DE SIGLAS

ANEEL - Agência Nacional de Energia Elétrica  
ANN – *Artificial Neural Network*  
CCEE – Câmara de Comercialização de Energia Elétrica  
COPEL - Companhia Paranaense de Energia  
CUST – Contrato de Uso do Sistema de Transmissão  
MUST – Montante de Uso do Sistema de Transmissão  
ONS - Operador Nacional do Sistema Elétrico  
R.B. - Rede Básica  
RBF – *Radial Basis Function Neural Network*  
SE – Subestação  
SVM – *Support Vector Machine*  
SVR – *Support Vector Machine to Regression*  
SV - Vetores de Suporte  
RN-MLP - Rede Neural *Multi Layer Perceptron*  
RNA – Rede Neural Artificial

## LISTA DE SÍMBOLOS

$C$	Custo, determina a troca entre a planicidade de $f$ e o tanto que os desvios maiores que $\varepsilon$ são tolerados nas SVMs.
$d$	<i>vetor resposta desejada</i> em redes neurais
$w_i$	índice de ponderação utilizado nas simulações de redes neurais (peso do vetor linear)
$\Phi$	matriz de interpolação usada em redes neurais
$(x_i, y_i)$	Conjunto de treino (SVR)
$\varepsilon$	Desvio (determina a precisão de interesse nas SVR)
$y_i$	Carga do $i$ -ésimo dia
$\ w\ ^2$	Norma Euclidiana
$\xi_i, \xi_i^*$	Variáveis de folga

## RESUMO

Com o advento da Lei Federal nº. 10.848/2004, onde os agentes de distribuição devem garantir a sua remuneração apenas por meio da gestão de seus ativos, existe necessidade de um maior controle e gerenciamento dos fluxos de potência passantes pelos pontos de conexão do sistema de distribuição com a rede básica em função dos montantes de uso do sistema de transmissão contratados, a fim de se evitar tanto as subcontratações que podem resultar em penalidades para a distribuidora, como os excessos indesejados que oneram o consumidor. Portanto, há a necessidade por parte do agente distribuidor de prever os fluxos passantes em cada ponto de conexão, com exatidão e tempo hábil suficiente para permitir que possam ser tomadas medidas operativas, a fim de se evitar a ultrapassagem dos valores contratados. Sendo assim, pretende-se implementar técnicas de previsão de séries temporais utilizando-se inteligência artificial, Redes Neurais Artificiais e Máquinas de Vetor Suporte, para previsão destes fluxos nos pontos de conexão.

Aplicando-se a técnica de máquinas de vetor suporte com regressão utilizando dados reais de uma distribuidora, depreende-se que os resultados obtidos alcançaram valores de erro médio percentual dentro dos limites cabíveis no problema real, com desempenho computacional suficiente para permitir a sua aplicação prática em casos reais

**PALAVRAS-CHAVE:** Demanda por ponto de conexão, Montantes de Uso do Sistema de Transmissão, *Support Vector Regression*, Redes de Função de Base Radial.

## ABSTRACT

With the advent of the Federal law no. 10.848/2004, where the agents of distribution should ensure their profits only through management of its assets, there is need for greater control and management of the flow of power by bystanders points of connection on the basis of the contracted amounts, and should be avoided the subcontracts that can result in penalties for the distributor, as the excesses that unwanted borne the costumers. So there is a need for the Distribution Agent to predict the flow bystanders at each point of connection, with accuracy and time enough to allow operational measures to be taken to prevent a breach of contract values. Therefore, it is intended to implement techniques for forecasting of time series, using artificial intelligence, Artificial Neural Network and Support Vector Machines, to forecast the flows at the point of connection. Applying to technical support vector machines with regression using actual data from an utility, it is clear that the results achieved value of average error percentage within the limits applicable in a real problem with performance computational enough to allow practical application in real cases.

KEY-WORDS: Power Flow, *Montantes de Uso do Sistema de Transmissão*, Support Vector Regression, Radial Basis Function.

# 1 INTRODUÇÃO

## 1.1 INTRODUÇÃO

O projeto RE-SEB – Reestruturação do Setor Elétrico Brasileiro (PAIXÃO, 1998), que se iniciou em 1995, teve sua efetiva implementação e produziu as primeiras regulamentações a partir de 1998, estabelecendo diretrizes que resultaram em um novo modelo para o setor elétrico brasileiro.

Uma das obrigаторiedades da nova regulamentação foi de que os custos incorridos pelo uso da rede básica em função da demanda dos fluxos de potência passantes (Montantes de Uso do Sistema de Transmissão – MUSTs), da demanda de carga e de geração de cada agente fossem rateados igualmente entre as partes tal qual um sistema condomínio.

Para atender este novo arcabouço legal, as distribuidoras passaram a contratar, junto ao Operador Nacional do Sistema - ONS, demandas de Rede Básica – RB, que consistem em demandas de potência máximas globais integralizadas a cada 15 minutos para o horário compreendido como de ponta (das 18 às 21 horas), pagando a denominada tarifa “selo”, por analogia com outros países que já haviam adotado tal sistemática.

Estes valores de potência máxima passantes entre a rede básica e os agentes conectados são então contratados, para períodos de no mínimo três anos, junto às empresas transmissoras com um ano de antecedência (até fins de outubro de um determinado ano informa-se os valores para o ano seguinte). A contratação se dá através do estabelecimento de Contratos de Uso do Sistema de Transmissão - CUST, para tanto, as transmissoras são representadas pelo ONS, que é o interveniente legal pelas transmissoras.

A partir do ano de 2003 por força de resolução da Agência Nacional de Energia Elétrica – ANEEL, órgão regulador do setor elétrico brasileiro, os valores de fluxo de potência demandados que pagavam tarifas “selo”, conforme Figura 1.1, passaram a ser descontratados, à razão de 25% ao ano. Os montantes faltantes para complementar a totalidade demandada pelas cargas (distribuidora + cliente livre), a partir de 1º de janeiro de

2007, passaram então a ser contratados pagando tarifa denominada “nodal” para cada Ponto de Conexão com a rede básica, para o horário de ponta do sistema elétrico nacional.

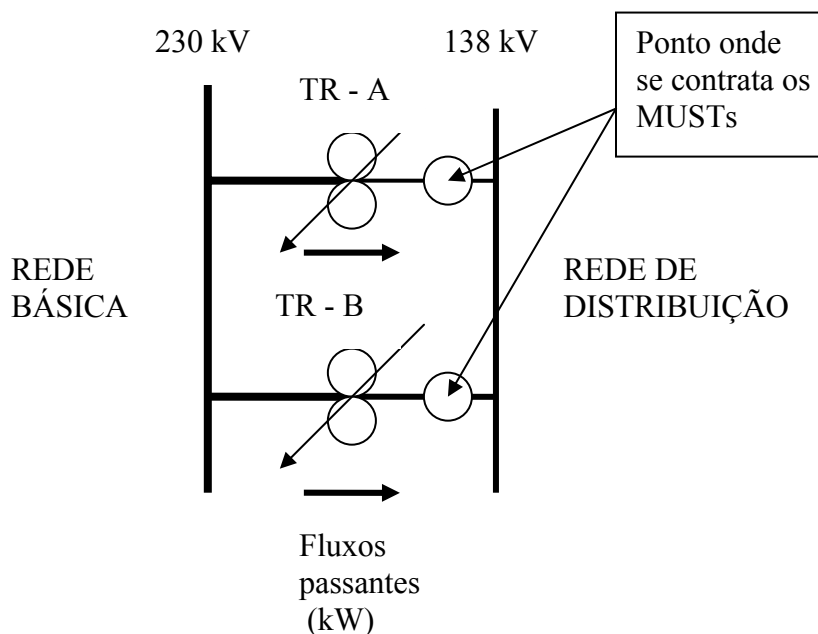


Figura 1.1 – Ponto de Conexão com a Rede Básica.

Para o estabelecimento do MUST a ser contratado, considera-se a maior demanda integralizada de 15 em 15 minutos, no horário de ponta do sistema, para cada ponto de conexão. Uma vez estabelecido o valor do MUST para um certo ponto de conexão, os usuários pagam um valor em reais obtido através da multiplicação do valor do MUST contratado em kW pela tarifa nodal, que é fornecida por resolução específica da ANEEL em R\$/kW.

Para o horário fora de ponta existe a necessidade de contratação dos fluxos passantes, mas não há penalizações por sub ou sobrecontratações, portanto análises para este horário não farão parte do escopo deste trabalho.

### 1.1.1 Monitoramento

É importante ressaltar que existe todo um sistema de medições para o monitoramento dos fluxos passantes pelos pontos de conexão onde, tanto a carga (distribuidora + cliente livre) ou geradora, quanto a Câmara de Comercialização de Energia Elétrica - CCEE e o ONS tem acesso ao valor medido em tempo real.



Esta dissertação foca exclusivamente o caso das distribuidoras, onde os pontos de conexão são os pontos de fronteira entre as instalações da rede básica e as de âmbito próprio do distribuidor, ou os pontos de fronteira disponibilizados exclusivamente a este através de Contratos de Conexão a Transmissão – CCT.

Os agentes têm a responsabilidade de controlar e verificar as informações de demanda por ponto de conexão, assim torna-se indispensável o monitoramento dos dados de fluxos passantes por ponto de conexão com a rede básica, de modo a permitir um gerenciamento dos valores contratados.

Os MUSTs por ponto de conexão contratados devem considerar a demanda máxima total prevista.

### **1.1.2 Contratação**

A regulamentação prevê a obrigatoriedade do agente distribuidor estabelecer o CUST, da mesma forma, determina que os encargos de uso do sistema de transmissão dos agentes de distribuição sejam aplicados ao mês pleno, ou seja, paga-se o valor do contrato ou a maior demanda verificada por ponto de conexão em cada mês para o horário de ponta.

A nova versão do modelo regulador implantado com o advento da Lei nº 10.848/2004, onde os agentes de distribuição devem garantir a sua remuneração por meio de gestão de seus ativos, torna muito importante o controle e gerenciamento dos montantes contratados. Deve-se evitar tanto excessos indesejados, com sobrecontratações que oneram os consumidores devido a premissa que a modicidade tarifária deve ser sempre buscada, quanto as subcontratações que podem resultar em multa para a distribuidora.

As demandas contratadas por ponto de conexão junto ao ONS devem refletir as previsões de demandas máximas integralizadas de 15 em 15 minutos e também a sazonalidade mensal dos fluxos passantes contratados que, não necessariamente são coincidentes ao longo do ano. É importante citar que estes montantes contratados não devem ser ultrapassados em mais do que 5%, sob o risco de se pagar três vezes a tarifa nodal no ponto de conexão como penalidade sobre todo o montante que excede o valor do MUST contratado, este custo é bancado pela distribuidora e é não repassável.

## 1.2 DESCRIÇÃO DO PROBLEMA

Conforme exposto na seção anterior, passa a existir a necessidade de se respeitar os valores de MUSTs pela distribuidora, para cada ponto de conexão com a rede básica, contratados junto ao ONS.

Para ultrapassagens de até 5% sobre o valor contratado, a distribuidora repassa aos consumidores o maior valor entre o contrato e o valor medido.

Para que a área de operação da distribuição da empresa obtenha, de antemão, informações que possibilitem a tomada da decisão quanto a modificações na configuração do sistema de distribuição, existe a necessidade de se desenvolver uma metodologia computacional e uma base de dados confiável que permitam a projeção destas demandas, por ponto de conexão com a rede básica, para o horizonte de uma semana a frente, de modo a se tentar evitar, penalizações por ultrapassagem dos 5% da demanda contratada.

Outro problema verificado é de que as concessionárias a cada ano acumulam mais e mais informações em seus bancos de dados. No entanto, apesar do enorme valor e do volume destes dados, a maioria das empresas é incapaz de aproveitar totalmente o que está armazenado em seus arquivos.

Estas importantes informações estão na verdade implícitas, faltando para o uso destas, ferramentas que aperfeiçoem os procedimentos de gerenciamento do objeto de interesse, tornando-os mais seguros e precisos sendo, portanto a base para a resolução do problema.

Os CUSTs que são assinados até 31 de dezembro do ano anterior ao primeiro ano de contratação, consideram valores de MUSTs por ponto de conexão, definidos de acordo com as Resoluções ANEEL nº 281/99, 358/02 e 655/02 e têm os seguintes critérios:

- Os agentes de distribuição devem contratar os MUSTs válidos para o período de 1º de janeiro a 31 de dezembro do primeiro, segundo e terceiro ano seguinte ao ano do estabelecimento do contrato;
- Pode-se efetuar novos aditivos contratuais para o ano em curso desde que decorrente de entradas de novos consumidores de porte ou aumentos

expressivos de carga de consumidores existentes, para isto o ONS exige justificativas e novos pareceres de acesso.

- Não podem ser solicitados novos aditivos contratuais para atendimento de erros de previsão da taxa de crescimento de demanda.

## 1.3 OBJETIVOS

### 1.3.1 Objetivo Geral

Desenvolver uma metodologia de previsão dos MUSTs, a qual seja capaz de prever um período de uma semana a frente para cada ponto de conexão com a rede básica, com exatidão e tempo hábil suficiente para permitir que possam ser tomadas medidas operativas de forma a se evitar ultrapassagens dos limites dos valores contratados. Assim, reduzir-se-á os impactos econômicos para a distribuidora pelo não pagamento de penalidades por ultrapassagens dos valores contratados.

### 1.3.2 Objetivos Específicos

Os objetivos específicos desta dissertação incluem:

- Levantar a base de dados disponível de uma concessionária;
- Realizar testes de consistência da base de dados verificando-se: dados faltantes, mudanças de configurações, sazonalidades, erros de leitura, etc.;
- Aplicar métodos para tratamento, limpeza e redução do volume de dados;
- Desenvolver metodologias de previsão com base em *support vector regression* e redes neurais dos valores dos MUSTs, que atendam ao objetivo geral;
- Testar as metodologias com dados reais de uma concessionária;
- Escolher a técnica mais robusta de previsão de forma que a estimativa da semana a frente obtenha valores com erros inferiores às margens admissíveis;
- Gerar subsídios para montar uma base de dados consistente para projeções e gestão das futuras contratações dos MUSTs das distribuidoras.

Dessa forma espera-se ter procedimentos mais vantajosos economicamente para o consumidor, ao mesmo tempo em que proporcionem uma substancial redução no nível de

risco de contratação e mesmo facilidade para o monitoramento dos valores de MUSTs para a distribuidora.

## 1.4 ESTADO DA ARTE

Existem algumas dificuldades para obtenção de modelos precisos de previsão de carga, pois estes exigem, entre tantos fatores:

- Especialistas com conhecimentos em sistemas elétricos e métodos matemáticos de previsão;
- Banco de dados consistentes e atualizados;
- Modelos que lidem com a natureza não estacionária da carga, presença de sazonalidade, periodicidade e tendências influenciadas por ruídos;
- Necessidade de estudo, análise e projeto de métodos numéricos eficientes de identificação do comportamento dinâmico e previsão de carga.

O objetivo desta revisão bibliográfica é o de dar suporte ao desenvolvimento de uma metodologia para a determinação dos MUSTs, de modo a minimizar os riscos de ultrapassagem dos requisitos de potência nos pontos de conexão e evitar penalidades.

Diversos modelos de previsão de carga estão baseados em técnicas de análise de séries através de: regressão múltipla, redes neurais artificiais – RNAs, sistemas de inferência por lógica nebulosa, diversos modelos híbridos, sistemas especialistas e atualmente os sistemas de máquinas de vetor suporte, do inglês *support vector machines* (SVMs).

Basicamente todas estas propostas requerem intervenção constante dos especialistas, ora selecionando a estrutura, ora adequando os dados de entrada para o modelo ou frequentemente ambos.

Sobre previsão de MUSTs por ponto de conexão, foi localizado um único trabalho publicado (SILVA, COSTA e MATTAR, 2005), no entanto não apresenta o problema da mesma forma que aqui se pretende.

Desta forma a escolha de uma metodologia robusta e confiável para previsão de fluxo passante pelos enrolamentos dos transformadores, nos pontos de conexão com a rede básica, ou seja, a carga dos níveis inferiores do sistema de potência, é de vital importância, dada a incerteza de que uma dada ferramenta seja efetiva, independente da base de dados.

O número elevado de pontos de conexão com a rede básica inviabiliza o tratamento individualizado de cada um destes pontos de conexão para fins de previsão, surge então a necessidade de desenvolvimento de uma metodologia robusta e autônoma que permita o tratamento generalizado para a previsão destes fluxos passantes pelos pontos de conexão da distribuidora com a rede básica.

A metodologia deve ser automatizada para evitar a necessidade de manipulação constante da base de dados pelos especialistas, reduzindo-se a interferência destes, como também se reduzindo os tempos para a obtenção dos resultados de interesse.

É importante frisar que a base de dados deve ser consistente, pois do contrário as previsões estarão baseadas em informações não confiáveis ou mesmo faltantes, o que não permite a segurança e a margem de erro pretendida das projeções de demanda.

Especificamente, os estudos recentes mencionados na literatura sobre previsão de carga têm focado a utilização das seguintes metodologias:

- modelos lineares baseados em modelos do tipo autoregressivo (AR), de média móvel (MA), autoregressivo de média móvel (ARMA) (NOGALES et al., 2002) e autoregressivo integrado de média móvel (ARIMA) (CALMARZA e FUENTE, 2000);
- modelos Box-Jenkins (MOGHRAM e RAHMAN, 1989; BARAKAT e AL-QASEM, 1998; AMJADY, 2001; WEON e MISIOREK, 2005);
- modelos GARCH (GARCIA et al., 2003);
- modelos baseados em abordagens de estatística Bayesiana (DOUGLAS et al., 1998);
- modelos baseados em estimativas de filtro de Kalman e espaço de estados (MOGHRAM e RAHMAN, 1989);

- modelos baseados em métodos de alisamento exponencial (alisamento exponencial simples, linear de Brown, sazonal de Holt-Winters, biparamétrico de Holt) (LOTUFO e MINUSSI, 1999);
- análise de grupos (*clustering*) (GERBEC et al., 2002; Sfetsos, 2003);
- modelos baseados em processo hierárquico analítico (REN-JUN e XIAN-ZHONG, 2002);
- sistemas especialistas (KANDIL et al., 2002);
- redes neurais artificiais (SWARUP e SATISH, 2002; YALCINOZ e EMINOGLU, 2005);
- redes neurais com função de base radial (GONTAR e HATZIARGYRIOU, 2001; MAMUN e NAGASAKA, 2004);
- redes neuro-evolutivas (PAI e HONG, 2005);
- redes de *wavelets* (OONSIVILAI e EL-HAWARY, 1999; HUANG e YANG, 2001);
- sistemas nebulosos (LIANG e CHENG, 2000; MASTOROCOSTAS et al., 2001; SHAKOURI et al., 2033; RASHIDI et al., 2003);
- sistemas nebuloso-evolutivos (CHEN et al., 2001);
- sistemas neuro-nebulosos (SRINIVASAN et al., 1998; KODOGIANNIS e ANAGNOSTAKIS, 2002);
- máquinas de vetor suporte – SVM (SMOLA e SCHÖLKOPF, 1998; CHEN, CHANG e LIN, 2004).
- Dentre esses trabalhos salientam-se os baseados nas máquinas de vetor suporte - SVM, onde Chen, Chang e Lin (2004) adotaram a técnica SVM na competição EUNITE em 2001 quando se sagraram campeões, apresentando então suas previsões, os menores erros entre as várias técnicas participantes.
- Existem ainda outros trabalhos utilizando-se outras técnicas de previsão com Inteligência Artificial – IA como a seguir são detalhados.

#### **1.4.1 Previsão por Técnicas de Inteligência Artificial**

O problema da previsão de demanda tem sido alvo de pesquisas já há bastante tempo devido à sua importância para as concessionárias. Técnicas clássicas de previsão, tais como Regressão Linear Modelo ARIMA de Box e Jenkins (1994), e diversos modelos não-lineares

têm sido propostos. Redes Neurais (HIPPERT, BUNN e SOUZA, 2001; BILINGS e ZHENG, 1995) vêm de encontro a um modelamento não-linear, não-paramétrico, aplicado a problemas cujo modelo analítico é de difícil, senão impossível obtenção, como é de frequente ocorrência nos casos de previsão de demanda de potência elétrica a curto, médio e longo prazo.

Zhang, Patuwo e Hu (1998) apresentam o estado da arte da previsão de séries temporais usando redes neurais, com análises de técnicas publicadas até o ano de 1997. O trabalho apresenta uma introdução às redes neurais definindo os principais pontos a serem definidos pelos projetistas que usam estas técnicas, tais como considerações sobre a arquitetura da rede (número de camadas, número de neurônios em cada camada, tipo de funções de ativação, interconectividade dos neurônios) e algoritmos de aprendizado e seus parâmetros (taxa de treinamento, uso do momento, critério de parada).

Os autores traçam comentários sobre o banco de dados, seu pré-processamento, normalização e uso para treinamento, teste e validação da rede neural. Uma breve revisão das medidas mais utilizadas para avaliar o desempenho da rede é também realizada, fatos estes que não fogem à regra para os estudos de previsões de carga e que também serão tratados neste trabalho.

Para diversos trabalhos com diferentes aplicações, os autores verificam que redes neurais se comparam ou são muito melhores que os métodos estatísticos convencionais, tais como modelos ARMA, ARIMA proposto por Box e Jenkins (1994) e regressão linear ou de ordem superior. Embora os autores não tenham focalizado em um problema específico, as conclusões foram bastante animadoras, e deixaram clara a adequação do uso de redes neurais em problemas de previsão, sem, no entanto deixar no ar as questões que sempre permeiam qualquer aplicação que envolva redes neurais, tais como, falta de um modelo físico, uma vez que são métodos baseados em dados. Dada a inerente não-linearidade das redes, pode ser que para problemas lineares não haja vantagem sobre uma regressão simples; pois existe o perigo de especialização da rede, devido ao grande número de parâmetros a serem calculados; inexistência de um procedimento que permita definir a arquitetura e algoritmos de treino adequados a um determinado problema.

Hippert, Bunn e Souza (2001) analisam o problema de *overfitting* que ocorre facilmente nas redes neurais e comparam com modelos clássicos que possuem um pequeno

número de parâmetros. Os experimentos com previsão de perfis de carga revelam que as redes neurais, mesmo tendo um grande número de parâmetros (pesos) a serem calculados são modelos adequados cujo desempenho é comparável, senão melhor, que os modelos convencionais.

Apesar destas vantagens, desde as primeiras propostas de modelos neurais para previsão de carga, a utilização prática destas estruturas vem enfrentando alguns empecilhos, a saber: elevado esforço computacional para estimação do modelo, ausência de intervalos de confiança das previsões, baixa interpretabilidade dos resultados, escolha adequada do espaço de entrada e controle de complexidade da estrutura.

O aumento da capacidade de processamento disponível viabilizou o desenvolvimento de algoritmos de treinamentos mais velozes e a utilização de modelos neurais para aplicações práticas em sistemas de potência diante do elevado grau de não-linearidade presente nestes modelos. O desenvolvimento de intervalos de confiança analíticos, de maneira análoga aos obtidos para modelos lineares foi deixado de lado em um primeiro momento, com o foco voltado para metodologias baseadas em técnicas de re-amostragem. Atualmente, a aplicação de técnicas de inferência bayesiana ao problema de treinamento de *Multi-Layer Perceptrons* - MLPs fornece uma estimativa analítica para os intervalos de confiança das previsões. A extração de conhecimento de modelos neurais, possibilitando o fornecimento de interpretações qualitativas sobre as previsões, constitui uma das suas principais vantagens (PENG, HUBELE e KARADY, 1992; PAPALEXOPOULOS, HAO e PENG, 1994).

O controle de complexidade ou regularização de modelos neurais visa adequar o nível de não-linearidade disponibilizado pela estrutura à regularidade apresentada pelos dados, evitando a modelagem desnecessária do ruído e a conseqüente redução da capacidade de generalização do modelo. Um dos procedimentos mais populares de regularização de modelos neurais, e um dos mais utilizados em aplicações de previsão de carga, têm como base técnicas de validação cruzada, sendo conhecido como parada antecipada do treinamento, *early stopping*. Esta técnica é heurística, devendo ser selecionados adequadamente tanto o ponto a partir do qual o erro para o conjunto de validação é degradado, quanto o próprio conjunto de validação, que deve representar de maneira fidedigna a função a ser aproximada.



No caso das SVMs, a otimização de estimativas analíticas para o limite superior do erro estimado por validação cruzada única, método também conhecido como *leave-one-out*, é utilizada tanto para seleção das variáveis de entrada, através da análise dos parâmetros do kernel, quanto para controle de complexidade, com a escolha dos parâmetros  $C$ .

#### 1.4.2 Previsão de Demanda a Curto Prazo

Vasconcelos e Araújo Jr. (1999) estudaram a utilização de redes MLP, TDNN (*Time Delay Neural Network*) e GMDH (*Group Method of Data Handling*) para a previsão três meses a frente da demanda mensal máxima de energia elétrica.

O trabalho de Mori e Yuihara (2001) apresenta um método de agrupamento a ser utilizado em uma etapa de pré-processamento dos dados para a previsão de demanda a curto prazo. Os autores propõem o uso de uma rede MLP para previsão da demanda um dia à frente, cujos dados são previamente classificados usando o algoritmo de agrupamento proposto.

Kim et al. (2000) propõe um sistema híbrido Neural-Nebuloso para a previsão de demanda de carga a curto prazo (um dia a frente) em condições anômalas, tais como ocorrem na previsão de demanda de energia em feriados. Barakat (1998) aborda o problema de previsão de demanda em feriados que é tratado de forma diferenciada, pois os autores propõem também variações no algoritmo de treinamento. Magd e Findlay (2003) propõem a utilização conjunta de 2 redes MLP, uma para a previsão de dias normais e outra para feriados e dias anômalos.

Também Magd e Findlay (2003) se preocupam com a questão da previsão da demanda máxima do dia. Problema este compartilhado por Saini e Soni (2002) que propuseram a utilização de 11 variáveis climáticas e uma variação do gradiente descendente para o treinamento das redes. O problema de previsão de montante de uso é bastante similar a estes, porém em uma escala de tempo maior.

Grande parte dos trabalhos que utilizam redes neurais para a predição de carga utiliza o modelo MLP, no entanto, em se tratando de um problema que trata de processamento de séries temporais, o uso de um janelamento limitado no tempo não parece ser uma solução natural, dado que o tempo é mutável. As redes TDNN vêm de encontro a este problema,

fazendo que uma rede *feed-forward* possa incorporar dados históricos de maior profundidade do sinal de entrada sem aumentar em demasia o número de neurônios da rede.

Uma estrutura que processe sinais temporais de forma natural parece ser a escolha mais apropriada. As redes parcialmente recorrentes, tais como as estruturas de Elman e Jordan (1998), são mais utilizadas, talvez pela facilidade de treinamento e pelos ótimos resultados atingidos. Vermaak e Botha (1998) apresentam um compêndio do uso de redes neurais recorrentes na previsão de demanda a curto prazo utilizando um modelamento dinâmico dos dados históricos e temporais de clima. Moghram e Rahman (1989) fazem um comparativo de técnicas para previsões de carga.

Galarniotis et al. (2003) fazem uma comparação de desempenho do uso de redes MLP, FIR (*Finite Impulse Response*) e Elman para a previsão de curto prazo. Os autores concluem que mais importante que a estrutura da rede, o algoritmo de treino é um ponto crítico para o projeto das redes. Nos experimentos apresentados, as redes FIR e ELMAN produziram melhores estimativas do que a rede MLP, com a desvantagem de requererem um tempo de treinamento duas a quatro vezes maior. Os resultados confirmam que as redes FIR e Elman são mais aptas à estimação de séries temporais do que a rede MLP.

Verifica-se, como consenso comum, que a inclusão de variáveis climáticas no preditor vem a agregar informação útil tanto para os modelos convencionais quanto para os modelos neurais (LAGE et al.,2003; TAYLOR e BUIZZA, 2000).

### **1.4.3 Previsão de Demanda a Longo Prazo**

O problema de previsão de demanda a longo prazo não é o foco deste trabalho, pois as previsões de interesse estão limitadas a um horizonte de uma semana a frente que a princípio é o tempo suficiente para que o corpo técnico da distribuidora possa tomar decisões operacionais para evitar as ultrapassagens dos MUSTs contratados.

### **1.4.4 Considerações Finais sobre Previsão**

As máquinas de vetor suporte - SVM (CHEN, CHANG e LIN, 2001), juntamente com modelos de redes neurais, formam o conjunto de técnicas testado neste trabalho. Além de sua

capacidade de modelamento de sistemas não-lineares e imunidade a ruído presente nas medidas de entrada, o tratamento da base de dados proporciona maior confiabilidade ao especialista para as projeções necessárias na determinação dos novos MUSTs que serão contratados a cada ano.

As condições climáticas interferem de maneira decisiva em qualquer modelo de previsão de demanda (PALACIO, 2001), mas dada a dificuldades no manuseio e obtenção da base de dados, para o estudo em foco, estas informações deixarão de ser aproveitadas podendo fazer parte de um trabalho futuro.

## 1.5 ESTRUTURA DA DISSERTAÇÃO

Esta dissertação está organizada em seis capítulos.

No capítulo 1 é apresentado o problema, os objetivos deste trabalho e descrição dos modelos e estudos realizados na área de previsão de carga.

No segundo capítulo é apresentada a caracterização e análise dos dados, os pontos de medição escolhidos aleatoriamente para análise, os parâmetros da média e desvio padrão, os resultados das autocorrelações obtidos com a base de dados já depurada, análises dos valores e mesmo a interpolação dos valores faltantes. Também é efetuada a remoção dos *outliers*, que são os dados com discrepâncias que ultrapassam a dois e meio desvios padrão do valor médio, tanto para cima como para baixo na base gráfica considerada.

No terceiro capítulo são descritas as metodologias e os conceitos básicos que são necessários ao longo do estudo. Também são apresentados os principais fundamentos dos métodos de previsão conhecidos na literatura, que são utilizados para a previsão de carga; seja ela previsão de carga horária, diária, semanal ou mensal, focando o objetivo maior que é o de previsão de sete dias à frente para o horário compreendido como de ponta do sistema.

Ainda no terceiro capítulo, entra-se em maiores detalhes, aprofundando os estudos e averiguações sobre os dois métodos que podem dar melhor performance ao estudo de previsão pretendido.

No capítulo quatro, são apresentados os testes e resultados com a utilização das duas técnicas selecionadas.

No capítulo cinco têm-se as conclusões parciais, onde da comparação do desempenho, eficiência e cumprimento do objetivo é endossado o uso de uma das técnicas. Também neste capítulo, são efetuadas as considerações finais e sugestões de trabalhos futuros.

## 2 CARACTERIZAÇÃO E ANÁLISE DOS DADOS

### 2.1 INTRODUÇÃO

As informações brutas para a elaboração deste trabalho foram obtidas da Companhia Paranaense de Energia – COPEL DISTRIBUIÇÃO, donde foi possível retirar, de uma base de dados estruturada, as informações de fluxos passantes em MW para os anos de 2004, 2005 e 2006, para todos os meses e dias destes anos integralizados de 15 em 15 minutos para cada ponto de conexão com a rede básica. Foram escolhidas aleatoriamente sete subestações das quarenta existentes que fazem fronteira com a rede básica e onde se contratam os MUSTs.

### 2.2 LEVANTAMENTO DA BASE DE DADOS

Da forma como apresentado no item 2.1 anteriormente, pretende-se nesta etapa do trabalho:

- Estabelecer as curvas de tendência sobre os dados brutos sem tratamento, como base de conhecimento;
- Visualizar e escolher quais anos que, a princípio, podem ser utilizados como base histórica para o estudo por Subestação – SE, calculando-se as médias e desvios padrão destes dados;
- Usar os conhecimentos dos analistas da distribuidora, na busca e correção de possíveis erros devido ao sistema de medição;
- Elaborar estudos de correlação com a base de dados existente para segregar períodos de interesse para possível inferência da influência do histórico nas previsões de fluxo do MUST;
- Ajustar os dados faltantes através de método de interpolação por *spline* cúbica (BOOR, 1978);
- Remover os *outliers*, que para efeito deste trabalho são definidos como os dados de medições que excedem dois e meio desvios padrão, calculando-se a média e os novos desvios padrão desta base de dados;

- Segregar os dias de feriados, sábados e domingos das bases de dados para permitir uma maior correlação nas previsões.

Para elaboração de uma base de dados consolidada foi necessário fazer uma busca da demanda máxima integralizada de 15 em 15 minutos, para cada ponto de conexão que compõe a SE de interesse, para em seguida somar as demandas coincidentes dos medidores de interesse retornando com a demanda máxima no horário de ponta.

Atendendo o objetivo desta dissertação, deve-se abordar primordialmente, mas não somente os assuntos abaixo:

- As informações das redes de sistemas de potência sob análise se tornam não lineares devido as alterações existentes na base de dados, tais como: a entrada de novas SEs, mudanças de configuração de operação, remanejamento de cargas, sazonalidades, desligamentos de circuitos das SEs, etc.;
- O próprio processamento dos dados exige a aplicação de técnicas de mineração dos dados, expurgo dos remanejamentos de carga entre SEs, análise e ajustes na base de dados devido a mudanças de configuração temporárias e definitivas, etc.;
- As redistribuições de fluxos entre SEs ocasionam a perda da base histórica para a previsão, o que exige interferência e tomada de decisão dos analistas;
- O conhecimento dos Procedimentos de Rede do ONS, e as condições de aplicação de penalidades aos distribuidores, para efeito de balizamento dos valores buscados, bem como a sensibilidade quanto à margem de erro admissível nas previsões;
- Impacto decorrente da influência dos dias de feriados, sábados e domingos nas previsões de curto prazo, bem como o impacto decorrente dos casos de feriados prolongados, ajustes do horário de verão, etc.;
- Deve-se, também conhecer e gerenciar os valores dos MUSTs contratados para cada mês do ano para as SEs.
- Uso das ferramentas do MATLAB<sup>®</sup> Versão R 2006 B
  1. Para o tratamento dos dados foram utilizadas as seguintes ferramentas do “*TOOLS DATA FUN*”: “*Corrcoef*” que determina a correlação entre os dados, tal qual a “*toolbox time series*”, a função “*Interp1*” que

elabora a interpolação entre dados faltantes de uma série histórica, bem como a substituição dos *outliers* escolhidos, que respectivamente são ajustados através do método de interpolação cúbica “spline” do MATLAB®;

2. Da Biblioteca para máquinas de vetor suporte – LIBSVM de Chen, Chang e Lin (2001) foi obtido o *software* para uso no MATLAB®.

Também foram utilizadas as informações conforme Hsu, Chang e Lin (2007) que traz um guia prático para o uso da ferramenta LIBSVM.

A previsão de carga em sistemas elétricos já é em si um desafio matemático da área de previsão de séries temporais, especialmente por que sobreposto a este desafio, na maioria das vezes, os dados históricos disponíveis apresentam acentuada sazonalidade e elevado nível de incerteza devido aos procedimentos de coleta dos dados, às inconsistências e falta de atualização na base de dados.

Sete SEs foram escolhidas aleatoriamente, dado ao tamanho da amostra em relação ao universo total julga-se serem estas significativamente representativas do universo das 40 SEs que compõe o total de conexões com a rede básica da COPEL Distribuição. Os dados usados nos testes são relativos aos anos de 2004, 2005 e 2006, todas as medidas estão em MW, para os seguintes pontos de medição:

- 1 Campo do Assobio-Sec-138 ATF–A e B (SE CAS);
- 2 Cascavel-Sec-138 ATF–A, B e C (SE CEL);
- 3 Londrina-Sec-138 ATF–A e B (SE LNA);
- 4 Foz do Chopin-Sec-138 ATF A (SE FOZ);
- 5 Ibiporã-Sec-138 ATF–A e B (SE IBI);
- 6 Pato Branco-Sec-138 ATF–1 e 2 (SE PTO);
- 7 Umbará-Sec-69 TF–A e B (SE UMB).

## 2.3 ANÁLISE GERAL DA BASE DE DADOS

Para análise geral da base de dados coletada na COPEL Distribuição há de fazer as seguintes observações:

- Utilizou-se como simplificação, o termo MUST para representar o fluxo passante entre o enrolamento primário e secundário ou mesmo o terciário dos transformadores da SE propriamente dita;
- Foram analisadas as situações onde, por qualquer razão, não foram efetuadas as leituras ou mesmo quando estas foram lidas e se apresentaram com valores nulos evitando-se distorções no gráfico das tendências;
- Utilizaram-se valores linearmente crescentes na abscissa, pois a utilização de outras unidades, tais como, dias da semana, semana do ano, etc., trouxe resultados incorretos na curva da tendência;
- Os valores de demandas fora da faixa de dois e meio desvios padrão para cima e para baixo da média (*outliers*) não foram corrigidos para melhor apreciação da robustez das ferramentas a serem utilizadas.

Feriados nacionais com datas fixas, tais como: 1º de janeiro, 21 de abril, 1º de maio, 7 de setembro, 12 de outubro, 2 de novembro, 15 de novembro e 25 de dezembro e os demais feriados com datas variáveis, tais como carnaval, páscoa/semana santa, Corpus Christi, dias de eleição, finados, etc., afetam as demandas dos dias da semana onde estes feriados ocorrem. Sendo esta a razão do estabelecimento de quatro alternativas diferentes para previsão, apresentadas neste estudo, pois dependendo do dia da semana em que estes feriados ocorrem, a semana fica atípica dificultando as estimativas de previsões de demanda máxima.

Normalmente os erros encontrados nas bases de dados são decorrentes de:

- Remanejamento de carga entre SEs, o que pode ser visualizado nos gráficos, mas nem sempre as informações são documentadas, dificultando as análises;
- Erro de relação dos equipamentos de medição, por montagem equivocada de terminais ou mesmo cálculos para definir os parâmetros de ajuste;
- Mudança de configuração da área de atendimento de SE;
- Falta do envio dos dados para os servidores da empresa, por falha no sistema de comunicação, devido às mais diversas causas, tais como danos a equipamentos ocasionadas por descargas atmosféricas, falha no *link* de comunicação, etc.



Muitas vezes os erros acima mencionados podem ser corrigidos ou eliminados, desde que sejam facilmente identificáveis na visualização dos gráficos obtidos através da base de dados, ou estejam devidamente documentados. Por outro lado, quando ocorrem remanejamentos ou mesmo montagem errada dos equipamentos de medição que não afetam seriamente os valores típicos dos fluxos passantes nestas SEs, os erros ficam inseridos na base de dados sendo impossível sua detecção ou mesmo expurgo, ocorrendo casos de erros que se perpetuam por muito tempo sendo então as leituras incorporadas na base de dados e consideradas como verdadeiras.

Os gráficos têm como abscissa a contagem dos dias do ano que se inicia em 1º de janeiro de 2004 e termina em 31 de dezembro de 2006, o qual é o dia 1096, na ordenada os valores em MW dos fluxos passantes medidos entre os enrolamentos de alta tensão dos transformadores interligadores da Rede Básica (tensão  $\geq 230$  kV) e a distribuição propriamente dita.

A seguir apresentam-se os gráficos dos dados das SEs em estudo:

a) SE Campo do Assobio

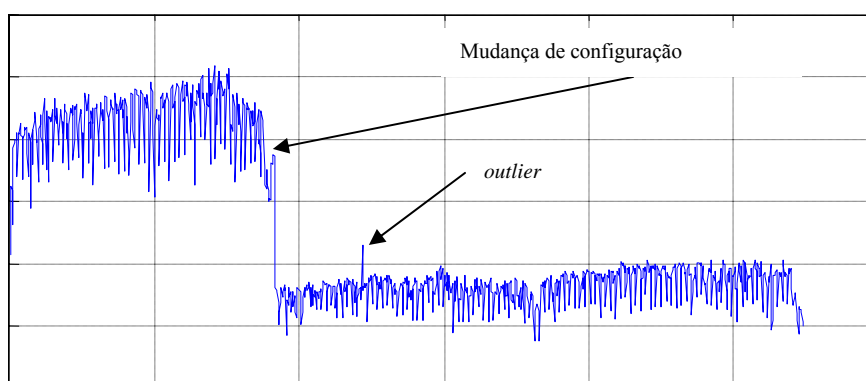


Figura 2.1 – Dados originais da SE CAS – Período 2004/2006.

Como observado, a Figura 2.1 mostra uma descontinuidade na tendência histórica entre os anos de 2004 e 2005/2006, e algumas não conformidades (mudança de configuração e *outlier*).

## b) SE Cascavel

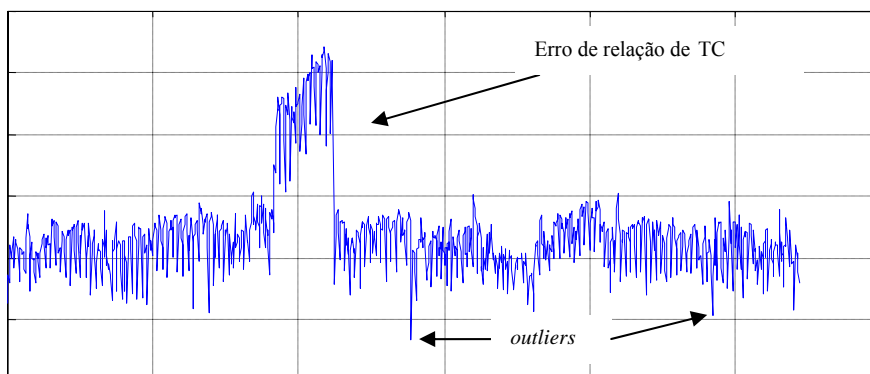


Figura 2.2 – Dados originais da SE CEL – Período 2004/2006.

Visualiza-se na Figura 2.2 que existe uma descontinuidade na tendência histórica no início do ano de 2005, que poderia ser atribuída as mais diversas razões, tais como remanejamento de circuitos, alteração de área de atendimento, erro de medição, etc. No entanto, junto aos técnicos da distribuidora foi obtida a informação de que neste caso ocorreu um erro de leitura, devido a problemas advindos de um erro no ajuste da relação dos TCs.

## c) SE Londrina

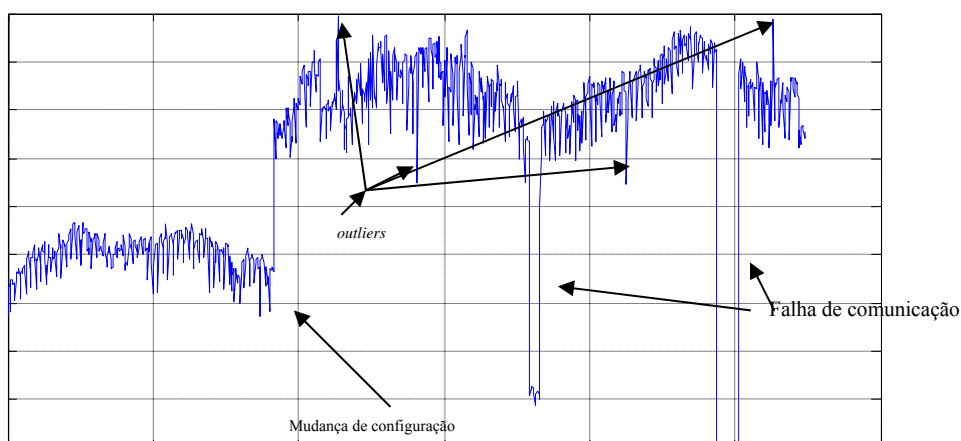


Figura 2.3 – Dados originais da SE LNA – Período 2004/2006.

A Figura 2.3 mostra uma descontinuidade na tendência histórica entre os anos de 2004 e 2005/2006 devido à mudança de configuração da área de atuação da SE, além de se verificar

a falta de dados para alguns períodos e mesmo dados, a princípio, incorretos para o início do ano de 2006, dado a problemas de comunicação.

d) SE Ibiporã

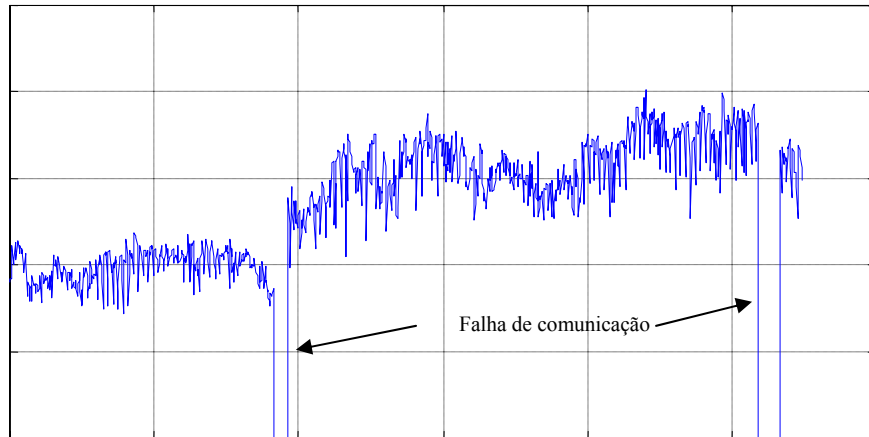


Figura 2.4 – Dados originais da SE IBI – Período 2004/2006.

Como observado a Figura 2.4 mostra falta de informações no início de 2005 e final de 2006, além de alteração na tendência histórica entre os anos de 2004 e 2005/2006.

e) SE Foz do Chopin

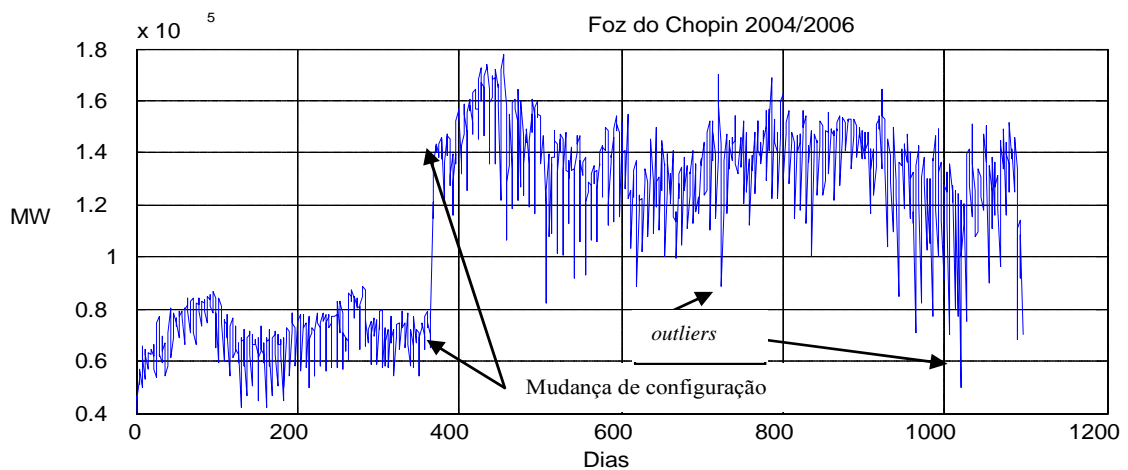


Figura 2.5 – Dados originais da SE FOZ – Período 2004/2006.

Na Figura 2.5 tem-se uma descontinuidade na tendência histórica entre os anos de 2004 e 2005/2006, bem como remanejamentos e alterações frequentes de configuração especialmente em 2006.

## f) SE Pato Branco

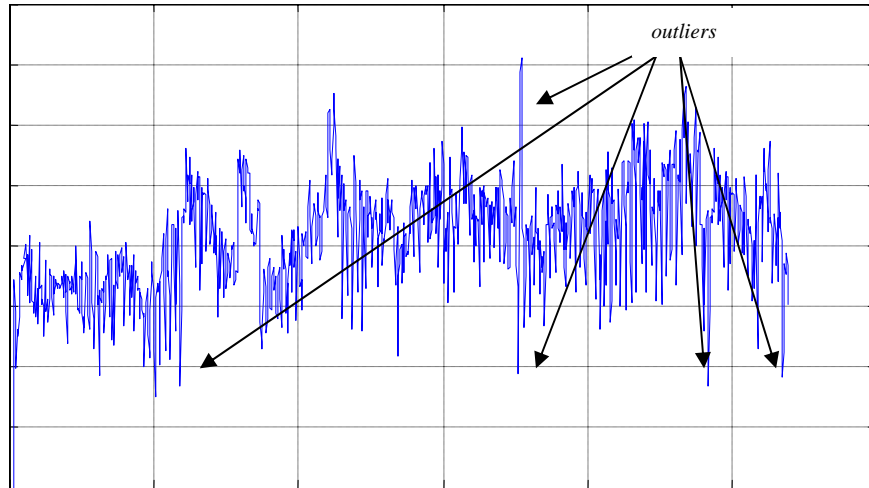


Figura 2.6 – Dados originais da SE PTO – Período 2004/2006.

Como visualizado na Figura 2.6 ocorre falta de dados em meados de 2004 e alguns *outliers* na observação da tendência histórica.

## g) SE Umbará

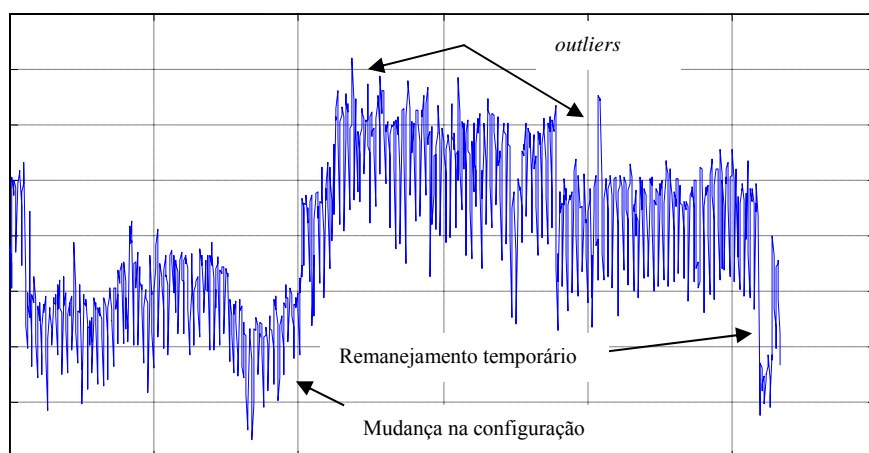


Figura 2.7 – Dados originais da SE UMB – Período 2004/2006.

Observa-se na Figura 2.7 uma descontinuidade na tendência histórica entre os anos de 2004 e 2005/2006 o mesmo ocorrendo em um pequeno período no fim de 2006 o primeiro motivado pela mudança de configuração e a segunda pelo remanejamento temporário de configuração.

Como um todo, verificou-se que a base de dados carece de tratamento para que os dados se tornem consistentes e permitam a observação de alguma tendência nas SEs.

## 2.4 ANÁLISE DA MÉDIA, TENDÊNCIAS E DESVIOS PADRÃO

A análise da média, tendência e desvios padrão foi executada para as sete SEs escolhidas de um total de 40 existentes no sistema de contratação da Distribuidora, sendo que serão apresentados nesta seção a metodologia e resultados para a SE CAS.

As informações das demais SEs constam no Apêndice A.

### a) SE Campo do Assobio período 2004/2006

Apresenta-se na Figura 2.8 a curva de tendência com os dados brutos.

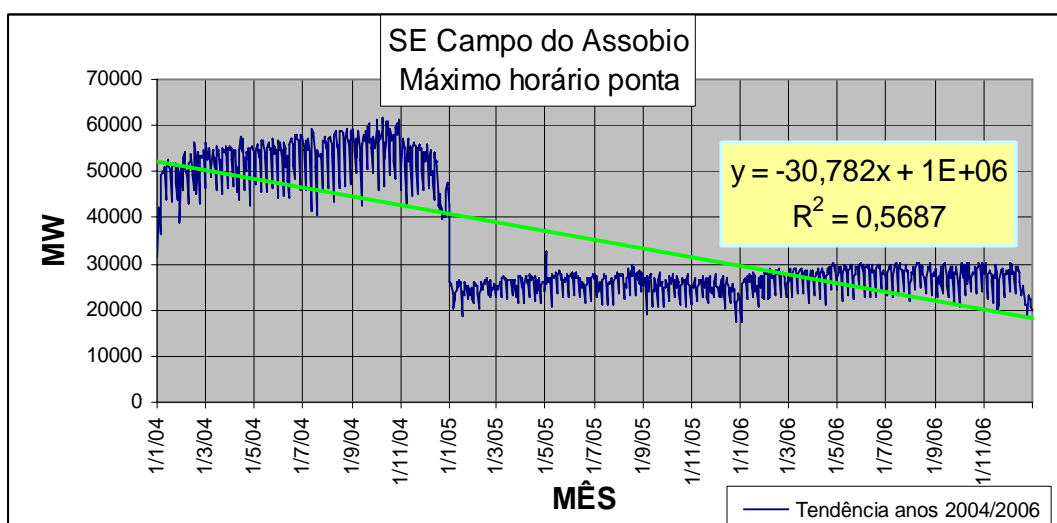


Figura 2.8 – Dados originais da SE CAS – Período 2004/2006 – Discretização plurianual.

A Figura 2.8 não permite generalização dado a mudança de patamar do fluxo passante por esta SE.

Observa-se na Figura 2.9 que existem dois patamares de cargas bem distintos para estes períodos, um para os anos de 2005 e 2006 e outro superior para o ano de 2004, o que permite inferir, em um primeiro momento, que houve alteração da configuração de atendimento às cargas da região, ocorrendo uma redução do fluxo passante do ano de 2004 para os anos de 2005 e 2006. Assim a princípio os dados de 2004 não devem ser aproveitados, pois trazem uma alteração significativa no estabelecimento de uma curva de tendência para a SE.

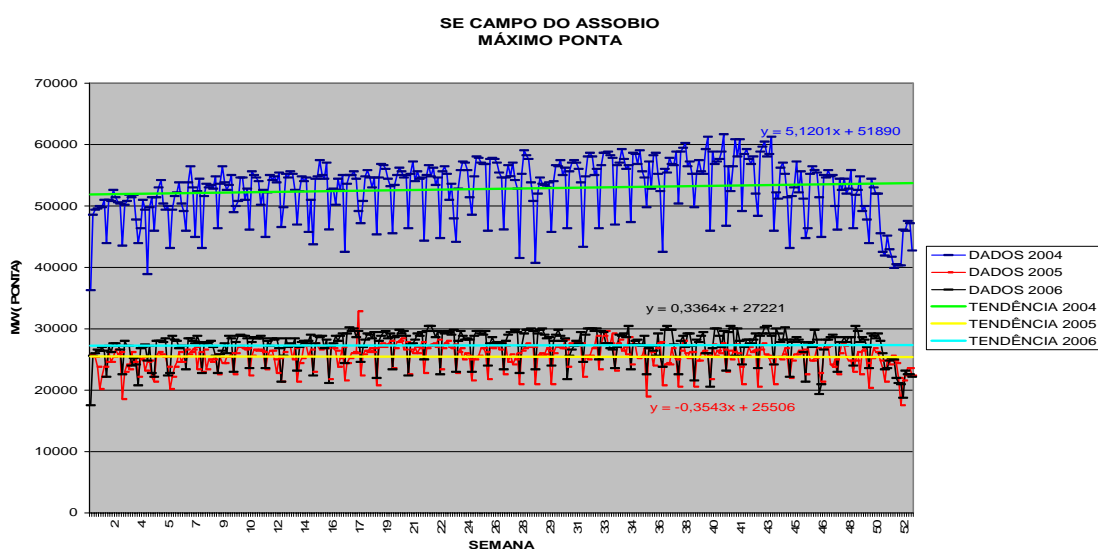


Figura 2.9 – Dados originais da SE CAS – Período 2004/2006 – Discretização anual.

Ainda observando a Figura 2.9, quando se compara semana a semana os anos de 2005 e 2006, observa-se que removida a influência dos feriados, sábados e domingos, tem-se uma correlação muito forte entre as curvas.

É importante frisar que, a ocorrência de feriados em determinados dias da semana afeta significativamente a forma de curva entre semanas correlatas, por exemplo, a última semana de um ano somada a semana de ocorrência do dia 1º do ano fica sempre prejudicada quando comparada com a do ano seguinte dado que, os feriados de Natal e Ano Novo normalmente não ocorrem nos mesmos dias da semana para ambos os anos, trazendo diferenças significativas nos valores diários das demandas máximas verificadas quando comparados os dias úteis destas semanas para cada ano.

Na Figura 2.10 são apresentados os dados de 2005 e 2006. Observa-se que a primeira semana dos anos de 2005 e 2006 se apresenta não conformes, dentro da sazonalidade

semanal, dado que dia 1º no ano de 2005 ocorreu num sábado e para 2006 num domingo, o mesmo ocorrendo nas últimas semanas do ano de 2006, devido ao feriado do Natal. Da Figura 2.10 pode-se obter ainda a taxa de crescimento anual da carga de aproximadamente 4,88% a.a.

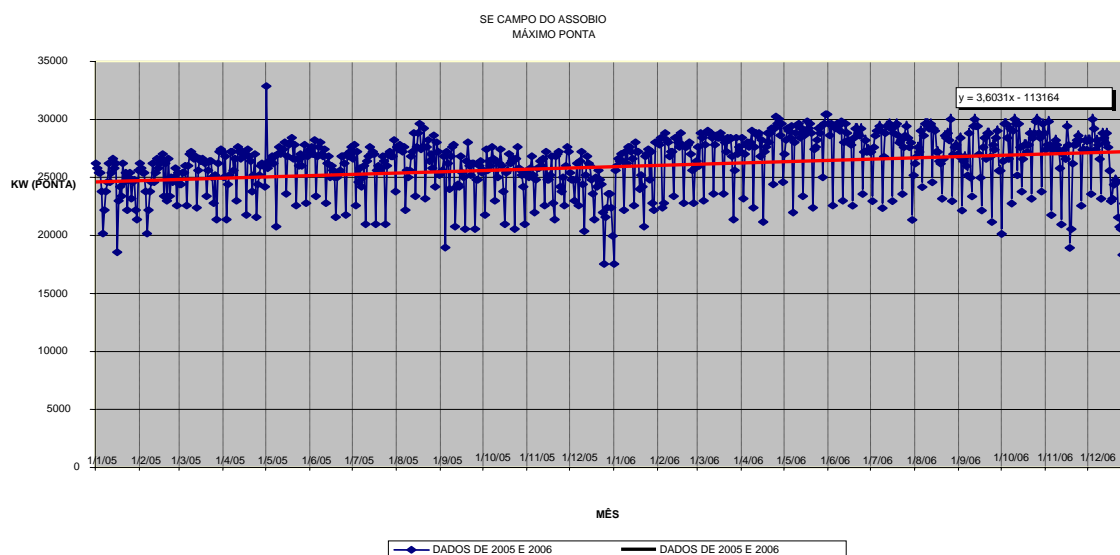


Figura 2.10 – Dados originais da SE CAS – Período 2005/2006 – Curva de tendência.

## 2.5 ANÁLISE E INTERPOLAÇÃO DOS DADOS FALTANTES

Para as SEs em que não ocorre a falta de informações na base de dados, a saber, SE CAS, SE FOZ, SE PTO e SE UMB, não foram efetuadas correções, mas quando observou-se, pelos gráficos que existiam dados faltantes, estes foram repostos utilizando-se o método de interpolação por *spline* cúbica (BOOR, 1978).

Nos gráficos a seguir “mu” significa média aritmética e “sigma” o desvio padrão.

### a) SE Cascavel

A SE apresentou necessidade de ajustes nos valores da base de dados, dada a descoberta de valores com erros de leitura, portanto foram corrigidos os dados de fluxos do início de 2005. Para efeito de comparação entre as situações pré e pós inserção dos dados faltantes tem-se que a média aritmética antes da interpolação no período 2004 a 2005 era de 348.894 MW para um desvio padrão de 82.675 MW. Já após a interpolação passou a 328.780 MW e 39.723 MW respectivamente.

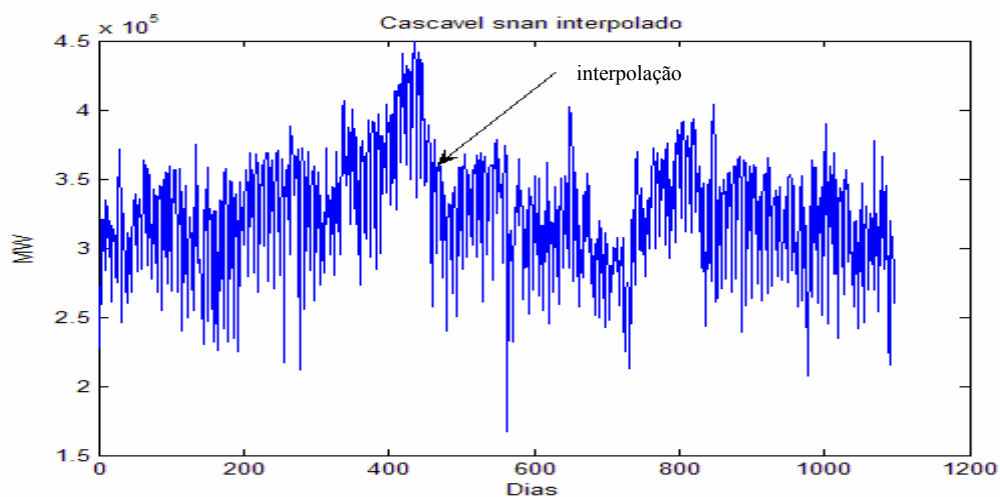


Figura 2.11 – Dados da SE CEL com Interpolação – Período 2004/2006.

b) SE Londrina

Como se observa na Figura 2.12 faltavam dados a serem repostos para a SE Londrina no período de 18 e 31 de dezembro de 2005, bem como no mês de setembro de 2006.

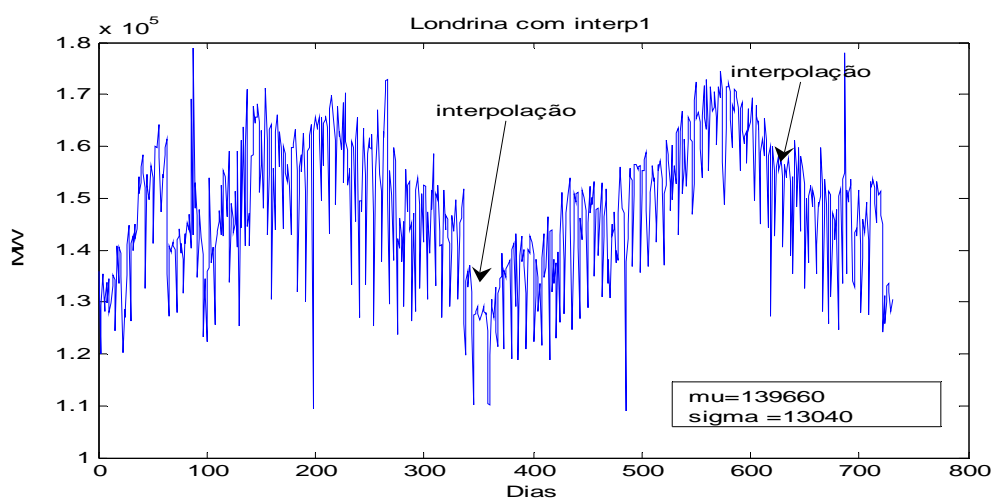


Figura 2.12 – Dados da SE LON com Interpolação – Período 2005/2006.

Como observado, a interpolação repôs satisfatoriamente a base de dados para estudos.

c) SE Ibiporã

Como se observa na Figura 2.13 da SE Ibiporã, existia necessidade de reposição dos dados dos dias 01 até dia 20 de janeiro de 2005 e todo período de novembro de 2006.



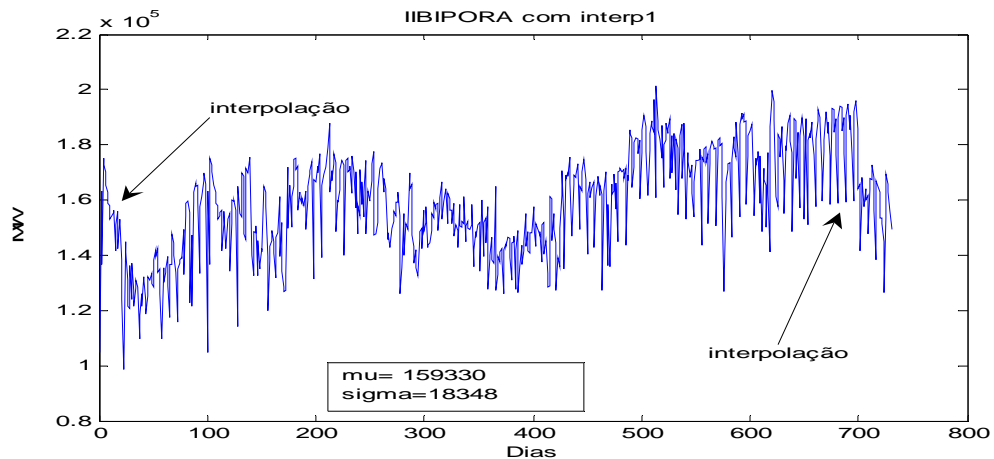


Figura 2.13 – Dados da SE IBI com Interpolação – Período 2005/2006.

## 2.6 REMOÇÃO DOS OUTLIERS

Numa série de dados, *outliers* são observações que são numericamente distantes do resto dos dados, que para efeito deste trabalho são definidos como os dados de medições que excedem dois e meio desvios padrão, as quais degradam a análise estatística de tais séries. Para exemplificar a preparação dos dados é apresentado na Figura 2.14 o processo de eliminação de *outliers*, o qual contribui para o melhor desempenho das técnicas de previsão.

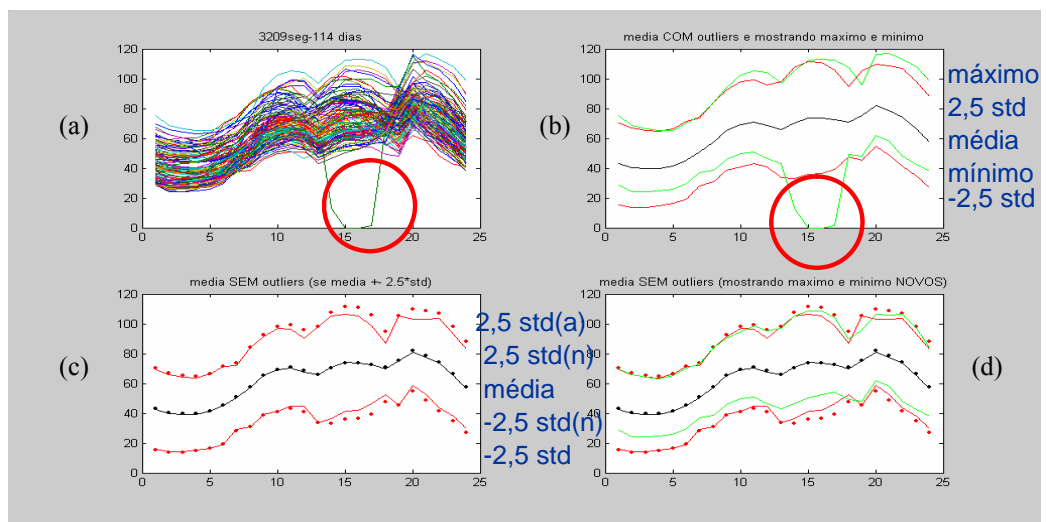


Figura 2.14 – Dados da SE CAS de Segunda-feira – Com e Sem *Outliers*.

A Figura 2.14 apresenta os resultados da aplicação dos métodos de depuração da base de dados (eliminação dos *outliers*, ajuste dos remanejamentos e mudanças de configuração e reposição dos valores faltantes por interpolação) na SE Cascavel para o caso simulado com os

dados de uma 2ª Feira. Inicialmente, quadro (a) apresenta-se todos os dados em conjunto, destacando-se um *outlier* por falta de medição. No quadro (b) visualiza-se a média, os limites de dois e meio desvios padrão para cima e para baixo (linhas em vermelho) e os valores máximos e mínimos em cada horário de medição (linhas em verde). O quadro (c) apresenta o resultado da média e dos desvios padrão antes (linhas pontilhadas) e após (linhas cheias) a eliminação dos *outliers*. No quadro (d) acrescentou-se novamente os novos valores de máximos e mínimos em cada horário de medição após a eliminação dos *outliers*.

Como regra geral para eliminação dos *outliers* foi feita a reposição dos dados por interpolação com *spline* cúbica. Nas Figuras 2.15 a 2.21 apresenta-se o comportamento dos dados das SEs após a eliminação dos *outliers*.

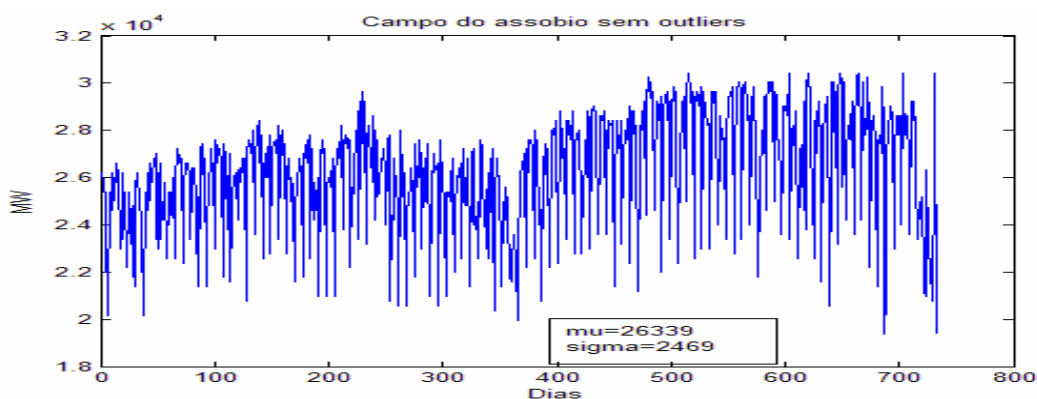


Figura 2.15 – Dados da SE CAS sem *Outliers* – Período 2005/2006.

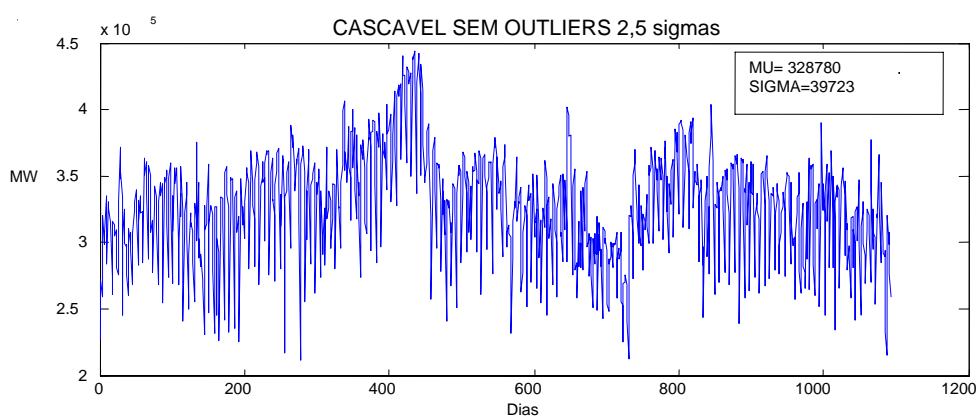


Figura 2.16 – Dados da SE CEL sem *Outliers* – Período 2005/2006.

Figura 2.17 – Dados da SE LNA sem *Outliers* – Período 2005/2006.

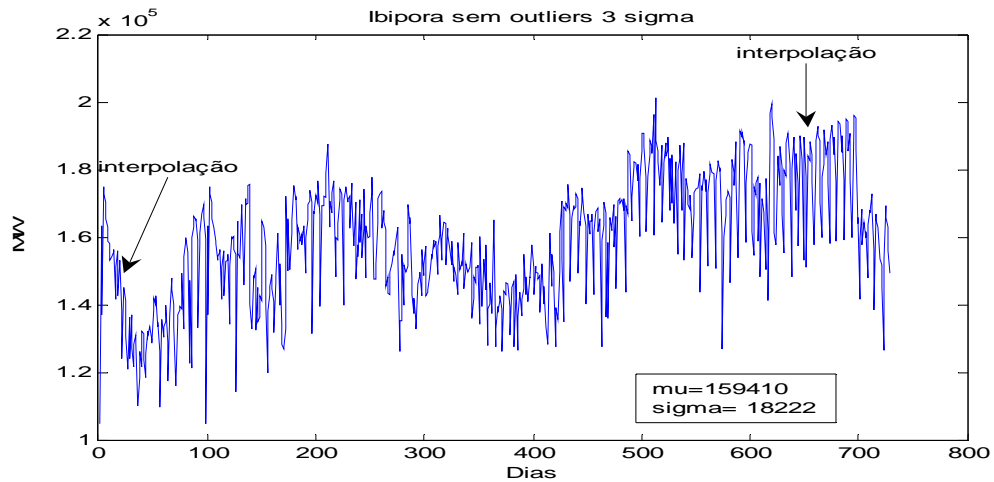


Figura 2.18 – Dados da SE IBI sem *Outliers* – Período 2005/2006.

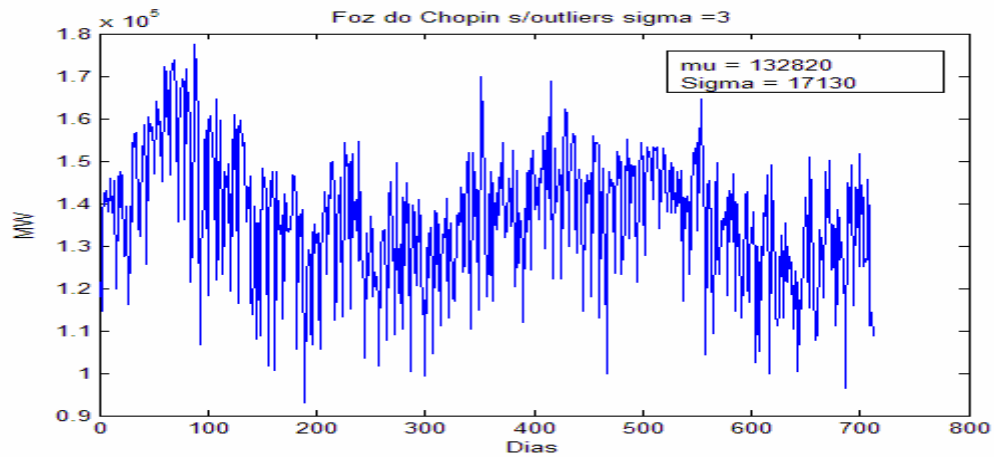


Figura 2.19 – Dados da SE FOZ sem *Outliers* – Período 2005/2006.

Para a SE PTO a média com *outliers* era de 44.157 MW e o desvio padrão era de 7.991,6 MW, passando para os valores que constam na Figura 2.20.

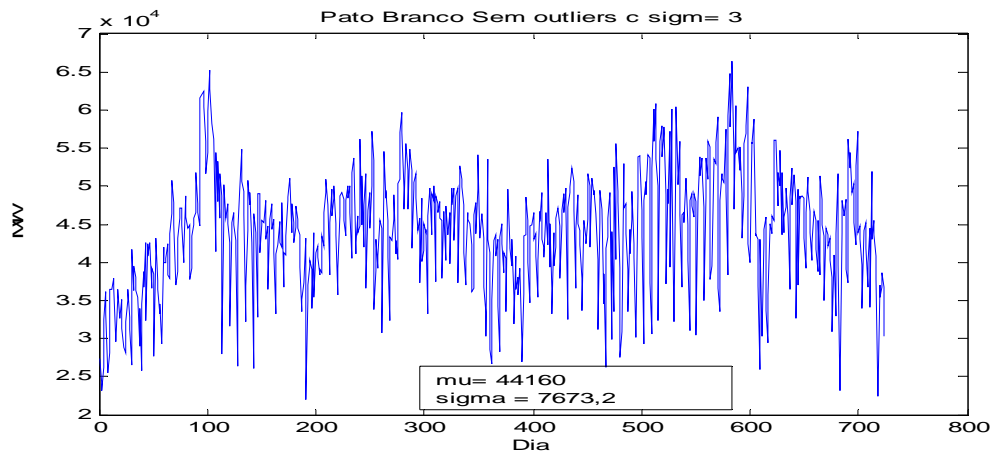


Figura 2.20 – Dados da SE PTO sem *Outliers* – Período 2005/2006.

Para a SE UMB, a média com *outliers* era de 171.610 MW e o desvio padrão era de 26.298 MW passando para os valores que constam na Figura 2.21.

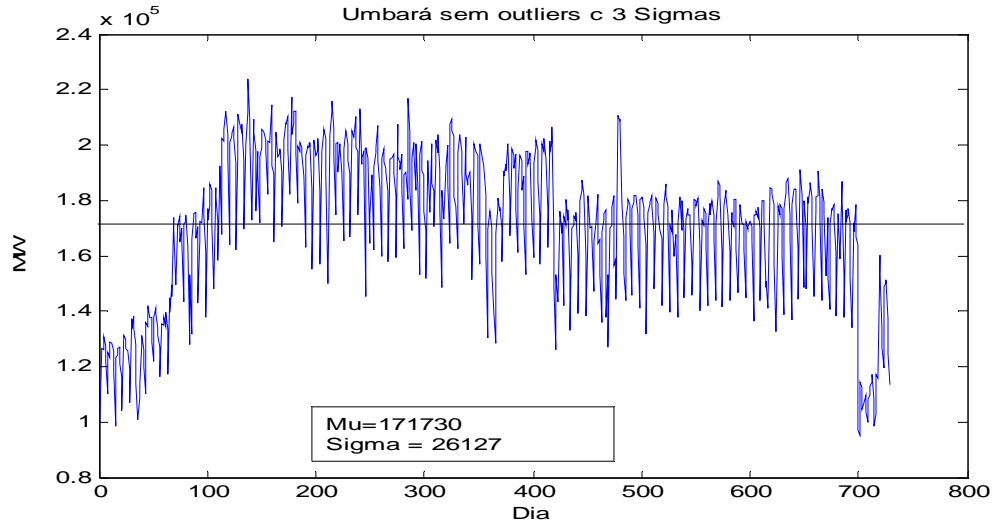


Figura 2.21 – Dados da SE UMB sem *Outliers* – Período 2005/2006.

## 2.7 AUTOCORRELAÇÃO

A seguir são apresentados os estudos de correlação por SE, depois do tratamento da base de dados por interpolação dos valores faltantes, ajuste dos *outliers* e demais correções das não conformidades encontradas na base de dados como: erro de relação de TC e a eliminação dos dados de mudanças de configuração (2004).

O coeficiente de autocorrelação calcula-se segundo a seguinte fórmula:

$$r = \frac{\sum_{k=1}^n (x_i - \bar{x})(x_{i+k} - \bar{x})}{\sigma^2} \quad (2.1)$$

Onde:  $r$  é o valor da autocorrelação para o dia  $i$ ;

$\bar{x}$  é a média da variável  $X$ ;

$\sigma^2$  é a variância da variável  $X$ ;

$x_i$  são os valores da variável  $X$ ;

$k$  é o deslocamento no tempo;

$n$  é nº máximo de dias sob estudo (no caso  $n = 730$  atrasos);

$i$  é o dia de referência.

A interpretação do coeficiente de autocorrelação é apresentada na Tabela 2.1.

Tabela 2.1 – Interpretação de Correlação.

<b>Coeficiente</b>	<b>Correlação</b>
$r < 0$	Relação Inversa
$r = 0$	Relação Nula
$0 < r \leq 0,3$	Relação Fraca
$0,3 < r \leq 0,7$	Relação Média
$0,7 < r \leq 0,9$	Relação Forte
$0,9 < r \leq 0,99$	Relação Fortíssima
$r = 1$	Relação Perfeita

O tratamento estatístico da base de dados ajustada resultou na Tabela 2.2.

Tabela 2.2 – Análise Estatística dos dados ajustados das SEs.

<b>SE</b>	<b>Média</b>	<b>Mediana</b>	<b>Desvio Padrão</b>	<b>Mínimo</b>	<b>Máximo</b>
CAS	26.394,00	26.813,0	2.352,86	19.354	30.442
CEL	328.780,00	351.360,00	39.723,10	235.268	447.436
LNA	148.162,48	149.184,0	12.665,44	118.742	179.021
IBI	159.276,63	159.869,0	18.562,91	61.992	201.398
FOZ	136.943,30	137.491,0	14.551,11	99.389	173.779
PTO	44.100,06	44.604,5	7.680,68	21.974	66.427
UMB	171.628,66	175.680,0	26.248,81	95.386	223.718

Na Figura 2.22 observa-se a autocorrelação da SE CAS, na qual verifica-se que os 100 últimos dias de histórico permitem uma taxa de correlação superior a 60%, o que pode ser um indicativo que para esta SE esta massa de dados seja suficiente para se obter previsões com margem de erro adequada. A correlação é maior quando considerado o dia correspondente da semana anterior, sendo este mais significativo na previsão, também se obtém nos gráficos de correlação casos de correlação inversa,

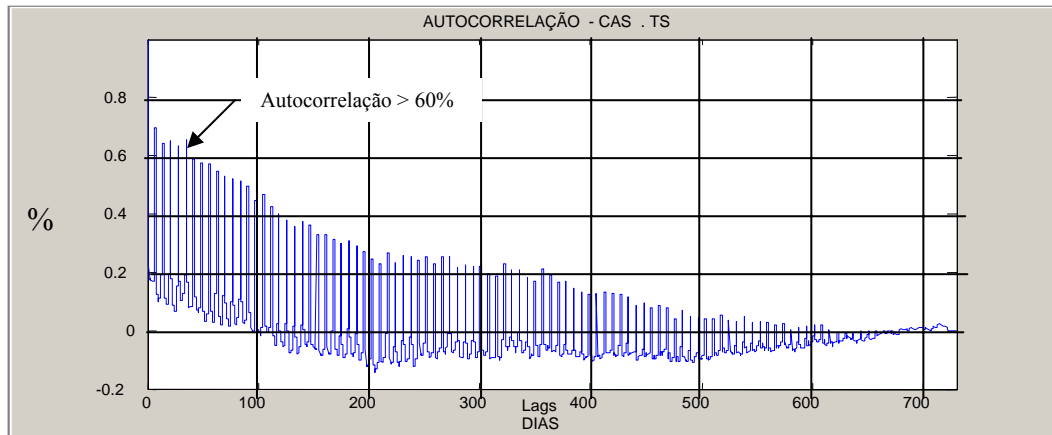


Figura 2.22 – Autocorrelação dos dados da SE CAS – Período 2005/2006.

Para a SE CEL verifica-se que a correlação é maior do que 60% somente para os 21 últimos dias de histórico o que pode ser um indicativo que para esta SE apenas as últimas semanas estão correlacionados efetivamente para os valores da semana seguinte (Figura 2.23).

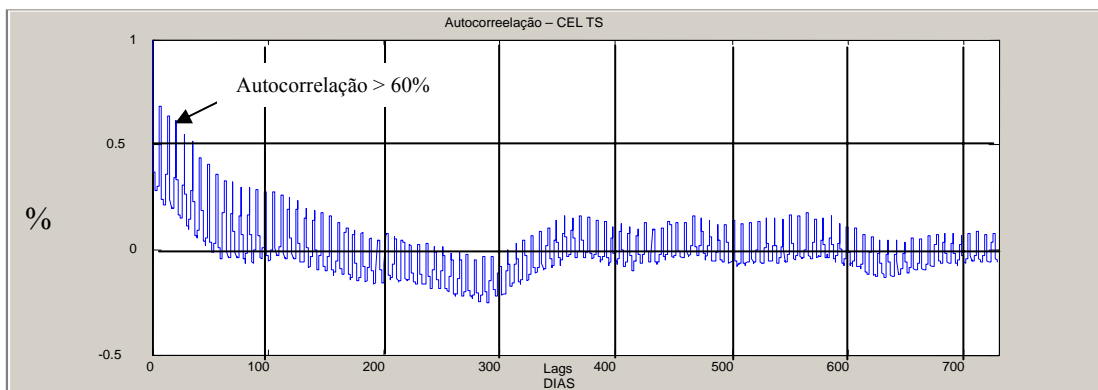


Figura 2.23 – Autocorrelação dos dados da SE CEL – Período 2005/2006.

Na Figura 2.24 observa-se a autocorrelação da SE LNA, na qual verifica-se que a correlação é maior do que 50%, somente para os 28 últimos dias de histórico, ao mesmo tempo em que, o mesmo período do ano anterior tem influência significativa na correlação para as projeções e que a semana anterior resulta em correlações muito altas maior do que 70%.

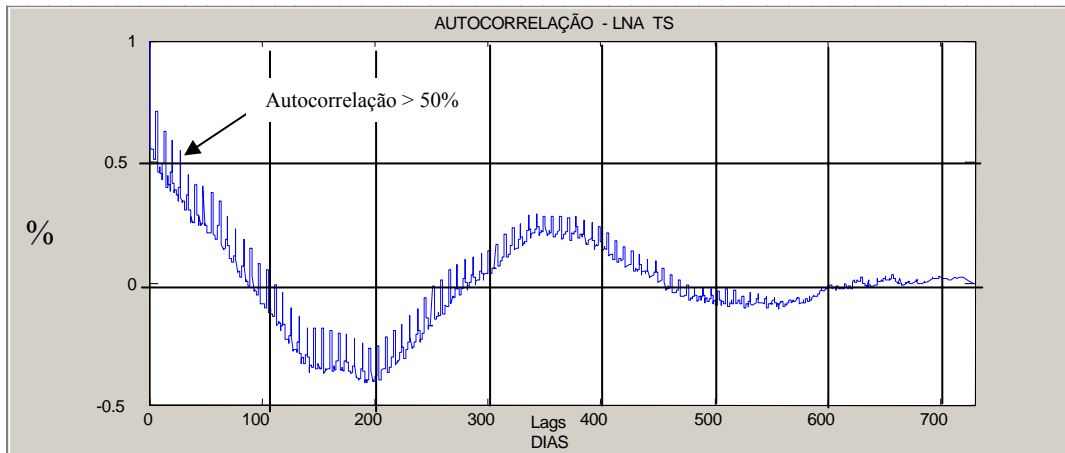


Figura 2.24 – Autocorrelação dos dados da SE LNA – Período 2005/2006.

Já na Figura 2.25 observa-se a autocorrelação da SE IBI, na qual verifica-se que a correlação é maior do que 60% para as três últimas semanas do histórico, também a correlações das semanas anteriores ainda é média, o que indica uma probabilidade de que as previsões serão adequadas.

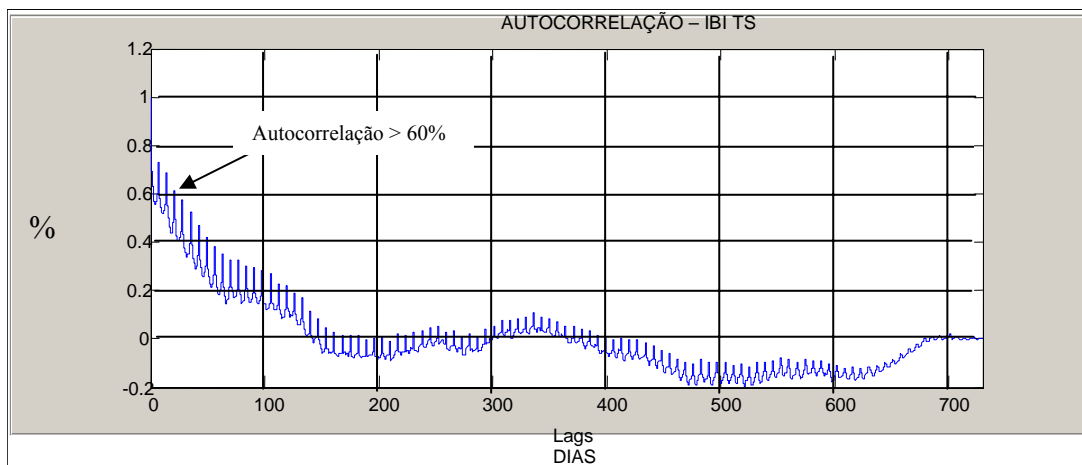


Figura 2.25 – Autocorrelação dos dados da SE IBI – Período 2005/2006.

A autocorrelação da SE FOZ apresentada na Figura 2.26 possibilita a observação de que a correlação é maior do que 50% somente para os 28 últimos dias de histórico (4 semanas), e que a semana anterior proporciona correlações não muito altas, indicando a necessidade de se tomar maiores cuidados na hora da obtenção dos resultados das projeções para a tomada da decisão.

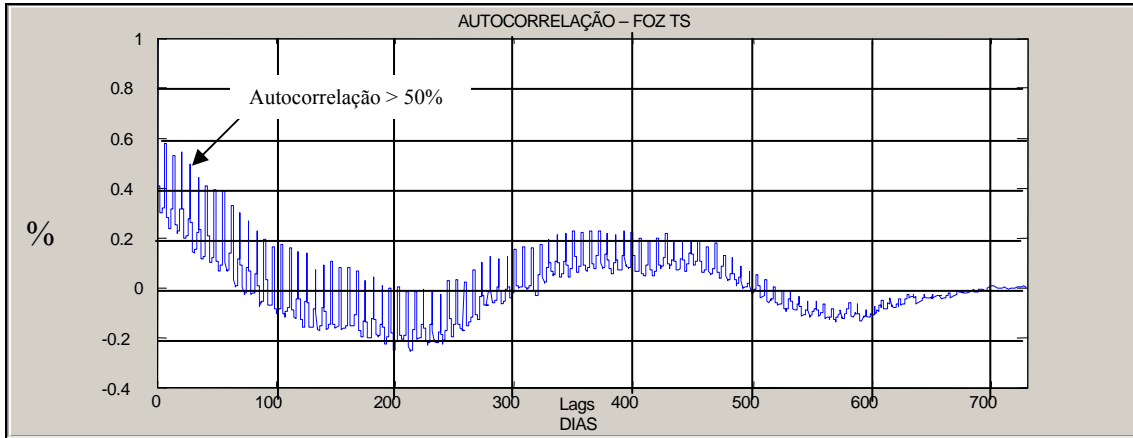


Figura 2.26 – Autocorrelação dos dados da SE FOZ – Período 2005/2006.

Na Figura 2.27 observa-se a autocorrelação da SE PTO, na qual verifica-se que a correlação é aceitável, maior do que 50% somente para os últimos sete dias de histórico, e que a semana anterior proporciona correlações não muito altas, o que permite inferir que as projeções de fluxo passante para estas subestações serão de elevada margem de erro ou mesmo com pequena exatidão, o que pode vir a ser corrigido no futuro com o incremento na base de dados, a medida que o histórico for sendo implementado com novos dados.

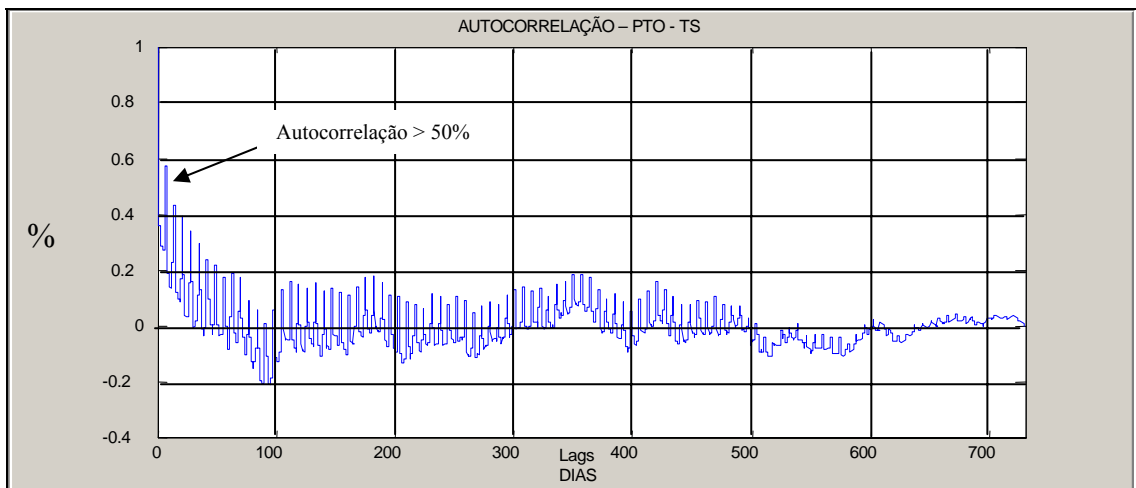


Figura 2.27 – Autocorrelação dos dados da SE PTO – Período 2005/2006.

Para a SE UMB observa-se a correlação é muito alta, maior que 80%, para os últimos sete dias de histórico (Figura 2.28), caindo rapidamente para as semanas anteriores, podendo-se inferir que os resultados das simulações de previsões não proporcionem margens de erros adequadas.



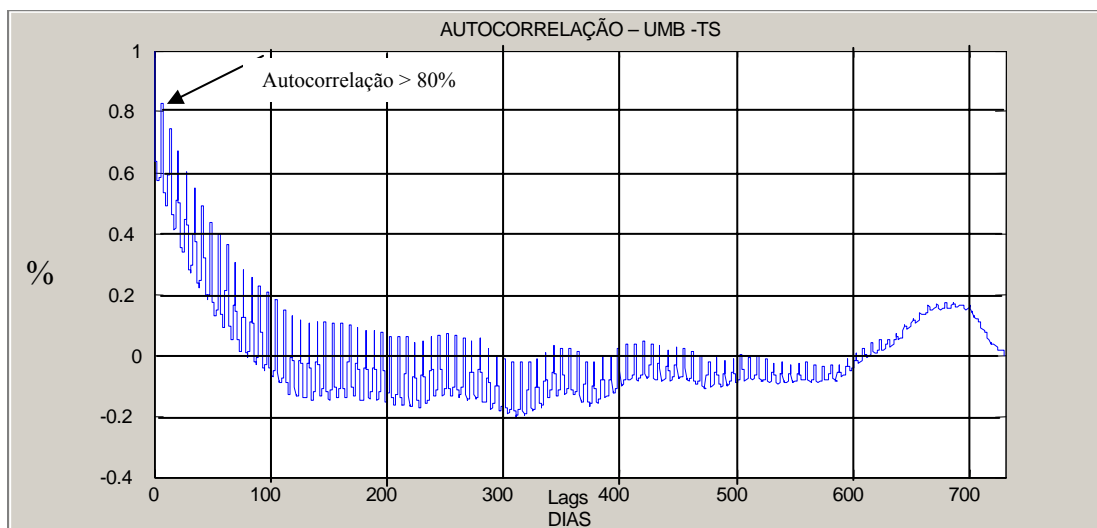


Figura 2.28 – Autocorrelação dos dados da SE UMB – Período 2005/2006.

## 2.8 CONSIDERAÇÕES FINAIS DO CAPÍTULO

Como observado, de modo geral, o tratamento da base de dados, utilizou-se de técnicas de simples aplicação, tais como:

- Retirada dos dados dos anos em que houve modificação significativa da base de dados, por alteração na configuração da área de atendimento da SE;
- Recomposição da base de dados faltantes entre dois determinados períodos através da técnica de interpolação por *spline* cúbica;
- Ajuste dos dados que ultrapassam os valores estabelecidos de dois e meio desvios padrão, que foram recompostos por interpolação *spline* cúbica.

Os ajustes e correções efetuadas na base de dados permitem inferir que os resultados de projeções a serem obtidos dependem das correções na base de dados.

Da mesma análise dos dados, visualiza-se que os anos atípicos precisam ser retirados da base de dados, pois distorcem as curvas de tendências e os resultados das correlações. Sendo assim, optou-se por utilizar os dados do período de 1º de janeiro de 2005 a 31 de dezembro de 2006 como base de dados tratada para os estudos.

Observa-se ainda que de modo geral, a informação da última semana tem correlações maiores do que 60%, fato observado para todas as SEs, permitindo inferir que quanto maior

for a correlação da última semana maior as probabilidades de se ter as projeções de fluxo passante nas SES com reduzida margem de erro, salvo quando esta correlação cai muito rapidamente para as semanas imediatamente anteriores, o que pode denotar sazonalidades semanais muito diferentes entre si.

As correlações obtidas a partir da análise dos resultados permitem deduzir que os dados da última semana do histórico ou no máximo da semana anterior é suficiente para a elaboração das previsões propriamente ditas (*lag* de uma semana) com as técnicas de interesse, pois possuem a melhor correlação para as previsões da semana seguinte.

No próximo capítulo será apresentada a formulação matemática do problema de previsão de demanda foco deste trabalho, para as técnicas adotadas.

## 3 METODOLOGIA DE PREVISÃO DE SÉRIES TEMPORAIS

### 3.1 INTRODUÇÃO

A previsão de demanda é um assunto muito importante no setor elétrico, além de se constituir um desafio matemático para a área de previsão, pois há diversas dificuldades a serem superadas, tais como:

- Os dados históricos disponíveis apresentam sazonalidades;
- Considerações de que as cargas não são estacionárias;
- Incertezas devido aos procedimentos na coleta dos dados;
- Mudanças de magnitude nas grandezas causadas principalmente por remanejamentos de cargas entre circuitos;
- Alteração de fluxos passantes nos pontos de medição que nem sempre podem ser detectados ou explicados a contento, que tem como origem:
  - Falta de registro durante as operações de manobras;
  - Mudanças de configuração de rede fora da fronteira sob análise; ou
  - Erros por falhas na instalação e aferição de medidores.

Decorre, então, a necessidade de que os dados, que se vai utilizar durante a previsão, sejam devidamente tratados, como o que foi feito no Capítulo 2, com o intuito de se evitar a entrada de dados que prejudiquem a obtenção de soluções confiáveis ou não atendam a margem de erro aceitável para os estudos. Uma base de dados não devidamente tratada pode impossibilitar que se avalie com isenção a eficiência da metodologia computacional escolhida.

As previsões de carga são usualmente elaboradas por construção de modelos e sobre informações relativas, tal como clima e dados históricos de demanda. Algumas importantes assertivas que serão adotadas neste trabalho foram descritas por Chen, Chang e Lin (2001) dos quais pode-se inferir que:

- a) O estudo desenvolvido tratou a previsão de carga para um agrupamento de dias à frente, semana à frente e mesmo um mês à frente, fazendo uso de técnica de *Support Vector Machine* – SVM, com sucesso, para as citadas previsões;

- b) Também concluem que a própria estimativa de temperatura não é tarefa fácil dada a grande gama de variáveis envolvidas, logo o uso da temperatura, por si só, pode ser um dilema. Neste trabalho optou-se em não fazer uso das informações de temperatura, dada a dificuldade inerente a própria previsão da temperatura e mesmo porque os dados reais de temperaturas necessários para os estudos não estão disponíveis para a área de abrangência das SEs escolhidas;
- c) A simples informação dos 5 dias da semana é mais importante para a classificação dos dados temporais do que os valores das informações numéricas em si;
- d) Ainda que, as informações de dias de feriados nacionais, eventos locais e dias de festividades, também afetam as demandas. Estes eventos podem levar as altas demandas devido ao uso extra de eletricidade, ou o contrário. A influência destes eventos pode ser local ou nacional. No caso do histórico apresentado para as SEs da COPEL as influências dos feriados são facilmente identificadas dado que as demandas normalmente diminuem nos feriados e finais de semana.

Com inspeções mais profundas observa-se que a carga também depende do que representa o feriado. Sobre alguns feriados de maior relevância como: Natal, Ano Novo, Carnaval (segunda e terça), Páscoa, *Corpus Christi*, Doze de Outubro - Padroeira do Brasil, Dia do trabalho, Dia da Independência, Finados e Proclamação da República, verifica-se que quando estes feriados ocorrem em dias úteis a demanda por eletricidade cai mais nestes dias do que em outros feriados de menor expressão.

Chen, Chang e Lin (2001) citam também que, quando se efetua uma inspeção mais aprofundada dos dados verifica-se a existência de um padrão de carga. Uma periodicidade de carga existe para cada semana e que a demanda em finais de semana é usualmente menor que nos dias de semana. Adicionalmente a demanda de eletricidade no sábado é pouco maior do que a do domingo. Segundo Chen, Chang e Lin (2001) para uma aproximação da série baseada no tempo, previsões não precisas de um dia podem afetar o sucesso das demais previsões.

A previsão da demanda de energia elétrica é apropriada se um critério de erro for definido previamente pelo projetista atendendo as necessidades do contrato do MUST de se ter até 5% de ultrapassagens sobre os valores contratados.

É importante salientar que as demandas que ocorrem no horário das 18 às 21 horas nos sábados, domingos e feriados nacionais, pela legislação do setor elétrico brasileiro, não são computadas para efeito de ultrapassagem dos valores contratados de MUST e, portanto, não serão consideradas como horário de ponta, para efeito deste estudo.

Foram selecionadas duas técnicas de previsões que pela literatura vêm se apresentando com bons resultados, redes neurais artificiais – RNA e máquinas de vetor suporte – SVM (do inglês *Support Vector Machine*), e que, portanto tiveram sua aplicação investigada na elaboração dos estudos de previsões que são o foco deste trabalho.

No estudo não se aborda com detalhes acerca do tipo de carga existente em cada SE de conexão, pois mesmo sendo o perfil da carga uma função da composição e representatividade de cada classe de carga (residencial, rural, comercial, industrial, serviços públicos, poder público e iluminação pública). Entende-se que estas já compõem a base histórica e não são dinâmicas o suficiente para que impactem substancialmente nas previsões, como por exemplo carga industrial ou serviço público de grande porte afetando a tendência da SE, salvo no caso de gerações pontuais, que afetam significativamente o perfil de MUST de algumas SEs.

O foco principal deste trabalho é a previsão de demanda por ponto de conexão com a rede básica, para o horizonte de uma semana à frente. Essa previsão permite a orientação do operador do sistema para que, quando possível, seja tomada alguma medida operativa para que não ocorra a ultrapassagem dos fluxos passantes além dos limites admissíveis dos valores contratados de MUST.

## **3.2 MÁQUINAS DE VETOR SUPORTE PARA REGRESSÃO**

A *Support Vector Regression* (SVR) é uma nova e promissora técnica de regressão não-linear para classificação dos dados e regressão proposta por Smola e Schölkopf (1998) e é uma técnica baseada em SVM. Ambas as técnicas estão firmemente baseadas na teoria de aprendizado estatístico, ou Teoria de Vapnik-Chervonenkis (VC), que tem sido desenvolvida nas últimas três décadas.

A teoria VC caracteriza as propriedades das máquinas de aprendizagem de modo a permitir a generalização ótima de dados não vistos. Na SVR, dado um conjunto de treino  $\{(x_1, y_1) \cdots (x_n, y_n)\} \subset \chi \in \mathfrak{R}$ , onde  $\chi$  denomina o espaço dos padrões de entrada, o objetivo é encontrar uma função  $f(x)$  que desvie no máximo  $\varepsilon$  dos valores alvo  $y_i$  para todo o conjunto de treinamento, e que, ao mesmo tempo, seja o mais plana possível. Para o caso linear, a função a ser obtida é:

$$f(x) = \langle w, x \rangle + b \text{ com } w \in \chi, b \in \mathfrak{R} \quad (3.1)$$

Onde  $\langle \bullet, \bullet \rangle$  denota o produto interno em  $\chi$ . Para que a função seja o mais plana possível, é necessário que  $w$  seja o menor possível. Uma maneira de se obter isso é minimizar a norma Euclidiana, isto é,  $\|w\|^2$ . Isto pode ser escrito como o seguinte problema de otimização convexa:

$$\begin{aligned} & \text{minimizar } \frac{1}{2} \|w\|^2 \\ & \text{sujeito a } \begin{cases} y_i - \langle w, x_i \rangle - b \leq \varepsilon \\ \langle w, x_i \rangle + b - y_i \leq \varepsilon \end{cases} \end{aligned} \quad (3.2)$$

A equação (3.2) assume que existe uma função  $f$  que aproxime todos os pares  $(x_i, y_i)$  com precisão  $\varepsilon$ . Como este não é sempre o caso, é necessário criar uma margem flexível (*soft margin*) através da introdução de variáveis de folga  $\xi_i, \xi_i^*$ . Desta maneira, temos o seguinte problema de otimização:

$$\begin{aligned} & \text{minimizar } \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n (\xi_i + \xi_i^*) \\ & \text{sujeito a } \begin{cases} y_i - \langle w, x_i \rangle - b \leq \varepsilon + \xi_i \\ \langle w, x_i \rangle + b - y_i \leq \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \geq 0 \end{cases} \end{aligned} \quad (3.3)$$

A constante  $C > 0$ , chamada custo, determina a troca entre a planicidade de  $f$  e o tanto que os desvios maiores que  $\varepsilon$  são tolerados. A formulação acima é correspondente a lidar com uma função de perdas  $|\xi|_\varepsilon$  insensível a  $\varepsilon$ , descrita por:

$$|\xi|_\varepsilon := \begin{cases} 0 & \text{se } |\xi| \leq \varepsilon \\ |\xi| - \varepsilon & \text{caso contrário} \end{cases} \quad (3.4)$$

O problema descrito na equação (3.3) pode ser resolvido mais facilmente na sua formulação dual. Essa formulação também é crítica para estender a SVR para funções não lineares. Após o desenvolvimento matemático, o problema de otimização dual fica:

$$\begin{aligned} & \text{maximizar} && \begin{cases} -\frac{1}{2} \sum_{i,j=1}^n (\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*) \langle x_i, x_j \rangle \\ -\varepsilon \sum_{i=1}^n (\alpha_i + \alpha_i^*) + \sum_{i=1}^n y_i (\alpha_i - \alpha_i^*) \end{cases} \\ & \text{sujeito a} && \begin{cases} \sum_{i=1}^n (\alpha_i - \alpha_i^*) = 0 \\ \alpha_i, \alpha_i^* \in [0, C] \end{cases} \end{aligned} \quad (3.5)$$

Sendo  $\alpha$  os multiplicadores de Lagrange. A função objetivo se torna, então:

$$f(x) = \sum_{i=1}^n (\alpha_i - \alpha_i^*) \langle x_i, x \rangle + b \quad (3.6)$$

Esta é a chamada expansão em vetores de suporte (*Support Vector Expansion*), isto é,  $w$  pode ser descrito como uma combinação linear dos padrões de treinamento  $x_i$ . A complexidade da representação da função pelos vetores de suporte (SVs) independe da dimensão do espaço  $\chi$ , dependendo apenas do número de SVs. Ainda mais importante é o fato que o algoritmo pode ser descrito em termos de produtos internos entre os dados. Mesmo ao avaliar  $f(x)$  não é necessário calcular  $w$  diretamente. Este fato é aproveitado para que a SVR seja capaz de realizar regressão não-linear.

Uma maneira de se obter a não-linearidade do algoritmo é através do pré-processamento dos padrões de treinamento  $x_i$ , através de um mapeamento  $\Phi: \chi \rightarrow F$  em um

espaço de características  $F$ , e aplicar o algoritmo padrão descrito acima. Isto, porém, rapidamente se torna impraticável com o aumento da dimensionalidade do mapeamento devido ao produto interno. A solução obtida é o mapeamento implícito através das chamadas funções Kernel. Como notado anteriormente, o algoritmo de SVs depende apenas dos produtos internos entre os diferentes padrões. Portanto é suficiente saber e usar Kernel  $k(x, x') := \langle \Phi(x), \Phi(x') \rangle$  ao invés de  $\Phi(x)$  explicitamente. Isto nos permite reescrever o algoritmo de SV como:

$$\begin{aligned} \text{maximizar} \quad & \begin{cases} -\frac{1}{2} \sum_{i,j=1}^n (\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*) k(x_i, x_j) \\ -\varepsilon \sum_{i=1}^n (\alpha_i + \alpha_i^*) + \sum_{i=1}^n y_i (\alpha_i - \alpha_i^*) \end{cases} \\ \text{sujeito a} \quad & \begin{cases} \sum_{i=1}^n (\alpha_i - \alpha_i^*) = 0 \\ \alpha_i, \alpha_i^* \in [0, C] \end{cases} \end{aligned} \quad (3.7)$$

e a função objetivo:

$$f(x) = \sum_{i=1}^n (\alpha_i - \alpha_i^*) (x_i, x) + b \quad (3.8)$$

Pode-se notar que a otimização busca, neste caso, a função mais plana no espaço de características  $F$ , e não no espaço de entradas.

### 3.2.1 Preparação dos Dados

No item 3.2 foi descrita a técnica do SVR e as informações a serem usadas para base de dados já foram descritas no Capítulo 2.

Nesta fase, é necessária a preparação da base de dados para construir o modelo SVM. Para isto, devem-se codificar as informações que serão usadas na entrada de dados, isto é  $x_i$ . Adicionalmente, a codificação de diferentes dados afeta a seleção dos modelos do projeto. Aqui serão discutidos estes assuntos em detalhes.



Quanto a seleção das características, cada componente dos dados de treinamento são chamados de características (atributos). Aqui, se considera as informações que devem ser incluídas. Assume-se que  $y_i$  é a carga do  $i$ -ésimo dia. Em geral incorporam-se informações do mesmo dia ou anteriores como características de  $x_i$ .

Há algumas informações básicas tais como: atributos do calendário, a periodicidade semanal da demanda de carga e que a carga de demanda de feriados é menor que a dos dias de não feriados. Por isto, a codificação destas informações (dias de semana e feriados) nas entradas do treinamento pode ser usada para modelar o problema.

Pode-se adotar o estilo de séries temporais ou não. Além de mês do ano, dias de semana, feriados e finais de semana, outras informações que devem ser codificadas como atributos são as demandas históricas. Isto introduz o conceito de séries temporais no modelo, se  $y_i$  é o valor alvo para previsão, o vetor  $x_i$ , inclui muitos valores alvo para previsão,  $y_{i-1}$ , ...,  $y_{i-\Delta}$  como atributos. Na fase de treinamento todos os  $y_i$  são conhecidos, mas para previsões futuras,  $y_{i-1}$ , ...,  $y_{i-\Delta}$  podem ser valores de previsões. Como sabido, a demanda passada pode afetar e implica na demanda futura. Por isso, inclui-se tal informação no modelo para auxiliar a previsão de demandas.

Se o treinamento dos dados está associado com diferentes séries, é melhor considerar somente a relação de dados segmentados para a mesma série do último segmento. O objetivo é prever os MUSTs para a semana imediatamente seguinte e o foco é o padrão de carga no período de nosso interesse.

Finalmente, depois de selecionar as informações a serem usadas e apropriar os dados a serem codificados, podem-se preparar muitas combinações de conjuntos de dados para treinamento. Nestes conjuntos de dados, foram codificadas os vetores das entradas, i.e.,  $x_i$ , para um particular  $i$ -ésimo dia, como segue:

$X_i = [\text{carga passada (kW), dia da semana, } flag \text{ de feriados, mês do ano}]$ .

Mês do ano = 1, ..., 12;

Carga passada ( $i$ -ésimo dia);

Dia da semana = 1, ..., 7 (domingo, ..., sábados);

*Flag* de feriados (1 ou 0);

### 3.2.2 Implementação

Uma série de experimentos são conduzidos com diferentes estruturas de construção de modelos. O modelo SVM para previsões de carga precisa ser treinado, validado e testado aplicando-se alguns parâmetros de escolha. Os parâmetros são influenciados pela performance do modelo SVM. Além disso, a fim de obter um bom modelo os parâmetros precisam ser selecionados adequadamente (Figura 3.1). Alguns aspectos importantes são:

- O custo do erro  $C$ ;
- A largura do  $\epsilon$ -insensível do tubo;
- A função de mapeamento  $\Phi$ ;
- A carga de muitos dias de histórico incluindo dados para treinamento.

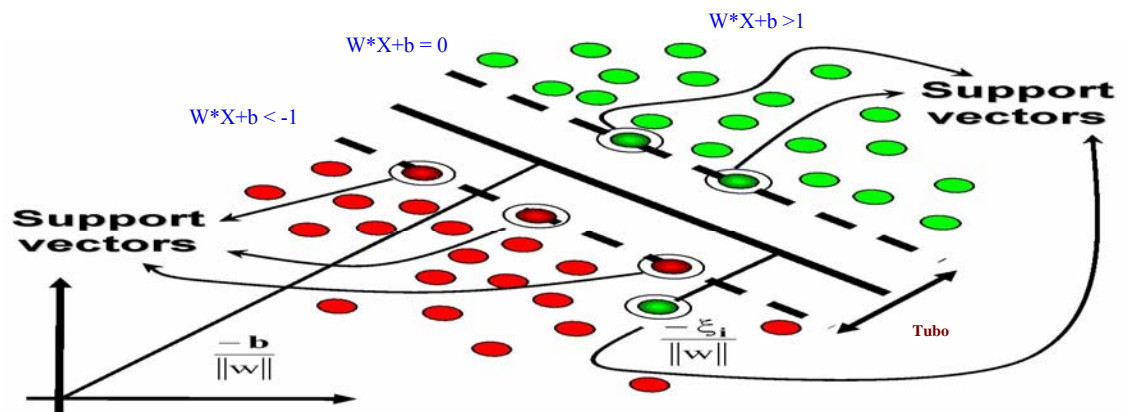


Figura 3.1 – Representação de Vetores de Suporte.

Então, para cada dado de treinamento pode-se incluir a carga máxima dos dias previstos. No caso do estudo, o parâmetro  $\epsilon$ , foi obtido através de validação cruzada (HSU. CHANG and LIN, 2007), onde se estabelece os limites de variações de um parâmetro e dentro da faixa de tolerância de variação escolhida, varia-se também os demais parâmetros  $C$  e  $\gamma$  como pares calculando-se para cada caso o conjunto de parâmetros que proporcionou o menor erro ou a melhor acurácia, reduzindo substancialmente o tempo computacional e o número de testes.

Adicionalmente, considerou-se somente o Kernell da função de base radial (RBF) definida como:

$$K(x_i, x_j) = e^{-\gamma |x_i - x_j|^2}, \quad \gamma > 0 \quad (3.9)$$

Observe que  $\gamma$  é um parâmetro associado com a função RBF. Mas os parâmetros a esquerda são  $C$  e  $\gamma$  e a busca espacial é reduzida.

Para se avaliar a performance do modelo deve-se dividir os dados de treino em dois conjuntos. Um deles é usado para treinar o modelo enquanto o outro, chamado de conjunto de validação, é usado para avaliação do modelo. De acordo com seu desempenho sobre os dados de validação tenta-se inferir valores próprios de  $C$  e  $\gamma$ .

Para aproximações baseadas em séries temporais, extraem-se, respectivamente, os dados de entrada de janeiro de 2005 a dezembro de 2006 formando um conjunto de treinamento e de validação dos modelos. A performance é decidida pelo erro médio destas duas validações. A performance do modelo é a média para estas previsões validadas.

### **3.3 REDES NEURAIIS ARTIFICIAIS**

Segundo Guerra (2006), as redes neurais (RNs) ou redes neurais artificiais são sistemas paralelos e distribuídos, compostos por unidades de processamento simples (neurônios), que podem mapear funções matemáticas não-lineares. Tais unidades são dispostas em uma, ou mais, camadas interligadas por conexões. Na maioria dos modelos, estas conexões estão associadas a pesos, os quais armazenam o conhecimento representado no modelo e servem para ponderar a entrada recebida por cada neurônio da rede. O funcionamento destas redes é inspirado em uma estrutura física concebida pela natureza: o cérebro humano.

O cérebro pode ser visto como um computador (sistema de processamento de informação) altamente complexo, não-linear e paralelo. Este tem a capacidade de organizar seus constituintes estruturais, conhecidos por neurônios, de forma a realizar certos processamentos (por exemplo, reconhecimento de padrões, percepção e controle motor) mais rapidamente que o melhor computador digital hoje existente.

O cérebro é composto por células especializadas denominadas neurônios. Cada neurônio possui um corpo celular, ou soma, dentro do qual encontram-se a maioria dos seus

orgânulos. Do corpo celular de cada neurônio partem prolongamentos que podem ser funcionalmente divididos em dendritos e axônios.

Conforme Figura 3.2, o neurônio é uma célula altamente especializada, capaz de receber impulsos elétricos a partir de seus dendritos, processá-los no corpo celular e, finalmente, transmiti-los através de seu axônio (geralmente único) para os dendritos (ou mesmo para o corpo celular) de outros neurônios.

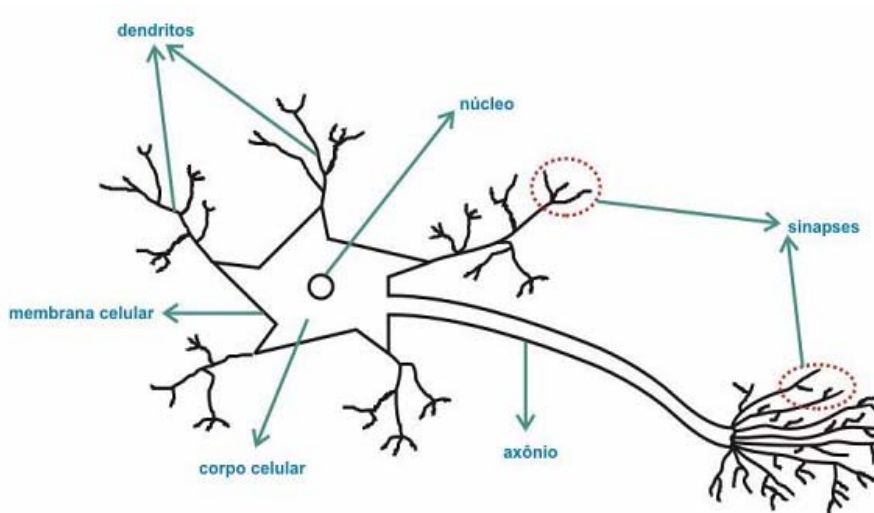


Figura 3.2 – Representação de um Neurônio.

A conexão entre um axônio de um neurônio e um dendrito de outro é denominada sinapse. A sinapse é a unidade funcional básica envolvendo as membranas plasmáticas de dois neurônios, de modo a formar uma junção pontual (o tamanho de uma junção sináptica é menor do que 1 mm) e orientada do neurônio pré-sináptico para o pós-sináptico.

As Redes Neurais Artificiais - RNAs são especificadas pela topologia, características dos neurônios e regras de treinamento. Estas regras de treinamento especificam como o conjunto de pesos iniciais deve ser adaptado para a rede aprender o comportamento de um sistema. Os algoritmos de treinamento (ou aprendizado) de uma RN ajustam iterativamente os pesos das conexões entre os neurônios até que os pares desejados de informações de entrada(s) e saída(s) sejam obtidos e as relações de causa e efeito possam ser estabelecidas. Se as condições mudarem de tal modo que o desempenho do modelo não seja mais adequado, pode-se submeter à RN a mais um treinamento sob estas novas condições de entrada(s) e

saída(s) para corrigir o seu desempenho. Assim sendo, podem-se apresentar dados a RN a uma atualização periódica, resultando num modelo auto-ajustável em linha.

Entre as aplicações usuais de RNs têm-se: reconhecimento e classificação de padrões, agrupamento (*clustering*), previsão de séries temporais, aproximação de funções, predição, otimização, etc.

Tipicamente, as redes neurais consistem de um conjunto de unidades sensoriais (nós de fonte) que compõem a camada de entrada, uma ou mais camadas ocultas de nós computacionais e uma camada de saída de nós computacionais. O sinal de entrada se propaga para frente através da rede, camada por camada. Segundo Guerra (2006) esta rede é normalmente chamada rede neural perceptron de múltiplas camadas (RN-MLP), conforme Figura 3.3.

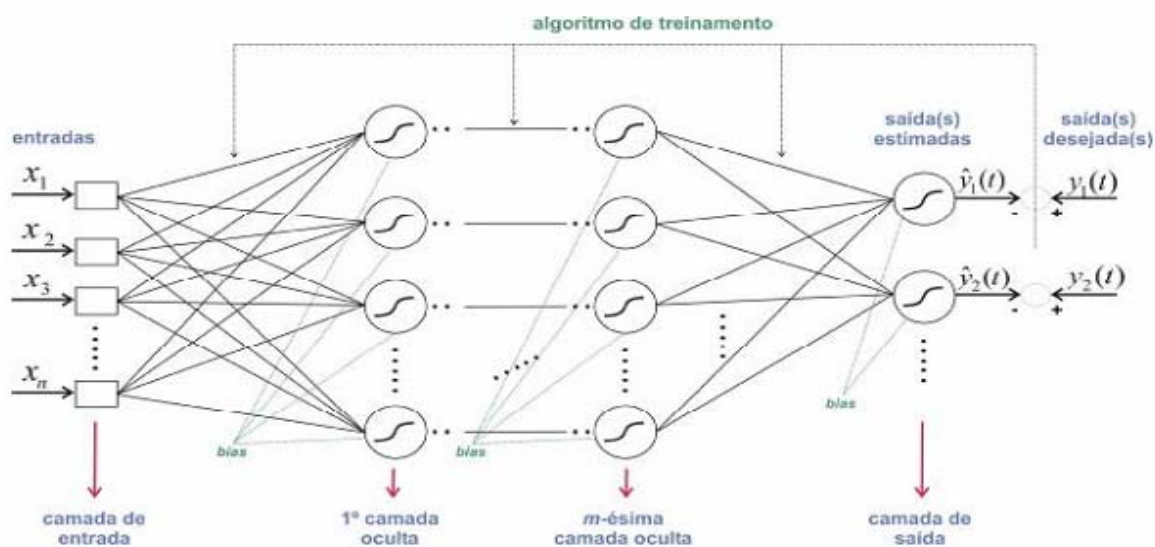


Figura 3.2 – Estrutura Geral de uma Rede Neural RN-MLP.  
Fonte: Guerra (2006).

As RN-MLP têm sido aplicadas com sucesso para resolver diversos problemas, através do seu treinamento de forma supervisionada com o algoritmo de retropropagação do erro (*error back-propagation*), muito popular. Este algoritmo é baseado na regra de aprendizagem por correção de erro. Como tal, pode ser visto como uma generalização de um algoritmo de filtragem adaptativa igualmente popular: o algoritmo de mínimos quadrados.

Basicamente, a aprendizagem por retropropagação do erro consiste de dois passos através das diferentes camadas da rede: um passo à frente, a propagação, e um passo para trás, a retropropagação. No passo para frente, um padrão de atividade (vetor de entrada) é aplicado aos nós sensoriais da rede e seu efeito se propaga através da rede, camada por camada. Finalmente, um conjunto de saídas é produzido como resposta real ou estimada da rede.

Durante o passo de propagação, os pesos sinápticos da rede são todos fixos. Durante o passo para trás, por outro lado, os pesos sinápticos são todos ajustados de acordo com uma regra de correção do erro. Especificamente, a resposta real da rede é subtraída de uma resposta desejada para produzir um sinal de erro. Este sinal é propagado para trás através da rede, contra a direção das conexões sinápticas, por isso o nome de retropropagação do erro. Os pesos sinápticos são ajustados para fazer com que a resposta real da rede se mova para mais perto da resposta desejada, em um sentido estatístico.

### 3.3.1 Rede Neural de Função de Base Radial

Segundo GUERRA (2006), a Rede Neural de Função de Base Radial (RN-RBF) pode ser considerada como uma rede neural para um problema de ajuste de curva (aproximação) em um espaço de alta dimensionalidade (Figura 3.4). Desta maneira, aprender é equivalente a encontrar uma superfície em um espaço multidimensional que forneça o melhor ajuste para os dados de treinamento, com o critério de “melhor ajuste” sendo medido em um sentido estatístico. Então o processo de generalização equivale a utilizar esta superfície multidimensional para interpolar outros pontos que não pertençam aos dados de treinamento, mas estejam em sua vizinhança.

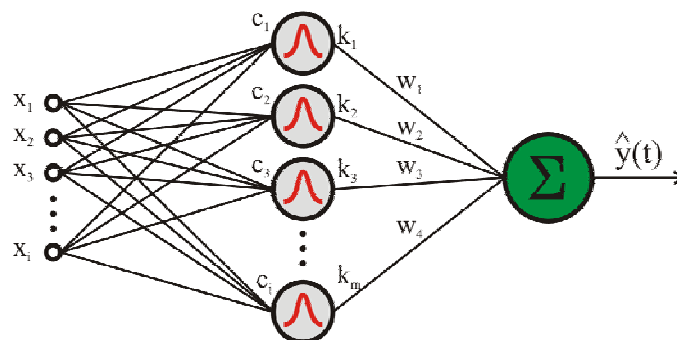


Figura 3.3 – Estrutura Geral de uma Rede Neural RN-RBF.

Fonte: Guerra (2006).

A saída estimada corresponde ao somatório da camada oculta ponderada pelos pesos das interconexões.

Para uma RN-RBF, as unidades ocultas fornecem um conjunto de “funções” que consistem em uma “base” arbitrária para os padrões (vetores) de entrada, quando estes são expandidos sobre o espaço oculto: estas funções são denominadas funções de base radial, das quais a rede deriva seu nome. Cada uma dessas funções é centrada em uma particular coordenada do espaço multidimensional dos pontos que compõem o espaço de dados de entrada. Cada uma destas coordenadas particulares caracteriza-se por definir o centro de uma (entre várias possíveis) região de maior aglomeração de pontos ou grupos (*clusters*), do espaço de dados de entrada.

Na Figura 3.5 observa-se graficamente as diferenças básicas entre redes MLP e RBF.

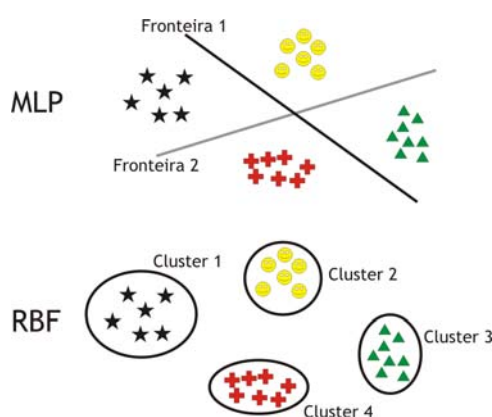


Figura 3.4 – MLP x RBF.  
Fonte: Guerra (2006).

Também desta Figura 3.5 é possível a comparação entre a forma de divisão do espaço de buscas em classes por redes neurais do tipo RBF e MLP. Enquanto redes MLP dividem o espaço através de hiperplanos, redes RBF o fazem através de hiperelipsóides, o que para o caso em tela se torna mais interessante sob o aspecto das previsões.

A arquitetura da RN-RBF, em sua estrutura mais básica, é constituída de três camadas, com propósitos totalmente diferentes:

- Camada de entrada: é composta por nós de fonte (unidades sensoriais) que conectam a rede ao seu ambiente;

- Camada oculta: é única, aplica uma transformação não-linear do espaço de entrada para o espaço oculto (na maioria das aplicações, o espaço oculto é de alta dimensionalidade);
- Camada de saída: é linear, fornecendo a resposta da rede ao padrão (sinal) de ativação aplicada à camada de entrada.

As funções de base radial são funções em geral não-lineares, cujo valor cresce ou decresce monotonicamente a medida que a distância a um ponto central aumenta. A esse ponto costuma-se denominar “centro” da função de base radial. A utilização destas funções, no contexto de aproximação de funções, tem sua origem na teoria da interpolação multivariada.

Para a demonstração deste problema, considere uma rede com uma camada de entrada, uma única camada oculta e uma camada de saída consistindo de uma única unidade. A escolha de uma unidade de saída é proposital para simplificar a exposição sem perda de generalidade. A rede é projetada para realizar um mapeamento não-linear do espaço de entrada para o espaço oculto, seguido de um mapeamento linear do espaço oculto para o espaço de saída. Considere que  $\rho$  represente a dimensão do espaço de entrada. Então, de uma maneira global a rede representa um mapa do espaço de entrada de dimensionalidade  $\rho$  em um espaço de saída unidimensional, escrito como:

$$s : \mathcal{R}^{\rho} \rightarrow \mathcal{R}^1 \quad (3.10)$$

Considera-se o mapa  $s$  como uma hipersuperfície  $\Gamma \subset \mathcal{R}^{\rho+1}$ . A superfície  $\Gamma$  é um gráfico multidimensional da saída como função da entrada. Em uma situação prática, a superfície  $\Gamma$  é desconhecida e os dados de treinamento estão normalmente contaminados com ruído.

A fase de treinamento e a fase de generalização do processo de aprendizagem podem ser entendidas, respectivamente, como:

- A fase de treinamento constitui a otimização de um procedimento de ajuste para a superfície  $\Gamma$ , baseada nos pontos dos dados conhecidos apresentados à rede na forma de exemplos (padrões) de entrada-saída;



- A fase de generalização é sinônimo de interpolação entre os pontos de dados, com a interpolação sendo formada ao longo da superfície restrita gerada pelo procedimento de ajuste, como a aproximação ótima à superfície verdadeira  $\Gamma$ .

A superfície de interpolação (isto é, a função  $F$ ) é obrigada a passar por todos os pontos de dados de treinamento. A técnica de funções de base radial consiste em escolher uma função  $F$  que tem a seguinte forma:

$$F(\mathbf{x}) = \sum_{i=1}^N w_i \varphi(\|\mathbf{x} - \mathbf{x}_i\|) \quad (3.11)$$

onde  $\varphi(\|\mathbf{x} - \mathbf{x}_i\|) | i = 1, 2, \dots, N$  é o conjunto de  $N$  funções (geralmente não-lineares) arbitrárias, conhecidas como funções de base radial, e  $\|\cdot\|$  representa uma norma, geralmente Euclidiana. Os pontos de dados conhecidos por  $\mathbf{x}_i \in \mathcal{R}^p, i = 1, 2, \dots, N$  são tomados como centros das funções de base radial.

Inserindo as condições de interpolação obtêm-se o seguinte conjunto de equações lineares simultâneas para coeficientes (pesos) desconhecidos da expansão  $\{w_i\}$ :

$$\begin{bmatrix} \varphi_{11} & \varphi_{12} & \cdots & \varphi_{1N} \\ \varphi_{21} & \varphi_{22} & \cdots & \varphi_{2N} \\ \vdots & \vdots & \vdots & \vdots \\ \varphi_{N1} & \varphi_{N2} & \cdots & \varphi_{NN} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_N \end{bmatrix} = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_N \end{bmatrix} \quad (3.12)$$

onde

$$\varphi_{ij} = \varphi(\|\mathbf{x}_j - \mathbf{x}_i\|), (j, i) = 1, 2, \dots, N \quad (3.13)$$

Considerando

$$\mathbf{d} = [d_1, d_2, \dots, d_N]^T \text{ e } \mathbf{w} = [w_1, w_2, \dots, w_N]^T \quad (3.14)$$

Os vetores  $d$  e  $w$  representam o vetor resposta desejada e o vetor de peso linear, respectivamente, onde  $N$  é o tamanho da amostra de treinamento. Considere que  $\Phi$  represente uma matriz  $N \times N$  com elementos  $\varphi_{ji}$ :

$$\Phi = \{\varphi_{ji} \mid (j, i) = 1, 2, \dots, N\} \quad (3.15)$$

A matriz  $\Phi$  é chamada de matriz de interpolação. Pode-se então reescrever a equação (3.11) na forma compacta:

$$\Phi w = x \quad (3.16)$$

Assume-se que  $\Phi$  seja não-singular e, portanto, que exista a matriz inversa  $\Phi^{-1}$ , pode-se então revolver a equação (3.16) para o vetor de peso  $w$  como mostrado abaixo:

$$w = \Phi^{-1} x \quad (3.17)$$

Um problema neste ponto é como se pode ter certeza que a matriz de interpolação  $F$  seja não-singular. Constata-se que para uma grande classe de funções de base radial e sob certas condições, a solução deste problema é dada pelo Teorema de Micchelli (1996), mostrado a seguir (HAYKIN, 2001).

### 3.3.2 Teorema de Micchelli

Existe uma grande classe de funções de base radial que é coberta pelo Teorema de Micchelli, de particular interesse no estudo de RNs-RBF, sendo que as funções gaussianas são as mais comumente utilizadas:

$$\theta_i(\bar{x}) = e^{-\frac{\|\bar{x} - \bar{c}_i\|^2}{\sigma_i^2}} \quad (3.18)$$

Neste tipo de função de base radial, o parâmetro padrão de escolha é o da função Gaussiana. Assim, define-se a distância Euclidiana média (raio médio) que mede o espalhamento dos dados representados pela função de base radial em torno de seu centro. Os raios de cada uma das funções de base radial de uma mesma RN-RBF podem assumir

diferentes valores, no entanto, para as RN-RBF usuais, o mesmo raio utilizado para cada neurônio não-linear já permite que a rede aproxime uniformemente qualquer função contínua, desde que exista número suficiente de funções de base radial. Em síntese, o valor do raio das funções de base radial afeta as propriedades numéricas dos algoritmos de aprendizado, mas não afeta a capacidade geral de aproximação das RN-RBF segundo Haykin (2001) e Bilings e Zheng (1995).

### **3.3.3 Aplicações**

As RNs são capazes de resolver, basicamente, problemas de aproximação, predição, classificação e otimização. As aplicações conhecidas de RNs como reconhecimento de caracteres, reconhecimento de voz, previsão de séries temporais, modelagem de processos, controle de processos, entre outras, são na verdade novas formas de apresentação destas classes de problemas (HAYKIN, 2001).

O problema de predição é de que se pretende prever estados seguintes de um determinado sistema baseado nos seus estados anteriores, podendo ser caracterizado como um problema de extrapolação. O problema de predição não requer uma estrutura de rede neural específica para a sua solução, podendo o mesmo ser tratado como um problema de aproximação de funções, considerando que a saída seguinte do sistema seja uma função das entradas e saídas anteriores.

## **3.4 CONSIDERAÇÕES SOBRE O CAPÍTULO**

Foram apresentados duas técnicas de inteligência artificial, SVR e RN- RBF, que estão sendo utilizadas na atualidade, juntamente com a forma da preparação dos dados e a implementação sobre a base de dados usando estas técnicas,

Tanto SVR quanto RN-RBF, foram testados para realização de previsões de demanda. A fim de analisá-los e quantificá-los, também foram efetuadas simulações onde computou-se o tempo de processamento para as estimativas para um determinado número de projeções de dias à frente, bem como, no capítulo 4 serão apresentadas Tabelas comparativas com os resultados dos testes para ambos os métodos. Estas Tabelas conterão o Erro Percentual e o

Erro Absoluto Percentual e valores de desvios padrão, que permitiram a escolha do melhor método.

O aspecto facilidade de manuseio e o tempo necessário para a preparação dos parâmetros para o uso da ferramenta também foi considerado no capítulo 4. Outro aspecto que se procurará buscar será a robustez do modelo para a utilização num caso real com e sem tratamento de parte da base de dados.

## 4 RESULTADOS OBTIDOS

### 4.1 INTRODUÇÃO

Os dados utilizados para os estudos foram os fluxos passantes – MUSTs, nos pontos de conexão entre a distribuidora e a rede básica das SEs Campo do Assobio, Cascavel, Londrina, Ibiporã, Foz do Chopin, Pato Branco e Umbará da Copel para o período histórico de 2005 a 2006. Além dessa base de dados, foram acrescentados valores entre 1º de janeiro a 10 de abril de 2007, valores estes sem tratamento, com o objetivo de uma melhor análise do comportamento da ferramenta comparando-se o erro percentual entre os valores históricos verificados e os valores previstos pela ferramenta escolhida para as previsões, quando numa simulação de um caso real. Buscou-se utilizar como base de dados para treinamento os dados compreendidos entre 1º de janeiro de 2005 e 31 de julho de 2006 e para os testes os dados compreendidos entre agosto de 2006 e 31 de dezembro de 2006.

### 4.2 PREVISÃO USANDO MÁQUINAS DE VETOR SUPORTE PARA REGRESSÃO – SVR

O uso de Máquinas de Vetor Suporte para Regressão – SVR foi aplicada no trabalho utilizando-se o programa computacional conforme Chang e Lin (2001), onde por sugestão dos autores e outros foi usado o modelo de *Radial Basis Function – RBF Kernel de biblioteca pública divulgada por estes autores, sobre a plataforma do MATLAB®*:

$$K(x_i, x_j) = e^{-\gamma |x_i - x_j|^2}, \gamma > 0 \quad (4.1)$$

Resultados recentes de Keerthi e Lin (2003) mostraram que o modelo RBF deve ser aquele de primeira escolha, não existindo a necessidade de se considerar o uso do modelo Kernel linear que não traria ganhos maiores.

Kernels Polinômiais normalmente trazem bons resultados, mas é uma alternativa que encontra dificuldades quando a gama de dados possui grandes diferenças entre os valores mínimos e máximos, podendo nestes casos se verificar instabilidades numéricas.

Então, neste trabalho, para se reduzir os riscos com problemas de instabilidade numérica adotou-se que os dados seriam normalizados.

Para a normalização adotou-se os limites superiores e inferiores entre -1 e 1, pois entende-se que os resultados entre a escolha dos limites de normalização entre -1 e 1 ou 0 e 1 são indiferentes em termos de performance, mas em termos computacionais podem trazer algumas diferenças, especialmente se considerando a esparsidade dos dados de entrada em função da quantidade de zeros em uma ou outra alternativa.

De modo genérico foi adotada a configuração conforme apresentado abaixo, tanto para a fase de treinamento da SVM quanto para a fase de testes do mesmo:

- Utilização da opção da Máquina de Vetor Suporte para Regressão *epsilon* “ $\epsilon$ -SVR”;
- Utilização de Kernel RBF gaussiano (KEERTHI e LIN, 2003), para se estabelecer o parâmetro *gamma*  $\gamma$ ;
- Estabelecimento do parâmetro custo especificado C ou também conhecido por parâmetro de penalidade;
- Especificação do parâmetro *epsilon*  $\epsilon$  (tolerância da função perda).

Como nas SVMs, a escolha do Kernel e seus parâmetros definem a fronteira de decisão reduzida (NALDI e CARVALHO, 2005), tendo efeito no desempenho do classificador obtido. Esta escolha dos parâmetros representa uma das deficiências dos SVMs (LORENA e CARVALHO, 2003).

Uma das formas para se encontrar valores para os parâmetros livres que permitam obter os melhores resultados com economia de tempo computacional é o uso da validação cruzada (*cross validation*). Para obtenção dos melhores valores, dada a gama possível de variação dos parâmetros de entrada, foram utilizadas as fórmulas de Cherkassky (1998) para acelerar o processo de busca dos parâmetros de interesse.

Esta técnica foi usada como padrão para todas as simulações propiciando uma grande redução do esforço para as obtenções dos melhores resultados para cada caso de interesse.

#### 4.2.1 Parâmetros Básicos que Variaram Durante o Estudo

- Função custo  $C$  variando entre 1 e 7000;
- Parâmetro *epsilon* ( $\epsilon$ ) teve variação entre 0,001 e 0,75;
- Função *gamma* ( $\gamma$ ) variou entre 0,0001 e 0,5.

Foram definidos os melhores valores como aqueles obtidos por validação cruzada, que permitem a minimização do erro absoluto percentual das fases de treinamento e permitiram bons resultados nos testes.

#### 4.2.2 Período de Interesse do Histórico para Análise

Para efeito dos estudos foram considerados dois períodos distintos para a elaboração das simulações:

- Com base de dados entre o dia 01/01/2005 e o dia 31/12/2006, com todos os dados depurados, utilizando-se os primeiros dois terços da massa de dados para treinamento e o terço restante dos dados para os testes;
- Idem, com base de dados entre 01/01/2005 e 28/02/2007, sendo que para este caso à base de dados entre 01/01/2005 e o dia 31/12/2006 foram acrescentados os dias 01/01/2007 até dia 28/02/2007 sem tratamento.

#### 4.2.3 Número de Simulações de Treinamentos e Testes Utilizados

Foram efetuados testes de sensibilidade para se obter o número de testes que permite os menores erros percentuais para cada ponto de conexão, resultando em um número de 10 testes, que na realidade representa  $2^{10} = 1024$  testes. Não foi constatado em nenhum caso problemas com *overfitting*, o que serviu como subsídio para o estabelecimento do número de testes, o que foi adotado para todos os estudos para efeito comparativo entre as alternativas apresentadas no item 4.2.7.

#### 4.2.4 Sazonalidade

Com o intuito de se contemplar a influência da sazonalidade nas previsões, foram efetuados estudos quanto ao número de *lags* (atrasos em dias) a serem considerados, variando-se entre 7, 14, 21, 28, 90, 180 e 365 dias de atraso, sendo adotado 7 *lags* (7 dias de atraso) para todos os estudos, dado que as simulações demonstraram que os melhores resultados são obtidos com este período de dias de atraso,.

#### 4.2.5 Número de Previsões

O número de previsões a frente foi estabelecido em 7 dias, mas o maior interesse do estudo é a obtenção do primeiro dia de previsão com a máxima precisão, ou seja, o menor erro. Da mesma forma foram efetuados estudos com previsões entre 1 e 40 dias à frente, para possibilitar comparações dos erros das previsões versus dados reais.

Com base no exposto e das diversas simulações para cada SE, com o uso do SVR, se estabeleceu os parâmetros de referências para o estudo das demais subestações, tais como:

- *Lags* (atrasos) mais interessantes sob o ponto de vista de menores erros nas previsões;
- Tempo de histórico a ser considerado como massa de dados para possibilitar os menores erros nas previsões;
- Meses de interesse para os estudos de previsão, por exemplo, previsões para todos os dias do mês de março de 2007 e agosto 2006;

#### 4.2.6 Tempo de Processamento

Para uma melhor qualificação do tempo de processamento constante nas Tabelas, é importante citar que este consiste em um tempo médio para se fazer a separação dos dados de forma diferente para cada modelo conforme descrito, mais o processamento propriamente dito do treinamento e da validação e ainda o tempo necessário para a execução da previsão de 7 dias a frente.

Além disto, está embutido neste tempo de processamento aquele necessário para a separação da base de dados em duas partes, sendo que 2/3 dos dados são utilizados como base



de entrada para o treinamento da ferramenta e 1/3 na validação, independente do modelo a ser utilizado, mais os tempos necessários nos cálculos das estatísticas, erro, erro absoluto, erro percentual, erro absoluto percentual, desvios padrão, máximos e mínimos de cada Tabela, etc.

Não estão computados os tempos para:

- Aquisição da base de dados das Tabelas do Excel<sup>®</sup> para o MATLAB<sup>®</sup>;
- Exportação dos resultados de cada simulação para Tabelas Excel<sup>®</sup>.

#### **4.2.7 Alternativas Usadas nos Estudos com SVR**

Foram estabelecidas quatro alternativas de aplicação da ferramenta SVR na base histórica, considerando-se, para cada alternativa, particularidades que possibilitassem uma maior sensibilidade e certeza no tipo de tratamento que se fez na base de dados para se obter os resultados de interesse.

- **Alternativa 1**

Esta alternativa considera um único modelo de previsão para todos os dias, com a sazonalidade semanal sendo modelada em sete variáveis binárias segundo codificação (1 de n), não considera feriados, os quais são tratados em um modelo dedicado, igual ao anterior, porém utilizando somente padrões associados a feriados.

Os meses também são modelados através de codificação 1 de n, ou seja, 12 variáveis binárias, bem como, os últimos sete valores medidos de carga (atrasos) como entradas.

- **Alternativa 2**

Esta alternativa desenvolve oito modelos, sendo um para cada dia da semana e mais um específico para os feriados. Somente o modelo para feriados utiliza variáveis binárias para o dia da semana (codificação 1 de n).

São adotadas 12 variáveis binárias para codificação mensal e os últimos sete valores medidos de carga como entradas (atrasos).

- **Alternativa 3**

Esta alternativa considera três modelos distintos, o primeiro tratando dias úteis, o segundo finais de semana e o terceiro feriado. O modelo de dias úteis utiliza cinco variáveis binárias (codificação 1 de n) para modelagem da sazonalidade semanal. Para o modelo de fins de semana adotam-se duas variáveis binárias para distinção entre sábados e domingos. No modelo de feriados são usadas sete variáveis binárias, para modelagem da ocorrência de feriados em cada dia da semana.

Empregam-se os últimos sete atrasos como entradas e 12 variáveis binárias para codificação do mês.

- **Alternativa 4**

Esta alternativa considera três modelos distintos, o primeiro tratando dias úteis, o segundo finais de semana e o terceiro feriados. Este modelo não utiliza variáveis binárias para codificação do dia da semana.

São empregados os últimos sete atrasos como entradas e 12 variáveis binárias para codificação do mês.

### **4.3 RESULTADOS DO SVR**

Para efeito de apresentação dos resultados optou-se em detalhar todas as simulações e resultados obtidos para a SE Campo do Assobio. É importante citar que o menor erro absoluto percentual, às vezes, por si só não significa que o resultado foi melhor que outro onde se obtém um maior erro absoluto percentual. Deve-se lembrar que o objetivo da previsão de fluxo passante, por determinado ponto de conexão, é diminuir o risco de ultrapassagem do valor contratado, otimizando-se o uso do limite de 5% de folga contratual existente para que não sejam pagas penalidades. Portanto, a adoção pura e simples de um erro absoluto percentual como melhor resultado pode impactar como risco adicional a impossibilidade em se saber se os valores previstos têm uma tendência de estar acima ou abaixo (positivos ou

negativos) em relação aos valores que irão se verificar, ou melhor, que estas projeções estão sub ou superestimadas.

As previsões superestimadas, neste caso são mais importantes para efeitos de tomada de decisão pela área técnica da empresa, pois em muitos casos é possível efetuar medidas preventivas nestas SEs, tais como: remanejamentos nos barramentos e cargas que afetam este ponto de conexão, mudanças de configuração, etc., para que os valores contratados dos MUSTs não sejam ultrapassados e conseqüentemente não ocorra penalização da distribuidora neste ponto.

#### **4.3.1 SE Campo do Assobio**

- **Primeira Análise**

Considerando-se nas estatísticas os finais de semana para a Alternativa 1, apresenta-se a seguir as simulações com os dados da SE CAS, com previsão do mês de março de 2007 (dia 01 até dia 31) e os primeiros 9 dias do mês de abril de 2007 (dia 32 até 34, mais D+2 até D+7) para o histórico de 1º janeiro de 2005 até 28 de fevereiro de 2007, sendo que a Tabela 4.1 apresenta os valores de erro em kW, a Tabela 4.2 os valores dos erros absolutos em kW, a Tabela 4.3 os valores de erro percentual e a Tabela 4.4 os valores de erros absolutos percentuais, para os dias D+1, ..., D+7.

As Tabelas 4.1 a 4.4 apresentam ainda os valores da média, desvios padrão, mínimos e máximos verificados nas previsões, considerando-se no cômputo destes valores as previsões nos finais de semana.

Tabela 4.1 – SE CAS – Erro em kW da Alternativa 1.

Dia	CAMPO DO ASSOPIO M1 ERRO (KW)							Tempo (s) processamento
	D+1	D+2	D+3	D+4	D+5	D+6	D+7	
1	1568.72	239.83	-1894.00	-2262.26	-2783.58	-1954.10	-1932.84	156.48
2	-769.47	-2369.05	-2519.40	-3089.50	-2190.17	-2167.26	-2835.52	157.46
3	-2174.30	-2409.71	-2942.82	-2020.34	-2026.96	-2672.67	-981.70	157.42
4	-2001.17	-2598.11	-1641.05	-1728.39	-2516.06	-725.65	-783.65	158.02
5	-2256.21	-1522.47	-1701.99	-2580.54	-1016.99	-966.34	259.67	159.95
6	-1273.94	-1623.70	-2575.43	-1166.54	-1086.27	261.73	70.83	162.69
7	-1431.56	-2474.82	-1107.03	-1122.50	195.54	49.58	-123.30	161.64
8	-2288.68	-947.82	-1110.61	160.99	-21.01	-134.95	-877.95	161.01
9	-505.58	-968.71	174.66	-125.93	-382.80	-1001.84	1275.20	160.51
10	-947.19	161.97	-177.34	-422.76	-1066.70	1198.08	147.22	161.00
11	373.97	62.84	-257.46	-989.30	1238.71	269.42	-85.71	162.33
12	-169.02	-440.56	-1076.31	1148.26	177.79	-175.44	-832.10	162.85
13	-410.34	-1054.65	1168.52	204.21	-153.29	-812.95	-1272.67	162.80
14	-936.68	1259.01	303.04	-79.16	-777.09	-1202.36	-572.62	164.36
15	1479.15	508.01	20.62	-780.77	-1212.88	-613.96	-1318.41	163.96
16	-409.18	-564.11	-1113.94	-1700.70	-961.88	-1715.23	-2106.73	164.85
17	-457.74	-1054.12	-1624.67	-893.51	-1669.40	-2049.54	-550.30	164.70
18	-884.88	-1488.01	-744.26	-1511.76	-1915.94	-419.25	-1182.52	165.68
19	-1062.32	-435.52	-1254.12	-1697.69	-278.25	-1082.45	-1009.86	166.02
20	-161.57	-1045.59	-1560.03	-161.02	-1067.62	-981.20	-837.37	166.87
21	-1074.71	-1584.99	-151.69	-1093.58	-1008.22	-820.42	-803.84	167.52
22	-1368.78	-32.05	-1041.33	-1009.95	-932.83	-878.68	-1057.18	166.66
23	282.15	-907.48	-973.69	-914.14	-953.33	-1104.36	-1523.38	168.76
24	-1044.44	-1067.36	-963.71	-994.15	-1109.97	-1548.30	1844.52	169.48
25	-886.66	-834.28	-936.21	-1114.79	-1614.04	1786.18	-928.23	169.34
26	-573.93	-777.99	-1060.58	-1608.99	1721.56	-938.60	-1513.20	170.95
27	-683.52	-994.20	-1591.24	1708.29	-965.31	-1518.49	-1122.98	171.26
28	-842.63	-1497.35	1787.97	-928.80	-1525.58	-1098.89	-1338.49	172.36
29	-1311.45	1940.24	-849.53	-1508.76	-1124.79	-1348.92	1021.57	173.70
30	2356.61	-624.35	-1421.03	-1034.21	-1349.13	1019.05	1854.73	173.08
31	-1386.20	-1810.63	-1409.33	-1707.44	909.45	1584.03	3756.60	173.83
32	-1588.00	-1178.05	-1595.85	890.15	1509.80	3757.49	4571.79	175.14
33	-788.29	-1303.76	999.90	1652.28	3758.06	4787.19	1902.23	175.29
34	-1192.09	1055.20	1741.88	3758.20	4835.32	1983.95	-2397.99	175.20
Media	-730.00	-834.77	-855.83	-727.21	-510.70	-330.45	-331.89	165.98
Desvio	1026.73	1047.33	1122.51	1355.70	1618.73	1623.32	1622.92	
Minimo	-2288.68	-2598.11	-2942.82	-3089.50	-2783.58	-2672.67	-2835.52	
Maximo	2356.61	1940.24	1787.97	3758.20	4835.32	4787.19	4571.79	

Tabela 4.2 – SE CAS – Erro Absoluto em kW da Alternativa 1.

Dia	CAMPO DO ASSOPIO M1 ERRO ABSOLUTO (KW)							Tempo (s) processamento
	D+1	D+2	D+3	D+4	D+5	D+6	D+7	
1	1568.72	239.83	1894.00	2262.26	2783.58	1954.10	1932.84	156.48
2	769.47	2369.05	2519.40	3089.50	2190.17	2167.26	2835.52	157.46
3	2174.30	2409.71	2942.82	2020.34	2026.96	2672.67	981.70	157.42
4	2001.17	2598.11	1641.05	1728.39	2516.06	725.65	783.65	158.02
5	2256.21	1522.47	1701.99	2580.54	1016.99	966.34	259.67	159.95
6	1273.94	1623.70	2575.43	1166.54	1086.27	261.73	70.83	162.69
7	1431.56	2474.82	1107.03	1122.50	195.54	49.58	123.30	161.64
8	2288.68	947.82	1110.61	160.99	21.01	134.95	877.95	161.01
9	505.58	968.71	174.66	125.93	382.80	1001.84	1275.20	160.51
10	947.19	161.97	177.34	422.76	1066.70	1198.08	147.22	161.00
11	373.97	62.84	257.46	989.30	1238.71	269.42	85.71	162.33
12	169.02	440.56	1076.31	1148.26	177.79	175.44	832.10	162.85
13	410.34	1054.65	1168.52	204.21	153.29	812.95	1272.67	162.80
14	936.68	1259.01	303.04	79.16	777.09	1202.36	572.62	164.36
15	1479.15	508.01	20.62	780.77	1212.88	613.96	1318.41	163.96
16	409.18	564.11	1113.94	1700.70	961.88	1715.23	2106.73	164.85
17	457.74	1054.12	1624.67	893.51	1669.40	2049.54	550.30	164.70
18	884.88	1488.01	744.26	1511.76	1915.94	419.25	1182.52	165.68
19	1062.32	435.52	1254.12	1697.69	278.25	1082.45	1009.86	166.02
20	161.57	1045.59	1560.03	161.02	1067.62	981.20	837.37	166.87
21	1074.71	1584.99	151.69	1093.58	1008.22	820.42	803.84	167.52
22	1368.78	32.05	1041.33	1009.95	932.83	878.68	1057.18	166.66
23	282.15	907.48	973.69	914.14	953.33	1104.36	1523.38	168.76
24	1044.44	1067.36	963.71	994.15	1109.97	1548.30	1844.52	169.48
25	886.66	834.28	936.21	1114.79	1614.04	1786.18	928.23	169.34
26	573.93	777.99	1060.58	1608.99	1721.56	938.60	1513.20	170.95
27	683.52	994.20	1591.24	1708.29	965.31	1518.49	1122.98	171.26
28	842.63	1497.35	1787.97	928.80	1525.58	1098.89	1338.49	172.36
29	1311.45	1940.24	849.53	1508.76	1124.79	1348.92	1021.57	173.70
30	2356.61	624.35	1421.03	1034.21	1349.13	1019.05	1854.73	173.08
31	1386.20	1810.63	1409.33	1707.44	909.45	1584.03	3756.60	173.83
32	1588.00	1178.05	1595.85	890.15	1509.80	3757.49	4571.79	175.14
33	788.29	1303.76	999.90	1652.28	3758.06	4787.19	1902.23	175.29
34	1192.09	1055.20	1741.88	3758.20	4835.32	1983.95	2397.99	175.20
Media	1086.50	1142.25	1220.33	1287.35	1354.60	1312.60	1314.50	165.98
Desvio	622.07	685.99	693.06	821.76	999.23	986.17	983.41	
Minimo	161.57	32.05	20.62	79.16	21.01	49.58	70.83	
Maximo	2356.61	2598.11	2942.82	3758.20	4835.32	4787.19	4571.79	

Tabela 4.3 – SE CAS – Erro Percentual da Alternativa 1.

Dia	CAMPO DO ASSOBIÓ M1 Erro Percentual							Tempo (s) processamento
	D+1	D+2	D+3	D+4	D+5	D+6	D+7	
1	5.895	0.838	-6.348	-8.906	-9.205	-6.462	-6.392	156.48
2	-2.688	-7.940	-9.918	-10.217	-7.243	-7.167	-9.253	157.46
3	-7.288	-9.486	-9.732	-6.681	-6.703	-8.722	-3.290	157.42
4	-7.878	-8.592	-5.427	-5.716	-8.211	-2.432	-2.662	158.02
5	-7.461	-5.035	-5.628	-8.421	-3.409	-3.283	1.110	159.95
6	-4.213	-5.369	-8.404	-3.910	-3.691	1.119	0.251	162.69
7	-4.734	-8.076	-3.710	-3.814	0.836	0.176	-0.425	161.64
8	-7.469	-3.177	-3.773	0.688	-0.074	-0.465	-2.943	161.01
9	-1.695	-3.291	0.747	-0.446	-1.319	-3.358	4.651	160.51
10	-3.218	0.693	-0.628	-1.456	-3.575	4.370	0.507	161.00
11	1.599	0.223	-0.887	-3.316	4.518	0.928	-0.295	162.33
12	-0.599	-1.518	-3.607	4.188	0.612	-0.604	-3.383	162.85
13	-1.414	-3.535	4.262	0.703	-0.528	-3.305	-4.266	162.80
14	-3.139	4.592	1.044	-0.273	-3.159	-4.030	-1.919	164.36
15	5.395	1.750	0.071	-3.174	-4.065	-2.058	-4.302	163.96
16	-1.410	-1.943	-4.529	-5.700	-3.224	-5.597	-6.875	164.85
17	-1.577	-4.286	-5.445	-2.995	-5.448	-6.688	-1.844	164.70
18	-3.598	-4.987	-2.495	-4.933	-6.252	-1.405	-3.910	165.68
19	-3.561	-1.460	-4.093	-5.540	-0.933	-3.580	-4.040	166.02
20	-0.542	-3.412	-5.091	-0.540	-3.530	-3.925	-2.807	166.87
21	-3.507	-5.172	-0.508	-3.616	-4.033	-2.750	-2.658	167.52
22	-4.467	-0.107	-3.444	-4.040	-3.127	-2.906	-3.450	166.66
23	0.946	-3.001	-3.895	-3.064	-3.153	-3.604	-4.971	168.76
24	-3.454	-4.270	-3.230	-3.288	-3.622	-5.053	6.630	169.48
25	-3.547	-2.796	-3.096	-3.638	-5.267	6.420	-3.070	169.34
26	-1.924	-2.573	-3.461	-5.251	6.188	-3.104	-6.053	170.95
27	-2.260	-3.244	-5.193	6.141	-3.192	-6.074	-3.789	171.26
28	-2.750	-4.886	6.427	-3.071	-6.103	-3.708	-4.397	172.36
29	-4.280	6.974	-2.809	-6.036	-3.796	-4.431	3.646	173.70
30	8.471	-2.065	-5.685	-3.490	-4.432	3.637	6.734	173.08
31	-4.584	-7.243	-4.756	-5.609	3.245	5.751	17.253	173.83
32	-6.353	-3.975	-5.243	3.177	5.482	17.257	19.770	175.14
33	-2.660	-4.283	3.568	5.999	17.259	20.701	8.962	175.29
34	-3.916	3.766	6.324	17.260	20.909	9.347	-7.825	175.20
Media	-2.467	-2.850	-2.900	-2.323	-1.419	-0.735	-0.744	165.98
Desvio	3.640	3.687	3.935	5.160	6.358	6.452	6.410	
Minimo	-7.878	-9.486	-9.918	-10.217	-9.205	-8.722	-9.253	
Maximo	8.471	6.974	6.427	17.260	20.909	20.701	19.770	

Tabela 4.4 – SE CAS – Erro Absoluto Percentual da Alternativa 1.

Dia	CAMPO DO ASSOBIÓ M1 Erro Absoluto Percentual							Tempo (s) processamento
	D+1	D+2	D+3	D+4	D+5	D+6	D+7	
1	5.895	0.838	6.348	8.906	9.205	6.462	6.392	156.48
2	2.688	7.940	9.918	10.217	7.243	7.167	9.253	157.46
3	7.288	9.486	9.732	6.681	6.703	8.722	3.290	157.42
4	7.878	8.592	5.427	5.716	8.211	2.432	2.662	158.02
5	7.461	5.035	5.628	8.421	3.409	3.283	1.110	159.95
6	4.213	5.369	8.404	3.910	3.691	1.119	0.251	162.69
7	4.734	8.076	3.710	3.814	0.836	0.176	0.425	161.64
8	7.469	3.177	3.773	0.688	0.074	0.465	2.943	161.01
9	1.695	3.291	0.747	0.446	1.319	3.358	4.651	160.51
10	3.218	0.693	0.628	1.456	3.575	4.370	0.507	161.00
11	1.599	0.223	0.887	3.316	4.518	0.928	0.295	162.33
12	0.599	1.518	3.607	4.188	0.612	0.604	3.383	162.85
13	1.414	3.535	4.262	0.703	0.528	3.305	4.266	162.80
14	3.139	4.592	1.044	0.273	3.159	4.030	1.919	164.36
15	5.395	1.750	0.071	3.174	4.065	2.058	4.302	163.96
16	1.410	1.943	4.529	5.700	3.224	5.597	6.875	164.85
17	1.577	4.286	5.445	2.995	5.448	6.688	1.844	164.70
18	3.598	4.987	2.495	4.933	6.252	1.405	3.910	165.68
19	3.561	1.460	4.093	5.540	0.933	3.580	4.040	166.02
20	0.542	3.412	5.091	0.540	3.530	3.925	2.807	166.87
21	3.507	5.172	0.508	3.616	4.033	2.750	2.658	167.52
22	4.467	0.107	3.444	4.040	3.127	2.906	3.450	166.66
23	0.946	3.001	3.895	3.064	3.153	3.604	4.971	168.76
24	3.454	4.270	3.230	3.288	3.622	5.053	6.630	169.48
25	3.547	2.796	3.096	3.638	5.267	6.420	3.070	169.34
26	1.924	2.573	3.461	5.251	6.188	3.104	6.053	170.95
27	2.260	3.244	5.193	6.141	3.192	6.074	3.789	171.26
28	2.750	4.886	6.427	3.071	6.103	3.708	4.397	172.36
29	4.280	6.974	2.809	6.036	3.796	4.431	3.646	173.70
30	8.471	2.065	5.685	3.490	4.432	3.637	6.734	173.08
31	4.584	7.243	4.756	5.609	3.245	5.751	17.253	173.83
32	6.353	3.975	5.243	3.177	5.482	17.257	19.770	175.14
33	2.660	4.283	3.568	5.999	17.259	20.701	8.962	175.29
34	3.916	3.766	6.324	17.260	20.909	9.347	7.825	175.20
Media	3.779	3.958	4.220	4.568	4.892	4.836	4.833	165.98
Desvio	2.192	2.413	2.408	3.270	4.223	4.254	4.193	
Minimo	0.542	0.107	0.071	0.273	0.074	0.176	0.251	
Maximo	8.471	9.486	9.918	17.260	20.909	20.701	19.770	

Como todas as Tabelas têm a mesma formatação e os dados de interesse estão resumidos nas últimas linhas, será adotada doravante, a apresentação simplificada destas Tabelas para apresentação dos resultados.

Analisando-se os resultados obtidos nas Tabelas 4.1 a 4.4 observa-se que, para a alternativa 1 o tempo de processamento para cada dia a frente variou entre 156 e 175 segundos, o erro absoluto percentual médio foi de 3,8% para o dia D+1 e 4,8% para dia D+7, também que o erro percentual médio foi de -2,5% com um desvio padrão de 3,6% para D+1 e de -0,7% com desvio padrão de 6,4% para o dia D+7.

- **Segunda Análise**

Da mesma forma agora para a Alternativa 2 considerando-se os finais de semana nas estatísticas, apresenta-se a seguir as simulações com os dados da SE CAS, com previsão do mês de março de 2007 para o histórico de 1º janeiro de 2005 até 28 de fevereiro de 2007. Os resultados estão apresentados nas Tabelas 4.5 a 4.8.

Tabela 4.5 – SE CAS – Erro em kW da Alternativa 2.

Erro	CAMPO DO ASSOBIO M2 Erro (KW)							Tempo (s) processamento
	D+1	D+2	D+3	D+4	D+5	D+6	D+7	
Media	-937.05	-1065.45	-1052.54	-933.62	-724.14	-546.90	-630.56	25.00
Desvio	1136.07	1082.31	1133.86	1394.17	1679.35	1700.80	1729.90	
Minimo	-2973.87	-2644.22	-2846.16	-3145.95	-3165.40	-2627.87	-2928.92	
Maximo	1908.78	1776.57	1622.29	3757.81	4865.09	4965.89	5007.08	

Tabela 4.6 – SE CAS – Erro Absoluto em kW da Alternativa 2.

Erro	CAMPO DO ASSOBIO M2 Erro Absoluto (KW)							Tempo (s) processamento
	D+1	D+2	D+3	D+4	D+5	D+6	D+7	
Media	1260.97	1301.28	1334.03	1441.83	1510.65	1483.38	1541.69	25.00
Desvio	746.39	772.23	770.38	836.59	1004.63	966.32	976.42	
Minimo	13.65	42.44	11.00	16.81	17.61	49.14	119.16	
Maximo	2973.87	2644.22	2846.16	3757.81	4865.09	4965.89	5007.08	

Tabela 4.7 – SE CAS – Erro Percentual da Alternativa 2.

Dia	CAMPO DO ASSOBIO M2 Erro Percentual							Tempo (s) processamento
	D+1	D+2	D+3	D+4	D+5	D+6	D+7	
Media	-3.201	-3.657	-3.603	-3.065	-2.185	-1.518	-1.803	25.00
Desvio	4.051	3.819	3.995	5.332	6.604	6.760	6.856	
Minimo	-9.834	-10.393	-10.466	-10.266	-10.468	-8.922	-9.558	
Maximo	7.173	6.386	5.831	17.258	21.038	21.474	21.652	

Tabela 4.8 – SE CAS – Erro Absoluto Percentual da Alternativa 2.

Dia	CAMPO DO ASSOBIÓ M2 Erro Absoluto Percentual							Tempo (s) processamento
	D+1	D+2	D+3	D+4	D+5	D+6	D+7	
Media	4.388	4.505	4.621	5.117	5.447	5.436	5.632	25.00
Desvio	2.670	2.730	2.708	3.335	4.239	4.198	4.203	
Minimo	0.058	0.148	0.038	0.058	0.062	0.169	0.410	
Maximo	9.834	10.393	10.466	17.258	21.038	21.474	21.652	

Analisando-se esta alternativa observa-se que o Erro Percentual ficou em -3,2% com um desvio padrão de 4% para o dia D+1 e de -1,8% com desvio padrão de 6,9% para dia D+7, Nesta alternativa o Erro Absoluto Percentual foi da ordem de 4,4 % para o dia D+1 e da ordem de 5,6% para o dia D+7, do mesmo modo o tempo de processamento ficou na ordem de 25 segundos.

- **Terceira Análise**

Para a Alternativa 3, considerando-se os finais de semana nas estatísticas, apresenta-se a seguir as simulações com os dados da SE CAS, com previsão do mês de março de 2007 para o histórico de 1º janeiro de 2005 até 28 de fevereiro de 2007. Os resultados estão apresentados nas Tabelas 4.9 a 4.12.

Tabela 4.9 – SE CAS – Erro em kW da Alternativa 3.

Dia	CAMPO DO ASSOBIÓ M3 Erro (KW)							Tempo (s) processamento
	D+1	D+2	D+3	D+4	D+5	D+6	D+7	
Media	-2172.69	-2197.25	-2150.05	-1956.31	-1745.75	-1590.26	-1601.78	91.01
Desvio	997.63	948.62	1011.10	1429.54	1798.63	1856.82	1864.64	
Minimo	-3699.84	-3713.13	-3730.26	-3719.92	-3731.98	-3751.15	-3739.58	
Maximo	331.50	-62.72	-56.45	3752.97	4484.01	4491.19	4473.66	

Tabela 4.10 – SE CAS – Erro Absoluto em kW da Alternativa 3.

Dia	CAMPO DO ASSOBIÓ M3 Erro Absoluto (KW)							Tempo (s) processamento
	D+1	D+2	D+3	D+4	D+5	D+6	D+7	
Media	2192.19	2197.25	2150.05	2177.08	2230.28	2170.75	2181.61	91.01
Desvio	952.67	948.62	1011.10	1050.44	1118.11	1094.71	1103.03	
Minimo	59.34	62.72	56.45	49.93	24.74	21.81	7.26	
Maximo	3699.84	3713.13	3730.26	3752.97	4484.01	4491.19	4473.66	

Tabela 4.11 – SE CAS – Erro Percentual da Alternativa 3.

Dia	CAMPO DO ASSOBIÓ M3 Erro Percentual							Tempo (s) processamento
	D+1	D+2	D+3	D+4	D+5	D+6	D+7	
Media	-7.413	-7.504	-7.340	-6.551	-5.671	-5.089	-5.123	91.01
Desvio	3.289	3.100	3.321	5.364	6.915	7.188	7.205	
Minimo	-12.235	-12.279	-12.336	-12.301	-12.284	-12.241	-12.203	
Maximo	1.246	-0.229	-0.206	17.236	19.390	19.421	19.346	

Tabela 4.12 – SE CAS – Erro Absoluto Percentual da Alternativa 3.

Dia	CAMPO DO ASSOBIÓ M3 Erro Absoluto Percentual							Tempo (s) processamento
	D+1	D+2	D+3	D+4	D+5	D+6	D+7	
Media	7.486	7.504	7.340	7.565	7.825	7.695	7.727	91.01
Desvio	3.113	3.100	3.321	3.745	4.226	4.163	4.175	
Minimo	0.216	0.229	0.206	0.179	0.089	0.078	0.026	
Maximo	12.235	12.279	12.336	17.236	19.390	19.421	19.346	

Analisando-se esta alternativa 3 observa-se que o Erro Percentual ficou em -7,4% com um desvio padrão de 3,3% para o dia D+1 e de -5,1% com desvio padrão de 7,2% para dia D+7. Nesta alternativa o Erro Absoluto Percentual foi da ordem de 7,5 % para o dia D+1 e da ordem de 7,7% para o dia D+7, do mesmo modo o tempo de processamento ficou na ordem de 91segundos.

- **Quarta Análise**

Para a Alternativa 4, considerando-se os finais de semana nas estatísticas, apresenta-se a seguir as simulações com os dados da SE CAS, com previsão do mês de março de 2007 para o histórico de 1º janeiro de 2005 até 28 de fevereiro de 2007. Os resultados estão apresentados nas Tabelas 4.13 a 4.16.

Tabela 4.13 – SE CAS – Erro em kW da Alternativa 4.

Dia	CAMPO DO ASSOBIÓ M4 ERRO (KW)							Tempo (s) processamento
	D+1	D+2	D+3	D+4	D+5	D+6	D+7	
Media	-2309.82	-2323.76	-2274.21	-2092.30	-1883.76	-1718.60	-1517.09	96.09
Desvio	979.18	955.95	1023.55	1454.13	1831.78	1885.44	2074.15	
Minimo	-3827.76	-3843.79	-3882.08	-3923.36	-3925.38	-3921.06	-3967.26	
Maximo	84.22	-166.13	-147.98	3752.85	4432.99	4437.12	4386.33	

Tabela 4.14 – SE CAS – Erro Absoluto em kW da Alternativa 4.

Dia	CAMPO DO ASSOBIÓ M4 ERRO ABSOLUTO (KW)							Tempo (s) processamento
	D+1	D+2	D+3	D+4	D+5	D+6	D+7	
Media	2314.77	2323.76	2274.21	2313.05	2365.28	2293.29	2307.48	96.09
Desvio	967.06	955.95	1023.55	1054.76	1116.88	1086.01	1089.77	
Minimo	84.22	166.13	147.98	149.99	169.24	143.32	157.00	
Maximo	3827.76	3843.79	3882.08	3923.36	4432.99	4437.12	4386.33	

Tabela 4.15 – SE CAS – Erro Percentual da Alternativa 4.

Dia	CAMPO DO ASSOBIÓ M4 Erro Percentual							Tempo (s) processamento
	D+1	D+2	D+3	D+4	D+5	D+6	D+7	
Media	-7.885	-7.933	-7.761	-7.013	-6.141	-5.530	-5.584	96.09
Desvio	3.183	3.092	3.327	5.422	7.003	7.262	7.263	
Minimo	-12.658	-12.711	-12.838	-12.974	-12.981	-12.796	-12.946	
Maximo	0.316	-0.606	-0.540	17.235	19.170	19.188	18.968	



Tabela 4.16 – SE CAS – Erro Absoluto Percentual da Alternativa 4.

Dia	CAMPO DO ASSOPIO M4 Erro Absoluto Percentual							Tempo (s) processamento
	D+1	D+2	D+3	D+4	D+5	D+6	D+7	
Media	7.903	7.933	7.761	8.027	8.282	8.110	8.153	96.09
Desvio	3.135	3.092	3.327	3.700	4.150	4.060	4.049	
Minimo	0.316	0.606	0.540	0.547	0.617	0.523	0.573	
Maximo	12.658	12.711	12.838	17.235	19.170	19.188	18.968	

Analisando-se esta alternativa 4 observa-se que o Erro Percentual ficou em -7,9% com um desvio padrão de 3,2% para o dia D+1 e de -5,6% com desvio padrão de 7,3% para dia D+7. Nesta alternativa o Erro Absoluto Percentual foi da ordem de 7,9 % para o dia D+1 e da ordem de 8,1% para o dia D+7, do mesmo modo o tempo de processamento ficou na ordem de 96 segundos.

#### • Análise Comparativa das Quatro Alternativas

Dos resultados anteriormente apresentados pode-se depreender as informações constantes na Tabela 4.17, onde são resumidos os resultados do erros percentuais para cada alternativa empregada, quando se considera nas estatísticas as previsões para finais de semana.

Tabela 4.17 – SE CAS – Erros Percentuais para as 4 Alternativas.

CAMPO DO ASSOPIO COMPARATIVO DAS 4 ALTERNATIVAS UTILIZADAS NAS SIMULAÇÕES																				
	ALTERNATIVA 1					ALTERNATIVA 2					ALTERNATIVA 3					ALTERNATIVA 4				
	EAM%	DPEAM%	EM%	DFEM%	tempo(s)	EAM%	DPEAM%	EM%	DFEM%	tempo(s)	EAM%	DPEAM%	EM%	DFEM%	tempo(s)	EAM%	DPEAM%	EM%	DFEM%	tempo(s)
D+1	3,78	2,19	-2,47	3,64	166,0	4,39	2,67	-3,20	4,05	25,0	7,49	3,11	-7,41	3,29	91,0	7,90	3,14	-7,88	3,18	96,1
D+2	3,96	2,41	-2,85	3,69	166,0	4,50	2,73	-3,66	3,82	25,0	7,50	3,10	-7,50	3,10	91,0	7,93	3,09	-7,93	3,09	96,1
D+3	4,22	2,41	-2,90	3,93	166,0	4,62	2,71	-3,60	4,00	25,0	7,34	3,32	-7,34	3,32	91,0	7,76	3,33	-7,76	3,33	96,1
D+4	4,57	3,27	-2,32	5,16	166,0	5,12	3,33	-3,06	5,33	25,0	7,56	3,74	-6,55	5,36	91,0	8,03	3,70	-7,01	5,42	96,1
D+5	4,89	4,22	-1,42	6,36	166,0	5,45	4,24	-2,18	6,60	25,0	7,83	4,23	-5,67	6,92	91,0	8,28	4,15	-6,14	7,00	96,1
D+6	4,84	4,25	-0,74	6,45	166,0	5,44	4,20	-1,52	6,76	25,0	7,70	4,16	-5,09	7,19	91,0	8,11	4,06	-5,53	7,26	96,1
D+7	4,83	4,19	-0,74	6,41	166,0	5,63	4,20	-1,80	6,86	25,0	7,73	4,18	-5,12	7,20	91,0	8,15	4,05	-5,58	7,26	96,1
Méda	4,38	-2,12		166,0	4,92	-2,87		25,0	7,57	-6,59		91,0	8,00	-7,04		96,1				
EAM% ERRO ABSOLUTO MÉDIO PERCENTUAL																				
DPEAM% DESMOPADRÃO DO ERRO ABSOLUTO MÉDIO PERCENTUAL																				
EM% ERRO MÉDIO PERCENTUAL																				
DFEM% DESMOPADRÃO DO ERRO MÉDIO PERCENTUAL																				

Desta tabela observa-se que as alternativas 1 e 2 tem os menores Erros Absolutos Médio Percentuais – EAM% com tempo de processamento de 166 segundos e 25 segundos respectivamente, da mesma forma as alternativas 3 e 4 apresentam resultados com os maiores EAM% com tempos de processamento entre 91 e 96 segundos em média, considerando-se os finais de semana nas estatísticas.

Da mesma forma obtém-se a Figura 4.1 resultante do erro absoluto médio percentual das quatro alternativas utilizadas, sendo:

- Resumo\_Resultados\_CAS1 é o resumo para a Alternativa 1;
- Resumo\_Resultados\_CAS2 é o resumo para a Alternativa 2;
- Resumo\_Resultados\_CAS3 é o resumo para a Alternativa 3;
- Resumo\_Resultados\_CAS4 é o resumo para a Alternativa 4.

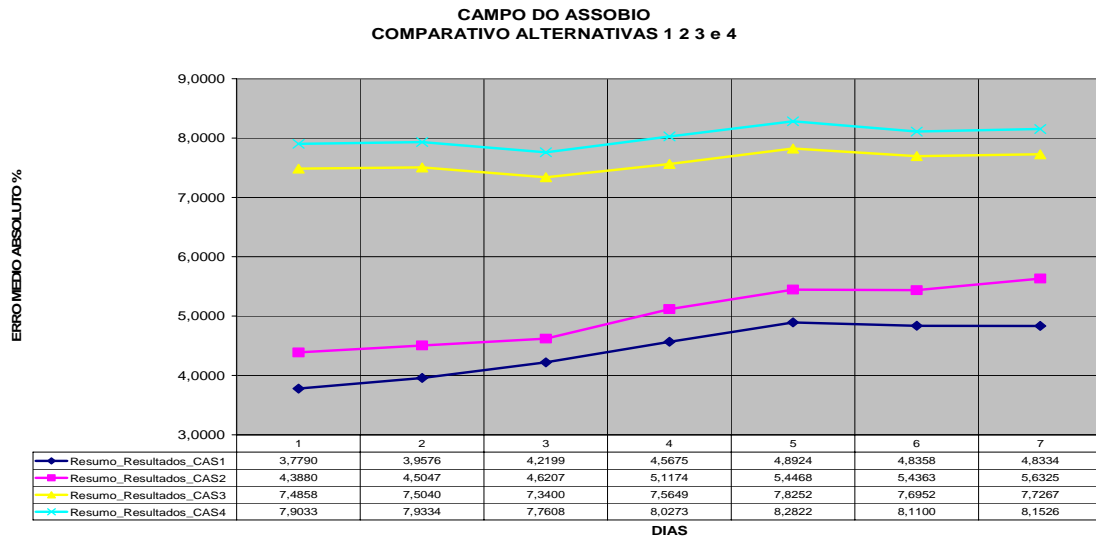


Figura 4.1 – Erro Absoluto Médio Percentual da SE CAS.

Desta figura observa-se que as alternativas 1 e 2 (linhas azul escuro e rosa) apresentam EAM% inferiores aos obtidos nas alternativas 3 e 4 (linhas azul claro e amarela)

A Figura 4.2 apresenta o resultado obtido dos EM% para as alternativas.

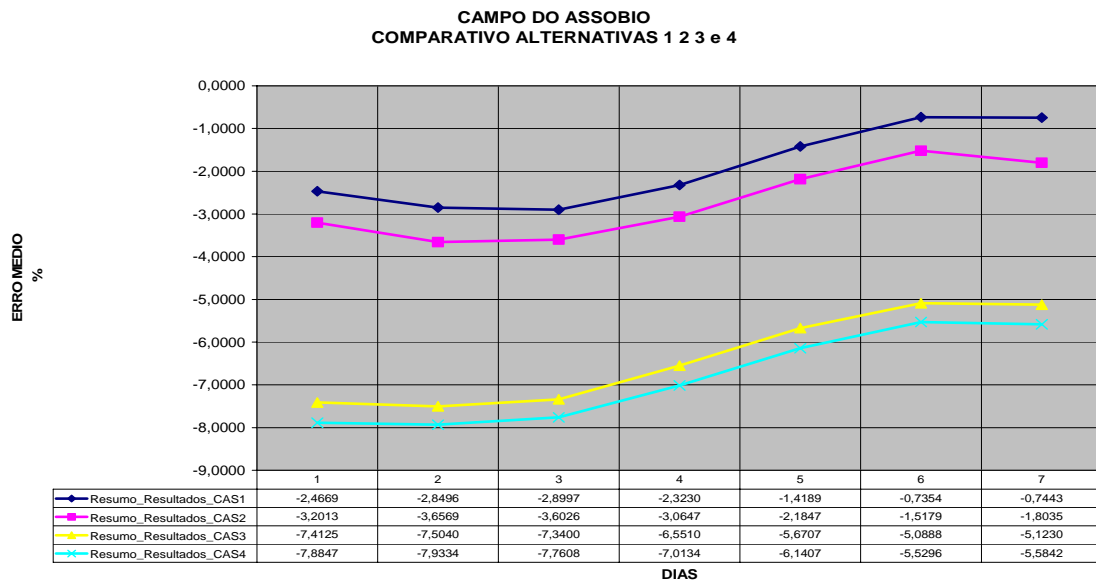


Figura 4.2 – Erro Médio Percentual da SE CAS.

Da mesma forma é observável que a alternativa 1 (linha Azul escuro) e a alternativa 2 (rosa) tem os melhores resultados em relação as alternativas 3 e 4 (linhas amarela e azul claro) quanto ao Erro Médio Percentual - EM% considerando-se os finais de semana nas estatísticas.

- **Análise das alternativas 1 e 2 descartando-se as previsões para finais de semana das estatísticas**

Doravante serão apresentados os valores resultantes das simulações desconsiderando-se nas estatísticas as influências dos finais de semana para as Alternativas 1 e 2 que foram as que demonstraram performance superior, conforme verificado na Tabela 4.17 e nas Figuras 4.1 e 4.2, tanto em termos de erros reduzidos quanto em termos de tempo computacional. A única diferença nessa análise é que se descarta no cômputo das estatísticas os valores das previsões para finais de semana.

O resultado obtido da Alternativa 1 está apresentado na Tabela 4.18 e da Alternativa 2 na Tabela 4.19.

Tabela 4.18 – SE CAS Erros Percentuais Sem Finais de Semana - Alternativa 1.

Dia	CAMPO DO ASSOBIO ALTERNATIVA 1 ERRO PERCENTUAL SEM CONSIDERAR OS FINAIS DE SEMANA							CAMPO DO ASSOBIO ALTERNATIVA 2 ERRO ABSOLUTO PERCENTUAL SEM CONSIDERAR OS FINAIS DE SEMANA							tempo(s) processamento
	D+1	D+2	D+3	D+4	D+5	D+6	D+7	D+1	D+2	D+3	D+4	D+5	D+6	D+7	
1	5,895	0,838	-6,348	-8,906	-9,205	-6,462	-6,392	5,895	0,838	6,348	8,906	9,205	6,462	6,392	156,48
2	-2,688	-7,940	-9,918	-10,217	-7,243	-7,167	-9,253	2,688	7,940	9,918	10,217	7,243	7,167	9,253	157,46
3	-7,285	-9,486	-9,732	-6,681	-6,703	-8,722	-3,290	7,288	9,486	9,732	6,681	6,703	8,722	3,290	157,42
4	-7,875	-8,592	-5,427	-5,716	-8,211	-2,432	-2,662	7,878	8,592	5,427	5,716	8,211	2,432	2,662	158,02
5	-7,461	-5,035	-5,628	-8,421	-3,409	-3,283	1,110	7,461	5,035	5,628	8,421	3,409	3,283	1,110	159,95
6	-4,213	-5,369	-8,404	-3,910	-3,891	1,119	0,251	4,213	5,369	8,404	3,910	3,891	1,119	0,251	162,69
7	-4,734	-8,076	-3,710	-3,814	0,836	0,176	-0,425	4,734	8,076	3,710	3,814	0,836	0,176	-0,425	161,64
8	-7,469	-3,177	-3,773	0,688	-0,074	-0,465	-2,943	7,469	3,177	3,773	0,688	0,074	0,465	2,943	161,01
9	-1,695	-3,291	0,747	-0,446	-1,319	-3,358	4,651	1,695	3,291	0,747	0,446	1,319	3,358	4,651	160,51
10	-3,218	0,693	-0,628	-1,456	-3,575	4,370	0,507	3,218	0,693	0,628	1,456	3,575	4,370	0,507	161,00
11	1,599	0,223	-0,887	-3,316	4,518	0,928	-0,295	1,599	0,223	0,887	3,316	4,518	0,928	0,295	162,33
12	-0,599	-1,518	-3,607	4,188	0,612	-0,604	-3,383	0,599	1,518	3,607	4,188	0,612	0,604	3,383	162,85
13	-1,414	-3,535	4,262	0,703	-0,528	-3,305	-4,266	1,414	3,535	4,262	0,703	0,528	3,305	4,266	162,80
14	-3,139	4,592	1,044	-0,273	-3,169	-4,030	-1,919	3,139	4,592	1,044	0,273	3,159	4,030	1,919	164,36
15	5,395	1,750	0,071	-3,174	-4,065	-2,058	-4,302	5,395	1,750	0,071	3,174	4,065	2,058	4,302	163,96
16	-1,410	-1,943	-4,529	-5,700	-3,224	-5,597	-6,875	1,410	1,943	4,529	5,700	3,224	5,597	6,875	164,85
17	-1,577	-4,286	-5,445	-2,995	-5,448	-6,688	-1,844	1,577	4,286	5,445	2,995	5,448	6,688	1,844	164,70
18	-3,598	-4,987	-2,495	-4,933	-6,252	-1,405	-3,910	3,598	4,987	2,495	4,933	6,252	1,405	3,910	165,68
19	-3,561	-1,460	-4,093	-5,540	-0,933	-3,580	-4,040	3,561	1,460	4,093	5,540	0,933	3,580	4,040	166,02
20	-0,542	-3,412	-5,091	-0,540	-3,530	-3,925	-2,807	0,542	3,412	5,091	0,540	3,530	3,925	2,807	166,87
21	-3,507	-5,172	-0,508	-3,816	-4,033	-2,750	-2,658	3,507	5,172	0,508	3,816	4,033	2,750	2,658	167,52
22	-4,467	-0,107	-3,444	-4,040	-3,127	-2,906	-3,450	4,467	0,107	3,444	4,040	3,127	2,906	3,450	166,66
23	0,946	-3,001	-3,895	-3,064	-3,153	-3,604	-4,971	0,946	3,001	3,895	3,064	3,153	3,604	4,971	168,76
24	-3,454	-4,270	-3,230	-3,288	-3,622	-5,053	6,630	3,454	4,270	3,230	3,288	3,622	5,053	6,630	169,48
25	-3,547	-2,796	-3,096	-3,638	-5,267	6,420	-3,070	3,547	2,796	3,096	3,638	5,267	6,420	3,070	169,34
26	-1,924	-2,573	-3,461	-5,251	6,188	-3,104	-6,053	1,924	2,573	3,461	5,251	6,188	3,104	6,053	170,95
27	-2,260	-3,244	-5,193	6,141	-3,192	-6,074	-3,789	2,260	3,244	5,193	6,141	3,192	6,074	3,789	171,26
28	-2,750	-4,886	6,427	-3,071	-6,103	-3,708	-4,397	2,750	4,886	6,427	3,071	6,103	3,708	4,397	172,36
29	-4,280	6,974	-2,809	-6,036	-3,796	-4,431	3,646	4,280	6,974	2,809	6,036	3,796	4,431	3,646	173,70
30	8,471	-2,065	-5,685	-3,490	-4,432	3,637	6,734	8,471	2,065	5,685	3,490	4,432	3,637	6,734	173,08
31	-4,584	-7,243	-4,756	-5,609	3,245	5,751	17,253	4,584	7,243	4,756	5,609	3,245	5,751	17,253	173,83
32	-6,353	-3,975	-5,243	3,177	5,482	17,257	19,770	6,353	3,975	5,243	3,177	5,482	17,257	19,770	175,14
33	-2,660	-4,283	3,568	5,999	17,259	20,701	8,962	2,660	4,283	3,568	5,999	17,259	20,701	8,962	175,29
34	-3,916	3,766	6,324	17,260	20,909	9,347	-7,825	3,916	3,766	6,324	17,260	20,909	9,347	7,825	175,20
Media	-1,832	-2,252	-2,459	-1,870	-1,830	-1,346	-1,322	3,558	3,764	4,261	4,867	4,814	4,557	4,628	165,98
Desvio	3,811	3,802	4,225	5,755	5,703	5,652	5,763	2,196	2,228	2,283	3,477	3,446	3,488	3,562	
Minimo	-7,469	-8,592	-9,732	-10,217	-9,205	-8,722	-9,253	0,542	0,107	0,508	0,446	0,074	0,176	0,251	
Maximo	8,471	6,974	6,427	17,260	17,259	17,257	17,253	8,471	8,592	9,732	17,260	17,259	17,257	17,253	

Tabela 4.19 – SE CAS Erros Percentuais Sem Finais de Semana - Alternativa 2.

Dia	CAMPO DO ASSOBIO ALTERNATIVA 2 ERRO PERCENTUAL SEM CONSIDERAR OS FINAIS DE SEMANA							CAMPO DO ASSOBIO ALTERNATIVA 2 ERRO ABSOLUTO PERCENTUAL SEM CONSIDERAR OS FINAIS DE SEMANA							tempo (s) process
	D+1	D+2	D+3	D+4	D+5	D+6	D+7	D+1	D+2	D+3	D+4	D+5	D+6	D+7	
1	7,173	0,148	-6,988	-10,201	-10,468	-6,714	-6,368	7,173	0,148	6,088	10,201	10,468	6,714	6,368	23,88
2	-2,061	-7,537	-10,466	-9,277	-6,337	-6,824	-8,717	2,061	7,537	10,466	9,277	6,337	6,824	8,717	24,07
3	-7,248	-10,393	-9,412	-6,418	-6,712	-8,575	-4,167	7,248	10,393	9,412	6,418	6,712	8,575	4,167	24,01
4	-9,826	-8,618	-5,939	-6,564	-8,281	-3,623	-3,738	9,826	8,618	5,939	6,564	8,281	3,623	3,738	24,14
5	-9,834	-7,229	-8,007	-10,266	-5,180	-4,653	-1,432	9,834	7,229	8,007	10,266	5,180	4,653	1,432	24,10
6	-6,313	-6,765	-9,103	-4,096	-4,771	-1,039	-0,906	6,313	6,765	9,103	4,096	4,771	1,039	0,906	24,31
7	-6,146	-8,629	-4,016	-4,185	-0,716	1,751	-2,157	6,146	8,629	4,016	4,185	0,716	1,751	2,157	24,22
8	-6,388	-3,525	-3,666	-0,382	0,062	0,350	-3,876	6,388	3,525	3,666	0,382	0,062	0,350	3,876	24,38
9	-3,454	-4,029	-0,414	-1,158	-0,972	-4,561	4,200	3,454	4,029	0,414	1,158	0,972	4,561	4,200	24,42
10	-3,805	-0,314	-0,657	-1,121	-4,580	4,422	-0,410	3,805	0,314	0,657	1,121	4,580	4,422	0,410	24,52
11	0,058	-1,016	-0,226	-4,764	4,562	-0,169	-1,522	0,058	1,016	0,226	4,764	4,562	0,169	1,522	24,71
12	-1,035	-0,182	-4,746	4,592	-0,169	-1,511	-5,905	1,035	0,182	4,746	4,592	0,169	1,511	5,905	24,58
13	0,097	-4,638	4,762	0,058	-1,433	-5,849	-5,652	0,097	4,638	4,762	0,058	1,433	5,849	5,652	24,70
14	-4,695	4,695	0,038	-1,472	-5,879	-5,676	-4,904	4,695	4,695	0,038	1,472	5,879	5,676	4,904	24,72
15	5,726	1,078	-1,299	-5,667	-5,132	-2,981	-4,870	5,726	1,078	1,299	5,667	5,132	2,981	4,870	24,96
16	-0,448	-2,144	-5,867	-5,844	-2,900	-5,952	-7,932	0,448	2,144	5,867	5,844	2,900	5,952	7,932	24,83
17	-2,119	-5,859	-5,790	-2,997	-5,931	-7,928	-3,052	2,119	5,859	5,790	2,997	5,931	7,928	3,052	25,01
18	-5,652	-5,529	-2,353	-5,531	-2,558	-5,651		5,652	5,529	2,353	5,531	2,558	5,651		25,08
19	-4,000	-1,189	-4,946	-7,060	-1,785	-5,267	-6,584	4,000	1,189	4,946	7,060	1,785	5,267	6,584	25,25
20	0,265	-4,477	-6,569	-1,325	-4,821	-6,311	-2,882	0,265	4,477	6,569	1,325	4,821	6,311	2,882	25,17
21	-4,545	-6,715	-1,322	-4,902	-6,351	-3,400	-3,981	4,545	6,715	1,322	4,902	6,351	3,400	3,981	25,15
22	-6,247	-0,829	-4,924	-6,181	-3,934	-3,178	-3,788	6,247	0,829	4,924	6,181	3,934	3,178	3,788	25,15
23	-0,123	-4,766	-6,103	-4,142	-3,486	-4,798	-6,442	0,123	4,766	6,103	4,142	3,486	4,798	6,442	25,28
24	-3,775	-5,999	-3,930	-3,019	-4,647	-5,004	5,689	3,775	5,999	3,930	3,019	4,647	5,004	5,689	25,33
25	-5,685	-3,553	-1,921	-4,650	-4,871	5,782	-3,564	5,685	3,553	1,921	4,650	4,871	5,782	3,564	25,25
26	-3,275	-2,842	-5,503	-5,826	5,415	-4,105	-9,021	3,275	2,842	5,503	5,826	5,415	4,105	9,021	25,37
27	-2,220	-5,270	-5,546	5,732	-4,101	-8,922	-4,661	2,220	5,270	5,546	5,732	4,101	8,922	4,661	25,43
28	-5,128	-5,340	5,831	-4,010	-8,858	-4,113	-4,671	5,128	5,340	5,831	4,010	8,858	4,113	4,671	26,98
29	-4,457	6,386	-3,486	-8,576	-5,477	-4,509	3,144	4,457	6,386	3,486	8,576	5,477	4,509	3,144	25,70
30	6,855	-3,291	-8,449	-4,792	-4,354	3,061	5,032	6,855	3,291	8,449	4,792	4,354	3,061	5,032	25,73
31	-4,306	-8,869	-5,525	-4,669	2,504	4,209	17,251	4,306	8,869	5,525	4,669	2,504	4,209	17,251	25,85
32	-8,411	-5,008	-3,522	2,704	4,616	17,255	21,652	8,411	5,008	3,522	2,704	4,616	17,255	21,652	25,85
33	-3,847	-4,193	1,580	4,552	17,256	21,474	8,125	3,847	4,193	1,580	4,552	17,256	21,474	8,125	25,87
34	-3,973	2,106	5,097	17,258	21,038	8,308	-9,558	3,973	2,106	5,097	17,258	21,038	8,308	9,558	25,83
Media	-3,201	-3,657	-3,603	-3,065	-2,185	-1,518	-1,803	4,388	4,505	4,621	5,117	5,447	5,436	5,632	25,00
Desvio	4,051	3,819	3,995	5,332	6,604	6,760	6,856	2,670	2,730	2,708	3,335	4,239	4,198	4,203	
Mínimo	-9,834	-10,393	-10,466	-10,266	-10,468	-8,922	-9,558	0,058	0,148	0,038	0,058	0,062	0,169	0,410	
Máximo	7,173	6,386	5,831	17,258	21,038	21,474	21,652	9,834	10,393	10,466	17,258	21,038	21,474	21,652	

- Comparativo Entre as Alternativas 1 e 2

Analisando-se as tabelas 4,18 e 4.20 apresentadas observam-se as influências da consideração ou não dos finais de semana para as simulações.

Neste caso o EAM% para as alternativas 1 e 2 foram de 3,6% para o dia D+1 e de 4,4% respectivamente, o EM% foi de -1,8% para um desvio padrão de 3,8% para a alternativa 1 e de -3,2% com desvio padrão de 4% para a alternativa 2.

Na Tabela 4.20 são apresentados os resultados das simulações para as duas alternativas onde não se considera as informações dos finais de semana nas estatísticas e se observa que a Alternativa 1 apresentou uma performance um pouco melhor do que a Alternativa 2, mas com uma média de tempo computacional superior a da Alternativa 2.

Tabela 4.20 – SE CAS Erros Percentuais Comparativos - Alternativas 1 e 2, desconsiderando-se os finais de semana das estatísticas

CAMPO DO ASSOPIO										
RESULTADOS ALTERNATIVAS 1 e 2										
ERRO (%)										
DIA	ALTERNATIVA 1					ALTERNATIVA 2				
	EA%	DP_EA%	EM%	DP_EM%	tempo(s)	EA%	DP_EA%	EM%	DP_EM%	tempo(s)
D+1	3,558	2,196	-1,832	3,811	165,98	4,096	2,565	-2,420	4,240	25,00
D+2	3,764	2,228	-2,252	3,802	165,98	4,165	2,565	-2,964	3,939	25,00
D+3	4,261	2,283	-2,459	4,225	165,98	4,431	2,600	-2,989	4,234	25,00
D+4	4,867	3,477	-1,870	5,755	165,98	5,137	3,476	-2,345	5,817	25,00
D+5	4,814	3,446	-1,830	5,703	165,98	5,089	3,493	-2,336	5,788	25,00
D+6	4,557	3,488	-1,346	5,652	165,98	4,891	3,354	-1,822	5,723	25,00
D+7	4,628	3,562	-1,322	5,763	165,98	5,180	3,364	-2,237	5,839	25,00
<b>MEDIA</b>	<b>4,350</b>		<b>-1,844</b>		<b>165,98</b>	<b>4,713</b>		<b>-2,445</b>		<b>25,00</b>
EA%	Erro Absoluto percentual									
DP_EA%	Desvio Padrão do Erro Absoluto %									
EM%	Erro percentual %									
DP_EM%	Desvio Padrão do Erro médio%									

Da mesma forma obtém-se a Figura 4.3 com os erros absoluto percentual e percentual para os testes das duas alternativas eleitas desconsiderando-se os finais de semana das estatísticas.

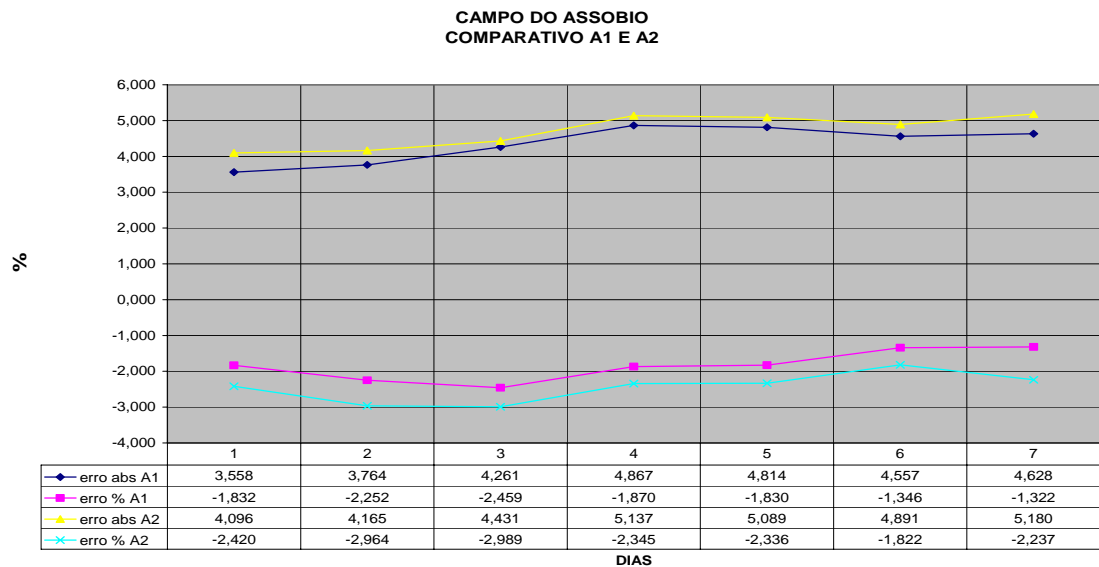


Figura 4.3 – Comparativo das Alternativas 1 e 2.

Observa-se que a alternativa 1 apresentou resultados de EAM% um pouco menores de previsão (linha azul escura) o mesmo se dando com o EM% (linha rosa) na comparação entre estas duas alternativas quando se desconsidera os finais de semana das estatísticas.

### 4.3.2 SE Cascavel

- **Comparativo Entre as Quatro Alternativas**

São apresentados os resultados em termos de erro absoluto percentual, erro percentual, e seus desvios padrão para a SE CEL para as quatro alternativas testadas. As simulações foram realizadas com os dados da SE CEL com previsão do mês de março de 2007 para o histórico de 1º janeiro de 2005 até 28 de fevereiro de 2007. Foram explicitados os resultados obtidos de erros percentual e erro absoluto percentual (Tabela 4.21), a média, desvios padrão, mínimas e máximas verificadas para os dias D+1, ..., D+7, considerando-se nas estatísticas os finais de semana.

Tabela 4.21 – SE CEL Comparativo entre as 4 Alternativas com Finais de Semana.

CASCVEL COMPARATIVO DOS 4 ALTERNATIVAS UTILIZADOS NAS SIMULAÇÕES																				
	ALTERNATIVA 1					ALTERNATIVA 2					ALTERNATIVA 3					ALTERNATIVA 4				
	EAM%	DPEAM%	EM%	DPEM%	tempo(s)	EAM%	DPEAM%	EM%	DPEM%	tempo(s)	EAM%	DPEAM%	EM%	DPEM%	tempo(s)	EAM%	DPEAM%	EM%	DPEM%	tempo(s)
D+1	6,74	5,35	3,70	7,83	156,70	3,58	7,58	6,11	5,68	23,20	8,86	7,59	5,75	10,22	81,20	8,83	7,55	5,32	10,40	86,40
D+2	8,18	5,69	4,14	9,14	156,70	4,00	9,27	7,87	6,20	23,20	9,14	7,49	5,35	10,62	81,20	9,17	7,50	5,09	10,78	86,40
D+3	8,76	6,06	4,55	9,72	156,70	4,46	9,58	8,46	6,21	23,20	9,06	7,45	5,27	10,56	81,20	9,28	7,41	5,06	10,83	86,40
D+4	8,72	6,61	4,70	9,96	156,70	4,31	9,94	8,34	6,79	23,20	9,02	7,55	4,89	10,78	81,20	9,37	7,36	4,49	11,13	86,40
D+5	8,81	6,56	4,97	9,88	156,70	4,59	10,02	8,64	6,73	23,20	9,08	7,56	4,96	10,80	81,20	9,48	7,25	4,45	11,18	86,40
D+6	8,87	7,00	4,85	10,28	156,70	4,54	10,70	9,03	7,19	23,20	9,47	7,58	4,66	11,29	81,20	9,80	7,24	4,15	11,56	86,40
D+7	8,83	7,55	4,85	10,64	156,70	4,72	10,93	9,24	7,38	23,20	9,95	7,37	4,26	11,73	81,20	10,26	7,02	4,15	11,99	86,40
Média	8,42	4,54	156,70	4,31	8,24	23,20	9,23	5,02	81,20	9,46	4,67	86,40								

EAM% ERRO ABSOLUTO MÉDIO PERCENTUAL  
DPEAM% DESVIO PADRÃO DO ERRO ABSOLUTO MÉDIO PERCENTUAL  
EM% ERRO MÉDIO PERCENTUAL  
DPEM% DESVIO PADRÃO DO ERRO MÉDIO PERCENTUAL

A Figura 4.4 mostra o comparativo das quatro alternativas empregadas, onde se pode observar uma ligeira vantagem da Alternativa 2 em termos de erros em relação a Alternativa 1 e uma grande diferença em relação as duas outras Alternativas.

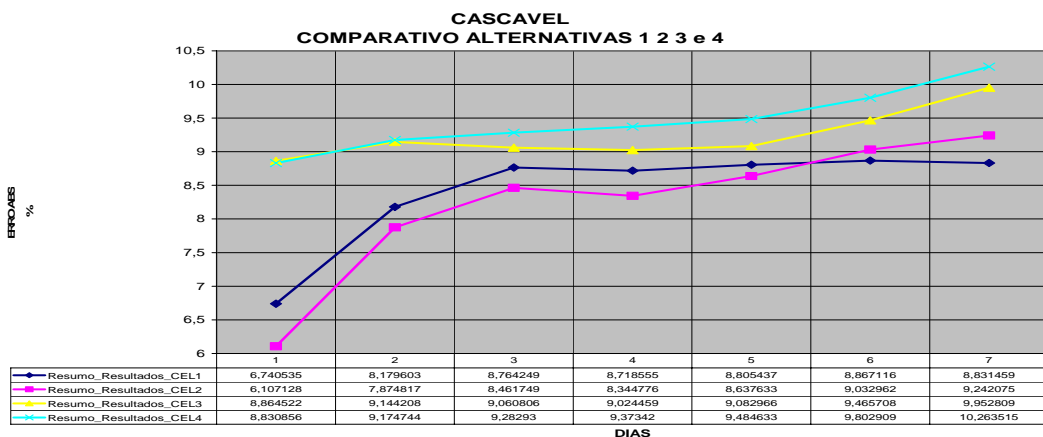


Figura 4.4 – SE CEL Comparativo entre as 4 Alternativas com Finais de Semana.

• **Comparativo Entre as Alternativas 1 e 2 sem finais de semana**

Nesta análise foram desconsiderados os valores previstos para os finais de semana. Dos resultados apresentados pode-se depreender as informações constantes na Tabela 4.22, onde se tem os resumos dos erros percentuais para cada alternativa empregada.

Tabela 4.22 – SE CEL Comparativo entre as Alternativas 1 e 2 sem Finais de Semana.

CASCAVEL										
RESULTADOS ALTERNATIVAS 1 2										
ERRO ABSOLUTO MEDIO PERCENTUAL										
DIA	MÉTODO 1					MÉTODO 2				
	EA%	DP_EA%	EM%	DP_E%	tempo(s)	EA%	DP_EA%	EM%	DP_E%	tempo(s)
D+1	5,74	4,43	1,69	7,15	155,7	5,30	4,35	1,76	6,71	22,2
D+2	7,16	4,70	1,47	8,56	155,7	6,73	4,60	1,33	8,16	22,2
D+3	7,72	4,92	1,76	9,12	155,7	7,44	4,57	1,90	8,65	22,2
D+4	7,78	5,73	2,31	9,51	155,7	7,48	5,76	2,23	9,29	22,2
D+5	8,07	5,77	2,84	9,63	155,7	7,84	5,89	2,57	9,59	22,2
D+6	7,92	6,05	3,72	9,36	155,7	8,08	6,27	3,31	9,80	22,2
D+7	8,06	6,53	3,79	9,77	155,7	8,61	6,59	3,84	10,26	22,2
<b>MÉDIA</b>	<b>7,49</b>		<b>2,51</b>		<b>155,7</b>	<b>7,35</b>		<b>2,42</b>		<b>22,2</b>
EAM% ERRO ABSOLUTO MÉDIO PERCENTUAL										
DPEAM% DESVIO PADRÃO DO ERRO ABSOLUTO MÉDIO PERCENTUAL										
EM% ERRO MÉDIO PERCENTUAL										
DPEM% DESVIO PADRÃO DO ERRO MÉDIO PERCENTUAL										

Pode-se observar desta tabela que para a SE CASCAVEL a alternativa 2 apresentou resultados melhores que a alternativa 1 tanto em termos de EAM% quanto EM%, o que pode-se observar na Figura 4.5 resultante do erro absoluto médio percentual para efeito de comparação visual das duas alternativas selecionadas.

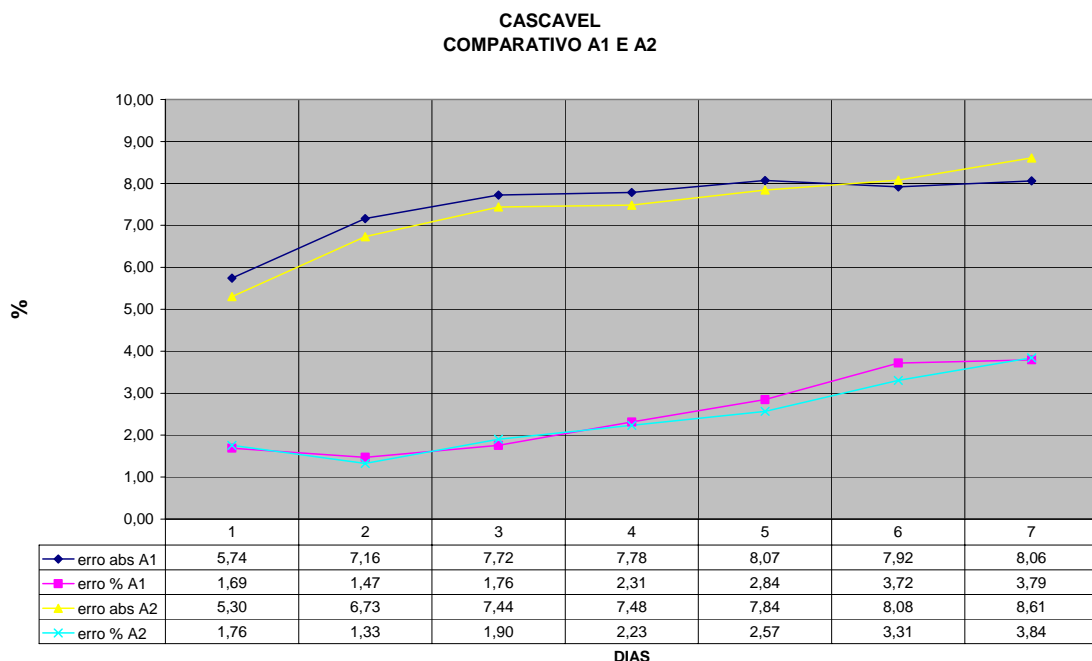


Figura 4.5 – SE CEL Comparativo entre as Alternativas 1 e 2 sem Finais de Semana.

### 4.3.3 SE Iporã

- **Comparativo Entre as Quatro Alternativas considerando-se nas estatísticas os finais de semana**

São apresentados os resultados em termos de erro absoluto percentual, erro percentual, e seus desvios padrões para a SE IBI para as quatro alternativas testadas. Foram realizadas simulações com os dados da SE IBIPORÃ com previsão do mês de março de 2007 para o histórico de 1º janeiro de 2005 até 28 de fevereiro de 2007, explicitados os resultados obtidos de erros percentual e erro absoluto percentual (Tabela 4.6.1), a média, desvios padrão, mínimas e máximas verificadas para os dias D+1, ..., D+7.

Tabela 4.23 – SE IBI Comparativo entre as 4 Alternativas com Finais de Semana.

IBIPORÃ COMPARATIVO DAS 4 ALTERNATIVAS UTILIZADAS NAS SIMULAÇÕES																					
	ALTERNATIVA 1					ALTERNATIVA 2					ALTERNATIVA 3					ALTERNATIVA 4					
	EAM%	DPEAM%	EM%	DPEM%	tempo(s)	EAM%	DPEAM%	EM%	DPEM%	tempo(s)	EAM%	DPEAM%	EM%	DPEM%	tempo(s)	EAM%	DPEAM%	EM%	DPEM%	tempo(s)	
D+1	5,27	3,96	-4,62	4,73	157,76	6,31	4,47	-5,52	5,44	24,87	14,18	3,65	-14,18	3,65	88,60	13,99	3,81	-13,99	3,81	99,08	
D+2	6,82	7,17	-6,01	7,88	157,76	7,66	7,34	-6,99	8,01	24,87	15,21	6,24	-15,21	6,24	88,60	15,00	6,35	-15,00	6,35	99,08	
D+3	6,79	7,15	-5,98	7,85	157,76	7,89	7,44	-7,18	8,15	24,87	15,20	6,24	-15,20	6,24	88,60	15,00	6,30	-15,00	6,30	99,08	
D+4	6,93	7,27	-6,13	7,98	157,76	8,19	7,39	-7,34	8,26	24,87	15,14	6,30	-15,14	6,30	88,60	14,97	6,25	-14,97	6,25	99,08	
D+5	6,91	7,22	-6,09	7,95	157,76	8,11	7,50	-7,58	8,05	24,87	14,88	6,41	-14,88	6,41	88,60	14,72	6,34	-14,72	6,34	99,08	
D+6	6,88	7,24	-6,07	7,96	157,76	7,71	7,53	-7,39	7,85	24,87	14,80	6,41	-14,80	6,41	88,60	14,64	6,37	-14,64	6,37	99,08	
D+7	6,90	7,18	-6,03	7,95	157,76	7,91	7,46	-7,59	7,79	24,87	14,61	6,50	-14,61	6,50	88,60	14,44	6,54	-14,44	6,54	99,08	
Média	6,64		-5,85		157,76	7,68		-7,08		24,87	14,86		-14,86		88,60	14,68		-14,68		3,81	99,08
EAM%	ERRO ABSOLUTO MÉDIO PERCENTUAL																				
DPEAM%	DESVIO PADRÃO DO ERRO ABSOLUTO MÉDIO PERCENTUAL																				
EM%	ERRO MÉDIO PERCENTUAL																				
DPEM%	DESVIO PADRÃO DO ERRO MÉDIO PERCENTUAL																				

A Figura 4.6 apresenta o comparativo das quatro alternativas empregadas, onde se pode observar mais uma vez uma ligeira vantagem da Alternativa 1 em termos de erros em relação a Alternativa 2 e uma grande diferença em relação as duas outras alternativas.

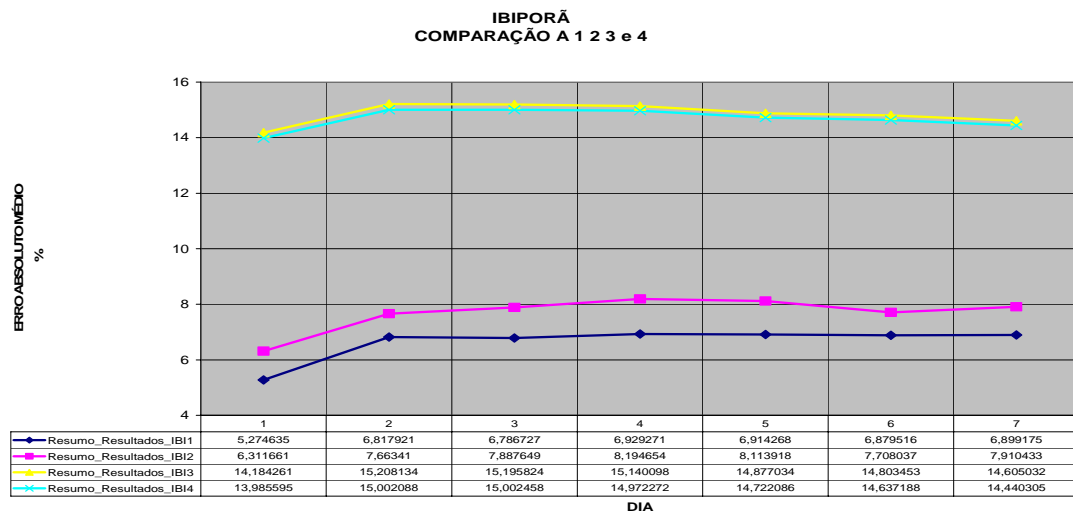


Figura 4.6 – SE IBI Comparativo entre as 4 Alternativas com Finais de Semana.



- **Comparativo Entre as Alternativas 1 e 2 desconsiderando-se os finais de semana das estatísticas**

Foram obtidas as informações constantes na Tabela 4.24, onde se tem os resumos dos erros percentuais para cada alternativa empregada.

Tabela 4.24 – SE IBI Comparativo entre as Alternativas 1 e 2 sem Finais de Semana.

IBIPORÃ														
RESULTADOS ALTERNATIVAS 1 2														
ERRO ABSOLUTO MEDIO PERCENTUAL														
DIA	ALTERNATIVA 1					ALTERNATIVA 2								
	EA%	DP	EA%	EM%	DP	E%	tempo (s)	EA%	DP	EA%	EM%	DP	E%	tempo (s)
D+1	4,50	3,29	-4,07	3,80	155,7	6,54	4,68	-6,07	5,29	23,3				
D+2	6,33	7,86	-5,86	8,20	155,7	8,39	8,22	-8,00	8,61	23,3				
D+3	6,48	7,97	-6,01	8,31	155,7	8,70	8,37	-8,32	8,76	23,3				
D+4	6,67	7,90	-6,26	8,22	155,7	9,03	8,15	-8,57	8,65	23,3				
D+5	6,85	7,93	-6,45	8,27	155,7	9,07	8,26	-8,80	8,56	23,3				
D+6	6,63	8,10	-6,19	8,45	155,7	8,56	8,52	-8,38	8,70	23,3				
D+7	6,47	8,05	-5,93	8,45	155,7	8,62	8,54	-8,47	8,69	23,3				
<b>MÉDIA</b>	<b>6,28</b>		<b>-5,82</b>		<b>158,0</b>	<b>8,42</b>		<b>-8,09</b>		<b>23,3</b>				
EAM%	ERRO ABSOLUTO MÉDIO PERCENTUAL													
DPEAM%	DESVIO PADRÃO DO ERRO ABSOLUTO MÉDIO PERCENTUAL													
EM%	ERRO MÉDIO PERCENTUAL													
DPEM%	DESVIO PADRÃO DO ERRO MÉDIO PERCENTUAL													

Desta tabela observa-se que a alternativa 1 apresentou EAM% de 6,28%, inferior ao erro da alternativa 2 que foi de 8,42%.

Da Figura 4.7 observa-se os erros absoluto médio percentual e médio percentual para efeito de comparação visual das duas alternativas selecionadas.

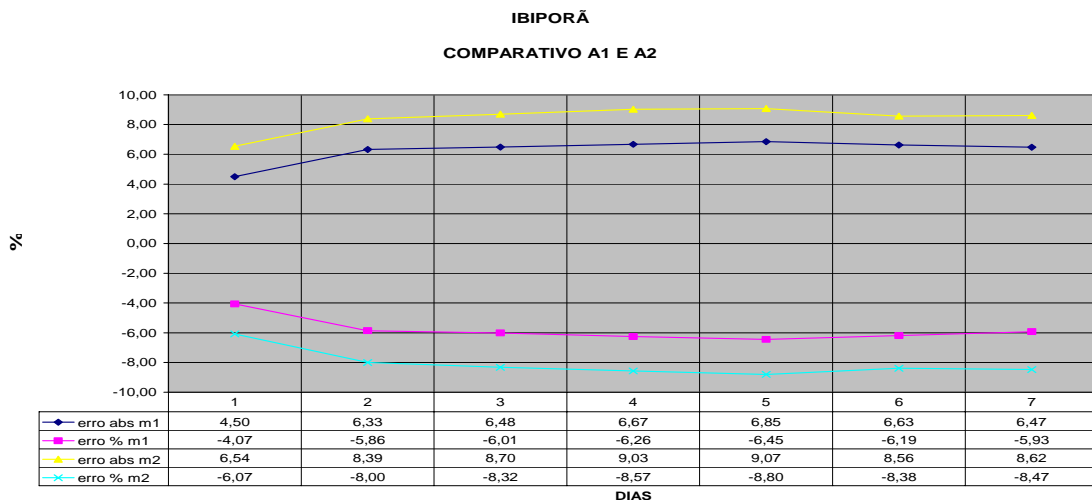


Figura 4.7 – SE IBI Comparativo entre as Alternativas 1 e 2 sem Finais de Semana.

#### 4.3.4 SE Foz do Chopin

- **Comparativo Entre as Quatro Alternativas, considerando-se nas estatísticas os finais de semana**

São apresentados os resultados em termos de erro absoluto percentual, erro percentual, e seus desvios padrões para a SE FOZ para as quatro alternativas testadas. Foram realizadas simulações com os dados da SE com previsão do mês de março de 2007 para o histórico de 1º janeiro de 2005 até 28 de fevereiro de 2007, explicitados os resultados obtidos de erros percentual e erro absoluto percentual (Tabela 4.25), a média, desvios padrão, mínimas e máximas verificadas para os dias D+1, ..., D+7.

Tabela 4.25 – SE FOZ Comparativo entre as 4 Alternativas com Finais de Semana.

DIA	ALTERNATIVA 1					ALTERNATIVA 2					ALTERNATIVA 3					ALTERNATIVA 4				
	EAM%	DPEAM%	EM%	DPEM%	tempo(s)	EAM%	DPEAM%	EM%	DPEM%	tempo(s)	EAM%	DPEAM%	EM%	DPEM%	tempo(s)	EAM%	DPEAM%	EM%	DPEM%	tempo(s)
D+1	15,56	18,47	12,46	20,75	151,41	15,22	15,63	10,81	19,04	23,32	26,84	27,79	26,83	27,79	86,40	25,66	27,91	25,09	28,43	89,61
D+2	17,41	20,31	13,18	23,36	151,41	17,75	18,44	13,03	22,13	23,32	26,22	27,87	26,07	28,01	86,40	24,91	28,21	24,56	28,53	89,61
D+3	19,27	20,90	14,81	24,37	151,41	19,63	19,77	15,55	23,22	23,32	26,72	28,11	26,56	28,26	86,40	25,25	28,30	24,88	28,63	89,61
D+4	19,42	20,96	15,68	23,97	151,41	19,82	19,47	16,63	22,33	23,32	26,58	28,04	26,58	28,04	86,40	25,34	27,97	24,73	28,52	89,61
D+5	18,48	20,45	15,62	22,77	151,41	18,97	19,31	17,06	21,06	23,32	25,86	27,00	25,85	27,00	86,40	24,63	26,72	24,00	27,30	89,61
D+6	19,47	21,01	16,18	23,71	151,41	20,44	20,70	18,26	22,71	23,32	26,16	27,08	26,15	27,08	86,40	24,85	26,74	24,21	27,34	89,61
D+7	19,83	21,10	16,55	23,83	151,41	21,64	22,59	19,71	24,34	23,32	26,12	27,21	26,08	27,25	86,40	24,93	27,09	24,26	27,71	89,61
Média	18,49		14,93		151,41	19,07		15,86		23,32	26,36		26,30		86,40	25,08		24,53		89,61
EAM%	ERRO ABSOLUTO MÉDIO PERCENTUAL																			
DPEAM%	DESVIO PADRÃO DO ERRO ABSOLUTO MÉDIO PERCENTUAL																			
EM%	ERRO MÉDIO PERCENTUAL																			
DPEM%	DESVIO PADRÃO DO ERRO MÉDIO PERCENTUAL																			

Como visto todas as alternativas da Tabela apresentaram erros não usualmente aceitáveis em previsões. Na Figura 4.8 observa-se um comparativo das quatro alternativas empregadas, onde existe uma significativa vantagem da Alternativa 2 em termos de erros em relação a Alternativa 1 e uma grande diferença em relação as duas outras alternativas.

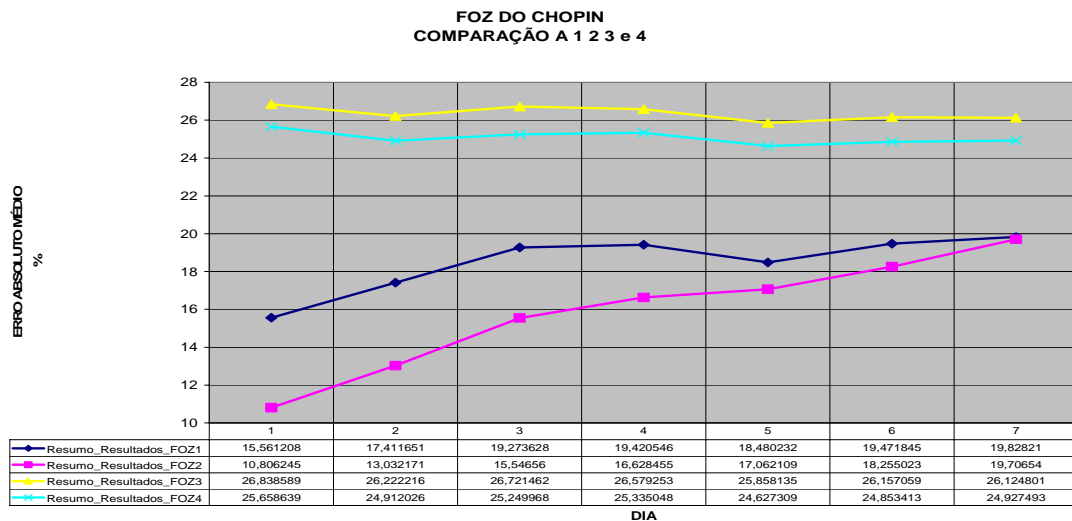


Figura 4.8 – SE FOZ Comparativo entre as 4 Alternativas com Finais de Semana.

- **Comparativo Entre as Alternativas 1 e 2 desconsiderando-se das estatísticas os finais de semana.**

Foram obtidas as informações constantes na Tabela 4.26, onde se tem os resumos dos erros percentuais para cada alternativa empregada.

Tabela 4.26 – SE FOZ Comparativo entre as Alternativas 1 e 2 sem Finais de Semana.

FOZ DO CHOPIN														
RESULTADOS ALTERNATIVAS 1 e 2														
ERRO PERCENTUAL														
	ALTERNATIVA 1						ALTERNATIVA 2							
DIA	EA%	DP	EA%	EM%	DP	E%	tempo(s)	EA%	DP	EA%	EM%	DP	E%	tempo(s)
D+1	8,21	7,14	5,75	9,31	155,7			9,73	7,46	5,91	10,87	23,3		
D+2	8,82	6,67	6,41	9,10	155,7			10,14	6,37	7,00	9,83	23,3		
D+3	10,30	8,48	8,56	10,30	155,7			11,38	7,68	9,63	9,87	23,3		
D+4	12,57	11,67	10,64	13,52	155,7			13,83	10,75	12,40	12,43	23,3		
D+5	12,56	12,01	10,62	13,83	155,7			14,18	11,33	12,56	13,16	23,3		
D+6	11,90	11,59	10,04	13,30	155,7			13,78	11,32	12,31	12,98	23,3		
D+7	12,34	11,59	10,49	13,36	155,7			14,25	11,66	13,42	12,65	23,3		
<b>MÉDIA</b>	<b>10,956</b>		<b>8,930</b>			<b>151,4</b>		<b>12,470</b>		<b>10,461</b>			<b>23,3</b>	
EAM%	ERRO ABSOLUTO MÉDIO PERCENTUAL													
DPEAM%	DESVIO PADRÃO DO ERRO ABSOLUTO MÉDIO PERCENTUAL													
EM%	ERRO MÉDIO PERCENTUAL													
DPEM%	DESVIO PADRÃO DO ERRO MÉDIO PERCENTUAL													

Observa-se da tabela que a Alternativa 1 apresentou resultados significativamente melhor que a alternativa 2

Da Figura 4.9 visualiza-se a resultante do erro absoluto médio percentual e erro médio percentual para efeito de comparação visual das duas alternativas selecionadas.

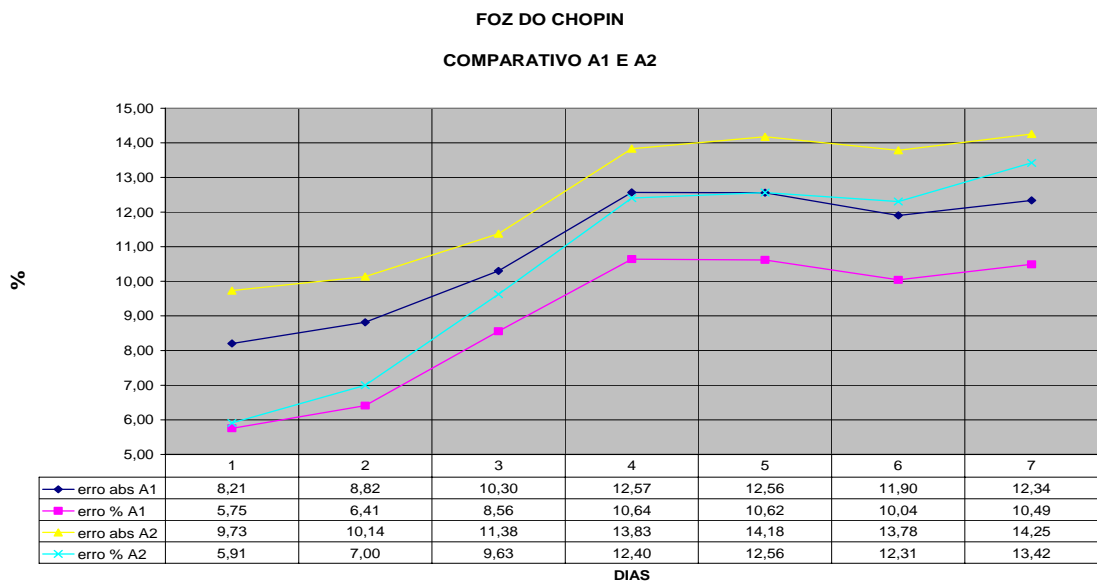


Figura 4.9 – SE FOZ Comparativo entre as Alternativas 1 e 2 sem Finais de Semana.

### 4.3.5 SE Londrina

- **Comparativo Entre as Quatro Alternativas, considerando-se nas estatísticas os finais de semana.**

São apresentados os resultados em termos de erro absoluto percentual, erro percentual, e seus desvios padrões para a SE LNA para as quatro alternativas testadas. Foram realizadas simulações com os dados da SE com previsão do mês de março de 2007 para o histórico de 1º janeiro de 2005 até 28 de fevereiro de 2007, explicitados os resultados obtidos de erros percentual e erro absoluto percentual (Tabela 4.27), a média, desvios padrão, mínimas e máximas verificadas para os dias D+1, ..., D+7.

Tabela 4.27 – SE LNA Comparativo entre as 4 Alternativas com Finais de Semana.

LONDRINA COMPARATIVO DAS 4 ALTERNATIVAS UTILIZADAS NAS SIMULAÇÕES																				
DIA	ALTERNATIVA 1					ALTERNATIVA 2					ALTERNATIVA 3					ALTERNATIVA 4				
	EAM%	DPEAM%	EM%	DPEM%	tempo(s)	EAM%	DPEAM%	EM%	DPEM%	tempo(s)	EAM%	DPEAM%	EM%	DPEM%	tempo(s)	EAM%	DPEAM%	EM%	DPEM%	tempo(s)
D+1	2,67	2,36	0,11	3,59	156,9	2,26	2,38	0,17	3,30	22,9	3,54	2,41	-2,17	3,72	77,9	3,54	2,43	-2,15	3,75	80,0
D+2	2,61	2,41	-0,25	3,57	156,9	2,23	2,30	-0,12	3,22	22,9	3,46	2,39	-2,28	3,56	77,9	3,46	2,44	-2,27	3,59	80,0
D+3	2,64	2,10	-0,22	3,40	156,9	2,19	1,92	-0,27	2,92	22,9	3,42	2,40	-2,32	3,50	77,9	3,42	2,44	-2,32	3,53	80,0
D+4	2,61	1,96	-0,31	3,29	156,9	2,33	2,04	-0,48	3,09	22,9	3,36	2,28	-2,26	3,40	77,9	3,36	2,32	-2,25	3,44	80,0
D+5	2,56	1,92	-0,19	3,23	156,9	2,14	1,81	-0,33	2,81	22,9	3,18	1,96	-2,08	3,12	77,9	3,17	1,99	-2,07	3,14	80,0
D+6	2,88	1,85	-0,11	3,46	156,9	2,32	1,69	-0,12	2,90	22,9	3,14	1,96	-2,04	3,11	77,9	3,14	1,99	-2,05	3,13	80,0
D+7	2,69	1,87	-0,49	3,27	156,9	2,23	1,78	-0,36	2,86	22,9	3,17	1,97	-2,30	2,96	77,9	3,17	2,02	-2,30	2,99	80,0
Média	2,67		-0,21		156,9	2,24		-0,22		22,9	3,32		-2,21		77,9	3,32		-2,20		80,0

Observa-se da tabela que as alternativas 1 e 2 apresentam os menores EAM%.

A Figura 4.10 apresenta o comparativo das quatro Alternativas empregadas, onde se pode observar uma ligeira vantagem da Alternativa 2 em termos de erros em relação ao Alternativa 1 e uma diferença significativa em relação as duas outras alternativas.

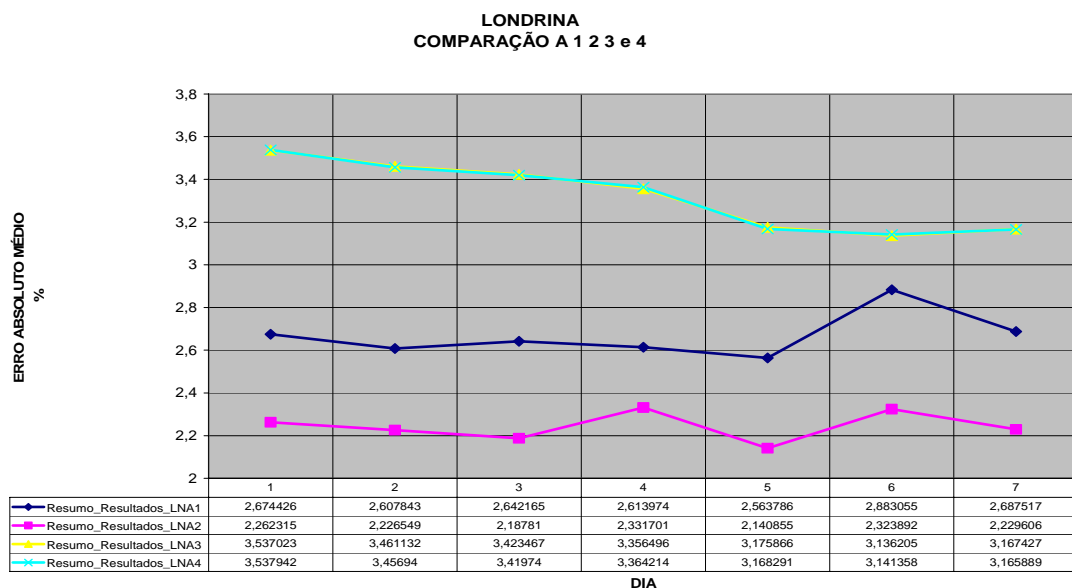


Figura 4.10 – SE LNA Comparativo entre as 4 Alternativas com Finais de Semana.

- **Comparativo Entre as Alternativas 1 e 2, desconsiderando-se das estatísticas os finais de semana.**

O resultado das simulações consta na Tabela 4.28, onde se tem os resumos dos erros percentuais para cada alternativa utilizada.

Tabela 4.28 – SE FOZ Comparativo entre as Alternativas 1 e 2 sem Finais de Semana.

LONDRINA															
RESULTADOS ALTERNATIVAS 1 2															
ERRO PERCENTUAL															
DIA	ALTERNATIVA 1						ALTERNATIVA 2								
	EA%	DP	EA%	E%	DP	E%	tempo(s)	EA%	DP	EA%	E%	DP	E%	tempo(s)	
D+1	2,39	2,37	0,32	3,38	155,7	2,05	2,52	0,71	3,20	22,9					
D+2	2,20	1,90	0,08	2,94	155,7	1,99	2,19	0,44	2,95	22,9					
D+3	2,31	1,77	-0,02	2,95	155,7	1,98	1,81	0,16	2,71	22,9					
D+4	2,44	1,95	-0,37	3,14	155,7	2,12	1,88	-0,20	2,86	22,9					
D+5	2,59	1,97	-0,39	3,28	155,7	2,09	1,81	-0,27	2,78	22,9					
D+6	2,96	1,98	-0,30	3,60	155,7	2,32	1,79	-0,15	2,96	22,9					
D+7	2,83	2,03	-0,66	3,47	155,7	2,27	1,84	-0,38	2,93	22,9					
MÉDIA	2,53		-0,19		156,9	2,12		0,04		22,9					
EAM%	ERRO ABSOLUTO MÉDIO PERCENTUAL														
DPEAM%	DESVIO PADRÃO DO ERRO ABSOLUTO MÉDIO PERCENTUAL														
EM%	ERRO MÉDIO PERCENTUAL														
DPEM%	DESVIO PADRÃO DO ERRO MÉDIO PERCENTUAL														

Observa-se uma vantagem da alternativa 2 em relação a alternativa 1.

Da mesma forma obtém-se a Figura 4.11 resultante do erro absoluto médio percentual e erro médio percentual para efeito comparação visual das duas alternativas selecionadas.

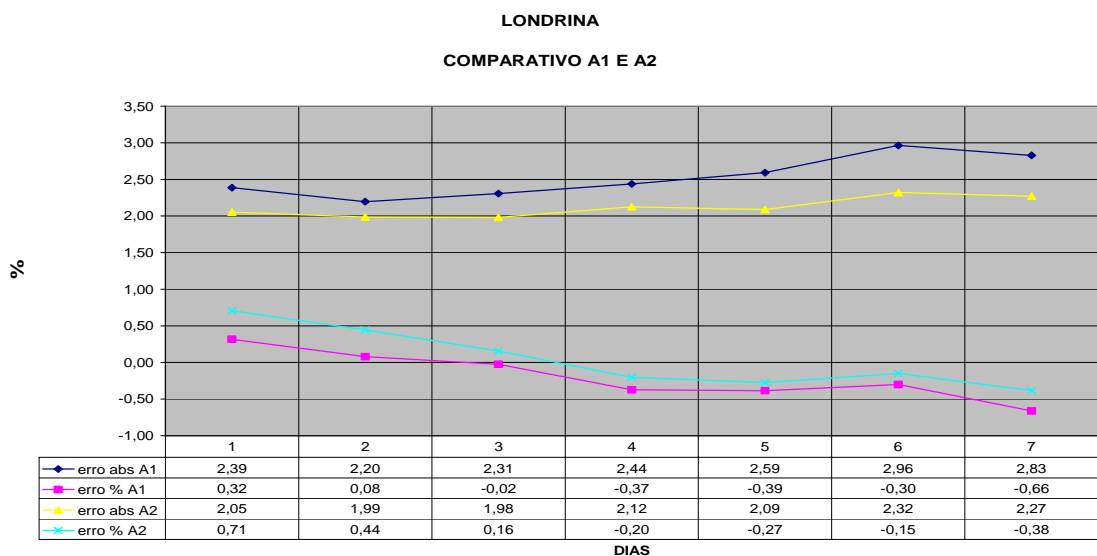


Figura 4.11 – SE LNA Comparativo entre as Alternativas 1 e 2 sem Finais de Semana.

#### 4.3.6 SE Pato Branco

- **Comparativo Entre as Quatro Alternativas, considerando-se nas estatísticas os finais de semana.**

São apresentados os resultados em termos de erro absoluto percentual, erro percentual, e seus desvios padrões para a SE PTO para as quatro alternativas testadas. Foram realizadas simulações com os dados da SE com previsão do mês de março de 2007 para o histórico de 1º janeiro de 2005 até 28 de fevereiro de 2007, explicitados os resultados obtidos de erros percentual e erro absoluto percentual (Tabela 4.29), a média, desvios padrão, mínimas e máximas verificadas para os dias D+1, ..., D+7.

Tabela 4.29 – SE PTO Comparativo entre as 4 Alternativas com Finais de Semana.

DIA	PATO BRANCO COMPARATIVO DAS 4 ALTERNATIVAS UTILIZADAS NAS SIMULAÇÕES																			
	ALTERNATIVA 1					ALTERNATIVA 2					ALTERNATIVA 3					ALTERNATIVA 4				
	EAM%	DPEAM%	EM%	DPEM%	tempo(s)	EAM%	DPEAM%	EM%	DPEM%	tempo(s)	EAM%	DPEAM%	EM%	DPEM%	tempo(s)	EAM%	DPEAM%	EM%	DPEM%	tempo(s)
D+1	7,86	5,44	-3,80	8,85	147,49	7,38	4,25	-4,38	7,38	22,73	8,55	5,31	-6,52	7,73	81,29	9,78	5,47	-7,93	7,98	83,98
D+2	9,10	6,77	-3,91	10,74	147,49	8,27	4,94	-4,85	8,40	22,73	8,70	5,45	-6,35	8,13	81,29	9,77	5,55	-7,87	8,08	83,98
D+3	9,54	6,45	-4,07	10,87	147,49	8,80	5,30	-4,97	9,08	22,73	9,26	5,77	-6,09	9,13	81,29	10,05	5,51	-7,59	8,67	83,98
D+4	9,61	6,06	-4,35	10,60	147,49	9,21	5,36	-5,73	9,08	22,73	9,61	5,60	-6,46	9,14	81,29	10,21	5,33	-7,76	8,59	83,98
D+5	10,48	6,61	-4,04	11,84	147,49	10,02	5,38	-5,42	10,10	22,73	10,36	5,90	-5,79	10,53	81,29	10,97	5,80	-7,03	10,34	83,98
D+6	10,25	6,27	-3,30	11,68	147,49	9,66	5,10	-4,84	9,90	22,73	10,28	5,92	-4,55	11,07	81,29	11,12	5,93	-5,78	11,32	83,98
D+7	9,60	6,56	-2,94	11,36	147,49	9,10	5,06	-4,61	9,44	22,73	10,05	5,84	-3,79	11,11	81,29	10,71	5,97	-5,18	11,23	83,98
Média	9,49		-3,77		147,49	8,92		-4,97		22,73	9,54		-5,65		81,29	10,37		-7,02		83,98
EAM%	ERRO ABSOLUTO MÉDIO PERCENTUAL																			
DPEAM%	DESVIO PADRÃO DO ERRO ABSOLUTO MÉDIO PERCENTUAL																			
EM%	ERRO MÉDIO PERCENTUAL																			
DPEM%	DESVIO PADRÃO DO ERRO MÉDIO PERCENTUAL																			

Observa-se desta tabela que a alternativa 2 trouxe resultados melhores que as outras alternativas, mas ainda dentro de uma faixa de erro não aceitável para efeito de previsões.

A Figura 4.12 apresenta o comparativo das quatro Alternativas empregadas, onde se pode observar esta ligeira vantagem da Alternativa 2 em termos de erros em relação ao Alternativa 1 e uma grande diferença em relação as duas outras alternativas.

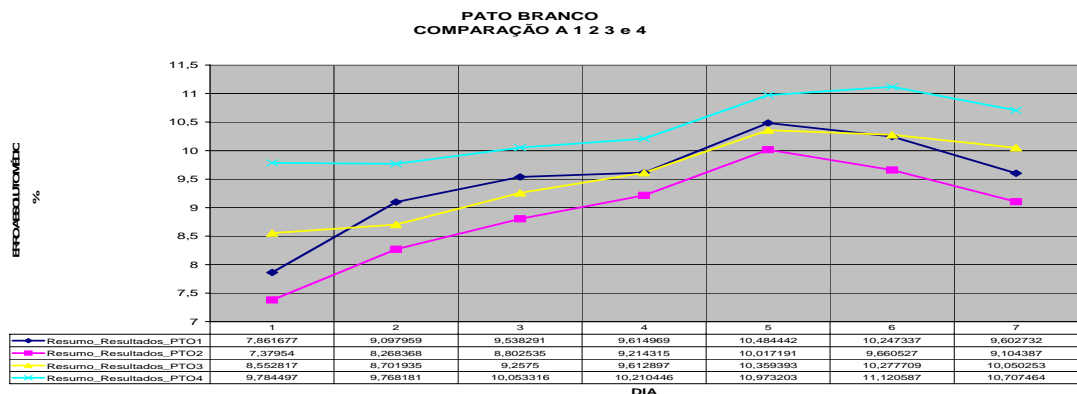


Figura 4.12 – SE PTO Comparativo entre as 4 Alternativas com Finais de Semana.

- **Comparativo Entre as Alternativas 1 e 2, desconsiderando-se nas estatísticas os finais de semana**

Dos resultados acima apresentados pode-se depreender as informações constantes na Tabela 4.30 onde se tem os resumos dos erros percentuais para cada alternativa empregada.

Tabela 4.30 – SE FOZ Comparativo entre as Alternativas 1 e 2 sem Finais de Semana.

PATO BRANCO													
RESULTADOS ALTERNATIVAS 1 2													
ERRO ABSOLUTO MEDIO PERCENTUAL													
DIA	ALTERNATIVA 1						ALTERNATIVA 2						
	EA%	DP	EA%	EM%	DP	E% tempo (s)	EA%	DP	EA%	EM%	DP	E% tempo (s)	
D+1	8,41		5,34	-6,78		7,38	155,7	7,81	4,56	-6,48		6,37	23,3
D+2	9,23		5,39	-7,52		7,67	155,7	8,71	4,96	-7,33		6,91	23,3
D+3	9,86		5,81	-7,37		8,86	155,7	9,42	5,49	-7,34		8,16	23,3
D+4	9,80		5,51	-7,23		8,71	155,7	9,75	5,79	-7,70		8,42	23,3
D+5	10,32		5,50	-7,32		9,24	155,7	10,35	5,13	-8,27		8,17	23,3
D+6	10,14		5,47	-6,54		9,62	155,7	9,84	4,89	-7,71		7,94	23,3
D+7	9,50		5,50	-6,01		9,31	155,7	9,02	5,04	-7,26		7,43	23,3
MÉDIA	9,61			-6,97			147,5	9,27		-7,44			22,7
EAM%	ERRO ABSOLUTO MÉDIO PERCENTUAL												
DPEAM%	DESVIO PADRÃO DO ERRO ABSOLUTO MÉDIO PERCENTUAL												
EM%	ERRO MÉDIO PERCENTUAL												
DPEM%	DESVIO PADRÃO DO ERRO MÉDIO PERCENTUAL												

Observa-se que a alternativa 2 apresenta resultados com menores EAM% e EM%.

A Figura 4.13 apresenta o resultado do EAM% e do EM% para efeito comparação visual das duas alternativas selecionadas.

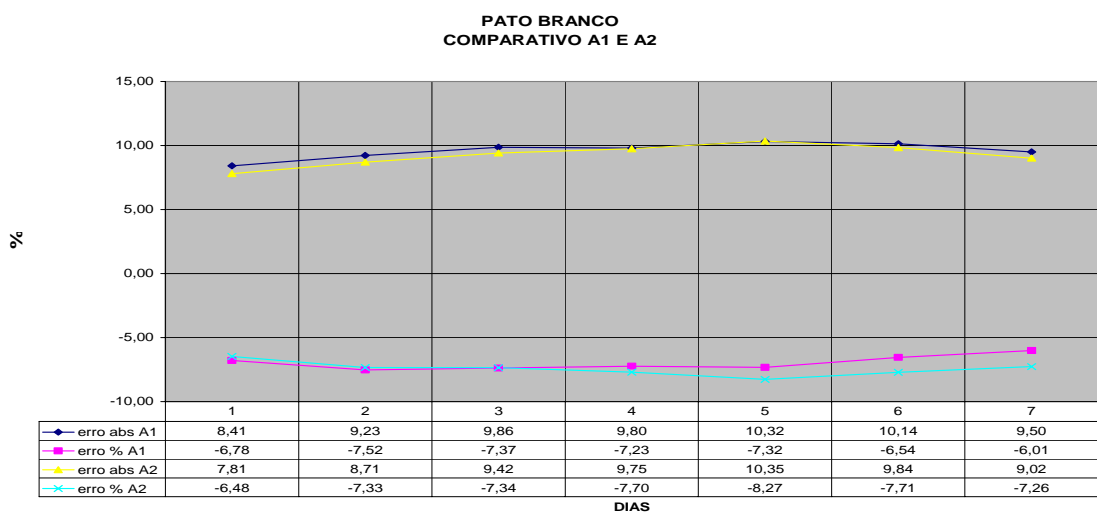


Figura 4.13 – SE PTO Comparativo entre as Alternativas 1 e 2 sem Finais de Semana.

### 4.3.7 SE Umbará

- Comparativo entre as quatro alternativas, considerando-se nas estatísticas os finais de semana.

São apresentados os resultados em termos de erro absoluto percentual, erro percentual, e seus desvios padrões para a SE UMB para as quatro alternativas testadas. Foram realizadas simulações com os dados da SE com previsão do mês de março de 2007 para o histórico de 1º janeiro de 2005 até 28 de fevereiro de 2007, explicitados os resultados obtidos de erros percentual e erro absoluto percentual (Tabela 4.31), a média, desvios padrão, mínimas e máximas verificadas para os dias D+1, ..., D+7.

Tabela 4.31 – SE UMB Comparativo entre as 4 Alternativas com Finais de Semana.

UMBARÁ COMPARATIVO DAS 4 ALTERNATIVAS UTILIZADAS NAS SIMULAÇÕES																				
DIA	ALTERNATIVA 1					ALTERNATIVA 2					ALTERNATIVA 3					ALTERNATIVA 4				
	EAM%	DPEAM%	EM%	DPEM%	tempo(s)	EAM%	DPEAM%	EM%	DPEM%	tempo(s)	EAM%	DPEAM%	EM%	DPEM%	tempo(s)	EAM%	DPEAM%	EM%	DPEM%	tempo(s)
D+1	6,06	4,81	0,63	7,78	184,31	6,30	4,55	0,32	7,84	27,38	11,43	4,41	7,83	9,53	77,87	10,65	4,68	6,45	9,80	112,30
D+2	6,82	5,49	0,23	8,84	184,31	7,54	4,82	-0,37	9,04	27,38	11,27	4,67	7,47	9,75	77,87	10,65	4,95	6,21	10,09	112,30
D+3	7,44	6,67	0,40	10,07	184,31	8,36	5,71	-0,82	10,20	27,38	11,27	4,82	7,58	9,74	77,87	10,64	4,77	6,25	9,96	112,30
D+4	7,91	7,15	0,28	10,75	184,31	8,34	6,70	-1,14	10,74	27,38	11,22	4,72	7,59	9,63	77,87	10,62	4,99	6,34	9,98	112,30
D+5	9,08	7,97	0,31	12,18	184,31	9,11	7,36	-0,92	11,79	27,38	11,56	5,27	7,90	10,06	77,87	11,17	5,62	6,88	10,55	112,30
D+6	9,71	7,89	0,33	12,62	184,31	8,93	7,49	-1,37	11,67	27,38	11,73	5,12	8,36	9,80	77,87	11,23	5,05	7,31	10,02	112,30
D+7	9,23	8,03	0,14	12,34	184,31	8,64	7,97	-2,08	11,66	27,38	11,65	5,43	8,25	9,96	77,87	10,94	4,84	6,94	9,86	112,30
Média	8,04		0,33		184,31	8,18		-0,91		27,38	11,45		7,85		77,87	10,84		6,63		112,30
EAM%	ERRO ABSOLUTO MÉDIO PERCENTUAL																			
DPEAM%	DESVIO PADRÃO DO ERRO ABSOLUTO MÉDIO PERCENTUAL																			
EM%	ERRO MÉDIO PERCENTUAL																			
DPEM%	DESVIO PADRÃO DO ERRO MÉDIO PERCENTUAL																			

Observa-se que a alternativa 1 apresentou resultados melhores de EAM e EM%. A Figura 4.14 apresenta o comparativo das quatro alternativas empregadas, onde se pode observar uma ligeira vantagem da Alternativa 1 em termos de erros em relação a Alternativa 2 e uma grande diferença em relação as duas outras alternativas.

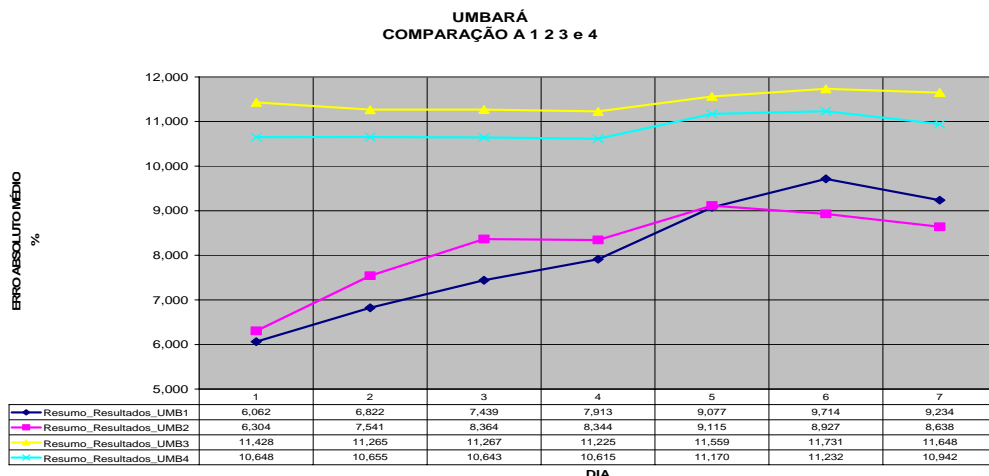


Figura 4.14 – SE UMB Comparativo entre as 4 Alternativas com Finais de Semana.



- **Comparativo Entre as Alternativas 1 e 2, desconsiderando-se nas estatísticas os finais de semana.**

As informações resultaram na Tabela 4.32, onde se tem os resumos dos erros percentuais para cada alternativa empregada.

Tabela 4.32 – SE UMB Comparativo entre as Alternativas 1 e 2 sem Finais de Semana.

UMBARÁ												
RESULTADOS ALTERNATIVAS 1 2												
ERRO												
DIA	ALTERNATIVA 1						ALTERNATIVA 2					
	EAM%	DP	EAM%EM%	DP	EM%	tempo (s)	EAM%	DP	EAM%EM%	DP	EM%	tempo (s)
D+1	6,868	5,340	-0,097	8,816	184,31		6,843	5,024	-0,486	8,594	23,32	
D+2	7,841	5,986	-0,503	9,986	184,31		8,355	5,394	-1,452	9,987	23,32	
D+3	8,599	7,328	0,096	11,439	184,31		9,425	6,332	-1,976	11,345	23,32	
D+4	9,002	7,779	0,251	12,036	184,31		9,299	7,325	-1,735	11,857	23,32	
D+5	9,607	8,521	0,031	12,991	184,31		9,347	7,835	-1,754	12,214	23,32	
D+6	10,263	8,197	-0,019	13,308	184,31		9,015	7,802	-2,107	11,876	23,32	
D+7	9,250	8,248	-0,505	12,532	184,31		8,585	8,254	-3,021	11,642	23,32	
MÉDIA	8,776		-0,107		184,31		8,696		-1,790		23,32	
EAM%	ERRO ABSOLUTO MÉDIO PERCENTUAL											
DPEAM%	DESVIO PADRÃO DO ERRO ABSOLUTO MÉDIO PERCENTUAL											
EM%	ERRO MÉDIO PERCENTUAL											
DPEM%	DESVIO PADRÃO DO ERRO MÉDIO PERCENTUAL											

Observa-se uma pequena vantagem para a alternativa 1 para os dias D+2, D+3 e D+4, e piores resultados para os dias D+1, D+5, D+6, e D+7 de previsão.

A Figura 4.15 apresenta os resultados para os EMA% e EM%. para as duas alternativas selecionadas.

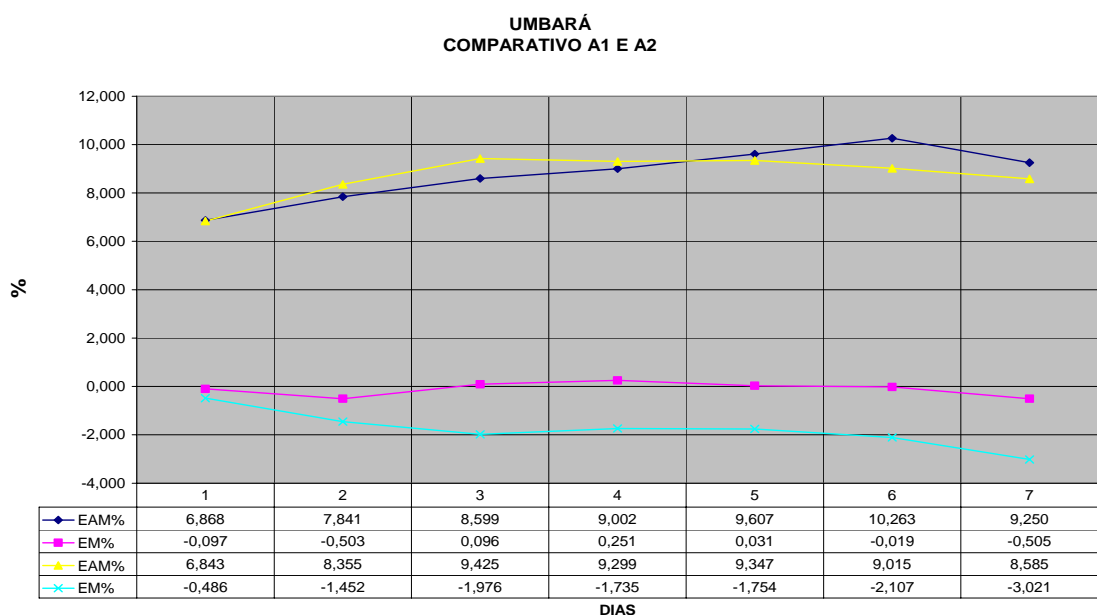


Figura 4.15 – SE UMB Comparativo entre as Alternativas 1 e 2 sem Finais de Semana.

#### 4.3.8 Resumo dos Resultados do SVR para Março/2007

Na Tabela 4.33 é apresentado um resumo das previsões com os valores médios das médias dos erros das quatro alternativas usando SVR para o mês de Março/2007, considerando-se os finais de semana nas estatísticas.

Tabela 4.33 – Resumo das Previsões Usando SVR para Março/2007 com Finais de Semana.

RESUMO COMPARATIVO DAS 4 ALTERNATIVAS UTILIZADAS NAS SIMULAÇÕES												
SE	ALTERNATIVA 1			ALTERNATIVA 2			ALTERNATIVA 3			ALTERNATIVA 4		
	EAM%	EM%	tempo(s)	EAM%	EM%	tempo(s)	EAM%	EM%	tempo(s)	EAM%	EM%	tempo(s)
CAS	4,38	-2,12	166,0	4,92	-2,87	25,0	7,57	-6,59	91,0	8,00	-7,04	96,1
CEL	8,42	4,54	156,7	8,24	4,31	23,2	9,23	5,02	81,2	9,46	4,67	86,4
IBI	6,64	-5,85	157,8	7,68	-7,08	24,9	14,86	-14,86	88,6	14,68	-14,68	99,1
FOZ	18,49	14,93	151,4	19,07	15,86	23,3	26,36	26,30	86,4	25,08	24,53	89,6
LNA	2,67	-0,21	156,9	2,24	-0,22	22,9	3,32	-2,21	77,9	3,32	-2,20	80,0
PTO	9,49	-3,77	147,5	8,92	-4,97	22,7	9,54	-5,65	81,3	10,37	-7,02	84,0
UMB	8,04	0,33	184,3	8,18	-0,91	27,4	11,45	7,85	77,9	10,84	6,63	112,3
Média	8,35	1,25	156,1	8,51	0,84	23,7	11,81	0,34	84,4	11,82	-0,29	89,2
EAM%	ERRO ABSOLUTO MÉDIO PERCENTUAL											
EM%	ERRO MÉDIO PERCENTUAL											

Observa-se que a alternativa 1 apresentou resultados médios superiores às alternativas 3 e 4 s dado ao menor EAM% e com uma pequena diferença quanto ao EAM% quando comparada com a alternativa 2 mas esta a de melhor resultado, com média de 23,7 segundos, quanto ao tempo de processamento.

Na Tabela 4.34 é apresentado um resumo das previsões das alternativas 1 e 2 usando SVR para o mês de Março/2007. Idem anterior, desconsiderando-se os finais de semana nas estatísticas.

Tabela 4.34 – Resumo das Previsões Usando SVR para Março/2007 sem Finais de Semana.

RESUMO COMPARATIVO DAS ALTERNATIVAS 1 E 2 DAS SIMULAÇÕES						
ERRO ABSOLUTO E ABSOLUTO PERCENTUAL						
SE	ALTERNATIVA 1			ALTERNATIVA 2		
	EAM%	EM%	tempo(s)	EAM%	EM%	tempo(s)
CAS	4,35	-1,84	166,0	4,71	-2,44	25,0
CEL	7,49	2,51	156,7	7,35	2,42	23,2
IBI	6,28	-5,82	158,0	8,42	-8,09	24,9
FOZ	10,96	8,93	151,4	12,47	10,46	23,3
LNA	2,53	-0,19	156,9	2,12	0,04	22,9
PTO	9,61	-6,97	147,5	9,27	-7,44	22,7
UMB	8,78	-0,11	184,3	8,70	-1,79	27,4
MÉDIA	7,14	-0,50	160,1	7,58	-0,98	24,2
EAM%	Erro Absoluto percentual					
EM%	Erro percentual %					

Observa-se do comparativo das duas tabelas que a não consideração dos finais de semana nas análises estatísticas trouxe melhores resultados no emprego da técnica para 4 dos

7 resultados, e mesmo para os casos onde não houve melhoria no resultado resultou em pequenas diferenças entre as tabelas 4.33 e 4.34 sob análise, se justificando a eliminação dos finais de semana das estatísticas para os casos em tela.

Na Tabela 4.35 são apresentados os resultados em termos de erro absoluto percentual médio e erro percentual médio, para as quatro alternativas testadas com SVR para todas as SEs verificadas para o dia D+1, ou seja, com base no histórico recente é feita a previsão do dia seguinte, considerando-se nas estatísticas os finais de semana.

Tabela 4.35 – Resumo das Previsões Usando SVR do dia D+1 com Finais de Semana.

RESUMO COMPARATIVO DAS 4 ALTERNATIVAS UTILIZADAS NAS SIMULAÇÕES																
CONSIDERANDO AS PREVISÕES PARA O DIA D+1 (DIA SEGUINTE)																
SE	ALTERNATIVA 1				ALTERNATIVA 2				ALTERNATIVA 3				ALTERNATIVA 4			
	EAM%	EM%	DPEM%	tempo(s)	EAM%	EM%	DPEM%	tempo(s)	EAM%	EM%	DPEM%	tempo(s)	EAM%	EM%	DPEM%	tempo(s)
CAS	3,78	-2,47	3,64	166,0	4,39	-3,20	4,05	25,0	7,49	-7,41	3,29	91,0	7,90	-7,88	3,18	96,1
CEL	6,74	3,70	7,83	156,7	3,58	6,11	5,68	23,2	8,86	5,75	10,22	81,2	8,83	5,32	10,40	86,4
IBI	5,27	-4,62	4,73	157,8	6,31	-5,52	5,44	24,9	14,18	-14,18	3,65	88,6	13,99	-13,99	3,81	99,1
FOZ	15,56	12,46	20,75	151,4	15,22	10,81	19,04	23,3	26,84	26,83	27,79	86,4	25,66	25,09	28,43	89,6
LNA	2,67	0,11	3,59	156,9	2,26	0,17	3,30	22,9	3,54	-2,17	3,72	77,9	3,54	-2,15	3,75	80,0
PTO	7,86	-3,80	8,85	147,5	7,38	-4,38	7,38	22,7	8,55	-6,52	7,73	81,3	9,78	-7,93	7,98	84,0
UMB	6,06	0,63	7,78	184,3	6,30	0,32	7,84	27,4	11,43	7,83	9,53	77,9	10,65	6,45	9,80	112,3
Média	6,85	0,86	160,1	6,49	0,61	24,2	11,56	1,45	83,46	11,48	0,70	92,49				
EAM%	ERRO ABSOLUTO MÉDIO PERCENTUAL															
EM%	ERRO MÉDIO PERCENTUAL															
DPEM%	DESVIO PADRÃO DO ERRO MÉDIO PERCENTUAL															

Observa-se neste caso, considerando-se nas estatísticas os finais de semana, que a alternativa metodológica 2 foi superior à alternativa 1 em 4 das 7 simulações.

Na Tabela 4.36 é apresentado um resumo das previsões das alternativas 1 e 2 usando SVR. Idem anterior, desconsiderando-se nas estatísticas os finais de semana.

Tabela 4.36 – Resumo das Previsões Usando SVR do dia D+1 sem Finais de Semana.

COMPARATIVO DOS ALTERNATIVAS 1 E 2 S/FDS												
CONSIDERANDO-SE AS PREVISÕES PARA O DIA D+1												
SE	ALTERNATIVA 1					ALTERNATIVA 2						
	EAM%	EM%	DPEM%	tempo(s)	EAM%	EM%	DPEM%	tempo(s)	EAM%	EM%	DPEM%	tempo(s)
CAS	3,56	-1,83	3,81	166,0	4,09	-2,42	4,24	25,0				
CEL	5,74	1,70	7,15	156,7	5,30	1,76	6,70	23,2				
IBI	4,49	-4,07	3,80	157,8	6,54	-6,08	5,30	24,9				
FOZ	8,21	5,75	9,31	151,4	9,37	5,91	10,90	23,3				
LNA	2,39	0,32	3,39	156,9	2,05	0,71	3,20	22,9				
PTO	8,41	-6,78	7,38	147,5	7,81	-6,48	6,37	22,7				
UMB	6,87	-0,10	8,81	184,3	6,84	-0,49	8,60	27,4				
Média	5,67	-0,72	160,1	6,00	-1,01	24,2						
EAM%	ERRO ABSOLUTO MÉDIO PERCENTUAL											
EM%	ERRO MÉDIO PERCENTUAL											
DPEM%	DESVIO PADRÃO DO ERRO MÉDIO PERCENTUAL											

Observa-se neste caso um empate entre as duas alternativas metodológicas em termos de EAM% quando desconsidera-se os finais de semana das estatísticas, com resultados

médios melhores para a alternativa 1, mas com tempo computacional que favorece a alternativa 2.

Na Tabela 4.37 é apresentado um comparativo entre as duas situações de análise, considerando-se ou não nas estatísticas os finais de semana para situação das previsões do Dia D+1.

Tabela 4.37 – Resumo Comparativo das Previsões Usando SVR do dia D+1 com e sem Finais de Semana.

RESUMO COMPARATIVO ALTERNATIVAS 1 e 2 COM E SEM FDS							
PREVISÕES DO DIA D+1							
SE	ALTERNATIVA 1			ALTERNATIVA 2			
	EAM%	EM%	tempo(s)	EAM%	EM%	tempo(s)	
CAS	6%	26%	166,0	7%	24%	25,0	
CEL	15%	54%	156,7	-48%	71%	23,2	
IBI	15%	12%	157,8	-4%	-10%	24,9	
FOZ	47%	54%	151,4	38%	45%	23,3	
LNA	11%	-195%	156,9	9%	-318%	22,9	
PTO	-7%	-78%	147,5	-6%	-48%	22,7	
UMB	-13%	116%	184,3	-9%	255%	27,4	
Média	10%	-2%	160,1	-2%	3%	24,2	
EAM%	% MELHORIA SEM FDS DO ERRO ABSOLUTO MÉDIO %						
EM%	% MELHORIA SEM FDS DO ERRO MÉDIO %						

Observa-se uma vantagem significativa na alternativa 1 quando se desconsidera nas estatísticas os finais de semana, com uma melhoria média de 10 % para a EAM% e de uma ligeira piora quando se considera a EM% de (-2 %), dado a grande influência dos EM% das SEs UMB e LNA.

Por outro lado para a alternativa 2 houve uma pequena piora desconsiderando-se os finais de semana das estatísticas (-2%) para o EAM% e uma pequena melhoria para o EM% de 3%.

#### 4.3.9 Previsões Efetuadas para o Mês de Agosto/2006

Para melhor análise do potencial que a técnica da SVR proporciona, bem como para observar a importância do tratamento da base de dados para se efetuar previsões, foram efetuados os mesmos testes até aqui apresentados para todas as SEs, agora para o mês de agosto de 2006, a base de dados é parte daquela até agora utilizada e com os dados já tratados segundo a metodologia apresentada.

Foram tomados para as simulações os dados compreendidos entre 1º de janeiro de 2005 e 31 de julho de 2006, com o intuito de se prever os dados já verificados de agosto de 2006.

Na Tabela 4.38 é apresentado um resumo comparativo entre as duas situações de análise, desconsiderando-se nas estatísticas os finais de semana, para situação das previsões dos Dias D+1, ..., D+7 para todas as SEs analisadas.

Tabela 4.38 – Resumo Comparativo de Todas as SEs - dias D+1, ..., D+7.

RESUMO GERAL PARA TODAS AS SES DESCONSIDERANDO-SE OS FINAIS DE SEMANA						
RESULTADOS ALTERNATIVAS 1 e 2 PREVISAO AGOSTO 2006						
ERRO (%)						
DIA	ALTERNATIVA 1			ALTERNATIVA 2		
	EAM%	EM%	tempo(s)	EA%	EM%	tempo(s)
CAS	4,30	-1,14	82,51	4,19	-2,10	16,47
CEL	6,87	3,51	82,79	7,16	2,81	16,96
IBI	4,39	-1,81	79,83	4,95	-3,16	16,39
FOZ	6,61	4,91	77,11	6,51	5,09	15,85
LNA	3,25	-1,69	77,05	3,75	-2,12	15,62
PTO	12,23	-1,24	77,40	12,42	-0,32	15,80
UMB	6,09	5,78	85,96	7,05	6,83	17,56
<b>MEDIA</b>	<b>6,25</b>	<b>1,19</b>	<b>85,96</b>	<b>6,57</b>	<b>1,01</b>	<b>16,38</b>
EAM%	Erro Absoluto percentual					
EM%	Erro percentual %					

Observa-se que os resultados obtidos com o EAM% apresentam uma vantagem no uso da alternativa 1 em 5 dos 7 casos com 6,25% em média contra 6,57% para a alternativa 2, para o caso de EM% o mesmo ocorre, mas com erro médio de 1,19% para a alternativa 1 contra 1,01 % para a alternativa 2.

Na Tabela 4.39 é apresentado um resumo comparativo para situação das previsões do Dia D+1 para todas as SEs.

Tabela 4.39 – Resumo Comparativo de Todas as SEs - dia D+1.

RESUMO GERAL PARA TODAS AS SES DESCONSIDERANDO-SE OS FINAIS DE SEMANA										
RESULTADOS ALTERNATIVAS 1 e 2 = DIA D +1 = PREVISAO AGOSTO 2006										
ERRO (%)										
DIA	ALTERNATIVA 1 (D+1)					ALTERNATIVA 2 (D+1)				
	EAM%	DP_EAM%	EM%	DP_EM%	tempo(s)	EA%	DP_EAM%	EM%	DP_EM%	tempo(s)
CAS	3,51	2,27	-0,74	4,18	82,51	3,43	2,37	-1,41	3,98	16,47
CEL	4,76	4,06	0,87	6,27	82,79	4,81	4,42	0,15	6,60	16,96
IBI	3,93	2,98	-1,66	4,71	79,83	3,96	3,28	-2,49	4,54	16,39
FOZ	6,16	5,88	3,74	7,72	77,11	6,30	5,63	3,77	7,64	15,85
LNA	3,72	4,09	-2,12	5,15	77,05	4,15	4,35	-2,37	5,57	15,62
PTO	8,88	6,93	-4,53	10,43	77,40	8,76	5,82	-3,62	10,01	15,80
UMB	5,45	4,54	5,15	4,89	85,96	5,55	4,61	5,03	5,19	17,56
<b>MEDIA</b>	<b>5,20</b>		<b>0,10</b>		<b>85,96</b>	<b>5,28</b>		<b>-0,14</b>		<b>16,38</b>
EAM%	Erro Absoluto percentual									
DP_EAM%	Desvio Padrão do Erro Absoluto %									
EM%	Erro percentual %									
DP_EM%	Desvio Padrão do Erro médio%									

Observa-se que os resultados obtidos com o EAM% apresentam uma pequena vantagem no uso da alternativa 1 em 5 dos 7 casos com 5,20% em média contra 5,28% para a alternativa 2, para o caso de EM% o mesmo se verifica com 4 casos onde a alternativa 1 apresentou resultados superiores à alternativa 2. Da mesma forma quando compara-se as tabelas 4.38 com a tabela 4.39 visualiza-se que o expurgo dos finais de semana das estatísticas permite uma melhora significativa dos resultados obtidos com a técnica.

#### **4.3.10 Considerações Sobre o Uso de SVR**

Observou-se que de modo geral existe uma grande aproximação entre os resultados obtidos pela Alternativa 1, quando comparados com a Alternativa 2, podendo-se, a princípio, desconsiderar o uso das Alternativas 3 e 4 que apresentaram desempenho inferior a das duas primeiras. Verificou-se ainda que o Alternativa 1 é ligeiramente superior a Alternativa 2 com relação a precisão das previsões executadas.

No entanto, a Alternativa 2 mostrou uma diferença significativa em termos computacionais em relação à Alternativa 1 onde, os resultados são obtidos em um tempo da ordem de seis vezes maior que o obtido com a Alternativa 2. Isto pode ser representativo, dependendo dos equipamentos disponíveis, da quantidade de subestações para previsão e da disponibilidade de tempo para se efetuar as mesmas. Portanto, a escolha entre as duas alternativas deve ser determinada através da análise da relação: desempenho versus custo computacional com relação a sua aplicação real.

Outra informação é decorrente das comparações dos resultados obtidos quando das análises de previsão para Março de 2007 versus a efetuada para Agosto de 2006, considerando-se que:

- As previsões para Março não sofreram ajustes da base de dados, no período de janeiro de 2007 até Abril de 2007;
- As previsões para Agosto de 2006 aproveitaram os ajustes já efetuados quando da correção da base de dados como apresentado no capítulo 2.

Destas simulações depreende-se que:

- Existe um ganho quando se trabalha adequadamente a base de dados, permitindo inferir que os dados depurados proporcionem, como esperado, uma melhor performance para uso da SVR;
- Um período histórico mais longo tende a trazer resultados de previsões mais confiáveis e aderentes.

Aqui cabe uma consideração sobre os resultados obtidos para as previsões da SE Foz do Chopin, pois segundo levantado junto ao pessoal técnico da Copel, foi observado que esta SE sofre constantes alterações em seus MUSTs devido a existência de uma usina hidrelétrica que injeta potência na barra de 138 kV da SE, usina esta que não é despachada centralizadamente pelo ONS, impactando significativamente nos valores do MUST deste ponto de conexão.

Como os dados de demanda entre janeiro e abril de 2007 não foram depurados, para efeito de se testar a robustez da alternativa, observaram-se resultados que não estão em conformidade com o objetivo do estudo, dada a grande margem de erro verificada para alguns casos.

Portanto, quando constatados estes casos de gerações, remanejamento de cargas, mudanças de configuração que afetem substancialmente os fluxos passantes em pontos de conexão, deve-se previamente fazer as considerações cabíveis, ou na base de dados ou nos resultados das previsões obtidas pela técnica SVR.

#### **4.4 RESULTADOS DAS RNA-RBF**

Para efeito exclusivamente de comparação entre as duas técnicas, SVR e RNA-RBF, efetuou-se estudos de previsão para todas as SEs para o período de Março/2007 e Agosto/2006, para as Alternativas 1 e 2, que foram as que melhor desempenho apresentaram até o momento.

Para o estudo, será apresentada exclusivamente a situação da previsão D+1, para efeito de comparação entre as duas técnicas acima citadas, importante observar que para estes casos

não se consideram nas estatísticas os finais de semana, pois os valores de interesse são os dias úteis.

O número de neurônios foi estabelecido em função dos menores erros obtidos na etapa de treinamento e testes, que também para os casos estudados redundaram em menores erros na fase de previsão.

Para a abertura da Gaussiana também foram efetuadas simulações de sensibilidade para determinação do número ideal, também em função do menor erro obtido nas previsões.

A Tabela 4.40 apresenta os valores de números de neurônios obtidos na camada escondida da RNA para as Alternativas 1 e 2 por SE para as previsões de Março de 2007.

Tabela 4.40 – Número de Neurônios na Camada Escondida – Março/2007.

NÚMERO DE NEURÔNIOS		
SE	ALTERNATIVA 1	ALTERNATIVA 2
CAS	8	4
CEL	7	7
IBI	8	7
FOZ	11	6
LNA	7	5
PTO	13	6
UMB	8	6
MÉDIA	9	6

Observa-se a grande diferença no número de neurônios na camada escondida para ambas as alternativas e entre cada caso em estudo.

A Tabela 4.41 apresenta um resumo comparativo entre as duas situações de análise, Alternativas 1 e 2, desconsiderando-se nas estatísticas os finais de semana, para situação das previsões dos Dias D+1 para todas as SEs analisadas para Março de 2007.



Tabela 4.41 – Comparativo para D+1 com RNA-RBF – Março/2007.

USO DE REDES NEURAIS ARTIFICIAIS										
COMPARATIVO DAS ALTERNATIVAS 1 e 2 PREVISAO MARÇO 2006										
SES PARA - D+1 - DESCONSIDERADOS FERIADOS E FINS DE SEMANA										
ERRO (%)										
SE	ALTERNATIVA 1					ALTERNATIVA 2				
	EAM%	DP_EA%	EM%	DP_EM%	tempo(s)	EAM%	DP_EA%	EM%	DP_EM%	tempo(s)
CAS	3,93	3,06	1,07	4,87	5,98	3,42	2,42	1,01	4,13	8,24
CEL	7,81	6,19	-0,28	9,96	5,64	4,29	3,62	0,79	5,63	8,14
IBI	4,12	3,27	1,28	5,10	8,18	2,57	1,97	1,43	2,94	5,49
FOZ	14,33	19,02	-7,97	22,44	10,16	6,79	6,55	-1,74	9,37	8,47
LNA	3,37	2,24	-0,02	4,05	12,16	2,04	1,79	0,27	2,73	6,41
PTO	7,27	6,12	0,76	9,42	9,54	5,37	4,31	2,54	6,47	6,32
UMB	6,07	4,92	1,78	7,61	17,50	5,58	4,57	0,62	7,28	5,36
MÉDIA	6,70		-0,48		9,88	4,29		0,70		6,92
EAM%	Erro Absoluto percentual									
DP_EA%	Desvio Padrão do Erro Absoluto %									
EM%	Erro percentual %									
DP_EM%	Desvio Padrão do Erro médio%									

Como se pode observar os resultados obtidos permitem inferir que para o caso da utilização de RNA, quando não se considera os finais de semana nas estatísticas, a Alternativa 2 apresentou resultados muito superiores aos obtidos com o uso da Alternativa 1 tanto na média dos resultados obtidos que foi de EAM% de 6,70% para esta alternativa e de 4,29% para a Alternativa 2, o mesmo se dando para o EM%, que conjuntamente com os desvios padrão verificados permitem deduzir que a alternativa 2 é vantajosa em relação a 1.

Da mesma forma para comparação entre as duas técnicas, SVR e RNA, foram efetuados também, estudos de previsão para todas as SEs para o Mês de Agosto de 2006, para as Alternativas 1 e 2, que foram as que melhor desempenho apresentaram até o momento.

Mais uma vez, para o estudo será apresentada exclusivamente a situação da previsão D+1, para efeito de comparação entre as duas técnicas acima citadas, também é importante observar que para estes casos não se consideram nas estatísticas os finais de semana, pois os valores de maior interesse são os dias úteis.

A Tabela 4.42 apresenta os valores de números de neurônios obtidos na camada escondida da RNA para as Alternativas 1 e 2 e por SE para as previsões de Agosto de 2006.

Tabela 4.42 – Número de Neurônios na Camada Escondida – Agosto/2006.

NÚMERO DE NEURÔNIOS (agosto)		
SE	ALTERNATIVA 1	ALTERNATIVA 2
CAS	8	4
CEL	13	4
IBI	8	5
FOZ	14	5
LNA	8	4
PTO	7	5
UMB	8	3
<b>MÉDIA</b>	<b>9</b>	<b>4</b>

Neste caso também a determinação do número de neurônios em cada alternativa é obtido por tentativa, e pode-se observar a variação é grande entre as alternativas e entre casos da mesma alternativa.

A Tabela 4.43 mostra o resumo comparativo entre as duas situações de análise, Alternativas 1 e 2, desconsiderando-se nas estatísticas os finais de semana, para situação das previsões dos Dias D+1 para todas as Subestações até aqui analisadas para Agosto de 2006.

Tabela 4.43 – Comparativo para D+1 com RNA-RBF – Agosto/2006.

USO DE REDES NEURAIS ARTIFICIAIS											
COMPARATIVO DAS ALTERNATIVAS 1 e 2 PREVISAO AGOSTO 2006											
SES PARA - D+1 - DESCONSIDERADOS FERIADOS E FINS DE SEMANA											
ERRO (%)											
SE	ALTERNATIVA 1					ALTERNATIVA 2					
	EAM%	DP_EA%	EM%	DP_EM%	tempo(s)	EAM%	DP_EA%	EM%	DP_EM%	tempo(s)	
CAS	5,05	3,26	1,70	5,82	2,66	3,13	2,41	2,25	3,27	3,27	
CEL	6,03	4,99	-2,12	7,60	7,38	3,73	3,33	-0,80	4,98	3,80	
IBI	5,32	3,88	1,38	6,50	2,79	2,78	3,34	-0,77	4,30	9,14	
FOZ	5,99	5,44	-1,06	8,09	2,77	4,49	4,71	-1,57	6,36	2,56	
LNA	2,98	2,66	0,57	3,98	2,53	1,83	2,31	0,26	2,95	3,44	
PTO	9,48	7,93	0,95	12,44	3,07	6,76	5,77	2,99	8,43	3,24	
UMB	3,34	2,50	-1,56	3,90	2,82	2,29	1,46	-1,49	2,29	2,93	
<b>MÉDIA</b>	<b>5,45</b>		<b>-0,02</b>		<b>3,43</b>	<b>3,57</b>		<b>0,12</b>		<b>4,05</b>	
EAM%	Erro Absoluto percentual										
DP_EA%	Desvio Padrão do Erro Absoluto %										
EM%	Erro percentual %										
DP_EM%	Desvio Padrão do Erro médio%										

Como observado a alternativa 2 apresentou resultados superiores com erros EAM% menores que os observados para a alternativa 1 e tempo de processamentos muito próximos entre si. No entanto a alternativa 1 apresentou resultados algo superior de EM% (-0,02%) comparativa a alternativa 2 (0,12%), mas que pelos Desvios padrões apresentados não permite a mesma conclusão trazendo novamente vantagens significativas para a alternativa 2.

## 4.5 COMPARAÇÃO SVR VERSUS RNA-RBF

A comparação entre as duas técnicas foi efetuada exclusivamente para a situação da previsão do dia D+1, ou seja, o dia seguinte. Foram realizadas simulações para as previsões dos meses de agosto de 2006 e março de 2007, pois para estes casos é possível a comparação entre todos os dados estatísticos apresentados neste trabalho, além de que a previsão de maior interesse para o estudo é a do dia D+1.

Também, optou-se por efetuar a comparação entre as técnicas para as previsões do mês de março de 2007, dado que entende-se que as previsões para este mês são mais críticas uma vez que:

- A base de dados entre o período de 1º de janeiro de 2007 até dia 10 de abril não sofreu tratamento;
- Os dados apresentam remanejamentos e *outliers* não expurgados entre 1º de janeiro a 28 de fevereiro de 2007.

A Tabela 4.44 apresenta a comparação das duas técnicas para ambas as alternativas quando das previsões para Março de 2007, para o dia D+1, portanto para o dia seguinte, desconsiderando-se nas estatísticas a influência dos finais de semana.

Tabela 4.44 – Comparativo das Alternativas 1 e 2 com SVR e RNA-RBF – Março/2007.

USO DE RNA X SVR																				
COMPARATIVO DAS ALTERNATIVAS 1 e 2 PREVISAO MARÇO 2006																				
SES PARA - D+1 - DESCONSIDERADOS FERIADOS E FINS DE SEMANA																				
ERRO (%)																				
	ALTERNATIVA 1									ALTERNATIVA 2										
	RNA					SVR				RNA					SVR					
SE	EAM%	EM%	DP	EM%	tempo(s)	EAM%	EM%	DP	EM%	tempo(s)*	EAM%	EM%	DP	EM%	tempo(s)	EAM%	EM%	DP	EM%	tempo(s)*
CAS	3,93	1,07	3,06		6,0	3,56	-1,83	3,81		166,0	3,42	1,01	4,13		8,2	4,09	-2,42	4,24		25,0
CEL	7,81	-0,28	6,19		5,6	5,74	1,70	7,15		156,7	4,29	0,79	5,63		8,1	5,30	1,76	6,70		23,2
IBI	4,12	1,28	3,27		8,2	4,49	-4,07	3,80		157,8	2,57	1,43	2,94		5,5	6,54	-6,08	5,30		24,9
FOZ	14,33	-7,97	19,02		10,2	8,21	5,75	9,31		151,4	6,79	-1,74	9,37		8,5	9,37	5,91	10,90		23,3
LNA	3,37	-0,02	2,24		12,2	2,39	0,32	3,39		156,9	2,04	0,27	2,73		6,4	2,05	0,71	3,20		22,9
PTO	7,27	0,76	6,12		9,5	8,41	-6,78	7,38		147,5	5,37	2,54	6,47		6,3	7,81	-6,48	6,37		22,7
UMB	6,07	1,78	4,92		17,5	6,87	-0,10	8,81		184,3	5,58	0,62	7,28		5,4	6,84	-0,49	8,60		27,4
MÉDIA	6,70	-0,48			9,9	5,67	-0,72			160,1	4,29	0,70			6,9	6,00	-1,01			24,2
EAM%	Erro Absoluto percentual																			
DP_EA%	Desvio Padrão do Erro Absoluto %																			
EM%	Erro percentual %																			
DP_EM%	Desvio Padrão do Erro médio%																			
*Deve-se observar que no tempo computado na SVR é efetuada a previsão dos Dias D+1,....,D+7, na RNA só para o dia D+1																				

Observa-se que para a Alternativa 1 a técnica usando SVR apresentou um resultado superior ao obtido com RNA com um EAM% de 5,67 % superior aos 6,70% obtido com RNA, ocorrendo praticamente um empate entre a EM% de ambos para a Alternativa 1, no entanto com uma vantagem para o desvio padrão obtido para a RNA.

Por outro lado observa-se uma vantagem para a RNA, quando se tem respectivamente, para a Alternativa 2 um EAM% de 4,29% versus 6,00% para a SVR, também com uma certa vantagem para a RNA quando compara-se a EM% de 0,70% contra -1,01% para a SVR, com maiores valores de desvios padrão para a SVR.

Também neste caso, comparando-se as técnicas para as Alternativas 1 e 2 versus SVR e RNA, ainda existe vantagem para a RNA de 4,29% contra 6,00% para a SVR.

É importante citar que, o tempo constante nas Tabelas quando das apresentações dos resultados da RNA devem ser multiplicados por 5 a 9 vezes, dado ao fato que são estes os números de simulações necessárias para se obter o número ótimo de neurônios na camada escondida mais 2 a 3 simulações adicionais para se obter a melhor abertura da gaussiana, que aqui (depois de testes de sensibilidade) se adotou como referência o valor 10.

Já a Tabela 4.45 apresenta a comparação das duas técnicas para ambas as alternativas quando das previsões para Agosto de 2006, para o dia D+1, portanto para o dia seguinte, desconsiderando-se nas estatísticas a influência dos finais de semana.

Tabela 4.45 – Comparativo das Alternativas 1 e 2 com SVR e RNA-RBF – Agosto/2006.

USO DE RNA X SVR																
COMPARATIVO DAS ALTERNATIVAS 1 e 2 PREVISAO AGOSTO 2006																
SES PARA - D+1 - DESCONSIDERADOS FERIADOS E FINS DE SEMANA																
ERRO (%)																
SE	ALTERNATIVA 1								ALTERNATIVA 2							
	RNA				SVR				RNA				SVR			
	EAM%	EM%	DP_EM%	tempo(s)	EAM%	EM%	DP_EM%	tempo(s)*	EAM%	EM%	DP_EM%	tempo(s)	EAM%	EM%	DP_EM%	tempo(s)*
CAS	5,05	1,70	5,82	2,7	3,51	-0,74	4,18	82,5	3,13	2,25	3,27	3,3	3,43	-1,41	3,98	16,5
CEL	6,03	-2,12	7,60	7,4	4,76	0,87	6,27	82,8	3,73	-0,80	4,98	3,8	4,81	0,15	6,60	17,0
IBI	5,32	1,38	6,50	2,8	3,93	-1,66	4,71	79,8	2,78	-0,77	4,30	9,1	3,96	-2,49	4,54	16,4
FOZ	5,99	-1,06	8,09	2,8	6,16	3,74	7,72	77,1	4,49	-1,57	6,36	2,6	6,30	3,77	7,64	16,4
LNA	2,98	0,57	3,98	2,5	3,72	-2,12	5,15	77,1	1,83	0,26	2,95	3,4	8,76	-2,37	5,57	15,6
PTO	9,48	0,95	12,44	3,1	8,88	-4,53	10,43	77,4	6,76	2,99	8,43	3,2	5,37	3,62	10,01	15,8
UMB	3,34	-1,56	3,90	2,8	5,45	5,15	4,89	86,0	2,29	-1,49	2,29	2,9	5,55	5,03	5,19	17,6
<b>MÉDIA</b>	<b>5,45</b>	<b>-0,02</b>	<b>3,4</b>	<b>3,4</b>	<b>5,20</b>	<b>0,10</b>	<b>80,4</b>	<b>80,4</b>	<b>3,57</b>	<b>0,12</b>	<b>4,1</b>	<b>4,1</b>	<b>5,45</b>	<b>-0,14</b>	<b>16,5</b>	<b>16,5</b>
EAM% Erro Absoluto percentual EM% Erro percentual % DP_EM% Desvio Padrão do Erro médio%																
*Deve-se observar que no tempo computado na SVR é efetuada a previsão dos Dias D+1.....D+7, na RNA só para o dia D+1																

Do mesmo modo pode-se constatar que o uso de SVR para a Alternativa 1 apresentou resultado superior ao obtido com RNA-RBF, onde a técnica de SVR apresentou um EAM% de 5,20% contra 5,45% da RNA, mesmo o EM% da SVR resultou em erro menor do que o resultado obtido com RNA, os desvios padrões para a SVR variaram entre 4,18 e 10,43% entre os resultados obtidos e os valores reais, contra uma variação para a RNA compreendido entre 3,98 e 12,44% para o EM%.

No entanto para o caso da Alternativa 2 temos a inversão dos resultados onde a RNA-RBF apresentou um resultado de EAM% de 3,57% contra 5,45% para a SVR. Para o resultado apresentado do EM% a SVR aparentemente traz uma média melhor do que o apresentado para a RNA, mas quando observado individualmente para cada SE, os erros encontrados de EM% para a RNA apresentam variações (entre -1,57 e 2,99%), que são muito pequenos quando comparados com os resultados obtidos com a SVR (entre -2,49 e 5,03%). O mesmo se dando com os desvios padrão para o EM%. No entanto, para a Alternativa 2 o tempo observado da simulação usando SVR apesar de ser da ordem de 16,5 segundos por simulação é inferior ao tempo de 6,9 segundos médio obtido para a RNA, pois ao tempo computado para a RNA deve-se multiplicar por 5 a 9 vezes dado a necessidade de determinação dos números de neurônios necessários para obtenção dos menores erros médios percentuais, o que não é necessário para a SVR que já computa o tempo total da simulação.

Portanto para efeito de simplificação de comparação pode-se concluir que os tempos computacionais para a escolha de SVR ou RNA não devem ser considerado e pode-se ignorar tal informação, dado que o tempo para simulação na RNA considerou apenas a projeção de um dia a frente e na SVR os resultados contemplam sete dias a frente. Outro fato constatado é de que quando se utiliza RNA é muito importante a observação dos resultados em loco, caso a caso, pois pode ocorrer *overfitting*, o que só pode ser notado quando da observação dos resultados.

## **5 CONCLUSÕES E DESENVOLVIMENTOS FUTUROS**

### **5.1 CONCLUSÕES FINAIS**

Esta dissertação está focada na necessidade que tem o agente distribuidor de energia elétrica em prever os MUSTs em cada ponto de conexão com a rede básica, para um período de uma semana a frente, com exatidão e tempo hábil suficientes. Estipulou-se que as projeções devem estar disponíveis antes da chegada dos técnicos no horário de expediente da manhã, para permitir que possam ser tomadas medidas operativas para evitar ultrapassagens dos limites dos MUSTs contratados. Dessa forma, busca-se reduzir os impactos econômicos para a distribuidora pelo não pagamento de penalidades por ultrapassagens dos valores contratados.

É importante salientar que mudanças de configuração, remanejamento de cargas, influência de feriados, intercâmbio de energia entre subsistemas, sazonalidades, falha no recebimento das informações, etc., afetam significativamente os resultados das previsões.

Foram apresentadas duas técnicas para previsão, SVR e RNA-RBF, sendo que nas simulações que foram efetuadas com SVR foram adotadas quatro alternativas de modelagem de tratamento da base de dados onde foi computado o tempo de processamento para as estimativas para projeções de sete dias à frente, salientando-se as previsões para um dia à frente (D+1) dado ao interesse específico. E ainda, foram apresentadas Tabelas comparativas com os resultados dos testes para os quatro modelos analisados.

Para as simulações com RNA-RBF foram implementadas as Alternativas 1 e 2 de modelos de tratamento da base de dados para as projeções, porém com previsões para somente um dia à frente (D+1) para efeito comparativo com a técnica SVR.

#### **5.1.1 Uso da SVR para Previsões do MUST**

Para efeito de apuração dos resultados, adotou-se o erro médio percentual e o erro médio absoluto percentual e seus respectivos desvios padrão.

Da análise dos resultados observou-se de modo geral uma grande aproximação entre os resultados obtidos, empregando-se as técnicas de SVR, pela alternativa de modelo 1, quando comparado com a alternativa de modelo 2, podendo-se a princípio desconsiderar o uso das alternativas 3 e 4 que demonstraram um desempenho inferior às outras duas primeiras.

É importante ressaltar que a Alternativa 2 demonstrou uma diferença significativa em termos de tempo de processamento para obtenção dos resultados, da ordem de seis a sete vezes mais rápido do que a Alternativa 1, o que em termos computacionais pode ser representativo, dependendo das características do *hardware* que vier a ser utilizado para as previsões.

Por outro lado, o aspecto da facilidade de manuseio e o tempo necessário para a preparação dos parâmetros para o uso da técnica, também foi considerado, como também outro aspecto que se procurou buscar foi a robustez do modelo para a utilização num caso real com e sem tratamento de parte da base de dados.

Os objetivos de estudar e analisar os resultados de simulação, tanto das fases de estimação quanto validação do modelo da SVR, alcançaram resultados promissores para previsão de séries temporais. Como testes, utilizando-se a SVR, foram efetuadas simulações de previsões para os meses de Agosto de 2006 e Março de 2007 (neste caso parte dos dados não sofreram tratamento).

Cabe citar que a análise do erro absoluto médio percentual (EAM%) é importante para se ter parâmetros de comparação entre as técnicas, também o erro médio percentual (EM%) juntamente com o desvio padrão permitem analisar se os erros das projeções são para maior ou menor.

Valores sobre-estimados nas projeções são menos danosos para a tomada de decisão do que quando se tem projeções com valores a menor (subestimados), pois nestes casos valores estimados a maior impactam em menores riscos para as distribuidoras que pode tomar medidas operativas (quando possível) para evitar ultrapassagens nos valores contratados, o que não ocorre quando as estimativas são favoráveis (previsão otimista quando não indica ultrapassagens).

### 5.1.2 Uso de RNA-RBF para Previsões do MUST.

Para validação dos resultados obtidos com a SVR foi implementada a técnica de RNA-RBF e algoritmos Kernel para previsões de um dia à frente (D+1), donde obtiveram-se resultados que demonstraram que se, por escolha, for adotada a técnica de SVR pode-se usar tanto a Alternativa 1 quanto a Alternativa 2, no entanto a Alternativa 1 foi a que melhores resultados apresentou, inclusive competitivamente similar aos resultados obtidos através de RNA-RBF para a Alternativa 2, salvo o tempo computacional que foi muito superior com o uso de SVR.

Por outro lado o uso de RNA apresentou o inconveniente da necessidade de se efetuar, para cada base de dados e cada SE, simulações com diversas quantidades de neurônios da camada escondida que devem ser escolhidos caso a caso, portanto que mudam em cada teste efetuado até a escolha do melhor número de neurônios na camada escondida que proporcione o menor erro médio percentual e erro médio absoluto percentual, atentando-se para o fato que deve-se cuidar para evitar os casos de *overfitting*.

É importante citar que, o tempo constante nas Tabelas quando das apresentações dos resultados da RNA devem ser multiplicados por 5 a 9 vezes, dado ao fato que são estes os números de simulações necessárias para se obter o número ótimo de neurônios na camada escondida mais 2 a 3 simulações adicionais para se obter a melhor abertura da gaussiana.

### 5.1.3 Análise Comparativa SVR versus RNA-RBF.

Para aplicação no problema real da concessionária observou-se que o uso de SVR é mais adequado visto a facilidade de ajuste, implementação e interpretação dos resultados obtidos, pois as RNAs-RBF exigem do analista um conhecimento mais aprofundado da técnica e experiência com a série de dados em questão. Isto sem considerar o tempo necessário para o treinamento do usuário no emprego da técnica de RNA-RBF.



## 5.2 SUGESTÃO PARA TRABALHOS FUTUROS

Como sugestão para trabalhos futuros podem-se implementar outras técnicas para testar o desempenho e velocidade da previsão das demandas nos pontos de conexão, tais como Borland Multi Layer Perceptron - BMLP (FERREIRA, 2007) e *Least Squares Support Vector Machines* - LS-SVM (CARVALHO E BRAGA, 2005), haja visto ser de interesse a adoção da técnica que melhores resultados produzam quando de aplicação nos casos reais da distribuidora.

Outro ponto importante para elaboração e análise em trabalho futuro pode ser a inserção das variáveis climáticas no modelo, uma vez que são informações que impactam nos resultados da carga das SEs, mas que não foram contempladas nos estudos dada a dificuldade na obtenção da base de dados.

## REFERÊNCIAS BIBLIOGRÁFICAS

ABU-EL-MAGD, M.A.; FINDLAY, R.D. (2003). A new approach using artificial neural network and time series models for short term load forecasting. **Canadian Conference on Electrical and Computer Engineering**, pp. 1723 – 1726, May 2003.

AGÊNCIA NACIONAL DE ENERGIA ELÉTRICA – ANEEL. (1999). **Resolução Normativa ANEEL Nº 281**, 1999.

AGÊNCIA NACIONAL DE ENERGIA ELÉTRICA – ANEEL. (2002a). **Resolução Normativa ANEEL Nº 358**, 2002.

AGÊNCIA NACIONAL DE ENERGIA ELÉTRICA – ANEEL (2002b). **Resolução Normativa ANEEL Nº 655**, 2002.

AGUIRRE, L.A. (2004). **Introdução à Identificação de Sistemas - Técnicas Lineares e Não-Lineares Aplicadas a Sistemas Reais**. Ed. UFMG, Belo Horizonte, 2º Ed., 2004.

AL-FUHAID, A.S.; EL-SAYED, M.A.; MAHMOUD, M.S. (1997). Cascaded Artificial Neural Networks For Short-Term Load Forecasting. **IEEE Transactions on Power Systems**, Vol. 12, No. 4, Nov. 1997, pp 1524-1529.

AL MAMUN, M.; NAGASAKA, K. (2004). Artificial Neural Networks Applied to Long-term Electricity Demand Forecasting. In: **Proceedings of the Fourth International Conference on Hybrid Intelligent Systems**, pp. 204-209, Dec. 2004.

AMJADY, N. (2001). Short-term hourly load forecasting using time-series modeling with peak load estimation capability, **IEEE Transactions on Power Systems**, vol. 16, no. 3, 2001. pp. 498-505.

AOKI, A.R. (2005). **Notas de Aula da Disciplina Inteligência Artificial**, Universidade Federal do Paraná, Curitiba, PR, Curso Inteligência Artificial Aplicada, 2005.

ARAÚJO JUNIOR, A.L.; VASCONCELOS, G.C. (1999). Previsão de Demanda Máxima Mensal Baseada Em Redes Neurais. In: **XV seminário Nacional de Produção e Transmissão de Energia Elétrica**, Outubro 1999.

BARAKAT, E.H.; AL-QASEM, J.M. (1998). Methodology for weekly load forecasting, **IEEE Transactions on Power Systems**. vol. 13, no. 4, 1998. pp. 1548-1555.

BEÊ, R.T. (2007). **Alocação de Bancos de Capacitores em Sistema de Distribuição de Energia Elétrica Utilizando Algoritmos Genéticos**. Dissertação de Mestrado em Engenharia Elétrica, Universidade Federal do Paraná, Curitiba, PR, 2007.

BILINGS, S.A.; ZHENG, G.L. (1995). Radial Basis Function Network Configuration Using Genetic Algorithms. Pergamon, **Neural Networks**, vol. 8, no. 6, 1995, p. 877-890.

BOOR, C. (1978), A Practical Guide to Splines, New York: Springer-Verlag Press. 1978.

BORS, A.G. (2001) Introduction of the Radial Basis Function (RBF) Networks. Online Symposium for Electronics Engineers, DSP Algorithms: Multimedia, <http://www.osee.net/>, 13/05/2004, vol. 1, 2001, p. 1-7.

BOX, G.; JENKINS, G.; REINSEL, G. (1994) Time Series Analysis Forecasting and Control. Ed. Prentice Hall, 3rd ed., Englewood Cliffs, New Jersey, USA, 1994.

BRAGA, A.P.; CARVALHO, A.C.P.L.F.; LURDEMIR, T.B. (2000). Redes Neurais Artificiais, **Teoria e Aplicações**. Rio de Janeiro, Ed. LTC, 2000

CARPINTEIRO, O.A.S.; DA SILVA; A.P.A.; REIS, A.J.R. (2001). Short Term Load Forecasting via a Hierarchical Neural Model. In: **Proceedings of the V Brazilian Conference on Neural Networks**, pp. 55–59, Rio de Janeiro, Brasil, Abril 2001.

CALMARZA, A.M.; FUENTE, J.I. de la. (2000). New forecasting method for the residual demand curves using time series (ARIMA) models, PMAPPS.

CARVALHO, B.P.R.; BRAGA, A.P. (2005) Novas Estratégias para treinamento de Least Squares Support Vector Machines. In: **XXV Congresso da Sociedade Brasileira de Computação**.

CASTRO, M.C.F. (2001) Predição Não-Linear de Séries Temporais Usando Redes Neurais RBF por Decomposição em Componentes Principais. Tese de Doutorado, UNICAMP, Dept. de Engenharia Elétrica, São Paulo, 2001.

CHEN, B.-J.; CHANG, M.-W.; LIN, C.-J. (2001). Load Forecasting Using Support Vector Machines: A Study on EUNITE Competition 2001, **IEEE Transactions on Power Systems**, Vol. 19, No. 4, Nov. 2004, pp. 1821-1830.

CHEN, B.-J.; CHANG, M.-W.; LIN, C.-J. (2001). LIBSVM a library for support vector machines, 2001, software available at [url{http://www.csie.ntu.edu.tw/~cjlin/libsvm}](http://www.csie.ntu.edu.tw/~cjlin/libsvm).

CHEN, G.J.; LI, K.K.; CHUNG, T.S.; SUN, H.B.; TANG, G.Q. (2001). Application of an innovative combined forecasting method in power system load forecasting, **Electric Power systems Research**, vol. 59, 2001. pp. 131-137.

CHERKASSKY, V.; MULIER, F. (1998). Learning from Data: Concepts, Theory and Methods, John Wiley & Sons, New York, USA, 1998.

DOUGLAS, A.P.; BREIPAHL, A.M.; LEE, F.N.; ADAPA, R. (1998). Risk due to load forecast uncertainty in short term power system planning, **IEEE Transactions on Power Systems**, vol. 13, no. 4, 1998. pp. 1493-1499.

ELMAN, J.L. (1998). Finding Structure in Time, CRL Tech. Report 8801, University of California at San Diego, **Center for Research in Language, also in Cognitive Science**, vol. 14, pp. 179-211, 1998.

FAYYAD, U.; PIATETSKY-SHAPIRO, G.; SMYTH, P. (1996) From data mining to knowledge discovery: an overview. *Advances in Knowledge Discovery & Data Mining*, 1996. pp. 1-34.

FERREIRA, V.H. Desenvolvimento de Modelos Neurais Autônomos para Previsão de Carga Elétrica, Documento de Qualificação da Tese de Doutorado, RJ, 2007.

GALARNIOTIS, A.I.; TSAKOUMIS, A.C.; FESSAS, P.; VLADOV, S.S.; MLADENOV, V.M. (2003). Using Elman and FIR neural networks for short term electric load forecasting, **International Symposium on Signals, Circuits and Systems**, p. 433 - 436, Jul. 2003.

GARCIA, R.C.; CONTRERAS, J.; VAN AKKEREN, M.; GARCIA, J.B.C. (2003). A GARCH forecasting model to predict, **Workshop of Applied Infrastructure**, Berlin, Germany, 2003.

GONTAR, Z.; HATZIARGYRIOU, N. (2001). Short term load forecasting with radial basis function network, **IEEE Porto Power Tech Conference**, Porto, Portugal, 2001.

GUERRA, F.A. (2006). Análise de Métodos de Agrupamento para o Treinamento de Redes Neurais de Base Radial aplicadas à Identificação de Sistemas, Curitiba, Dissertação de Mestrado, 2006.

HAYKIN, S. (1998). *Neural Networks: a Comprehensive Foundation*, Prentice Hall PTR **Upper Saddle River**, NJ, USA, 1998.

HAYKIN, S. (2001). *Redes Neurais: Princípios e Prática*. 2ª Edição, Ed. Bookman, Porto Alegre, RS, 2001.

HERNANDEZ, E.D.M. (2003). Inteligência Computacional e Redes Neurais em Engenharia Elétrica. PSI2222 **Práticas de Eletricidade e Eletrônica II**, Seminário 6, 2003.

HIPPERT, H.S.; BUNN, D.W.; SOUZA, R.C. (2001) Previsão de Cargas Elétricas por Redes Neurais: Uma Investigação Empírica do Problema de Overfitting”, Proceedings of the **V Brazilian Conference on Neural Networks**, pp. 55–59, Rio de Janeiro, Brasil, Abril 2001.

HSU, C.-W.; CHANG, M.-W.; LIN, C.-J. (2007). A Practical Guide to Support Vector Classification, **Department of Computer Science National Taiwan University**, Taipei 106, Taiwan 18, Jul. 2007.

HUANG, C.-M.; YANG, H.-T. (2001). Evolving wavelet-based networks for short-term load forecasting, **IEE Proceedings Generation Transmission and Distribution**, vol. 148, no. 3, 2001. pp. 222-228.

KANDIL, M.S.; EL-DEBEIKY, S.M.; HASANIEN, N.E. (2002). Long-term load forecasting for fast developing utility using a knowledge-based expert system, **IEEE Transactions on Power Systems**, vol. 17, no. 2, 2002. pp. 491-496.

KHOTANZAD, A.; AFKHAMI-ROHANI, R.; MARATUKULAM, D. ANNSTLF. (1998). Artificial Neural Network Short-Term Load Forecaster -Generation Three. **IEEE Transactions on Power Systems**, Vol. 13, No. 4, November 1998, pp. 1413-1422.

KIM, K.-H.; YOUN, H.-S.; KANG, Y.-C. (2000). Short-Term Load Forecasting for Special Days in Anomalous Load Conditions Using Neural Networks and Fuzzy Inference Method, **IEEE Transactions on Power Systems**, Vol. 15, N<sup>o</sup>. 2, Maio 2000, pp. 559-565.

KODOGIANNIS, V.S.; ANAGNOSTAKIS, E.M. (2002). **Soft computing based techniques for short-term load forecasting**, **Fuzzy Sets and Systems**, vol. 128, 2002. pp. 413-426.

LAGE, W.F.; DA SILVA, A.P.A.; RODRIGUES, U.P.; MOREIRA, G.; MARRA, W.; DE ANDRADE, W.A. (2003). Previsor: Sistema de Previsão de Carga On-Line. In: **Anais V Simpósio de Automação de Sistemas Elétricos**, Recife, Brasil, 2003.

LAVRAČ, N.; FLACH, P.; ZUPAN, R. (1999). Rule evaluation measures: a unifying view. Proceedings of the **Ninth International Workshop on Inductive Logic Programming**, Volume 1634, pp. 74-185. Springer-Verlag, 1999. Lectures Notes in Artificial Intelligence.

LIANG, R.-H.; CHENG, C.-C. (2000). Combined regression-fuzzy approach for short-term load forecasting, **IEE Proceedings Generation Transmission and Distribution**, vol. 147, no. 4, 2000. pp. 261-266.

LOTUFO, A.D.P.; MINUSSI, C.R. (1999). Electric power systems load forecasting: a survey, **IEEE Power Tech Conference**, Budapest, 1999. Hungary.

MICCHELLI, C.A. (1986). **Interpolation of scattered data: Distance matrices and conditionally positive definite functions**. *Constructive Approximation*, vol. 2, 1986,

MOGHRAM, I.; RAHMAN, S. (1989). Analysis and evaluation of five short-term load forecasting techniques. **IEEE Transactions on Power Systems**, vol. 4, no. 4, pp. 1484-1491.

MORI, H.; YUIHARA, A. (2001). Deterministic Annealing Clustering for ANN-Based Short-Term Load Forecasting, **IEEE Transactions on Power Systems**, Vol. 16, Nº. 3, Ago. 2001, pp. 545-551.

NAGASAKA, K.; AL MAMUN, M. (2004). Long-term peak demand prediction of 9 Japanese power utilities using radial basis function networks, **IEEE Power Engineering Society General Meeting**, pp. 315-322, vol. 1, Jun. 2004.

NOGALES, F.J.; CONTRERAS, J.; CONEJO, A.J.; ESPÍNOLA, R. (2002). Forecasting next-day electricity prices by time series models, **IEEE Transactions on Power Systems**, vol. 17, no. 2, 2002. pp. 342-348.

OONSIVILAI, A.; EL-HAWARY, M.E. (1999). Wavelet neural network based short term load forecasting of electric power system commercial load, **Proceedings of the IEEE Canadian Conference on Electrical and Computer Engineering**, Edmonton, Alberta, Canada, 1999. pp. 1223-1228.

PAI, P.-F.; HONG, W.-C. (2005). Forecasting regional electricity load based on recurrent support vector machines with genetic algorithms, **Electric Power Systems Research**, vol. 74, 2005. pp. 417-425.

PAIXAO, L.E. (2000) **Memórias do Projeto - RESEB – Reestruturação do Setor Elétrico Brasileiro – A História da Nova Ordem Institucional do SEB**. 2000

PALACIO, D.R. Simple model for load forecast weather sensitive. In: **16th International Conference and Exhibition on Electricity Distribution**. Part 1: Contributions, Amsterdam, Jun. 2001.

PENG, T.M.; HUBELE, N.F.; KARADY, G.G. (1992). Advancement in the Application of Neural Networks for Short-Term Load Forecasting. **IEEE Transactions on Power Systems**, Vol. 7, No. 1, Fev. 1992, pp. 250-257.

RASHIDI, M.; RASHIDI, F.M.R.; MONAVAR, H. (2003). Peak load forecasting in power systems using emotional learning based fuzzy logic, **IEEE International Conference on Systems, Man and Cybernetics**, vol. 2, Washington, DC, USA, 2003. pp. 1985-1988.

REN-JUN, Z.; XIAN-ZHONG, R. (2002). Optimal combined load forecast based on the improved analytic hierarchy process. In: **Proceedings of the International Conference on Power System Technology**, pp. 1096-1100, vol. 2, Out. 2002.

REZENDE, S.O. (2003). **Sistemas Inteligentes - Fundamentos e Aplicações**. Volume 1. Manole, Bariri - SP. 2003.

ROCCO, S.M., AOKI, A.R., LAMAR, M.V. (2007). Load Forecasting Using Artificial Neural Networks and Support Vector Regressions. In: **Proceedings of the 7th WSEAS International Conference on Power Systems**, Beijing, Setembro 2007.

RUAS, G.I.S.; BRAGATTO, T.A.C.; LAMAR, M.V.; AOKI, A.R.; ROCCO, S.M. (2007). Previsão de Demanda de Energia Elétrica Utilizando Redes Neurais Artificiais e Support Vector Regression. In: **Anais do Encontro Nacional de Inteligência Artificial**, Rio de Janeiro, 2007.

SAINI, L.M.; SONI, M.K. (2002). Artificial Neural Network-Based Peak Load Forecasting Using Conjugate Gradient Methods. **IEEE Transactions on Power Systems**, Vol. 17, Nº. 3, Ago. 2002, pp. 907-912.

SFETSOS, A. (2003). Short-term load forecasting with a hybrid clustering algorithm, **IEE Proceedings Generation Transmission and Distribution**, vol. 150, no. 3, (2003). pp. 257-262.

SILVA, A.M.L.; COSTA, J.G.C.; MATTAR, C.M. (2005) Metodologia para a Determinação dos Montantes de Uso do Sistema de Transmissão por Pontos de Conexão. 2005.



SMOLA, A.J.; SCHÖLKOPF, B. (1998). A Tutorial on Support Vector Regression, NeuroCOLT2 Technical Report Series TR NC-TR-98-030, **Royal Holloway College, University of London**, UK, 1998.

SWARUP, K.S.; SATISH, B. (2002). Integrated ANN: approach to forecast load. **IEEE Computer Applications in Power**, April, 2002. pp. 46-51.

TAYLOR, J.W.; BUIZZA, R. (2000). Neural Network Load Forecasting With Weather Ensemble Predictions. **IEEE Transactions on Power Systems**, Vol. 17, N° 3, Ago. 2000, pp. 626-632.

YALCINOZ, T.; EMINOGLU, U. (2005). Short term and medium term power distribution load forecasting by neural networks. **Energy Conversion and Management**, vol. 46, 2005. pp. 1393–1405.

VERMAAK, J.; BOTHA, E.C. (1998). Recurrent neural networks for short-term load forecasting, **IEEE Transactions on Power Systems**, Vol. 13, No. 1, Feb. 1998. pp. 126 – 132.

# A CARACTERIZAÇÃO E ANÁLISE DOS DADOS

## A.1 ANÁLISE DA MÉDIA, TENDÊNCIAS E DESVIOS PADRÕES

A seguir apresenta-se para as SEs Cascavel, Londrina, Iporã, Foz do Chopim, Pato Branco e Umbará. A análise da média, tendência e desvios padrões foi executada para as demais seis SEs, conforme a metodologia apresentada no Capítulo 2.

### a) SE Cascavel período 2004/2006

Na Figura A.1 traçou-se a curva de tendência com os dados brutos, após a eliminação de erro de relação de Transformadores de Corrente – TC de medição observado na montagem de dados para o período de 1º de janeiro de 2005 e 22 de março de 2005, este ajuste foi pontual dado ao conhecimento da relação correta versus a relação utilizada no TC.

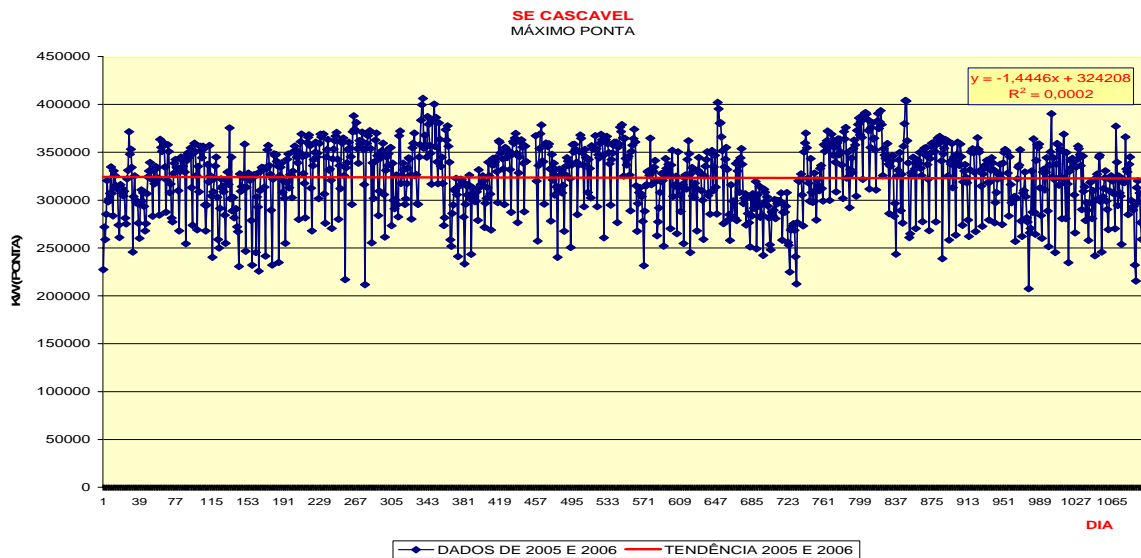


Figura A.1 – SE CEL Dados dos Anos de 2004 a 2006.

Pelo observado tem-se que, devido a remanejamentos constantes de carga ao longo do ano a curva de tendência não apresenta crescimento nestes anos sob análise, carecendo, portanto de um maior cuidado a análise das sazonalidades anuais das demandas desta SE.

Da Figura A.2 a seguir, representando dia a dia os fluxos passantes na SE Cascavel quando analisada a influência semanal tem-se que: para o caso das comparações semana a semana para os anos de 2004, 2005 e 2006, que retirada a influência dos feriados, sábados e domingos, a correlação é muito forte entre as curvas.

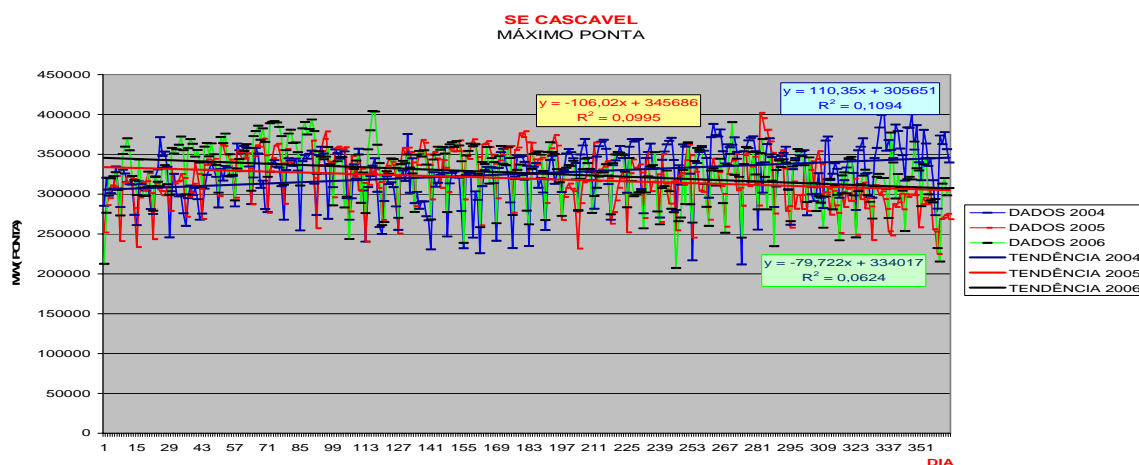


Figura A.2 – SE CEL Dados dos Anos de 2004 a 2006 e Curva de Tendência.

Da Figura A.2 acima, observa-se uma taxa de crescimento vegetativa do fluxo passante por esta SE praticamente nulo, o que não permite deduções sobre a tendência desta SE, também o fluxo passante por esta SE sofre influência do intercambio internacional (ANDE para COPEL) obrigando a um maior cuidado na análise de carregamento desta SE.

A ocorrência de mudança na configuração operativa desta SE do ano de 2004 para o ano de 2005 impõe que seja adotada como boa prática a exclusão dos valores de 2004 do escopo do estudo, dada a alteração significativa que provocaria nas projeções.

#### b) SE Londrina período 2004/2006

A Figura A.3 representa os dados dos anos de 2004, 2005 e 2006 da SE Londrina, observa-se que a curva tem tendência de crescimento positiva.

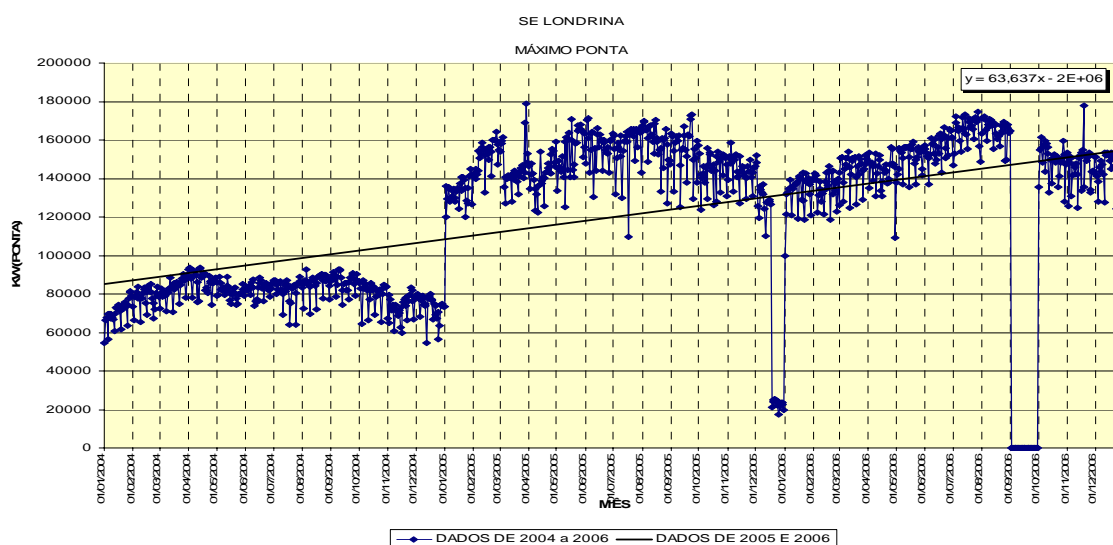


Figura A.3 – SE LNA Dados dos Anos de 2004 a 2006 e Curva de Tendência.

No entanto é facilmente observável que o ano de 2004 não possui a mesma base de dados dos dois anos seguintes, donde se pode inferir que é devido a mudança de configuração de atendimento da SE ou outra razão que fez com que o patamar de fluxo passando pelos enrolamentos dos transformadores desta SE tivesse um incremento considerável.

Para um melhor entendimento da tendência do MUST da SE, foi elaborado a Figura A.3 retirando-se o ano de 2004 do histórico.

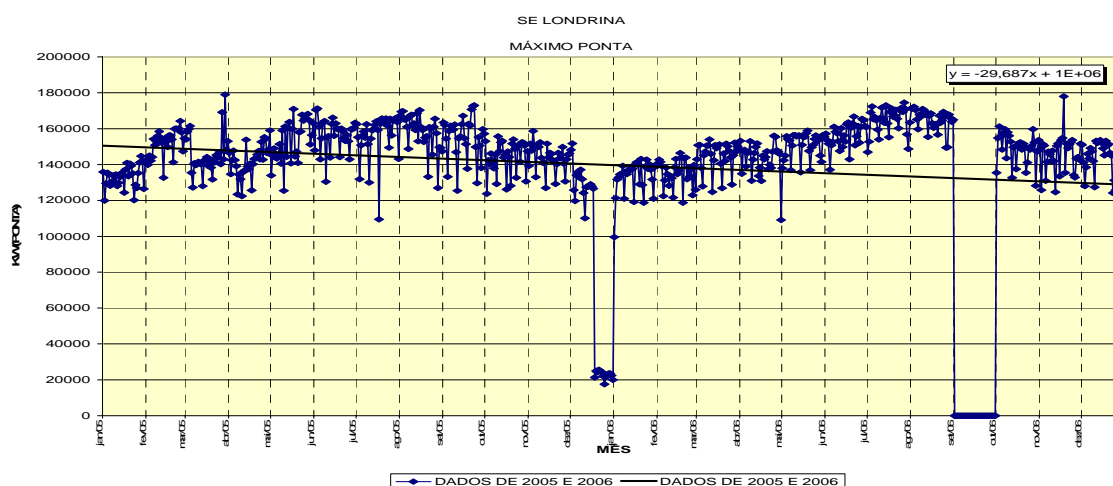


Figura A.4 – SE LNA Dados dos Anos de 2005 a 2006 e Curva de Tendência.

A Figura A.4 mostra que a curva de tendência tende a apresentar crescimento negativo ao longo destes anos, por outro lado observa-se que os dados entre os dias 18 e 31 de dezembro de 2005 e os de setembro de 2006, que correspondem respectivamente a 51<sup>a</sup> e 52<sup>a</sup>

semana de 2005 e entre a 35ª e 40ª semanas de 2006, não estão conformes e precisam ser analisados e depurados.

Para o caso das comparações semana a semana para os anos de 2004, 2005 e 2006 observa-se que de fato o ano de 2004 também poderia ser expurgado da base de dados.

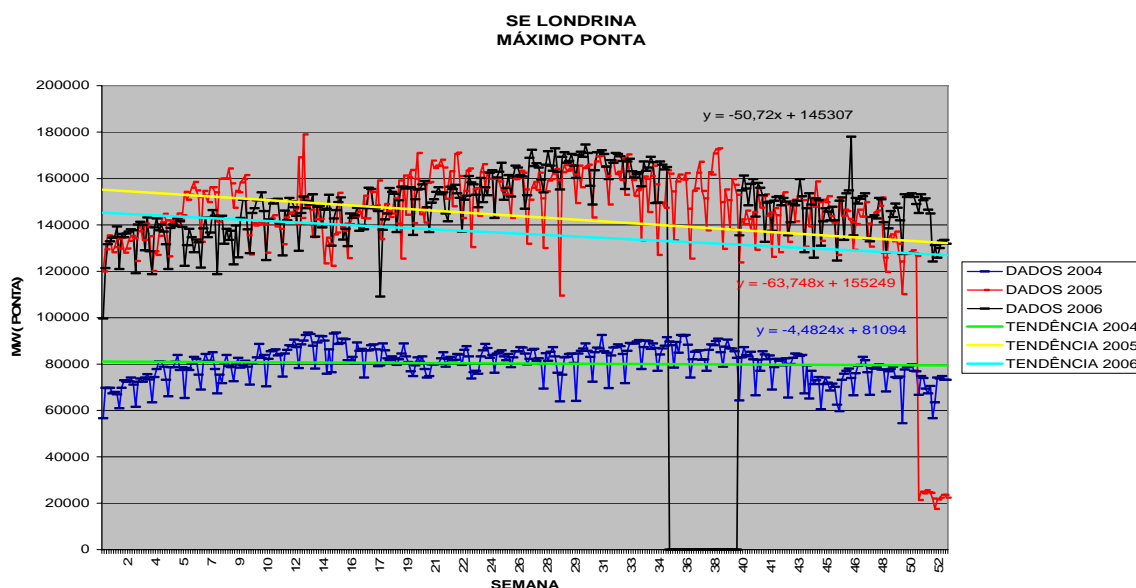


Figura A.5 – SE LNA Dados dos Anos de 2004 a 2006 e Curva de Tendência.

Dada a alteração significativa que a manutenção nas projeções do ano de 2004 implica, pode-se a princípio excluir estas informações da base de dados, por outro lado pode-se inferir, da análise da sazonalidade da SE, que agosto e setembro são os meses em que ocorrem a maior frequência de demandas altas, em relação aos outros meses e que mesmo quando se retira a influência dos feriados, sábados e domingos, tem-se uma correlação muito forte entre as curvas.

#### c) SE Ibiporã período 2004/2006

A Figura A.6 representa os dados dos anos de 2004, 2005 e 2006 da SE Ibiporã, observa-se que o primeiro ano (2004) se apresenta não conforme, provavelmente devido à mudança de configuração.

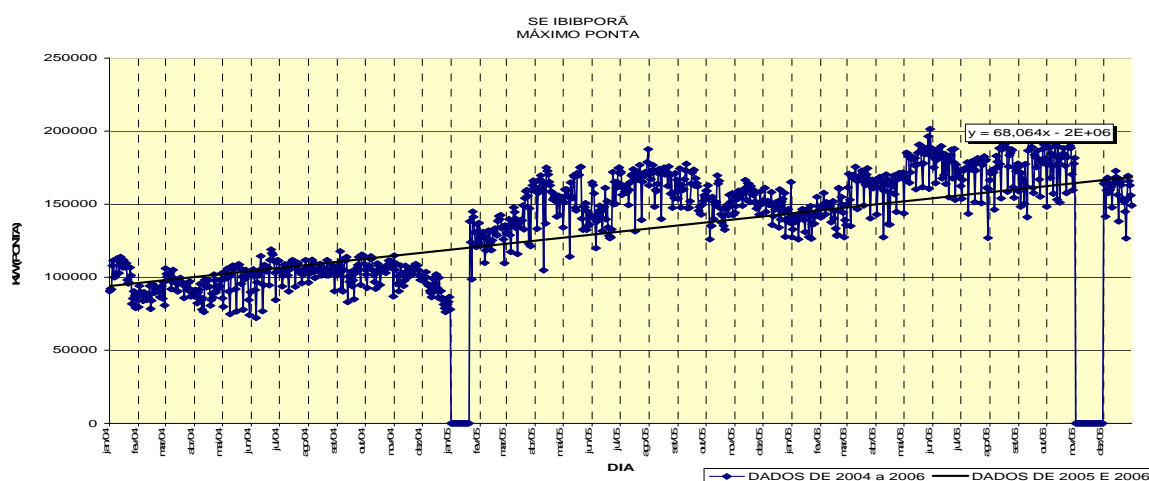


Figura A.6 – SE IBI Dados dos Anos de 2004 a 2006 e Curva de Tendência.

Para um melhor entendimento da tendência do MUST da SE, foi elaborado A Figura A.7 abaixo com a exclusão deste ano de 2004 da base de dados.

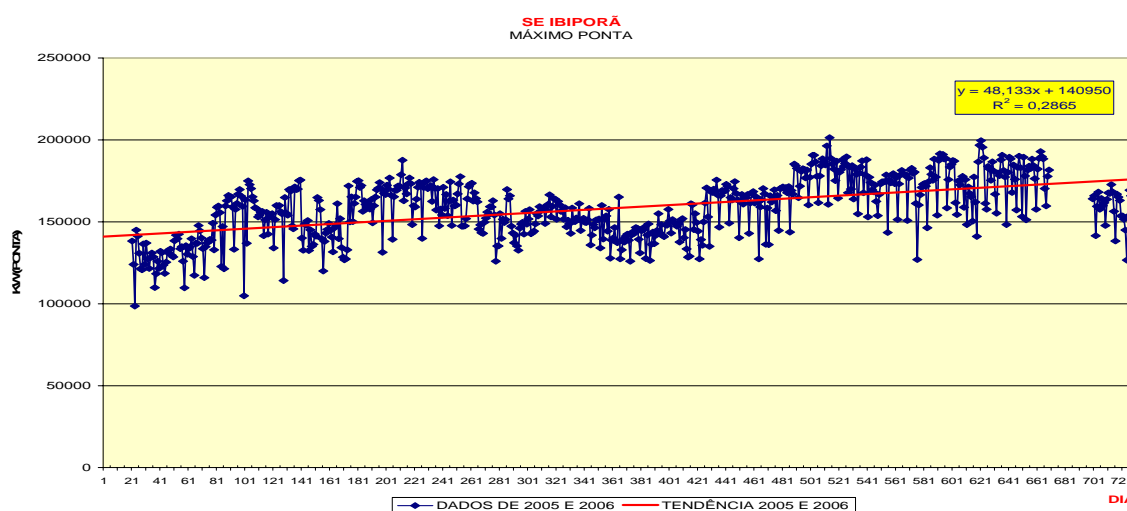


Figura A.7 – SE IBI Dados dos Anos de 2005 a 2006 e Curva de Tendência.

A retirada do ano de 2004 do histórico mostra que a curva de tendência tende a apresentar crescimento constante ao longo destes anos, por outro lado observa-se que os dados do mês de Novembro de 2005 estão faltando, vindo a prejudicar a possibilidade de observação da tendência correta de crescimento dos MUSTs desta SE.

Para o caso das comparações semana a semana para os anos de 2004, 2005 e 2006 observa-se que de fato o ano de 2004 pode ser expurgado da base de dados (Figura A.8).

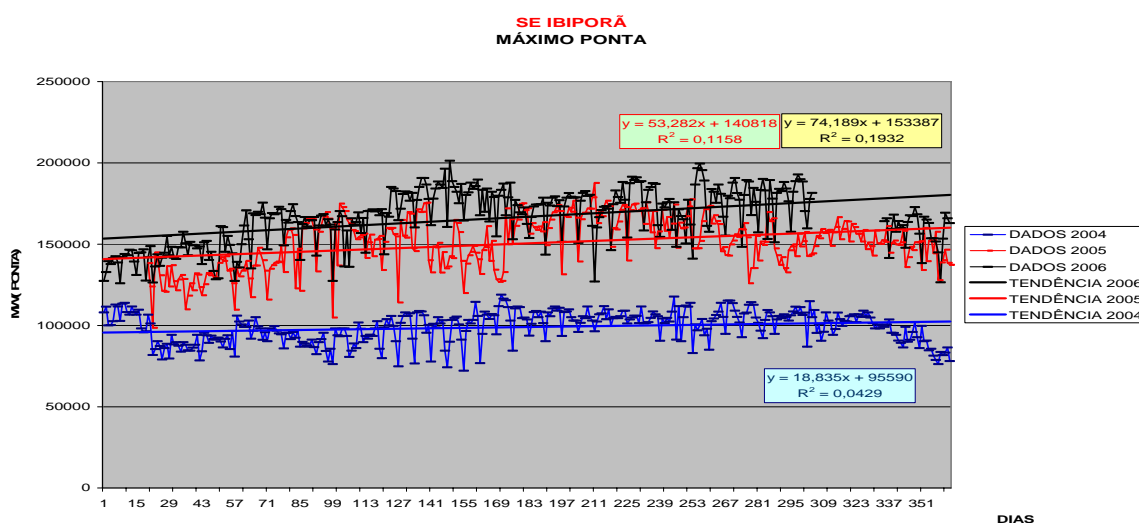


Figura A.8 – SE IBI Dados dos Anos de 2004 a 2006 e Curva de Tendência.

Dada à alteração significativa que a manutenção nas projeções do ano de 2004 implica, pode-se a princípio excluir estas informações da base de dados, por outro lado pode-se inferir, da análise da sazonalidade da SE, que setembro e outubro são os meses em que ocorrem a maior frequência de demandas altas, em relação aos outros meses e que mesmo quando se retira a influência dos feriados, sábados e domingos, tem-se uma correlação muito forte entre as curvas.

#### d) SE Foz do Chopin período 2004/2006

A Figura A.8 representa os dados dos anos de 2004, 2005 e 2006 da SE Foz do Chopin, tem-se que o primeiro ano (2004) se apresenta não conforme devido a mudança de configuração.

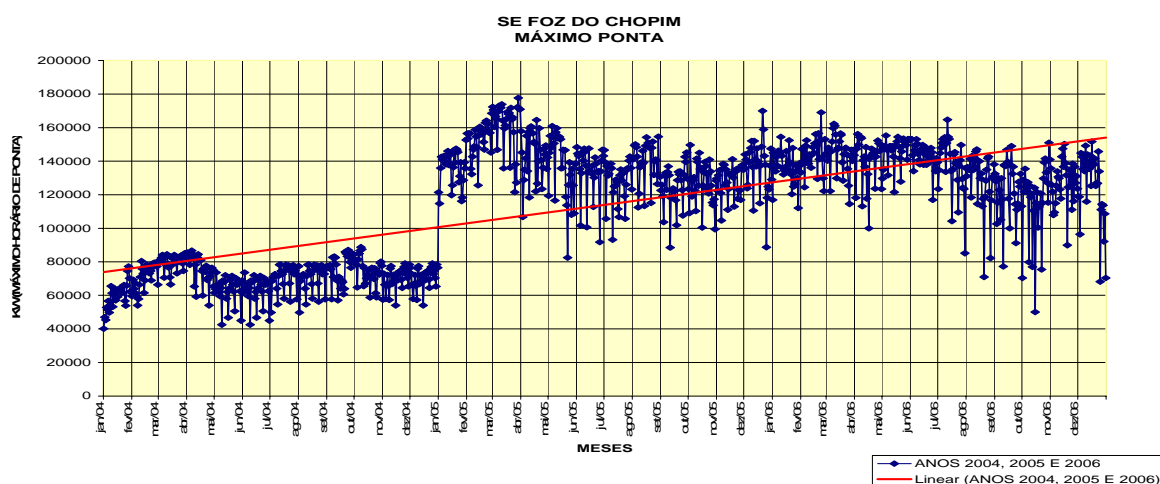


Figura A.9 – SE FOZ Dados dos Anos de 2004 a 2006 e Curva de Tendência.

Da Figura A.9 entende-se que a remoção em uma primeira análise dos dados de 2004 podem proporcionar uma tendência mais aderente ao MUST da SE.

A Figura A.10 tem-se observa-se a influências da mudanças na SE.

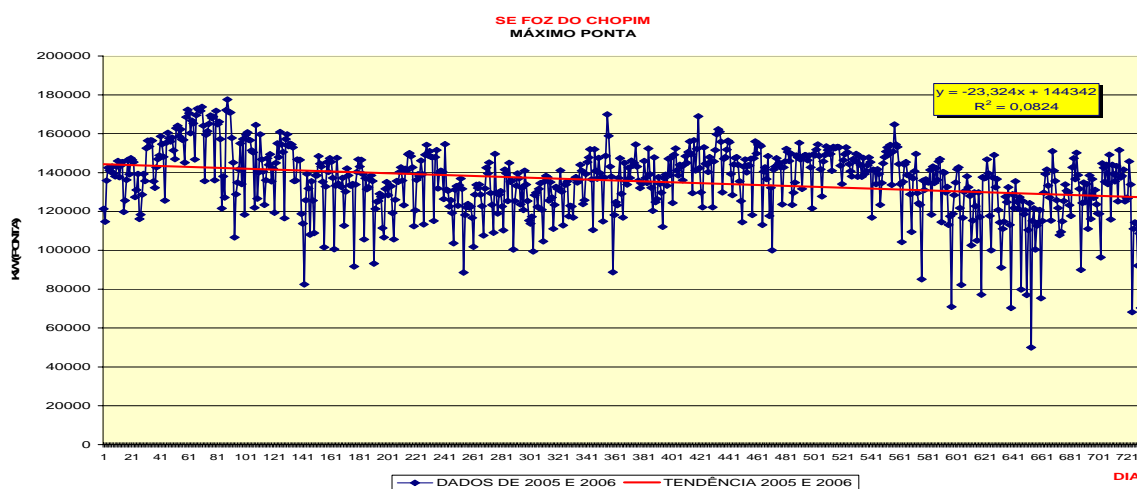


Figura A.10 – SE FOZ Dados dos Anos de 2005 a 2006 e Curva de Tendência.

A retirada do ano de 2004 do histórico faz com que a curva de tendência tenda a apresentar crescimento negativo ao longo destes anos o que pode demonstrar que a influência da SE ao longo deste período sob análise está diminuindo, ou por que alguma SE vizinha está absorvendo maior fluxo de carga, o que justificaria a redução verificada, ou mesmo, por mudança de configuração na área de influência desta SE.



Para o caso das comparações semana a semana para os anos de 2004, 2005 e 2006 observa-se que de fato o ano de 2004 pode ser expurgado da base de dados (Figura A.11).

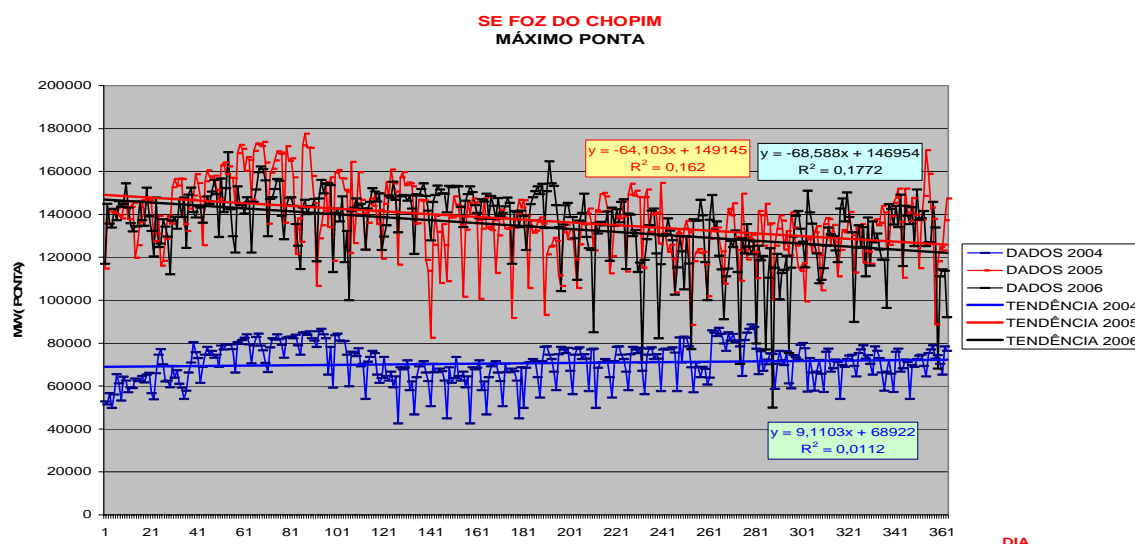


Figura A.11 – SE FOZ Dados dos Anos de 2004 a 2006 e Curva de Tendência.

Observando-se a sazonalidade da SE, tem-se que Março é o mês em que ocorre a maior frequência de demandas altas em relação aos outros meses e que mesmo com a retirada da influência dos feriados, sábados e domingos, tem-se uma correlação muito forte entre as curvas.

Importante frisar que a ocorrência de feriados em determinados dias da semana afeta significativamente a forma de curva entre semanas correlatas, por exemplo, a semana da ocorrência do dia 1º do ano fica sempre prejudicada quando comparada com outro ano.

#### e) SE Pato Branco período 2004/2006

Na Figura A.12 traçou-se a curva de tendência com os dados brutos. Observa-se que esta SE não apresenta, a princípio, uma boa correlação entre os valores anuais, carecendo de maiores análises para se verificar as tendências e sazonalidades da SE, mas ressalta-se que existe uma tendência de crescimento do fluxo passando pelos enrolamentos dos transformadores.

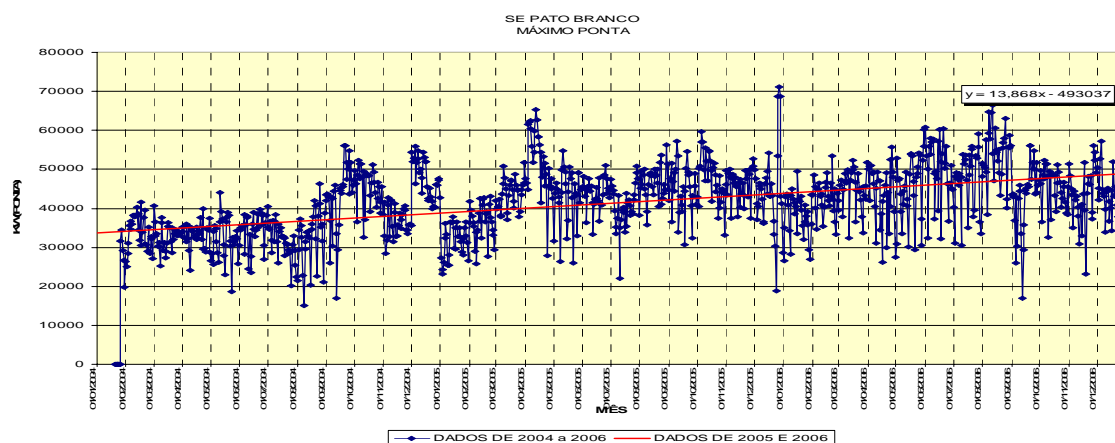


Figura A.12 – SE PTO Dados dos Anos de 2004 a 2006 e Curva de Tendência.

A Figura A.12 acima não permite generalização dada a variação no patamar do fluxo passante por esta SE. Da mesma forma representando a mesma SE, agora com situação de visualização anual (Figura A.13).

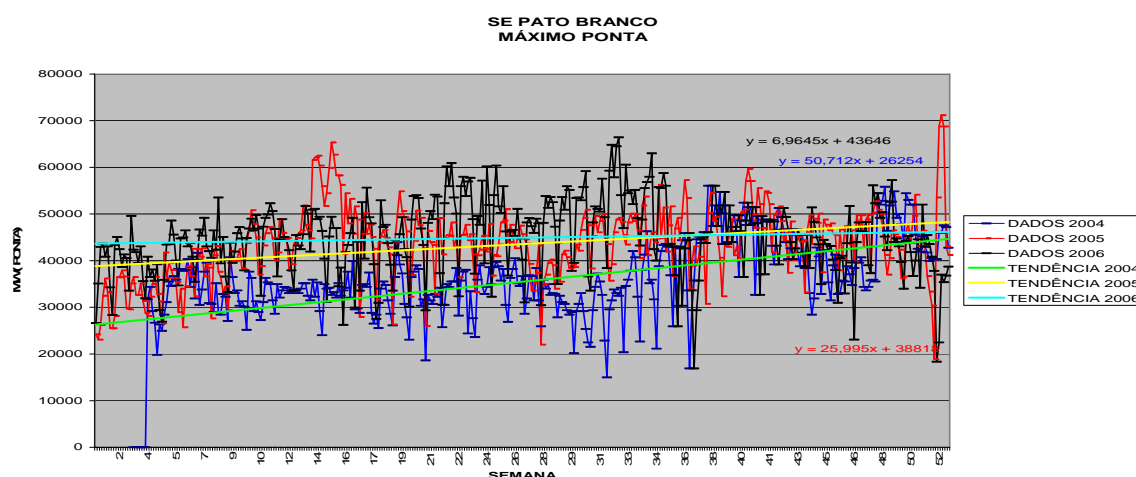


Figura A.13 – SE PTO Dados dos Anos de 2004 a 2006 e Curva de Tendência.

Observa-se da Figura A.13 que não existe patamares bem distintos para estes períodos, com algumas modificações mais acentuadas na tendência do ano de 2004. Da mesma forma que para as SEs analisadas anteriormente elaborou-se a Figura A.14 com as tendências somente para os anos de 2005 e 2006.

Para o caso das comparações semana a semana para os anos de 2004, 2005 e 2006 observa-se que a maior frequência de demandas altas em relação aos outros meses é visualizada nos meses de setembro e outubro, salvo no ano de 2005, onde o mês de março foi

o que apresentou as maiores demandas e que mesmo quando se retira a influência dos feriados, sábados e domingos, tem-se uma correlação fraca entre as curvas.

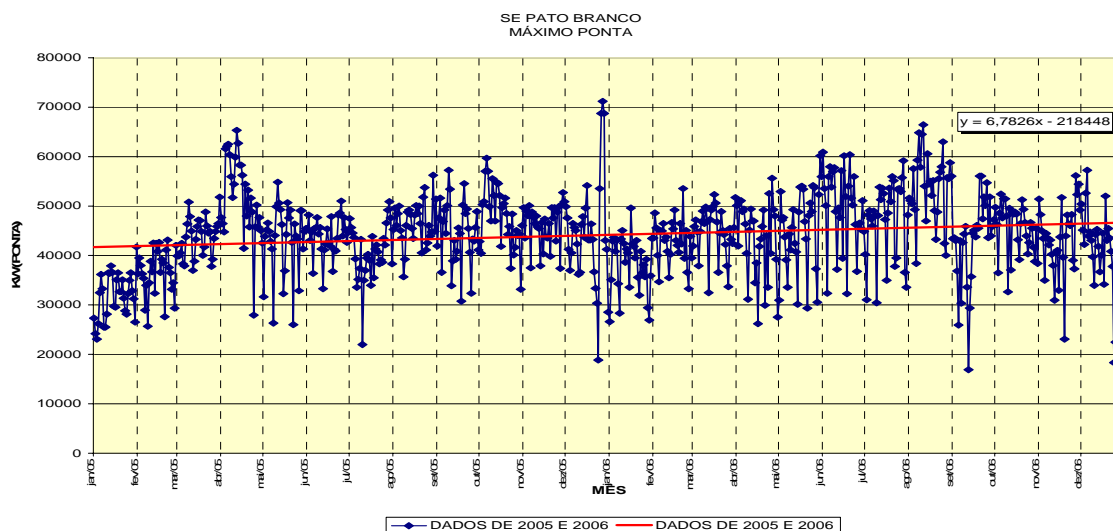


Figura A.14 – SE PTO Dados dos Anos de 2005 a 2006 e Curva de Tendência.

Mesmo com a retirada do ano de 2004 do histórico a curva de tendência, tende a apresentar um pequeno crescimento ao longo destes anos, no entanto, não se constata visualmente uma sazonalidade regular quando se compara mês a mês os dois anos.

Pelo observado tem-se que ocorreram algumas não conformidades *outliers*, que merecem algum tipo de tratamento da base de dados.

Importante frisar que a ocorrência de feriados em determinados dias da semana afeta significativamente a forma de curva entre semanas correlatas, por exemplo, a semana da ocorrência do dia 1º do ano fica sempre prejudicada quando comparada ano a ano.

#### f) SE Umbará período 2004/2006

A Figura A.15 representa os dados dos anos de 2004, 2005 e 2006 da SE Umbará, tem-se que o primeiro ano (2004) se apresenta não conforme devido, provavelmente a mudança de configuração, não apresenta a princípio uma boa correlação entre os valores anuais, carecendo de maiores análises para se verificar as tendências e sazonalidades da SE.

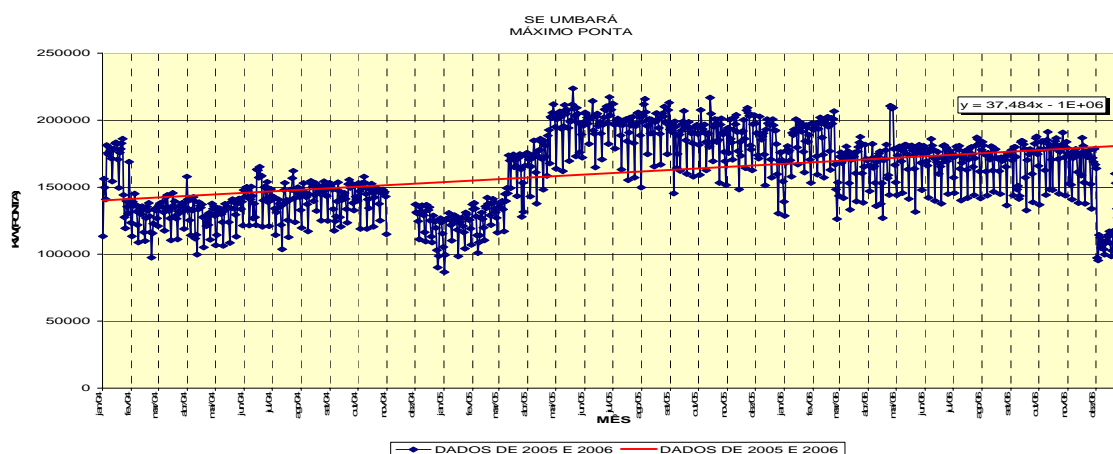


Figura A.15 – SE UMB Dados dos Anos de 2004 a 2006 e Curva de Tendência.

Da Figura A.15 entende-se que a remoção em uma primeira análise dos dados de 2004 pode proporcionar uma tendência mais aderente ao MUST da SE. Abaixo é apresentado a Figura A.16 de tendência de crescimento da SE Umbará sem os dados do ano de 2004.

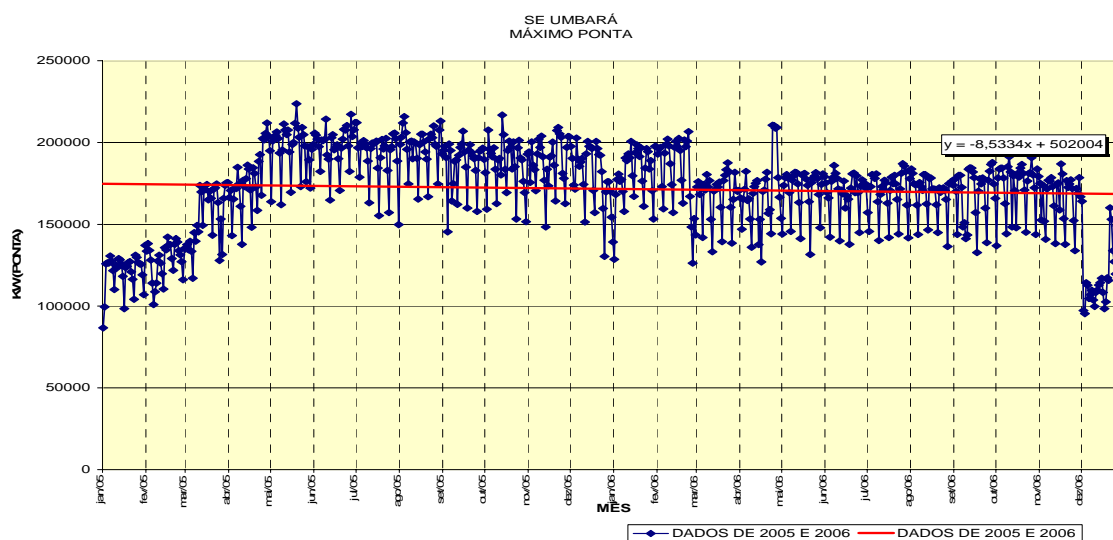


Figura A.16 – SE UMB Dados dos Anos de 2004 a 2006 e Curva de Tendência.

Observa-se que, com a retirada do ano de 2004 do histórico a curva de tendência tende a apresentar crescimento negativo ao longo destes anos o que pode demonstrar que a influência da SE ao longo deste período sob análise está diminuindo, ou por que alguma SE vizinha está absorvendo maior fluxo de carga, o que justificaria a redução verificada, ou mesmo, por mudança de configuração na área de influência desta SE.

Além do crescimento negativo ao longo destes anos, também não se apresenta uma sazonalidade regular entre o mesmo mês dos anos analisados.

Os meses de janeiro e fevereiro de 2005 ainda provocam distorção na base de dados dada a ter ocorrido mudança de configuração em março de 2005. Para o caso das comparações semana a semana para os anos de 2004, 2005 e 2006 observa-se que de fato o ano de 2004 pode ser expurgado da base de dados (Figura A.17).

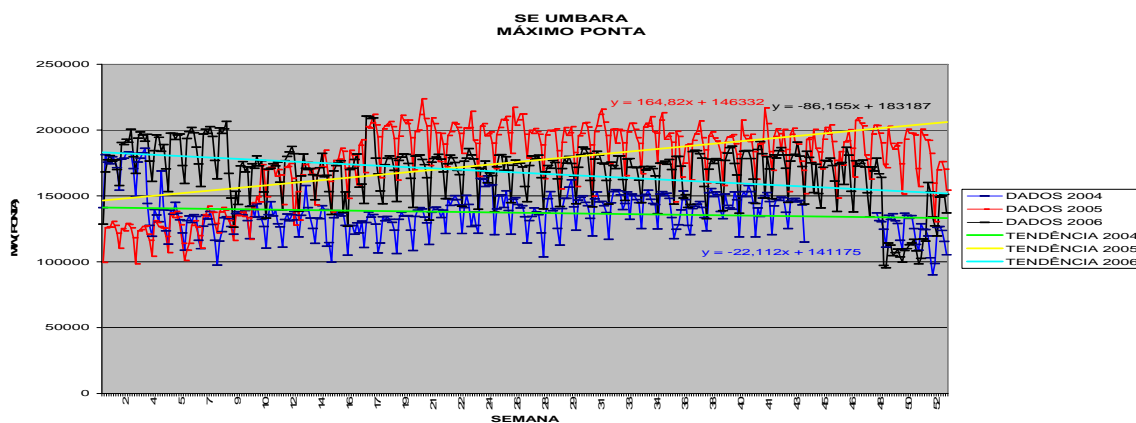


Figura A.17 – SE UMB Dados dos Anos de 2004 a 2006 e Curva de Tendência.

Observando-se a SE não apresenta uma tipicidade de sazonalidade de demandas altas em relação aos outros meses e que mesmo com a retirada da influência dos feriados, sábados e domingos, tem-se uma correlação muito forte entre as curvas.

Pode-se observar que de modo geral ano após ano a tendência é de redução para os fluxos passantes pelos enrolamentos dos transformadores 230/69 kV desta SE nos anos de 2004 e 2006 e o crescimento verificado em 2005, está mascarado pelos meses de janeiro e fevereiro que ainda estão incorporando a mudança de configuração no Anel 69 kV da cidade de Curitiba de um ano para o outro, sendo praticamente impossível ajustar a base de dados dada à falta de informações destas ocorrências. Uma análise mais aprofundada excluindo-se estes dois meses da base de dados de 2005 permite inferir que o crescimento do fluxo de potência ao longo de 2005 tem uma leve tendência de crescimento, mesmo por que a tendência é de os fluxos passantes pela SE tendem a aumentar devido a saturação das SEs adjacentes.