



Universidade Federal de Campina Grande
Centro de Engenharia Elétrica e Informática
Unidade Acadêmica de Engenharia Elétrica
Grupo de Sistemas Elétricos

Trabalho de Conclusão de Curso

Título: Estudos de Previsão de Carga

Aluno: Luiz Alberto Leite Filho

Orientadora: Núbia Silva Dantas Brito

Campina Grande, dezembro de 2005.



Biblioteca Setorial do CDSA. Fevereiro de 2021.

Sumé - PB

Título
Estudos de Previsão de Carga

Aluno:

Luiz Alberto Leite Filho
Luiz Alberto Leite Filho

Orientadora:

Núbia Silva Dantas Brito

Campina Grande, dezembro de 2005

Agradecimentos

Agradeço primeiramente a Deus, por ter me dado apoio espiritual para realizar as atividades deste trabalho, aos meus pais que foram as pessoas que mais me apoiaram e deram condição para que eu fizesse este curso, à minha irmã que sempre me apoiou em tudo, e a minha noiva, pela confiança, compreensão e companheirismo em todas as horas.

Agradeço a professora Núbia Silva Dantas Brito e ao colega Antônio Carlos Moreira de Melo Filho, por terem me guiado durante este trabalho, pela sua confiança, seus conselhos e pelo rigor que exigiram de mim.

Aos meus amigos em geral, Felipe Gadelha, Leonardo Lopes, Paulo Marcelo Lopes, Paulo Eduardo Lopes, George Barreto, Tiago Felinto, Alencar Junior, Mário Araújo Neto, dentre outros, pela amizade sólida construída e apoio em minha vida acadêmica.

Índice

1	Objetivos.....	5
2	Introdução	6
3	Revisão Bibliográfica.....	7
4	Fundamentação Teórica.....	10
4.1	Séries Temporais	10
4.2	Análise de Regressão.....	11
4.2.1	Funções de Regressão.....	12
4.2.2	Regressão Linear	13
4.2.3	Regressão Quadrática	14
4.2.4	Regressão Polinomial	15
4.2.5	Linearização.....	15
5	Implementação Computacional e Análise dos Resultados	16
5.1	O PCBase.....	17
5.1.1	Aquisição de Dados.....	18
5.1.2	Identificação de Defeitos.....	18
5.1.3	Estimação de Dados.....	19
5.2	Resultados.....	21
6	Conclusões	27
7	Referências Bibliográficas	28
	Anexos.....	289

1 Objetivos

Este Trabalho de Conclusão de Curso (TCC) tem como objetivos:

- Fazer um estudo sobre os métodos usuais de previsão de carga em sistemas de energia elétrica.
- Implementar o método de regressão usando diversos tipos de função. Para isso será usado o software Microsoft Excel[®].
- Fazer um estudo comparativo entre os resultados obtidos pelo método de regressão e por redes neurais artificiais (RNA), através de um caso específico.

2 Introdução

Prever a demanda de energia elétrica é uma atividade essencial para as concessionárias, dada à exigência cada vez maior por qualidade e continuidade do serviço prestado. A previsão da demanda de energia elétrica, também denominada *previsão de carga*, tem como principal benefício o planejamento da expansão do sistema elétrico e, conseqüentemente, dos recursos financeiros. Se a previsão é feita de forma muito conservativa, a capacidade instalada provavelmente se esgotará em pouco tempo, acarretando problemas de continuidade de serviço, regulação de tensão e até mesmo, de racionamento de energia. Por outro lado, previsão de carga bastante otimista pode conduzir a instalação de um sistema com capacidade excessiva (Souza, 1997).

Os estudos de previsão de carga são classificados com relação ao tempo à frente que se deseja prever: curtíssimo prazo, curto prazo, médio prazo e longo prazo:

- Curtíssimo prazo: previsões de poucos minutos até uma hora à frente. Usadas para operações do sistema, como despacho econômico e avaliação da segurança.
- Curto prazo: período de até uma semana à frente. Utilizadas na programação da manutenção e auxílio à operação do sistema.
- Médio prazo: período de semanas a meses à frente. Utilizado para intercâmbio de energia elétrica entre as concessionárias e no planejamento da produção.
- Longo prazo: período de anos à frente. É importante para o planejamento estratégico e prever possíveis construções de estruturas geradoras de energia.

O estudo realizado neste TCC diz respeito à *previsão de carga de curtíssimo prazo*, considerando-se um caso-teste bastante simples. Como já apresentado, foram avaliados os desempenhos dos métodos baseados em regressão e em RNA. O estudo e implementação do método baseado em RNA foram feitos por Melo Filho (2005). Neste TCC implementou-se o método baseado em regressão, considerando diversas funções de regressão.

3 Revisão Bibliográfica

O monitoramento da carga é um requisito fundamental para o controle e operação de sistemas elétricos. A adequação da energia gerada em função da variação da carga deve ser realizada de modo a se ter um controle efetivo e uma operação economicamente eficiente. Portanto, para se obter uma operação segura e econômica do sistema elétrico é necessário o estudo e análise das funções que estimam (fazem a previsão) o valor da carga no sistema.

De um modo geral, previsão de carga de curto prazo tem como objetivos (Gross & Galiana, 1987):

- Direcionar as funções de programação que determinam o comprometimento das fontes de geração de energia economicamente mais vantajosas. Para sistemas de geração hidráulicos, a previsão de carga é necessária para determinar o nível ótimo de geração e dos reservatórios nas hidrelétricas. Para sistemas de geração térmicos, a previsão se faz necessária para determinar o custo mínimo por hora para ligar e desligar as unidades de geração para atender as cargas do sistema. Para sistemas mistos (hidráulicos e térmicos), a previsão é requerida para programar a operação com os diferentes recursos existentes para minimizar o custo da produção.
- Prever a segurança de operação do sistema elétrico. A previsão da carga fornece informações necessárias que permitem detectar condições futuras as quais demonstram a vulnerabilidade do sistema. Assim, o operador pode tomar medidas corretivas para evitar que o sistema falhe futuramente.
- Informar a carga do sistema em qualquer tempo em função das últimas previsões de tempo e de outros fatores aleatórios que podem modificar a previsão.

A carga de um sistema é a soma de todas as demandas individuais de todos os nós do sistema elétrico. Em princípio, pode-se determinar a demanda de um sistema se todos os consumidores individuais fossem conhecidos, entretanto, a carga do sistema é um processo estocástico não-estacionário, tornando esse consumo individual imprevisível. Felizmente, a soma da carga de alguns consumidores tem uma característica que pode ser prevista.

Na prática, os estudos de previsão de carga de um sistema são realizados através de modelos e técnicas e a carga é influenciada por fatores: econômicos, temporais, climáticos e efeitos aleatórios. Como consequência, tais estudos dependem da forma como esses fatores influenciam na modelagem da carga (Gross & Galiana, 1987).

Avaliação do estado da arte evidencia uma grande diversidade de modelos propostos. De um modo geral, esses modelos são classificados conforme o modelo de carga utilizado (Gross & Galiana, 1987):

- Modelo de pico de carga: considera somente o pico diário ou semanal da carga. O clima é uma variável importante. A carga base é adicionada de um fator que é dependente do clima. Os fatores climáticos a serem considerados são: umidade, intensidade da luz, velocidade do vento e precipitação. As principais vantagens deste tipo de modelo são sua simplicidade estrutural e a pouca quantidade de dados exigida. Os parâmetros do modelo são estimados através de uma regressão linear ou não-linear. As desvantagens deste modelo são: não determina o tempo em que o pico ocorre, nem existe informação sobre a forma da curva de carga.
- Modelo de forma de carga: descreve a carga como uma série temporal sobre o período da previsão. O intervalo de amostragem é de alguns minutos ou horas. Esse modelo pode ser subdividido em duas classes básicas:
 - Modelo tempo do dia: define a carga durante uma amostragem de tempo por uma série temporal. Este modelo é obtido baseado em dados observados anteriormente, os quais servirão para fazer a previsão. A partir dos dados, um conjunto de curvas é gerado, divididas por sua característica climática. As vantagens deste modelo são sua estrutura simples, determinação fácil de seus parâmetros e pouco esforço computacional. Como desvantagem principal tem-se a não representação da correlação estocástica da natureza do processo, ou seja, da sua relação com as condições climáticas. Quando as condições climáticas são modificadas, os coeficientes calculados não são mais apropriados.
 - Modelo dinâmico: considera que a carga não é somente uma função do tempo do dia, mas também de causas como condições

climáticas e efeitos aleatórios. Os modelos dinâmicos são de dois tipos:

- Modelo auto-regressivo com média móvel: leva em conta uma parte do processo como uma função periódica (representado por séries temporais) e outra parte como função dos efeitos aleatórios. Este tipo de modelo representa, de uma forma adequada, as variações climáticas. Em contrapartida, ele requer maior esforço computacional. A literatura de previsão de carga apresenta uma vasta variação deste modelo.
- Modelo espaço-estado: é bem parecido com o modelo auto-regressivo. A diferença está no fato das técnicas disponíveis para o modelo espaço-estado assumirem que os parâmetros periódicos da carga são processos estocásticos. Isso permite que se faça o uso de uma informação *a priori* sobre seus valores, o que pode ajudar na estimação dos parâmetros do modelo.

É importante destacar o uso crescente de técnicas baseadas em Inteligência Artificial em estudos de previsão de carga, em particular das RNA (Al-Alawi & Islam, 1996). Para realizar a previsão, os métodos baseados em RNA geralmente não utilizam o conhecimento das relações funcionais entre as variáveis climáticas e as de carregamento elétrico. Na verdade, as RNA sintetizam o mapeamento entre as variáveis de entrada e de saída, a partir da extração do conhecimento de um conjunto de treinamento (base de dados), previamente tratado e analisado, apresentado durante a fase de aprendizagem da RNA.

4 Fundamentação Teórica

Apresenta-se a seguir uma breve fundamentação teórica sobre séries temporais e sobre os métodos baseados em regressão.

4.1 Séries Temporais

Uma *série temporal*, também denominada de *série histórica*, é uma seqüência de dados obtidos em intervalos regulares de tempo durante um período específico. Este conjunto pode ser obtido através de observações periódicas do evento de interesse ou através de processos de contagem. Se a série for denominada como Z , o valor da série no momento t pode ser escrito como $Z_t (t = 1, 2, \dots, n)$ (Latorre & Cardoso, 2001). Qualquer evento ou processo que é função do tempo, como a carga do sistema de potência de potência, é dito temporal.

Denomina-se *trajetória de um processo*, a curva obtida no gráfico da série histórica. O conjunto de todas as possíveis trajetórias é considerado um *processo estocástico*. Diz-se que uma série temporal é uma *amostra* deste processo (Latorre & Cardoso, 2001).

O conjunto de observações ordenadas no tempo pode ser *discreto* ou *contínuo*. Pode-se obter uma série temporal discreta a partir de uma amostra de pontos de uma série contínua ou por meio de um parâmetro como, por exemplo, a média de períodos fixos de tempo.

O processo de análise de uma série temporal segue, de modo geral, os seguintes passos (Esteves, 2003):

1. Modelagem do fenômeno estudado.
2. Descrição do comportamento da série.
3. Estimação de valores.
4. Avaliação dos fatores que influenciam o comportamento da série.

O método mais adequado de análise de séries temporais depende do modelo definido (ou estimado para a série), do tipo de série e do objetivo do trabalho.

Uma série temporal pode ser composta de três componentes não observáveis: tendência (T_t), sazonalidade (S_t) e variação aleatória, também denominada de ruído branco (a_t). Pode-se elaborar:

- Um modelo aditivo: $Z_t = T_t + S_t + a_t$;
- Um modelo multiplicativo: $Z_t = T_t S_t a_t$;
- Um modelo log-linear: faz-se uma transformação log no modelo multiplicativo.

Ao se analisar uma série temporal, deve-se avaliar a influência individual de cada um destes componentes.

A previsão de séries temporais é um processo de estimação de valores futuros da série, a qual é feita com base em informações atuais e passadas da série. O horizonte de previsão é o comprimento do tempo, contado a partir de uma origem predeterminada, ou seja, da origem da previsão (Esteves, 2003).

Para que se possam prever valores futuros de uma dada série, é necessário formular um modelo matemático capaz de representar o *comportamento* e as *características* da série temporal que se deseja prever. *Essas informações são extraídas dos dados disponíveis*. Existe uma grande quantidade de modelos de previsão na *literatura estatística* para exercer tal tarefa, os quais são classificados como (Esteves, 2003):

Modelos univariados: Os valores futuros de uma série são previstos considerando-se apenas os seus valores passados. Isso acontece, por exemplo, em geral, com os dados de consumo de carga de energia elétrica.

Modelos causais ou modelos de função de transferência: Os valores futuros de uma série são previstos considerando-se os valores passados e os valores de séries, que possuam relação com ela. No caso do consumo de carga de energia elétrica, o preço relativo poderia ser uma série relacionada com a série de carga.

Modelos multivariados: São modelos capazes de realizarem várias previsões ao mesmo tempo. Um exemplo de modelo multivariado é o modelo capaz de prever ao mesmo tempo, o consumo de energia em diversas concessionárias prestadoras do serviço no país.

4.2 Análise de Regressão

Análise de regressão é o estudo do passado de um evento temporal de modo que seu comportamento futuro possa ser estimado (previsto) por extrapolação. Análise de regressão faz parte de um conjunto de técnicas numéricas que tem como objetivo ajustar

curvas a um conjunto de dados para obter estimativas intermediárias. No caso particular de regressão, os dados possuem erros significativos e a curva obtida não intercepta, necessariamente, todos os pontos.

4.2.1 Funções de Regressão

O princípio da análise de regressão é, com base nos valores históricos de um evento, se poder definir uma função matemática para descrevê-lo. A partir desta função, o comportamento futuro do evento pode ser conhecido aproximadamente, por antecipação (Souza, 1997).

Qualquer curva $y = f(x)$ pode se ajustar a um conjunto de n pontos $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$. O critério usual de ajuste é o dos *mínimos quadrados*, segundo o qual uma curva se ajusta a um conjunto de pontos dados quando a soma dos quadrados dos erros em cada ponto é mínima. Isto é, se:

$$\sum_{i=1}^m [f(x_i) - y_i]^2 = E^2 \quad (4.1)$$

e E^2 é mínimo, então a curva $f(x)$ versus x se ajusta à seqüência $S = \{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$.

O erro quadrático mínimo, E^2 , é um indicador de qualidade de ajuste: quanto menor seu valor, mais fino é o ajuste. Ele é utilizado quando é necessário fazer a comparação de duas ou mais funções de regressão para efeito de encontrar a função mais adequada para uma dada seqüência (Souza, 1997).

Existem diversas outras maneiras de se quantificar o ajuste, como por exemplo, utilizar:

- O erro quadrático médio:

$$\bar{E} = \frac{1}{n} \sqrt{\sum_{i=1}^n [y_i - f(x_i)]^2} . \quad (4.2)$$

- O coeficiente de correlação:

$$R = \pm \sqrt{1 - \frac{\sum_{i=1}^n [f(x_i) - y_i]^2}{\sum_{i=1}^n [y_i - \bar{y}]^2}} , \quad (4.3)$$

sendo:

$$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i. \quad (4.4)$$

A faixa de valores possíveis do coeficiente de correlação é $-1 \leq R \leq 1$. Se $R = \pm 1$, a curva se ajusta perfeitamente aos pontos dados. Se $R = 0$, não há correlação entre os dados. Na prática, utiliza-se R^2 .

Apresentam-se na Tabela 4.1, as regressões comumente utilizadas em previsão de carga.

Tabela 4.1 – Regressões usuais em previsão de carga.

Regressão	$f(x)$
Linear	$a_0 + a_1 x$
Exponencial	$a_0 A^{a_1 x}$
Logarítmica	$a_0 + a_1 \log_A x$
Potencial	$a_0 x^{a_1}$
Polinomial	$a_0 + a_1 x + a_2 x^2 + \dots + a_n x^n$
Racional	$\frac{a_0 + a_1 x + a_2 x^2 + \dots + a_n x^n}{b_0 + b_1 x + b_2 x^2 + \dots + b_n x^n}$

4.2.2 Regressão Linear

Esse método consiste em determinar uma *função linear* (reta) que se ajuste aos pares de dados $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$:

$$y = a_0 + a_1 x + e. \quad (4.5)$$

Os coeficientes a_0 e a_1 representam o ponto de interseção e a inclinação da função, respectivamente. O termo e representa o erro ou resíduo, entre o modelo desejado e os dados:

$$e = y - a_0 - a_1 x. \quad (4.6)$$

Uma estratégia para determinar os coeficientes que melhor se ajustam ao modelo desejado é minimizar a soma dos quadrados dos resíduos:

$$S = \sum_{i=1}^n e_i^2 = \sum_{i=1}^n (y_i - a_0 - a_1 x_i)^2. \quad (4.7)$$

Para se determinar os valores de a_0 e a_1 que melhor se ajustam a reta, diferencia-se a expressão (4.7) em relação à cada coeficiente:

$$\frac{\partial S}{\partial a_0} = -2 \sum_{i=1}^n (y_i - a_0 - a_1 x_i). \quad (4.8a)$$

$$\frac{\partial S}{\partial a_1} = -2 \sum_{i=1}^n [(y_i - a_0 - a_1 x_i) x_i]. \quad (4.8b)$$

Para simplificar a notação, considera-se: $\sum = \sum_{i=1}^n$. Igualando as equações (4.8)

a zero, tem-se:

$$\sum y_i - \sum a_0 - \sum a_1 x_i = 0. \quad (4.9a)$$

$$\sum y_i x_i - \sum a_0 x_i - \sum a_1 x_i^2 = 0. \quad (4.9b)$$

Resolvendo o sistema formado por essas duas últimas expressões, obtém-se:

$$a_1 = \frac{n \sum x_i y_i - \sum x_i \sum y_i}{n \sum x_i^2 - (\sum x_i)^2} \quad (4.10a)$$

$$a_0 = \bar{y} - a_1 \bar{x}, \quad (4.10b)$$

sendo \bar{x} e \bar{y} médias de x e y , respectivamente.

4.2.3 Regressão Quadrática

Uma regressão quadrática consiste em determinar um polinômio de segundo grau que se ajuste aos dados disponíveis, ou seja:

$$y = a_0 + a_1 x + a_2 x^2 + e. \quad (4.11)$$

O erro ou resíduo entre o modelo desejado e os dados é obtido por:

$$e = y - a_0 - a_1 x - a_2 x^2. \quad (4.12)$$

Utilizando a mesma estratégia adotada anteriormente, obtém-se:

$$S = \sum_{i=1}^n e_i^2 = \sum_{i=1}^n (y_i - a_0 - a_1 x_i - a_2 x_i^2)^2. \quad (4.13)$$

Diferenciando a equação (4.13), tem-se:

$$\frac{\partial S}{\partial a_0} = -2 \sum (y_i - a_0 - a_1 x_i - a_2 x_i^2). \quad (4.14a)$$

$$\frac{\partial S}{\partial a_1} = -2 \sum x_i (y_i - a_0 - a_1 x_i - a_2 x_i^2). \quad (4.14b)$$

$$\frac{\partial S}{\partial a_1} = -2 \sum x_i^2 (y_i - a_0 - a_1 x_i - a_2 x_i^2). \quad (4.14c)$$

Igualando a zero as equações (4.14):

$$(n)a_0 + (\sum x_i)a_1 + (\sum x_i^2)a_2 = \sum y_i \quad (4.15a)$$

$$(\sum x_i)a_0 + (\sum x_i^2)a_1 + (\sum x_i^3)a_2 = \sum y_i x_i \quad (4.15b)$$

$$(\sum x_i^2)a_0 + (\sum x_i^3)a_1 + (\sum x_i^4)a_2 = \sum y_i x_i^2. \quad (4.15c)$$

4.2.4 Regressão Polinomial

Os coeficientes do polinômio de ordem m que se ajuste aos n pontos dados, são obtidos resolvendo-se o seguinte sistema de equações:

$$\begin{aligned} C_{00}a_0 + C_{01}a_1 + \dots + C_{0m}a_m &= B_0 \\ C_{10}a_0 + C_{11}a_1 + \dots + C_{1m}a_m &= B_1 \\ \vdots & \\ C_{m0}a_0 + C_{m1}a_1 + \dots + C_{mm}a_m &= B_m \end{aligned} \quad (4.16)$$

Sendo $n > m$ e:

$$C_{ij} = \sum_{k=1}^n x_k^{i+j}. \quad (4.17a)$$

$$B_i = \sum_{k=1}^n x_k^i y_k. \quad (4.17b)$$

4.2.5 Linearização

Essa técnica consiste em transformar as regressões exponencial, logarítmica e potencial em regressão linear. O procedimento é sintetizado na Tabela 4.2.

Tabela 4.2 – Linearização.

Regressão	Função	Amostra transformada	Coeficientes		
			b_0	b_1	
Exponencial	$b_0 A^{b_1 x}$	x_i	$\log_A y_i$	A^{a_0}	a_1
Logarítmica	$b_0 + \log_A x$	x_i	A^{y_i}	0	A^{a_0}
Potencial	$b_0 x^{b_1}$	$\log_A x_i$	$\log_A y_i$	$\log_A a_0$	a_1

5 Implementação Computacional e Análise dos Resultados

O estudo de previsão de carga realizado neste TCC, diz respeito ao alimentador que supre o shopping Iguatemi de Alagoas, pertencente ao sistema CEAL (Companhia Energética de Alagoas), portanto, uma carga de características bem definidas.

Os dados fornecidos se referem aos valores de carga registrados em intervalos de 10 minutos ao longo do dia, nos períodos de junho a setembro de 2004 (Melo Filho, 2005). Na Figura 5.1 mostra-se o gráfico da curva média de carga, a qual foi obtida da Base de Dados de quatro meses fornecida.

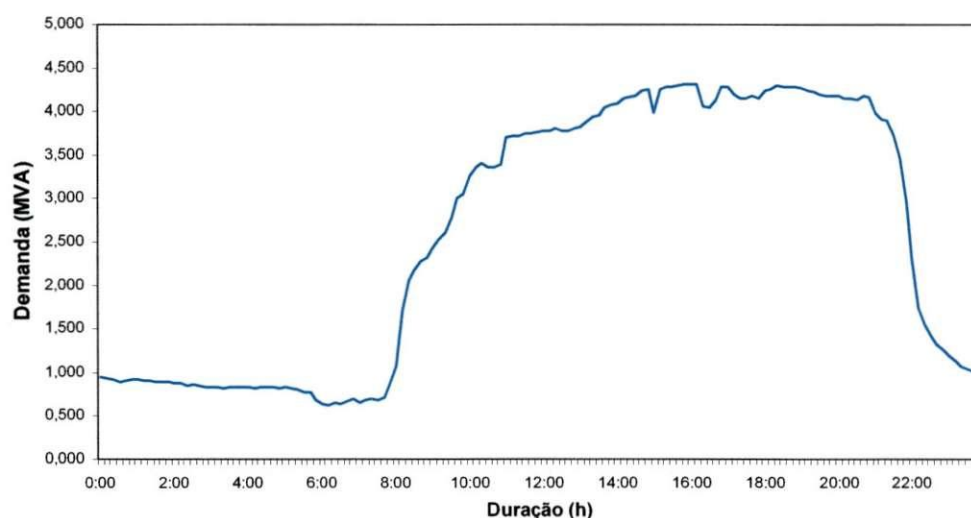


Figura 5.1 – Curva média da carga durante quatro meses.

Avaliando a Figura 5.1, observa-se que os equipamentos elétricos começam a serem ligados a partir das 8:00 horas e a partir das 10:00 horas começam a chegar clientes ao shopping, aumentando o consumo de energia. O consumo se mantém relativamente constante entre 14:00 horas e 21:00 horas. A partir das 21:00 horas o consumo começa a declinar. Às 22:00 horas o shopping fecha e os equipamentos começam a serem desligados (Melo Filho, 2005).

Como é comum ocorrer em problemas de previsão de carga, os dados fornecidos pela CEAL continham erros, tais como: ausência de leitura, leituras incorretas e leituras repetidas. Desta forma, foi necessário identificar e tratar esses erros. Para isto, utilizou-se o software PCBase (PCBase, 2005).

5.1 O PCBase

O PCBase é um software que tem como objetivo tratar e formatar dados para estudos de previsão de carga via redes neurais artificiais. O diagrama de blocos do PCBase é mostrado na Figura 5.2. Conforme se observa, após execução manual do bloco em vermelho, o restante do processo é feito de forma automatizada.

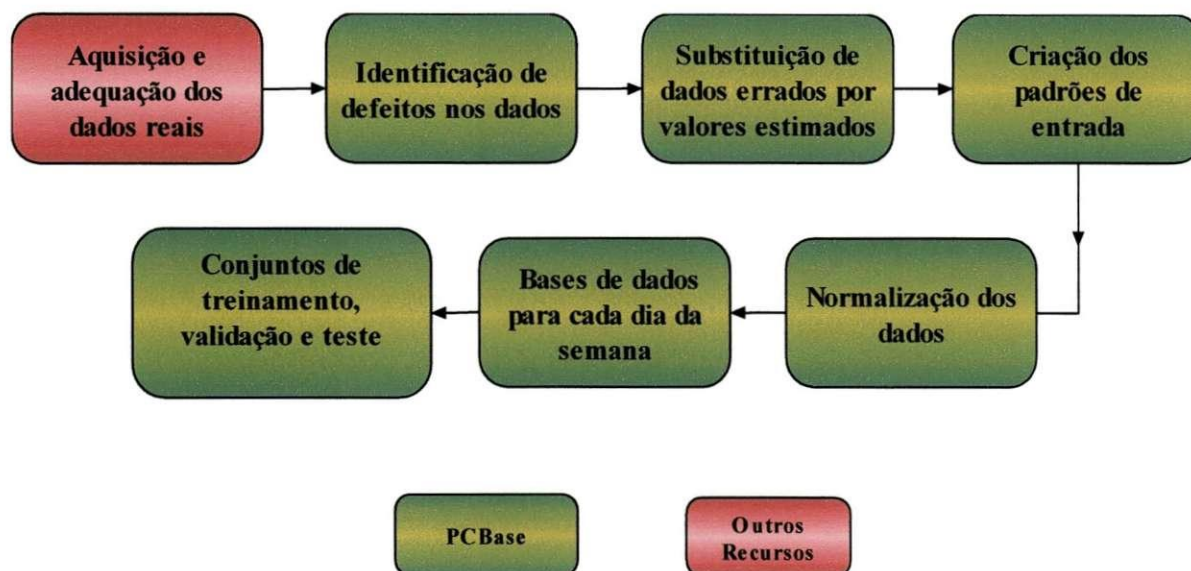


Figura 5.2 – Processo de construção de Bases de Dados do PCBase.

Algumas modificações foram realizadas no código do PCBase para adequá-lo às necessidades deste trabalho, resultando no diagrama mostrado na Figura 5.3. Detalhes sobre cada módulo são apresentados a seguir.



Figura 5.3 – Processo de construção da Base de Dados utilizada neste TCC.

5.1.1 Aquisição de Dados

O PCBase necessita de um arquivo no formato .txt. Como os dados fornecidos pela CEAL estavam no formato Excel, foi necessária a sua conversão para o formato desejado. A Figura 5.4 mostra como estes dados estão dispostos no arquivo.

```

      ○
      ○
5  06  2004  01  40  880.627
5  06  2004  01  30  888.867
5  06  2004  01  20  872.676
5  06  2004  01  10  866.123
5  06  2004  01  00  887.784
5  06  2004  00  50  917.072
5  06  2004  00  40  908.744
5  06  2004  00  30  889.671
5  06  2004  00  20  927.426
      ○
      ○

```

Figura 5.4 – Exemplo do documento fornecido pela CEAL.

A primeira coluna refere-se ao dia, a segunda coluna ao mês, a terceira coluna ao ano, a quarta coluna a hora, a quinta coluna aos minutos e a sexta coluna ao valor de potência aparente (em kVA) no respectivo horário.

5.1.2 Identificação de Defeitos

Após aquisição dos dados, segue-se para a etapa de busca de defeitos do conjunto de dados, que são: ausência de leitura, leitura zero na coluna correspondente à potência ou leitura repetida. O procedimento é o seguinte:

- Identificar os dados com defeitos.
- Substituir esses dados por valores nulos, considerando intervalo de 10 minutos, contados a partir do primeiro e último horários que possuem leituras corretas (Figura 5.5).
- Calcular a média de cada horário. Nesse cálculo não são contabilizados os dados substituídos por zeros.
- Identificar as leituras incorretas (valores muito acima ou muito abaixo da média). Essas leituras são identificadas tomando-se a média daquele horário

e em seguida, observando se os valores naquele horário estavam acima ou abaixo de 15% do valor médio, denominado de *faixa de segurança*.

5	06	2004	01	40	880.627
5	06	2004	01	30	888.867
5	06	2004	01	20	872.676
5	06	2004	01	10	866.123
5	06	2004	01	00	887.784
5	06	2004	00	50	917.072
5	06	2004	00	40	908.744
5	06	2004	00	30	889.671
5	06	2004	00	20	927.426
5	06	2004	00	10	942.561
4	06	2004	23	50	1025.779
4	06	2004	23	40	1051.247
4	06	2004	23	30	1076.172
4	06	2004	23	20	1119.723
4	06	2004	23	10	1170.006
4	06	2004	23	00	1266.904
4	06	2004	22	50	1340.633
4	06	2004	22	40	1413.439
4	06	2004	22	30	1472.703

Figura 5.5 – Exemplo de erro encontrado no PCBase.

5.1.3 Estimação de Dados

Os dados errados são substituídos por dados estimados através de interpolação quadrática. Supondo que haja ausência de leitura durante um dado intervalo, o procedimento adotado é o seguinte:

- Selecionar o primeiro valor registrado.
- Selecionar o último valor registrado.
- Calcular o valor médio desses valores, o qual já tem o valor médio do horário calculado previamente.
- Fazer interpolação para achar os pontos que faltam no intervalo.

O resultado desse processo é ilustrado na Figura 5.6, a qual refere-se a um dos piores casos encontrados. Observa-se que o PCBase localiza os pontos que estão fora do comportamento padrão e faz a interpolação, utilizando os pontos que são considerados coerentes.

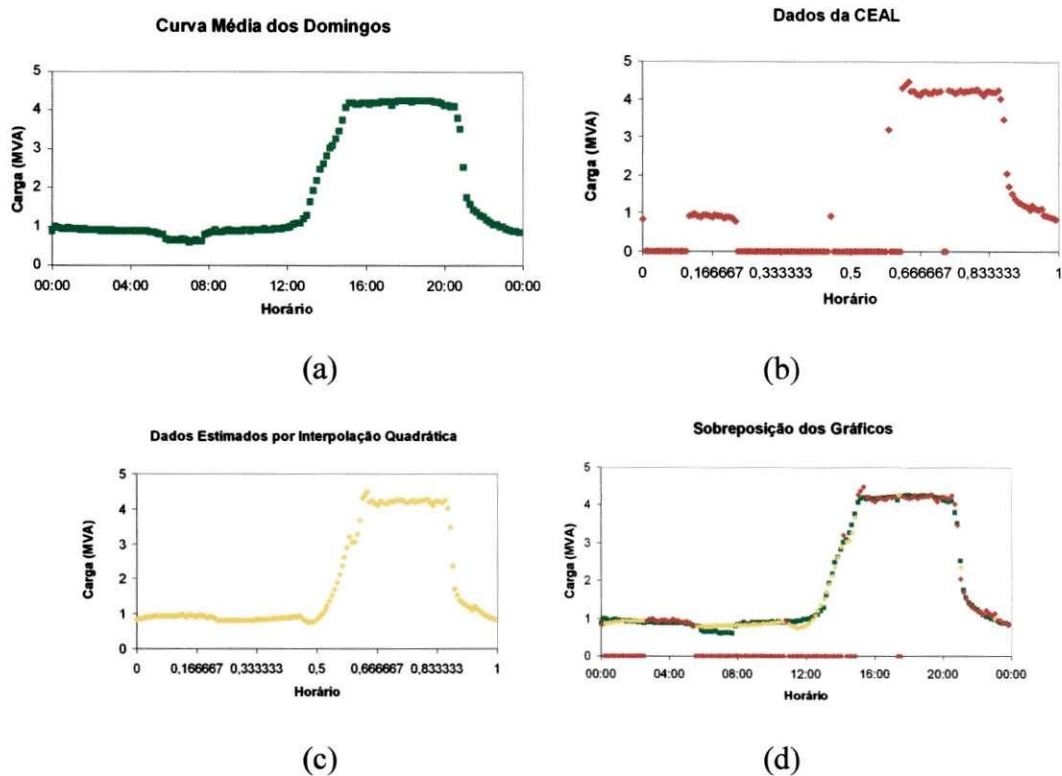


Figura 5.6 – Procedimento de interpolação.

(a) Curva média dos domingos. (b) Dados de um domingo fornecido pela CEAL. (c) Valores interpolados pelo PCBase. (d) Gráficos *b* e *c* superpostos.

A etapa final consiste no envio dos dados para o Microsoft Excel®, os quais são organizados, manualmente, por dia da semana. Em seguida, faz-se média de cada 10 minutos do dia, eliminando o dia da última semana, o qual será utilizado para a previsão. Na figura 5.7 é mostrada a disposição final dos dados.

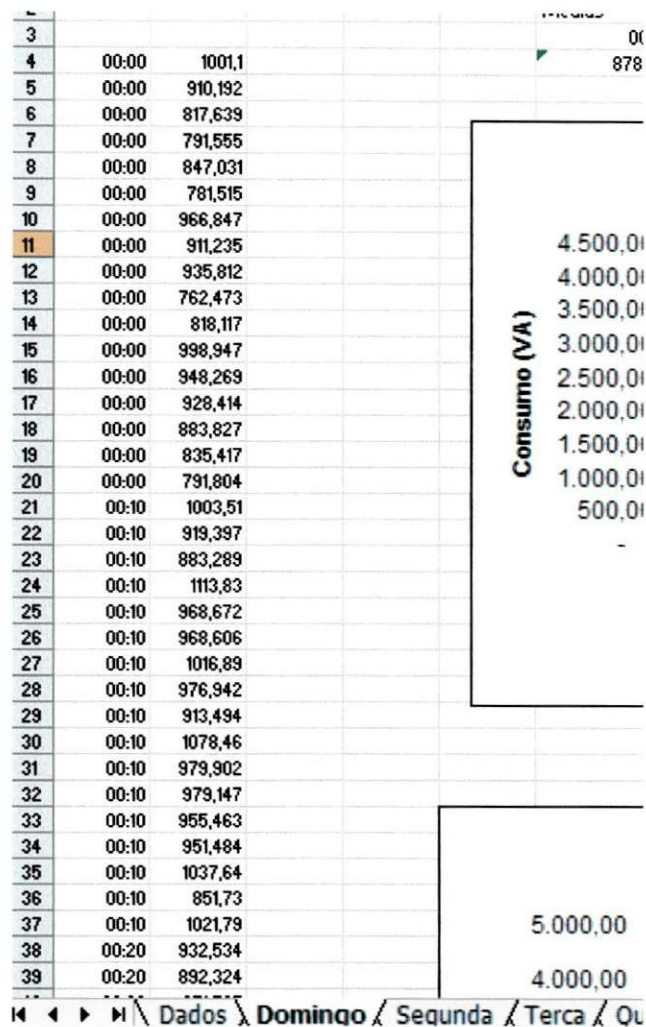


Figura 5.7 – Dados Importados para o Excel®.

5.2 Resultados

Após adequação, os dados foram separados por dias da semana, obtendo-se um total de: 17 domingos, 17 segundas-feiras, 18 terças-feiras, 18 quartas-feiras, 18 quintas-feiras, 17 sextas-feiras e 17 sábados. Em seguida, computou-se a média da leitura de cada dia, totalizando 144 leituras diárias. Escolheu-se a última semana para fazer a previsão de carga.

A partir das curvas médias dos dias da semana foram obtidas as equações para cada função de regressão através da linha de tendência, funcionalidade do Excel (os detalhes estão descritos no anexo).

As funções de regressão escolhidas foram as seguintes: linear, polinomial (até 6ª ordem), logarítmica, potencial e exponencial. Todas as funções foram calculadas no

Excel® através das linhas de tendência, as quais retornavam a equação correspondente a cada função desejada.

Os erros médios obtidos na semana selecionada para se fazer a previsão de carga são apresentados na Tabela 5.1. As equações de regressão e as curvas obtidas estão disponíveis no anexo.

Tabela 5.1 - Erro médio obtido na previsão.

Função	Erro Médio (%)
Linear	66
Logarítmica	77
Polinomial (2ª ordem)	65
Polinomial (3ª ordem)	29
Polinomial (4ª ordem)	26
Polinomial (5ª ordem)	28
Polinomial (6ª ordem)	21
Potencial	60
Exponencial	58

Observa-se que o melhor resultado foi obtido com regressão polinomial de 6ª ordem. Como era de se esperar, o esforço computacional acompanha a ordem de crescimento do polinômio. Portanto, na prática, é necessário se fazer uma avaliação da questão precisão versus esforço computacional.

Naturalmente, selecionou-se regressão polinomial de 6ª ordem para se fazer o estudo comparativo com o método baseado em RNA. Os resultados obtidos são apresentados na Tabela 5.2 e nas figuras apresentadas a seguir.

Análise dos gráficos e dos erros, evidencia a superioridade das RNAs. De um modo geral, constatou-se que os métodos baseados em regressão apresentam muita dificuldade em acompanhar as curvas de carga nas regiões relacionadas com os instantes de abertura e fechamento do *shopping center*, ou seja, nos instantes de mudança de tendência. Em tais momentos, o erro é significativo. (Melo Filho, 2005).

Tabela 5.2 - Erro médio obtido nas previsões por Regressão Polinomial e por RNA's
(Melo Filho, 2005).

Dia da Semana	Erro Médio	
	Regressão Polinomial (6ª ordem)	RNA
Domingo	19,03%	3,91%
Segunda	21,47%	3,71%
Terça	20,99%	3,92%
Quarta	21,24%	2,62%
Quinta	20,99%	3,07%
Sexta	20,53%	3,43%
Sábado	22,05%	3,37%

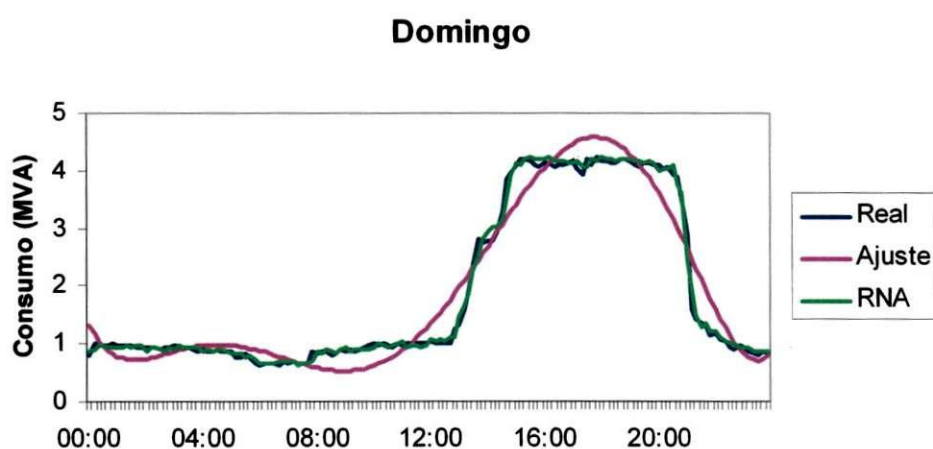


Figura 5.8 – Previsão de carga do domingo via regressão polinomial de 6ª ordem; rede neural artificial.

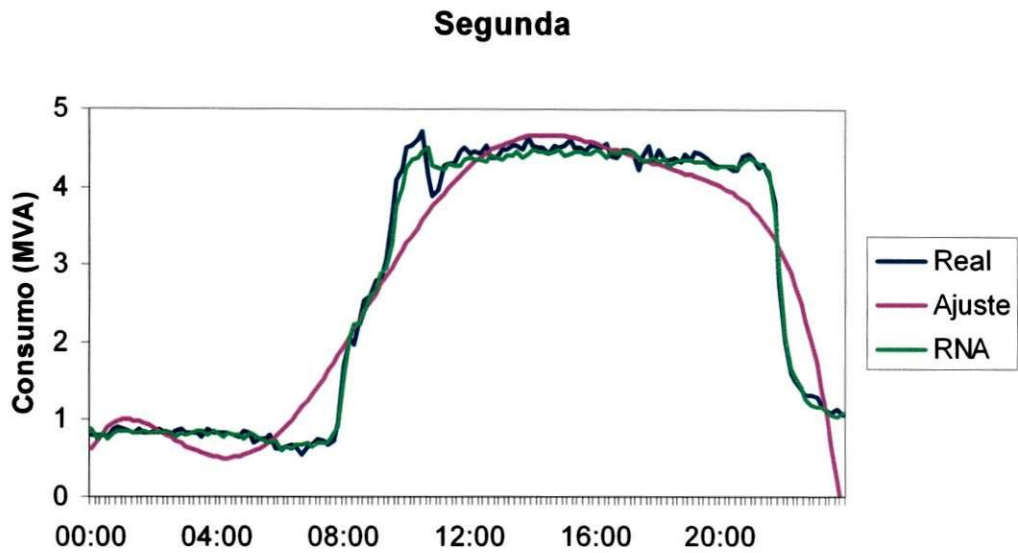


Figura 5.9 – Previsão de carga da segunda-feira via regressão polinomial de 6ª ordem; rede neural artificial.

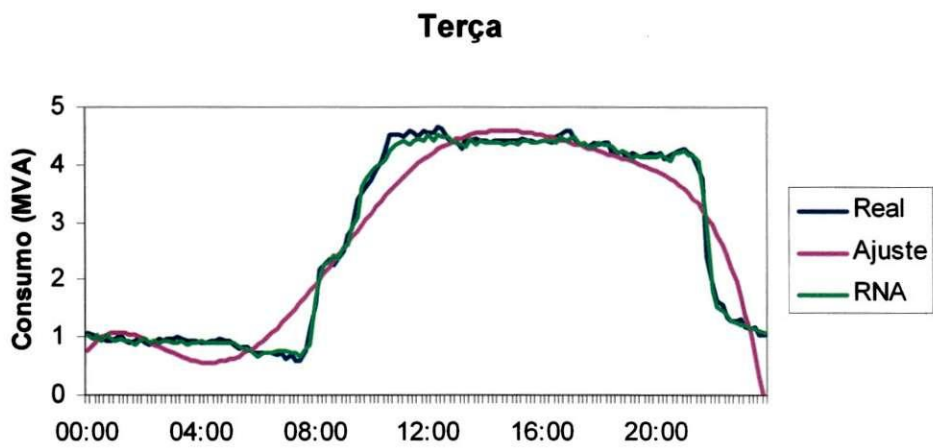


Figura 5.10 – Previsão de carga da terça-feira via regressão polinomial de 6ª ordem; rede neural artificial.

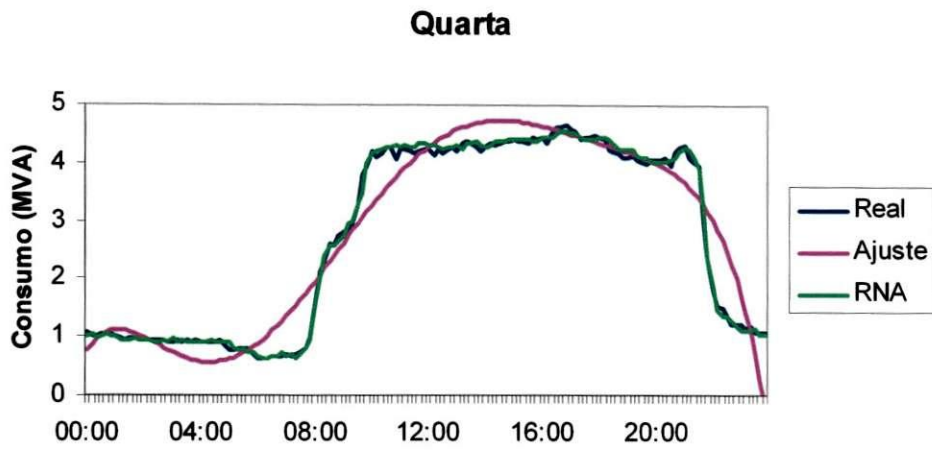


Figura 5.11 – Previsão de carga da quarta-feira via regressão polinomial de 6ª ordem; rede neural artificial.

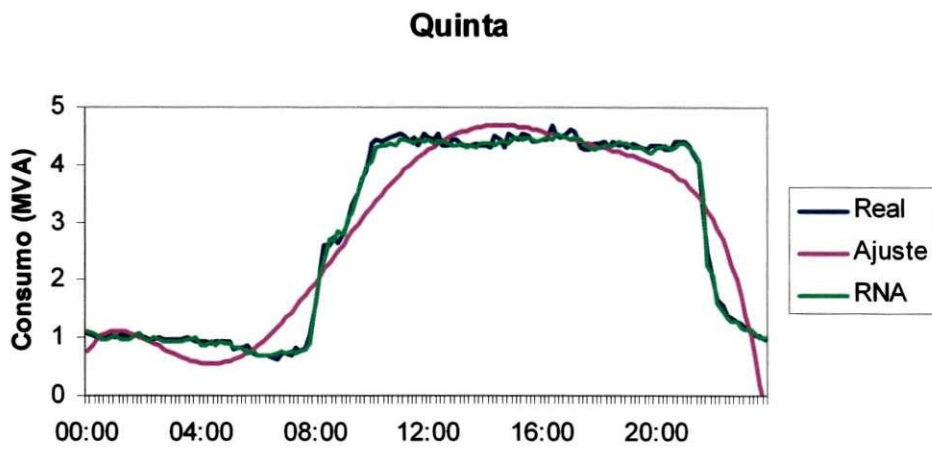


Figura 5.12 – Previsão de carga da quinta-feira via regressão polinomial de 6ª ordem; rede neural artificial.

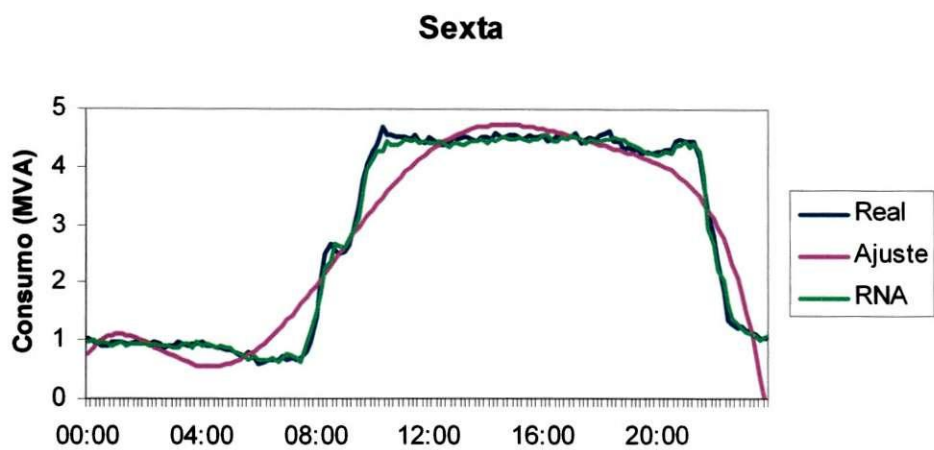


Figura 5.13 – Previsão de carga da sexta-feira via regressão polinomial de 6ª ordem; rede neural artificial.

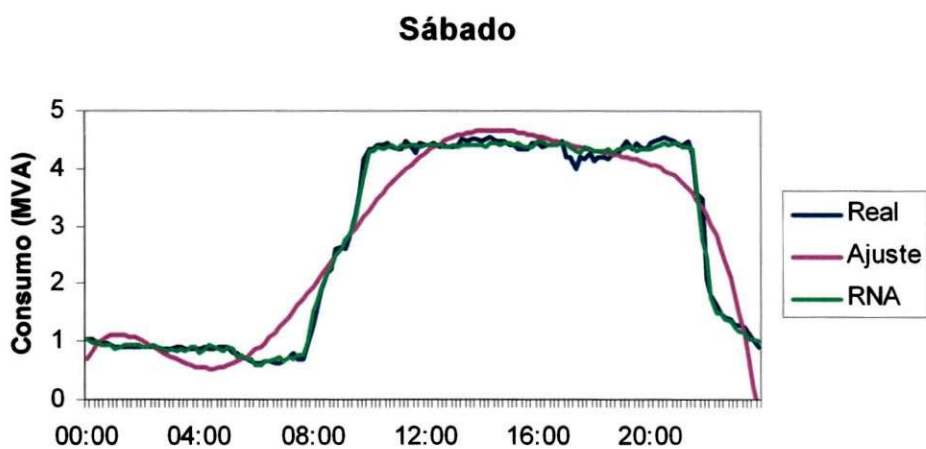


Figura 5.14 – Previsão de carga do sábado via regressão polinomial de 6ª ordem; rede neural artificial.

6 Conclusões

O estudo efetuado neste TCC ampliou os conhecimentos adquiridos durante o Curso de Graduação, visto que se avaliou um problema prático e atual das empresas de energia elétrica, em particular, das distribuidoras de energia. Além disso, conceitos adquiridos em diversas disciplinas foram aplicados conjuntamente, como por exemplo: Cálculo Numérico, Distribuição de Energia Elétrica e Programação.

Apesar da simplicidade do problema tratado, as simulações realizadas evidenciaram a superioridade das redes neurais artificiais em relação aos métodos baseados em regressão.

Finalmente, a ferramenta utilizada para fazer as implementações dos métodos de regressão, ou seja, o Microsoft Excel[®], mostrou-se ser bastante adequada para esse tipo de estudo.

7 Referências Bibliográficas

AL-ALAWI, S. M & ISLAM, S. M. Principles of electricity demand forecasting - Part I Methodologies. IEEE Power Engineering Journal, June, 1996.

ESTEVES, G. R. T. Modelos de previsão de carga de curto prazo. Dissertação de Mestrado, PUC-Rio, 2003.

GROSS, G. & GALIANA, F.D. Short-term load forecasting, IEEE, 1987.

LATORRE, M. R. D. O. & CARDOSO, M. R. A. Análise de séries temporais em epidemiologia: uma introdução sobre os aspectos metodológicos. Revista Brasileira de Epidemiologia, v. 4, n. 3, São Paulo, novembro de 2001.

MELO FILHO, A. C. M. Previsão de carga utilizando redes neurais. Relatório PIBIC, Universidade Federal de Campina Grande, 2005.

MICROSOFT, Guia do Usuário do MicroSoft Excel, 2003.

SOUZA, B. A. Distribuição de Energia Elétrica (apostila), 1997.

PCBase, versão 1.0, Manual do Usuário, Universidade Federal de Campina Grande, 2005.

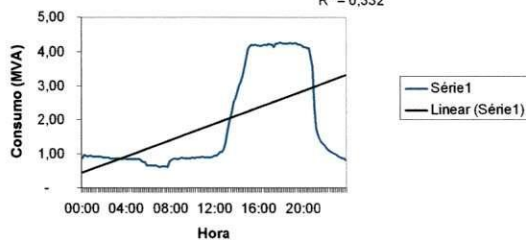
Anexos

Domingo

Domingo - Linear

$$y = 19,872x + 435,55$$

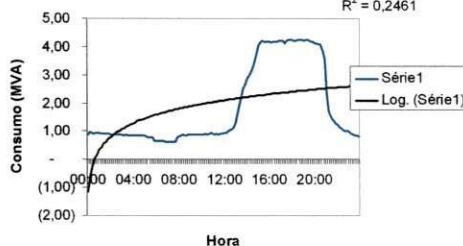
$$R^2 = 0,332$$



Domingo - Logarítmico

$$y = 756,54 \ln(x) - 1144,9$$

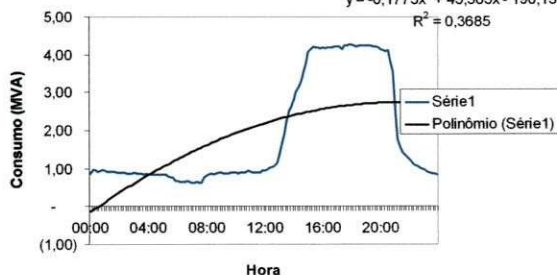
$$R^2 = 0,2461$$



Domingo - Polinomial 2a

$$y = -0,1773x^2 + 45,585x - 190,13$$

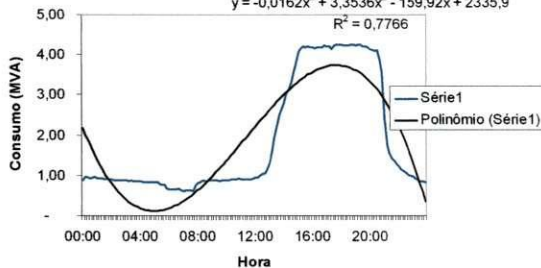
$$R^2 = 0,3685$$



Domingo - Polinomial 3a

$$y = -0,0162x^3 + 3,3536x^2 - 159,92x + 2335,9$$

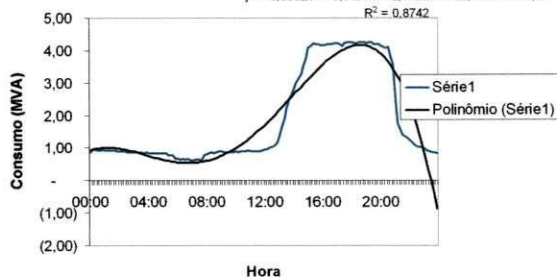
$$R^2 = 0,7766$$



Domingo - Polinomial 4a

$$y = -0,0002x^4 + 0,0473x^3 - 2,5789x^2 + 32,793x + 895,27$$

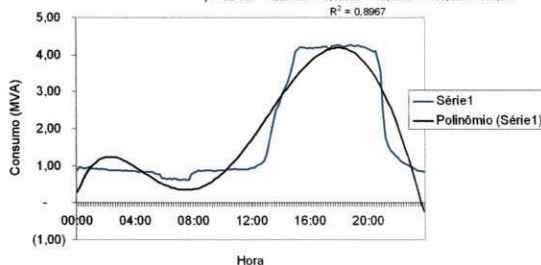
$$R^2 = 0,8742$$



Domingo - Polinomial 5a

$$y = 3E-06x^5 - 0,0013x^4 + 0,1833x^3 - 10,017x^2 + 189,02x + 103,54$$

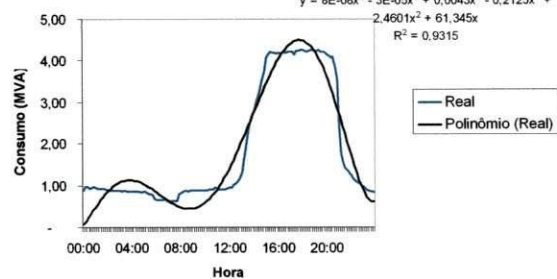
$$R^2 = 0,8967$$



Domingo - Polinomial 6a

$$y = 8E-08x^6 - 3E-05x^5 + 0,0043x^4 - 0,2125x^3 + 2,4601x^2 + 61,345x$$

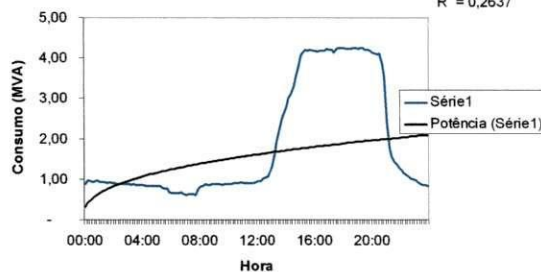
$$R^2 = 0,9315$$



Domingo - Potencial

$$y = 310,59x^{0,3837}$$

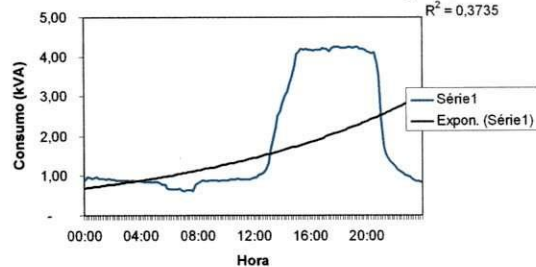
$$R^2 = 0,2637$$



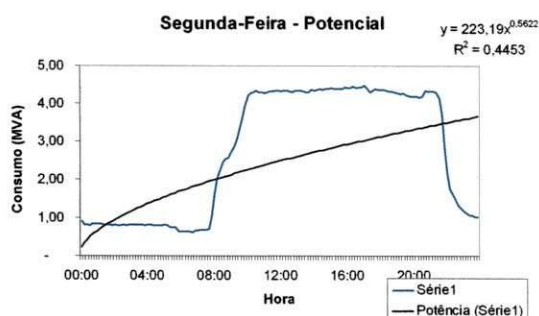
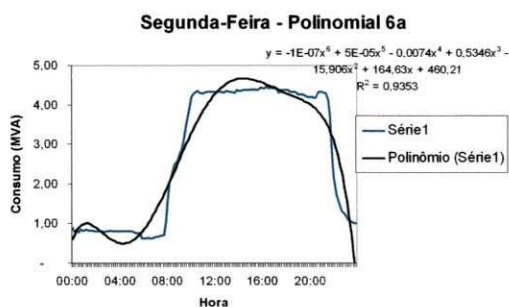
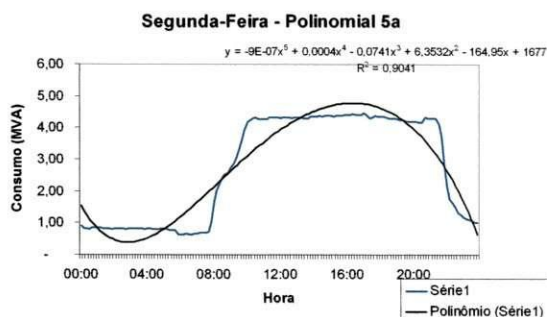
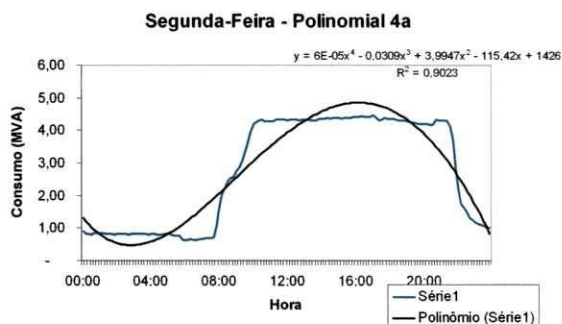
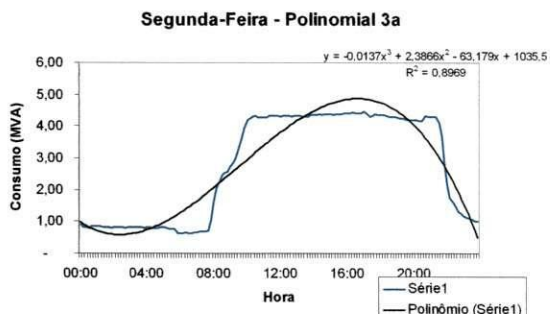
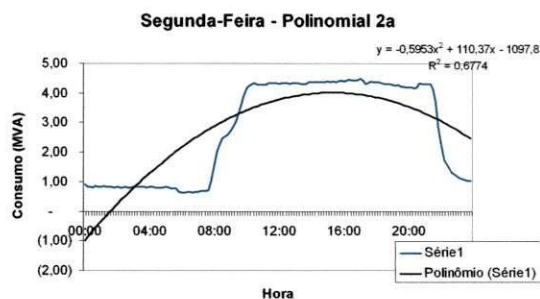
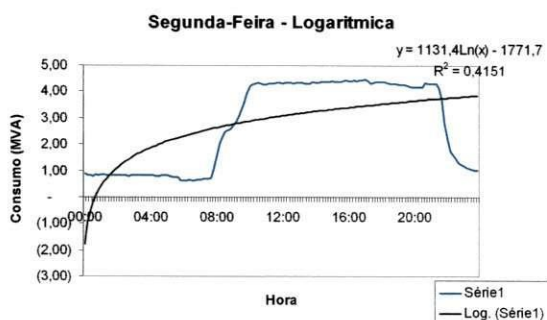
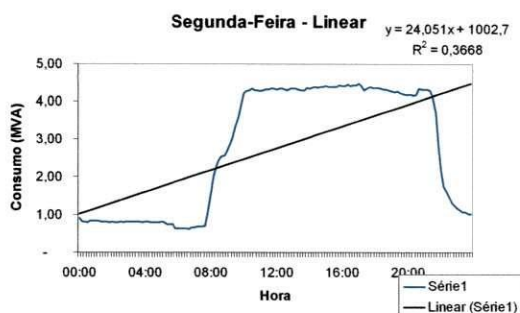
Domingo - Exponencial

$$y = 680,03e^{0,0103x}$$

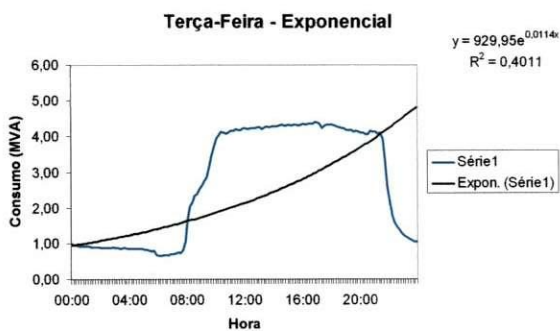
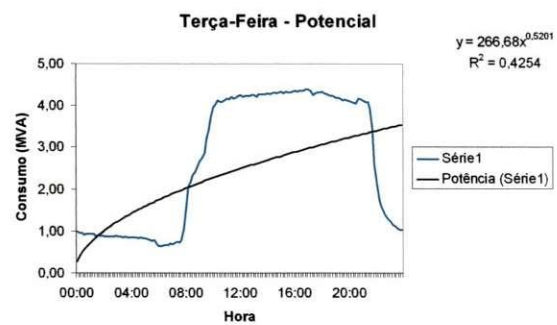
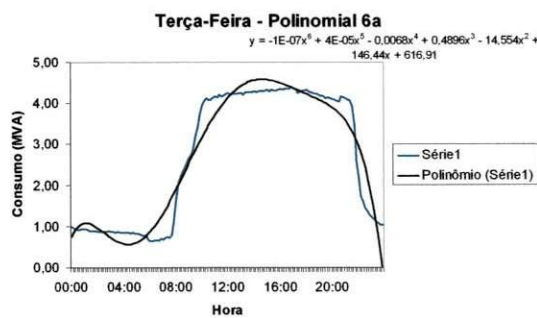
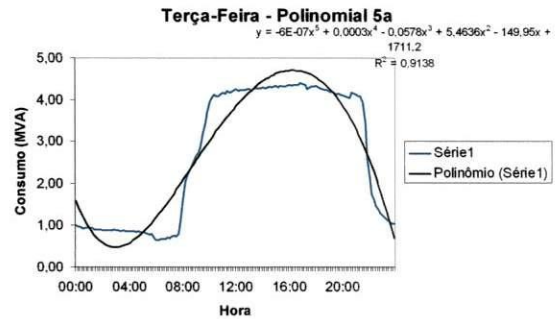
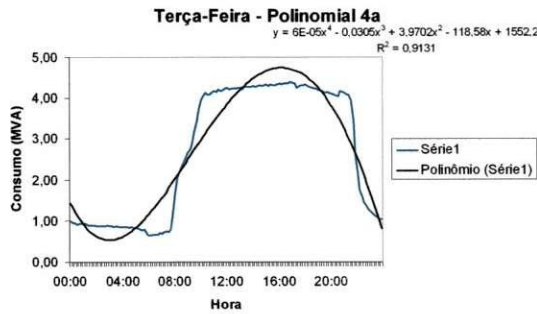
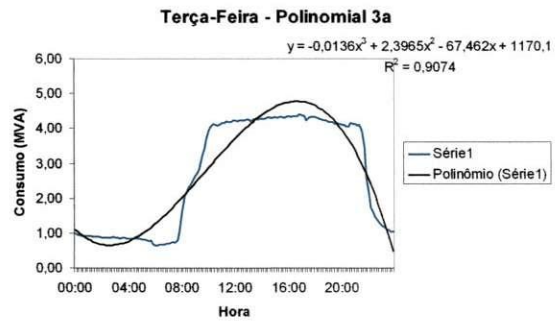
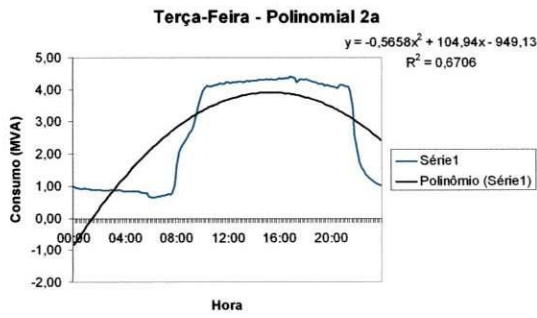
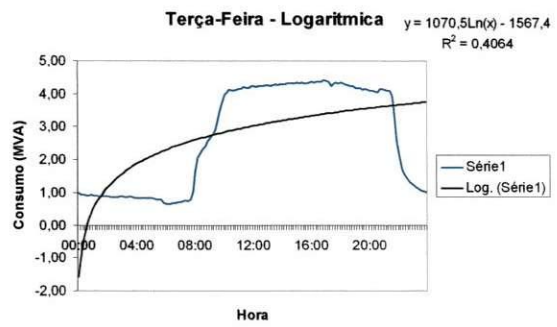
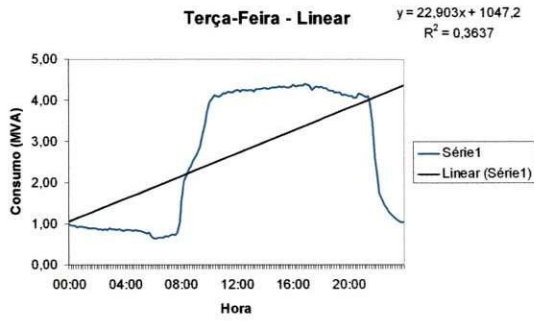
$$R^2 = 0,3735$$



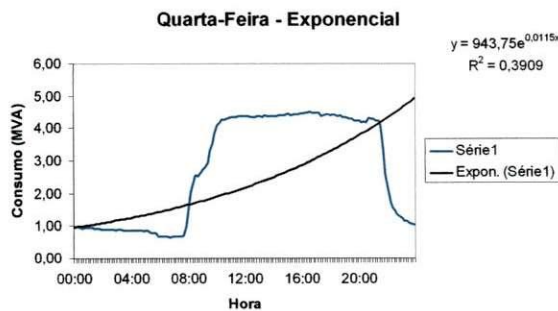
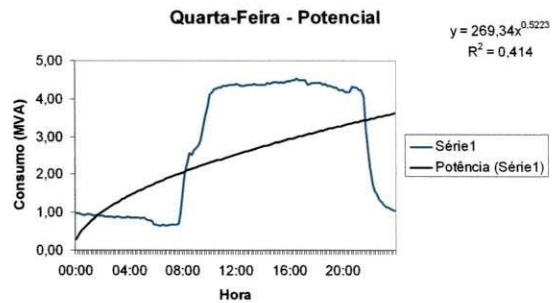
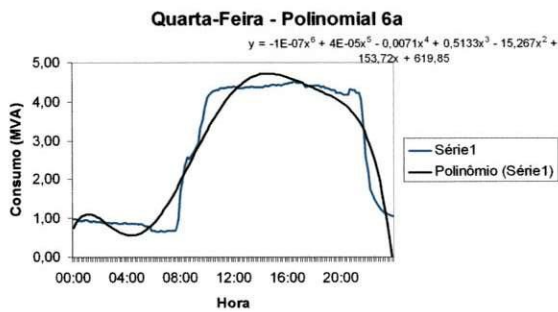
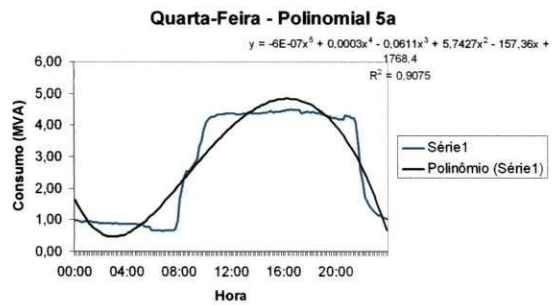
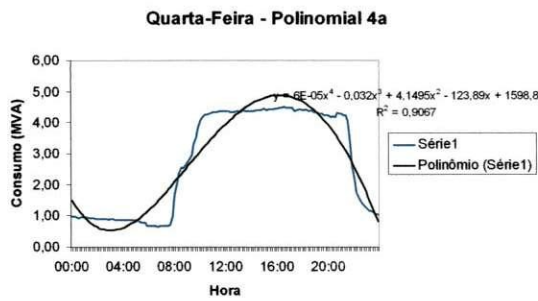
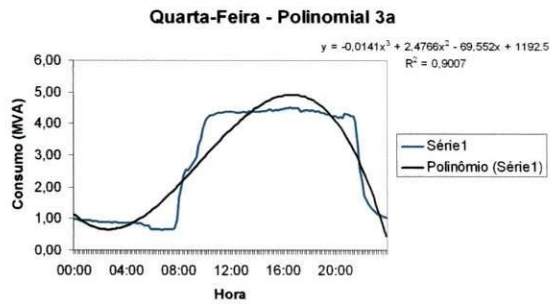
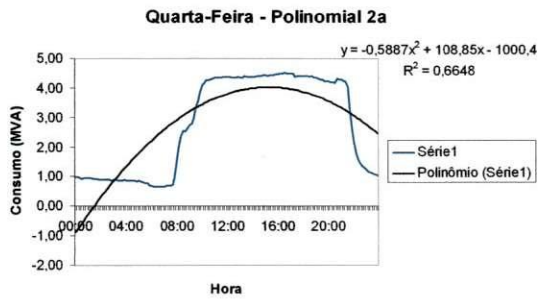
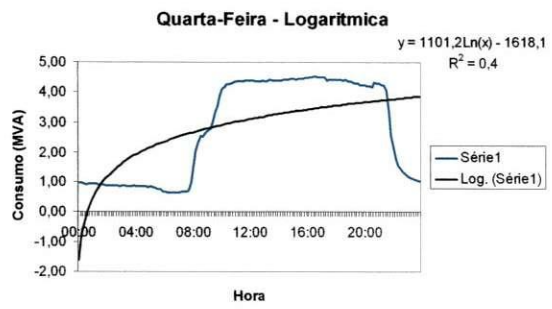
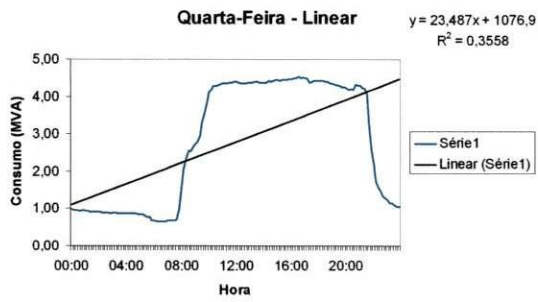
Segunda-Feira



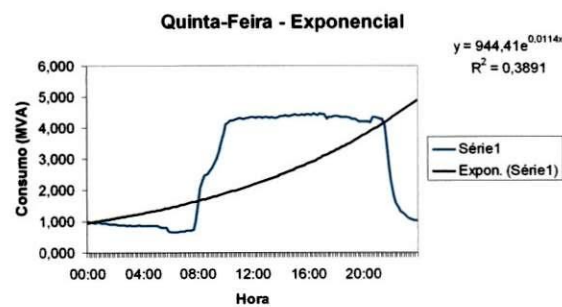
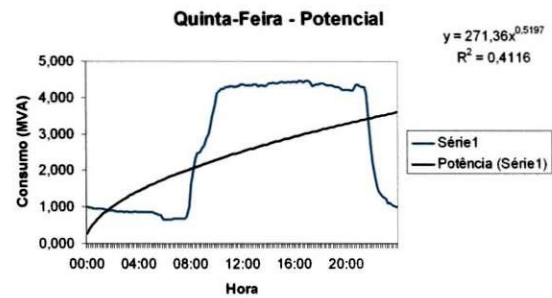
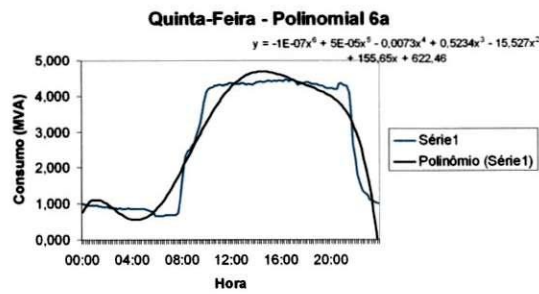
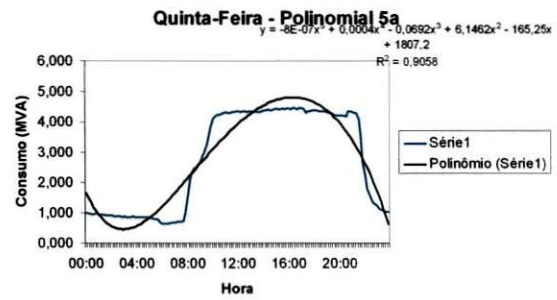
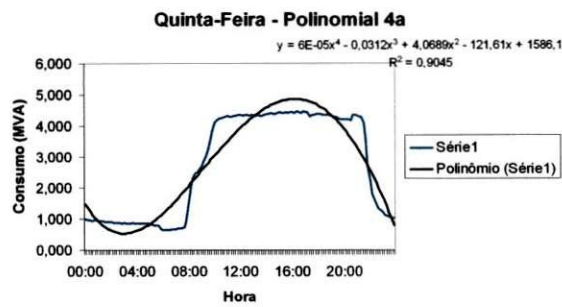
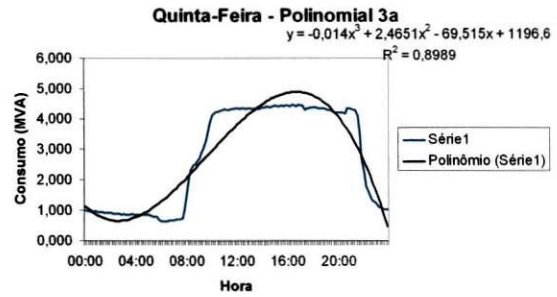
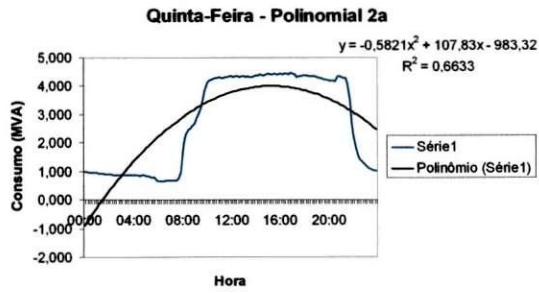
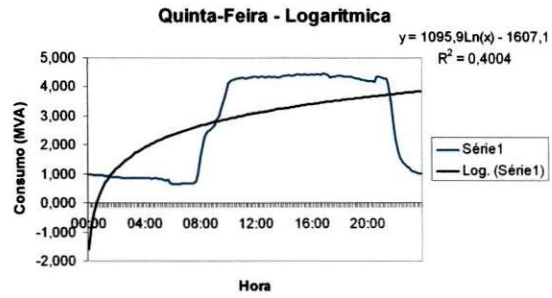
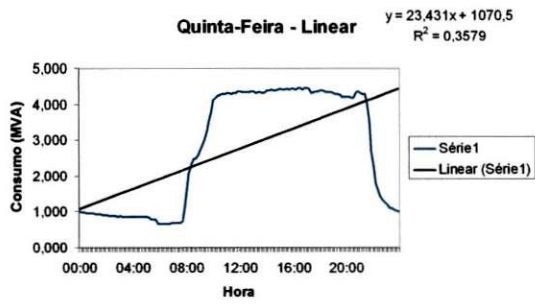
Terça-Feira



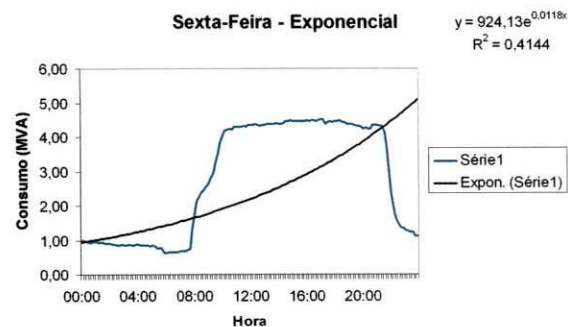
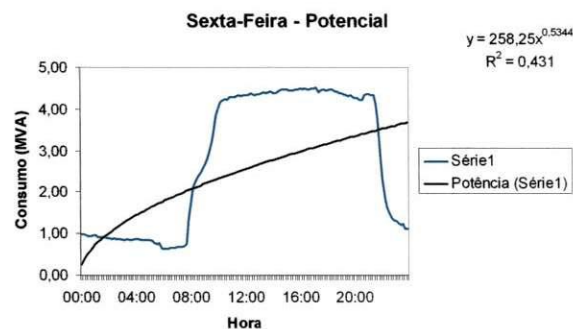
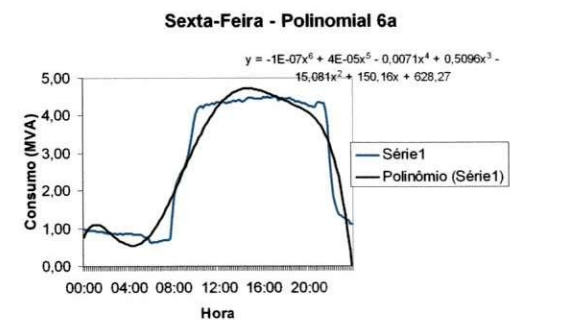
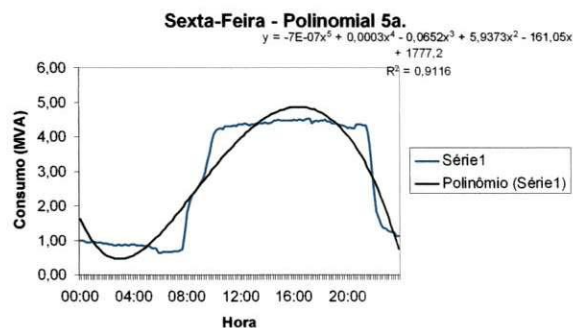
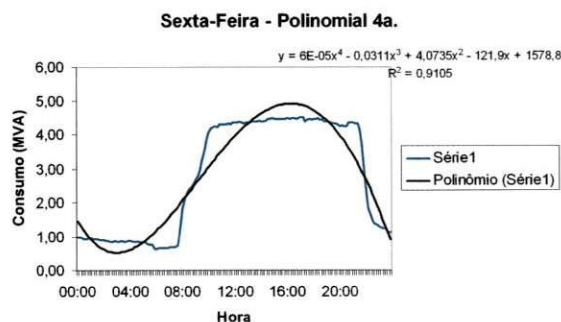
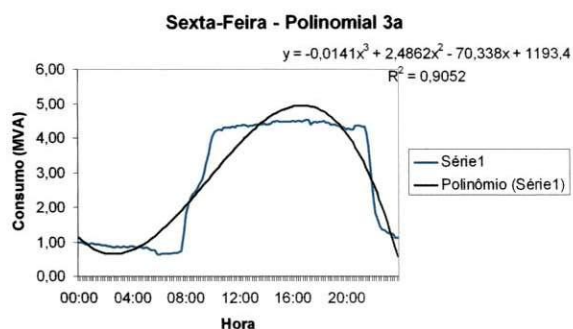
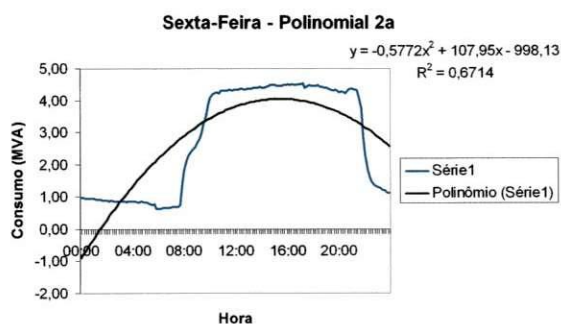
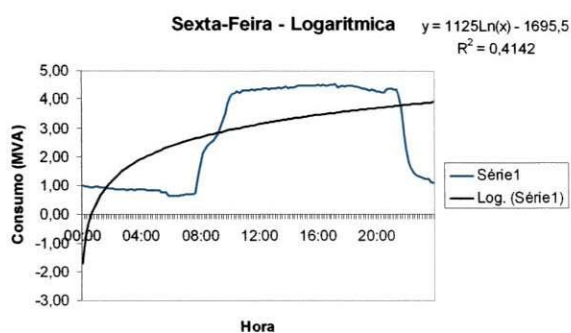
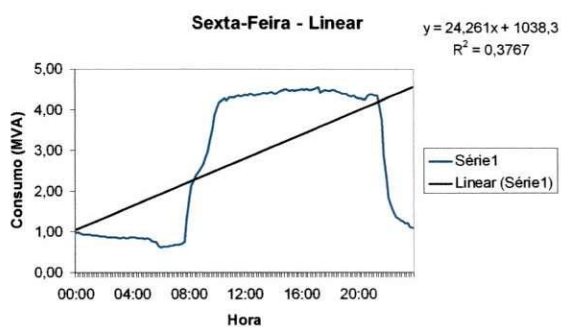
Quarta-Feira



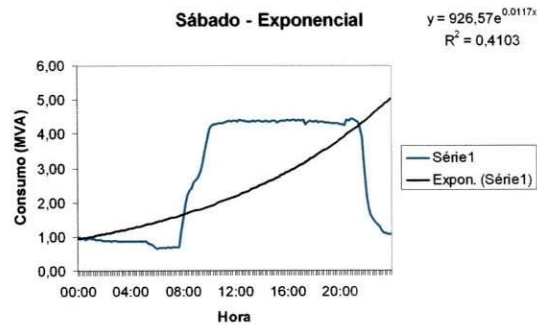
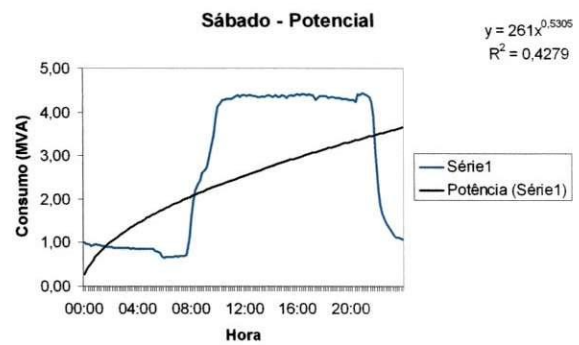
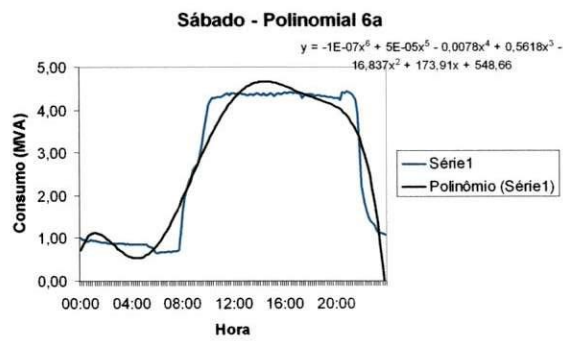
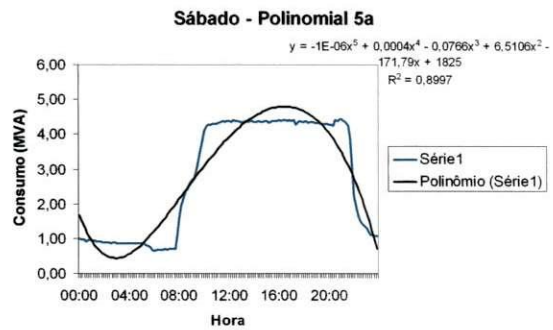
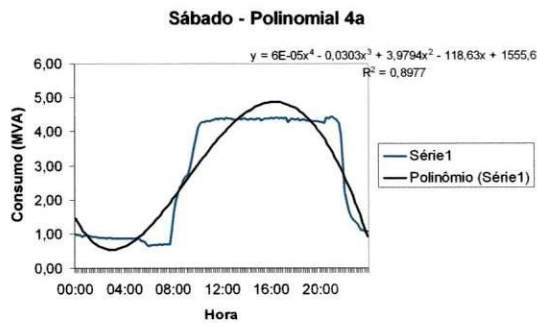
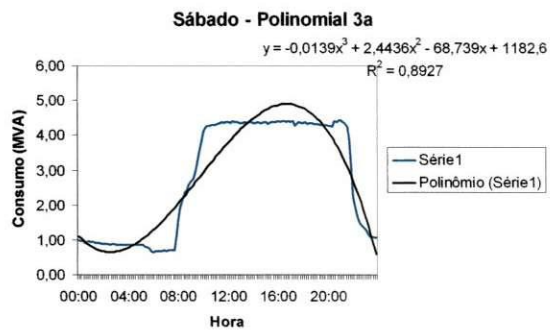
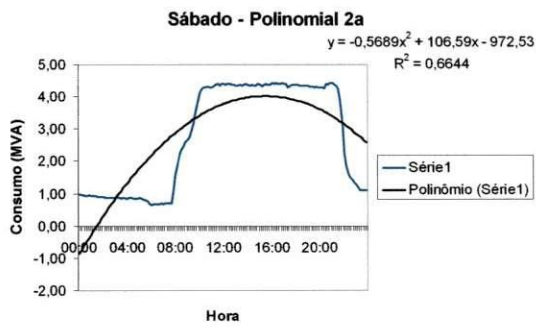
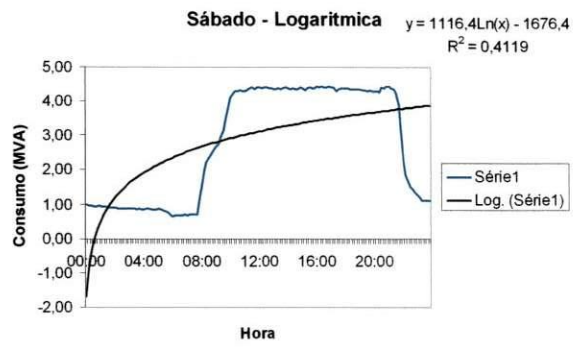
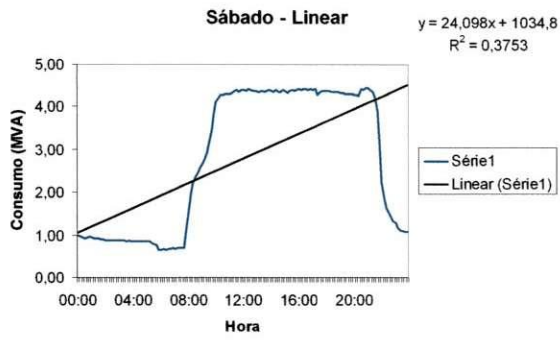
Quinta-Feira



Sexta-Feira



Sábado



Semana de previsão - 24/9/2004 - 30/9/2004

Domingo - 26/9/2004		Previsão												
Hora	Consumo real	Linear	Ero	Logaritmica	Ero	Polinomial 2a Ordem	Ero	Polinomial 3a. Ordem	Ero	Polinomial 4a. Ordem	Ero	Polinomial 5a. Ordem	Ero	
00:00	0,00	791,804	435,55	44,89%		-190,13	124,01%	2335,90	195,01%	925,53	16,89%	282,72	64,23%	
00:10	1,00	1021,79	455,42	55,43%	-1144,90	212,05%	-144,72	14,16%	2179,32	113,26%	950,91	6,94%	442,96	56,65%
00:20	2,00	990,592	475,29	52,02%	-620,51	162,64%	-99,67	110,06%	2029,34	104,86%	971,70	1,91%	585,30	40,91%
00:30	3,00	969,895	495,17	48,95%	-313,76	132,26%	-54,97	105,67%	1895,89	94,42%	988,15	1,87%	710,77	26,72%
00:40	4,00	967,288	515,04	46,75%	-96,11	109,94%	-10,63	101,10%	1748,94	80,80%	1000,53	3,44%	820,36	15,19%
00:50	5,00	991,283	534,91	46,04%	72,70	92,67%	33,36	96,63%	1618,12	63,23%	1009,11	1,80%	915,04	7,63%
01:00	6,00	990,826	554,79	43,44%	210,64	78,52%	77,00	92,16%	1493,61	52,28%	1014,14	3,40%	995,75	1,82%
01:10	7,00	993,611	574,65	41,89%	327,26	66,73%	120,28	87,77%	1375,23	39,81%	1015,86	3,26%	1063,39	8,11%
01:20	8,00	973,245	594,53	38,91%	428,28	55,99%	163,20	83,23%	1262,88	29,76%	1014,53	4,24%	1118,85	14,96%
01:30	9,00	961,657	614,40	36,12%	517,39	46,21%	205,77	78,61%	1166,45	20,23%	1010,38	5,04%	1162,98	20,91%
01:40	10,00	945,386	634,27	32,91%	597,10	36,84%	247,99	73,77%	1059,86	11,69%	1003,64	6,18%	1196,60	26,57%
01:50	11,00	947,299	654,14	30,95%	669,20	29,36%	289,85	69,40%	961,00	1,45%	994,05	4,99%	1220,51	28,84%
02:00	12,00	942,564	674,01	28,49%	735,03	22,02%	331,36	64,84%	871,78	7,51%	983,32	4,32%	1235,49	31,08%
02:10	13,00	947,665	693,89	26,76%	795,59	16,05%	372,91	60,69%	788,11	16,84%	970,18	2,38%	1242,28	31,09%
02:20	14,00	937,364	713,76	23,85%	851,65	9,14%	413,31	55,91%	709,87	24,27%	955,34	1,92%	1241,60	32,46%
02:30	15,00	904,333	733,63	18,93%	903,85	0,12%	453,75	49,86%	636,99	29,61%	939,00	3,76%	1234,15	36,38%
02:40	16,00	922,993	753,50	16,36%	952,67	3,22%	493,84	46,50%	569,35	38,32%	921,36	0,18%	1220,59	32,24%
02:50	17,00	926,448	773,37	16,52%	998,54	7,78%	533,58	42,41%	506,86	45,29%	902,63	2,57%	1201,56	29,70%
03:00	18,00	969,886	793,25	18,21%	1041,78	7,41%	572,95	40,93%	449,43	53,66%	882,99	8,96%	1177,68	21,43%
03:10	19,00	941,187	813,12	13,61%	1082,69	15,03%	611,98	34,98%	396,95	57,82%	862,64	8,35%	1149,56	22,14%
03:20	20,00	905,23	832,99	7,98%	1121,49	23,89%	650,65	28,12%	349,34	61,41%	841,75	7,01%	1117,76	23,48%
03:30	21,00	927,642	852,86	8,06%	1158,40	24,88%	688,97	25,73%	306,49	66,96%	820,49	11,55%	1082,82	16,73%
03:40	22,00	893,089	872,73	2,28%	1193,60	33,65%	726,93	18,61%	268,30	69,96%	799,05	10,53%	1045,27	17,04%
03:50	23,00	873,756	892,61	2,16%	1227,23	40,45%	764,53	12,50%	234,69	73,14%	777,58	11,01%	1005,62	15,09%
04:00	24,00	914,33	912,48	0,20%	1259,42	37,74%	801,79	12,31%	205,54	77,52%	756,24	17,29%	964,34	5,47%
04:10	25,00	861,693	932,35	8,20%	1290,31	49,74%	838,68	2,67%	180,78	79,02%	735,19	14,68%	921,88	6,98%
04:20	26,00	867,864	952,22	9,72%	1319,98	52,10%	875,23	0,85%	160,28	81,53%	714,59	17,66%	878,69	1,25%
04:30	27,00	912,883	972,09	6,49%	1349,53	47,72%	911,41	0,16%	143,97	84,23%	694,56	23,92%	835,17	8,51%
04:40	28,00	878,373	991,97	12,93%	1376,05	56,66%	947,25	7,84%	131,74	85,00%	675,26	23,12%	791,71	9,87%
04:50	29,00	874,474	1011,84	15,71%	1402,59	60,39%	982,73	12,38%	123,50	85,88%	656,81	24,89%	748,69	14,38%

Exemplo da Planilha

Construindo gráficos com linhas de tendência no Microsoft Excel[®] para estudos de previsão de carga via regressão.

1. Inicialmente, calculam-se as médias das demandas para cada hora da aquisição.
2. Em seguida, clica-se no ícone do *Assistente de gráfico* na barra de ferramentas (Figura 1).

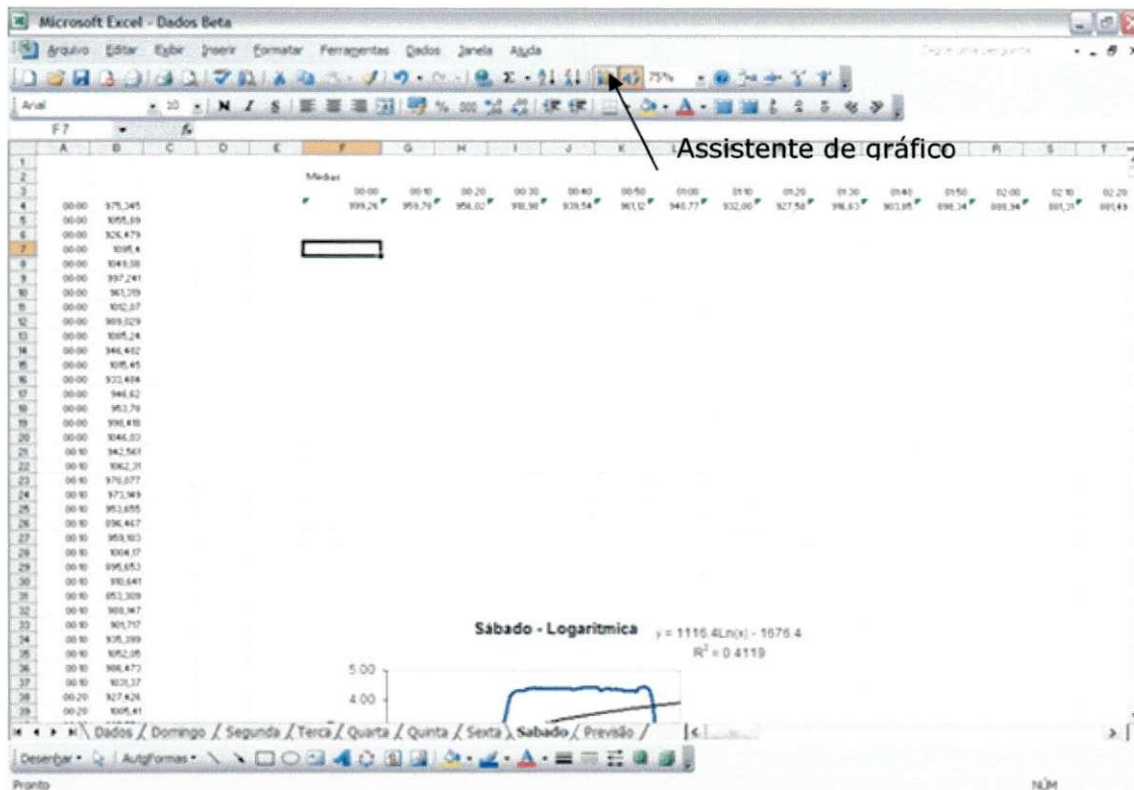


Figura 1.

3. Segue-se para a janela da Figura 2, a qual mostra os tipos de gráficos disponíveis. Neste caso, utilizam-se os gráficos do tipo *linha com tendências de tempo ou categoria*.

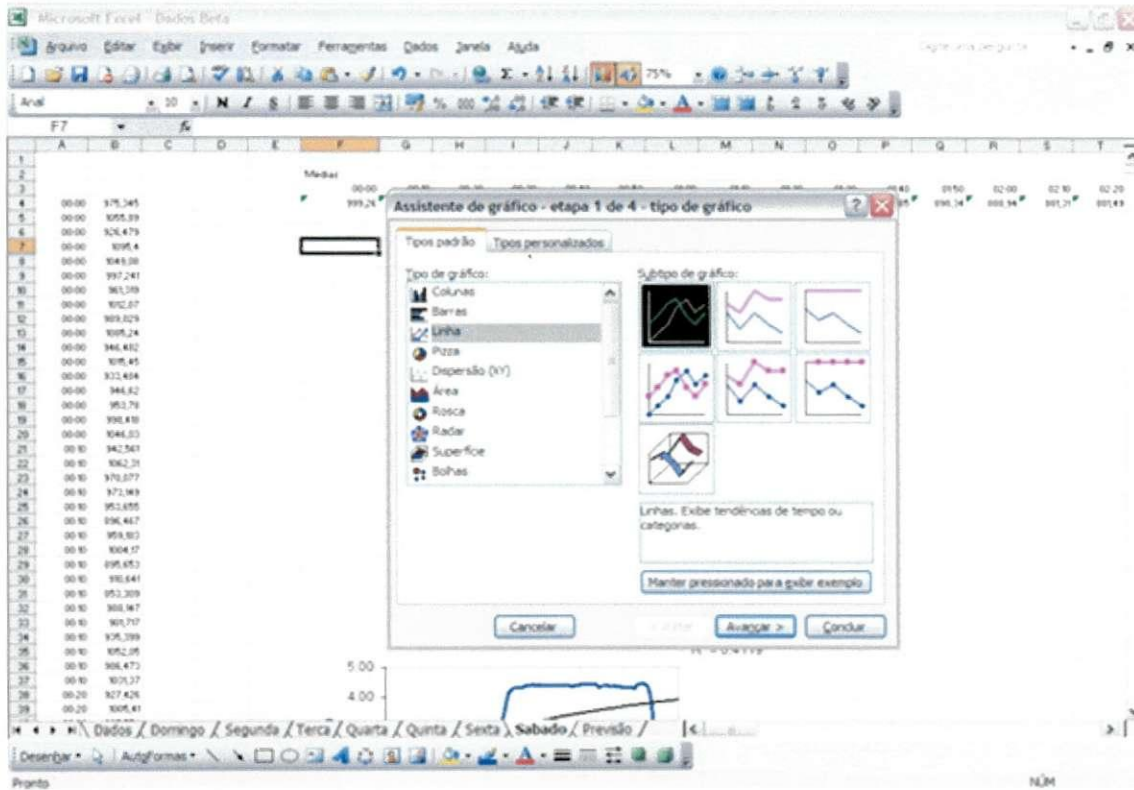


Figura 2.

4. Clica-se em avançar. Surge a janela da Figura 3, na qual seleccionam-se os dados de origem. No exemplo, clica-se em adicionar uma nova série.

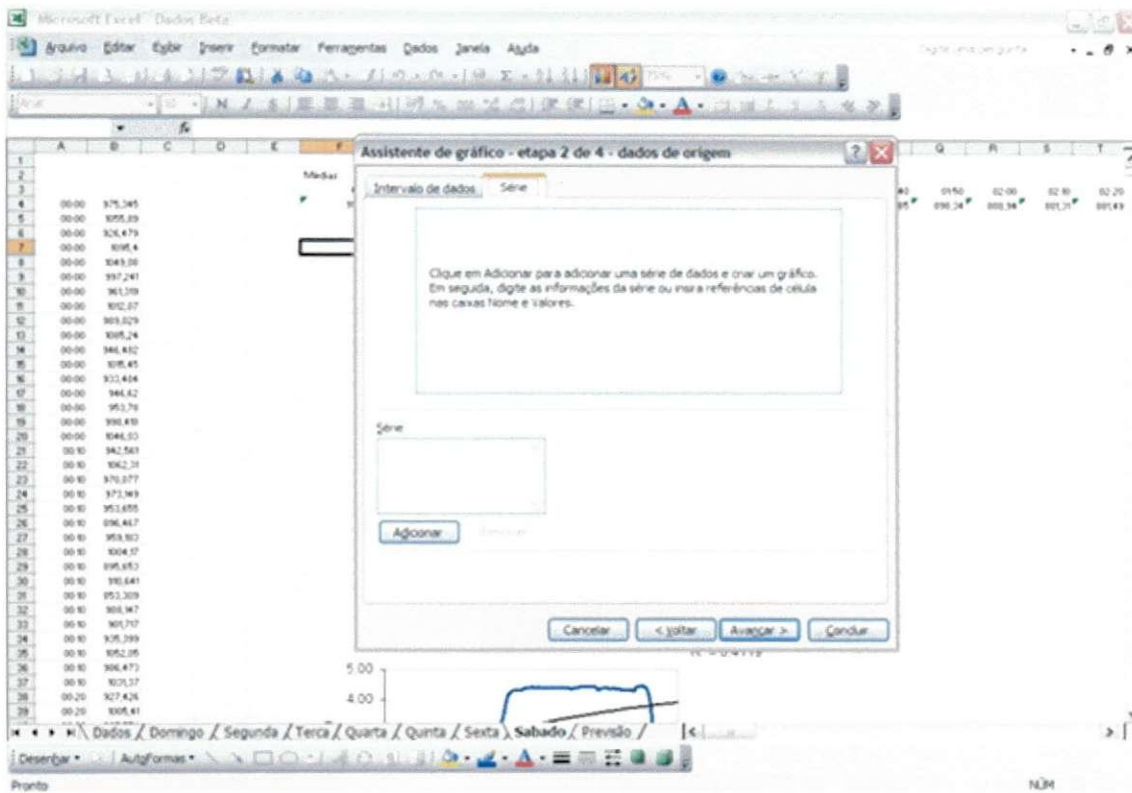


Figura 3.

- Escolhe-se o campo *Rótulo do eixo das categorias (X)* como sendo os valores das horas do dia e o campo *Valores*, como sendo os valores das demandas médias. No campo *Nome* coloca-se o nome que desejado na legenda do gráfico (Figura 4).

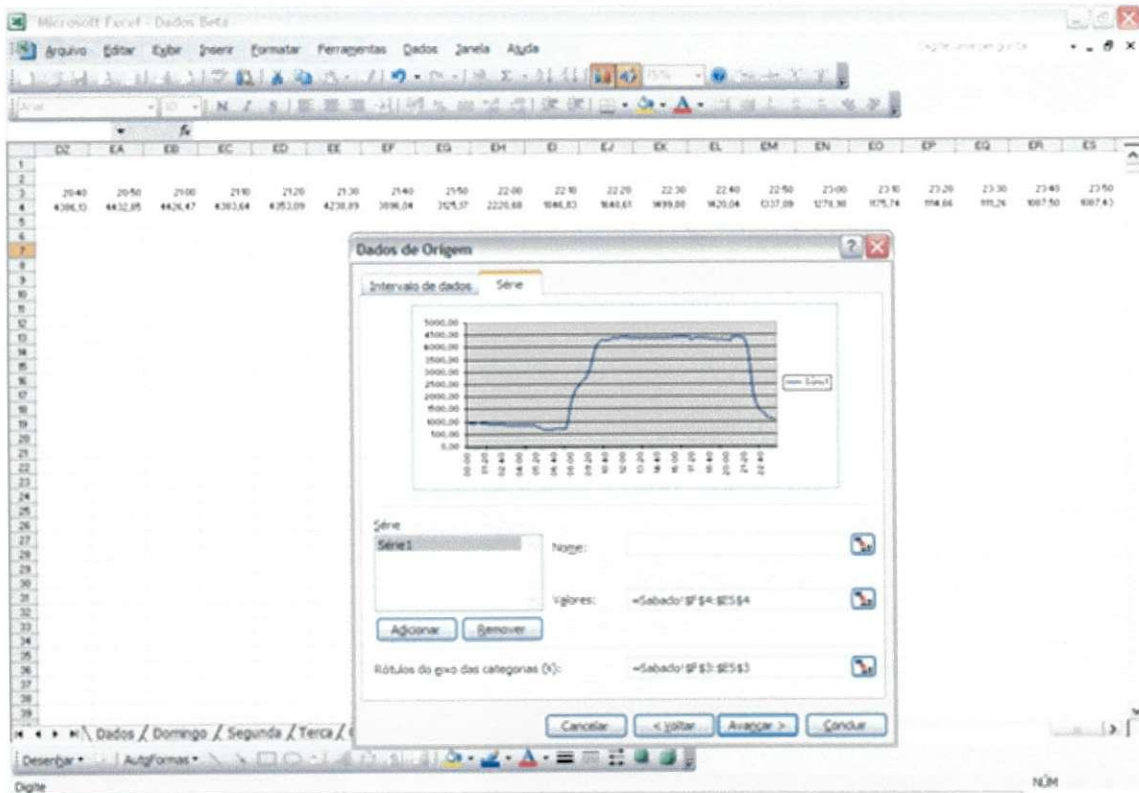


Figura 4.

6. Feita a seleção dos dados de origem, clica-se em *concluir*. Em seguida, podem-se manipular as opções visuais do gráfico: modificar cores, adicionar títulos, atribuir nomes aos eixos etc. (Figura 5).

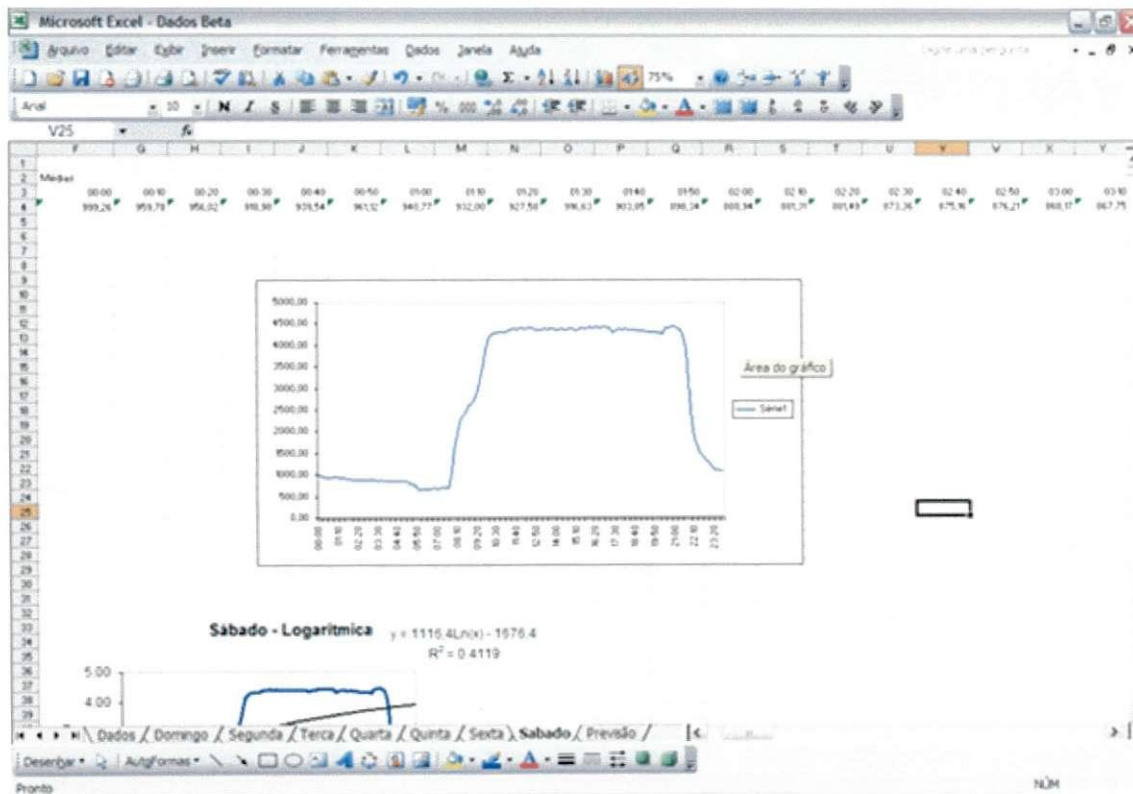


Figura 5.

7. Em seguida, clica-se com o botão esquerdo do mouse sobre a linha do gráfico. Depois, clica-se no botão direito do mouse. Aparecerão as opções mostradas na Figura 6. Seleciona-se a opção *Adicionar linha de tendências*.

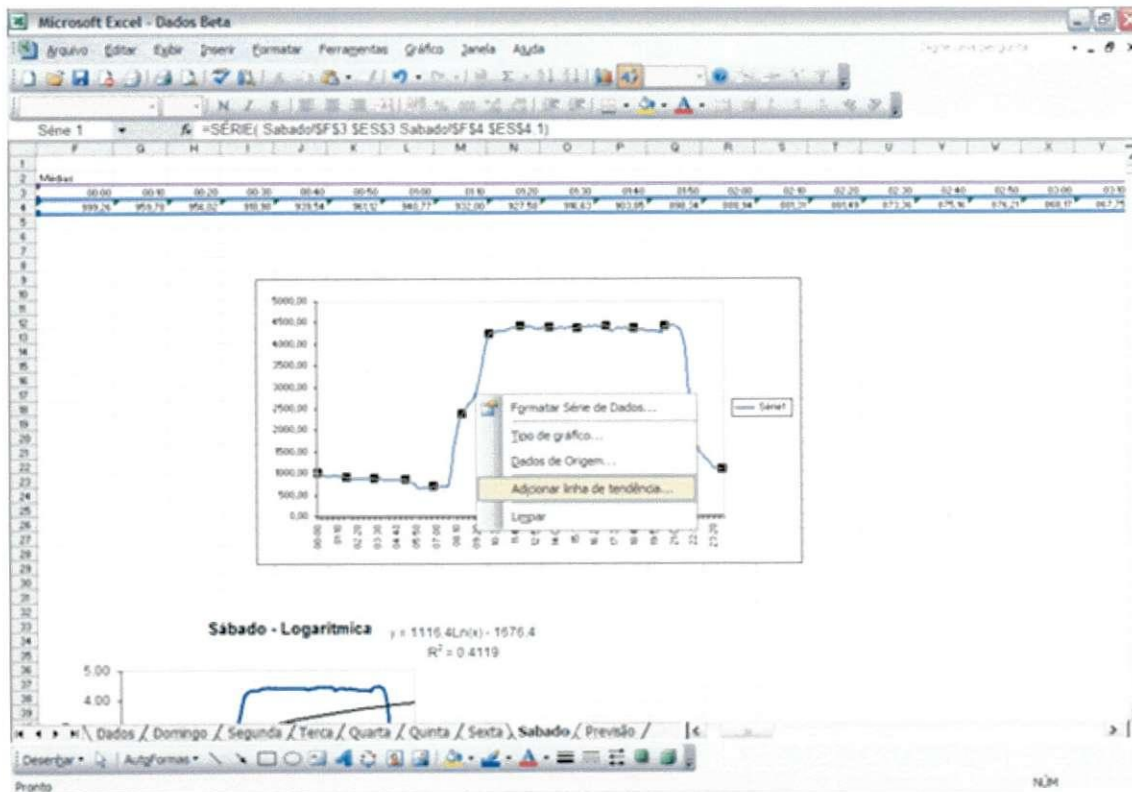


Figura 6.

8. A tela seguinte mostra os tipos de função de regressão que se pode utilizar. No exemplo utilizou-se a função linear.

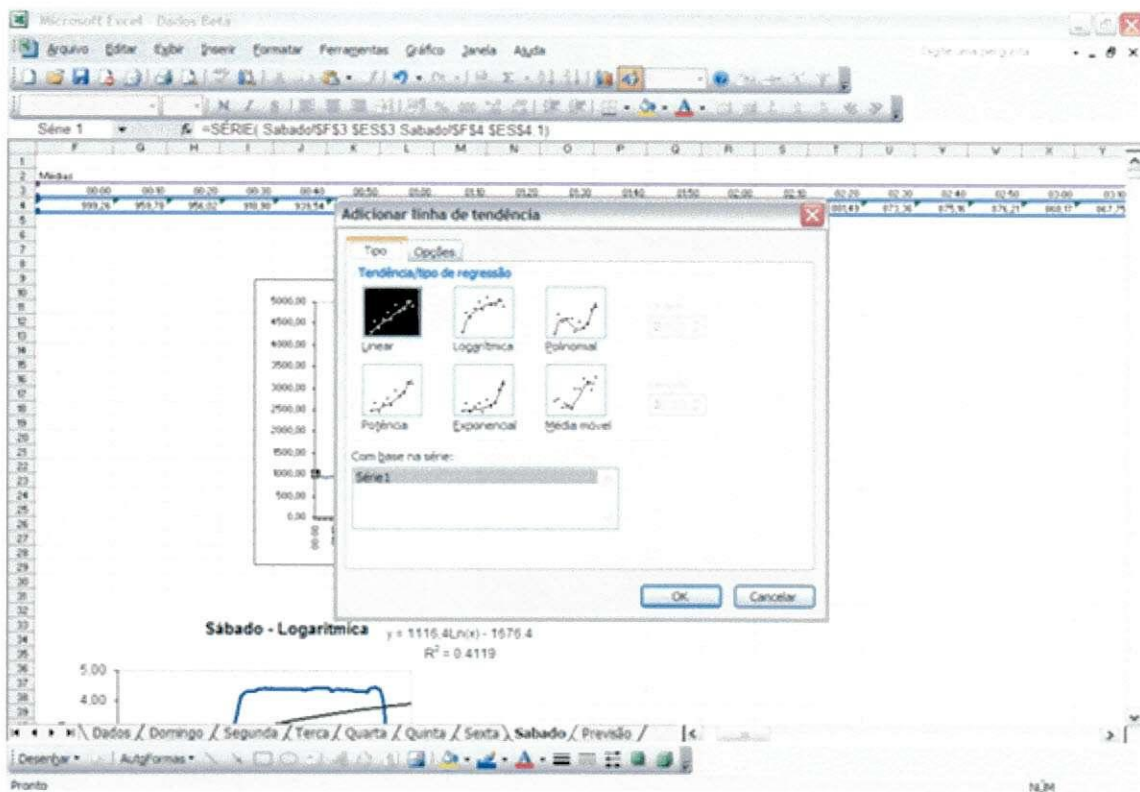


Figura 7.

9. No campo *Opções* podem-se escolher as opções da linha de tendência. No exemplo selecionou-se a opção *Exibir equação no gráfico* (Figura 8). Tem-se então, a equação para a função de regressão linear mostrada no gráfico.

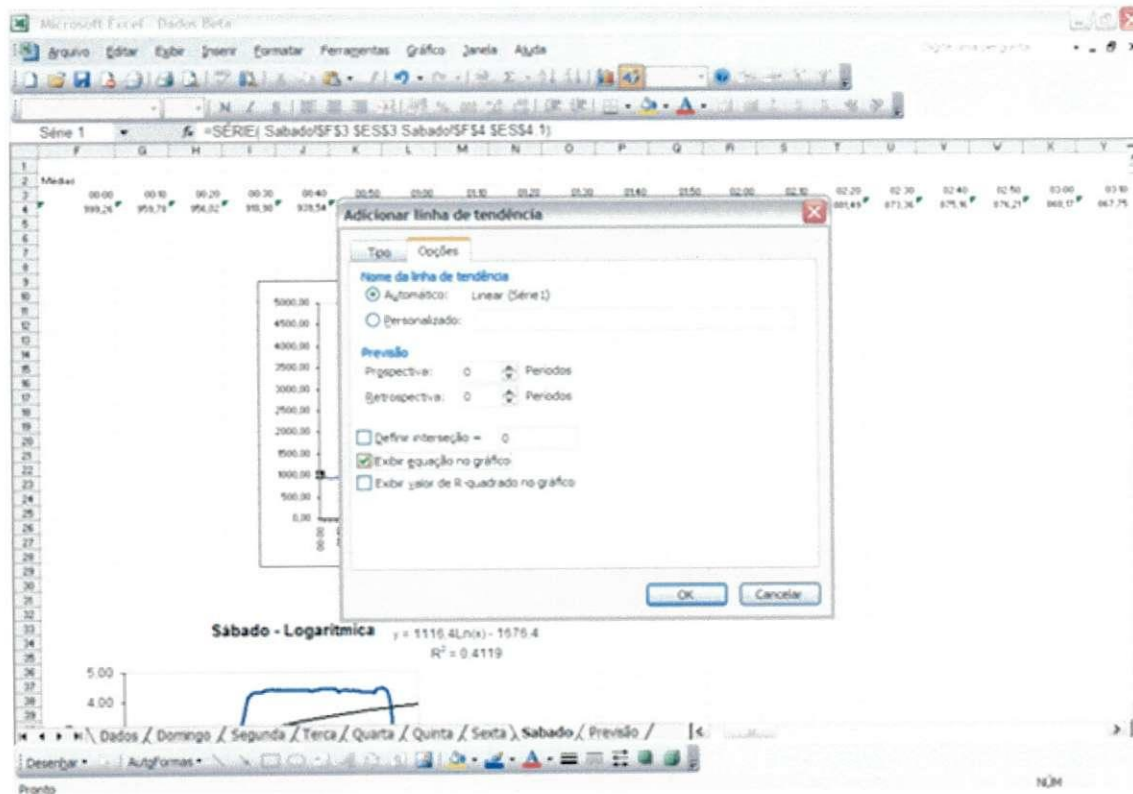


Figura 8.

10. O gráfico final é mostrado na Figura 9, junto com a linha de tendência linear e a respectiva equação para uso da previsão.

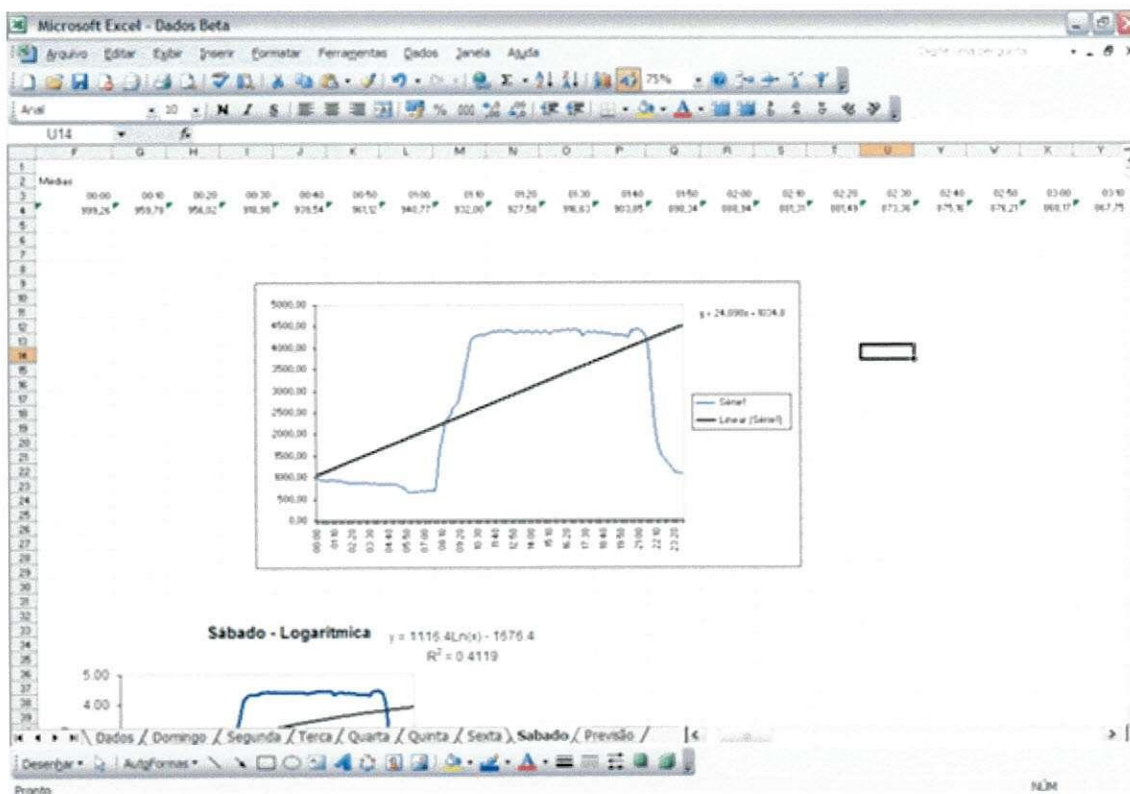


Figura 9.