

## 4

### Aplicação do Modelo

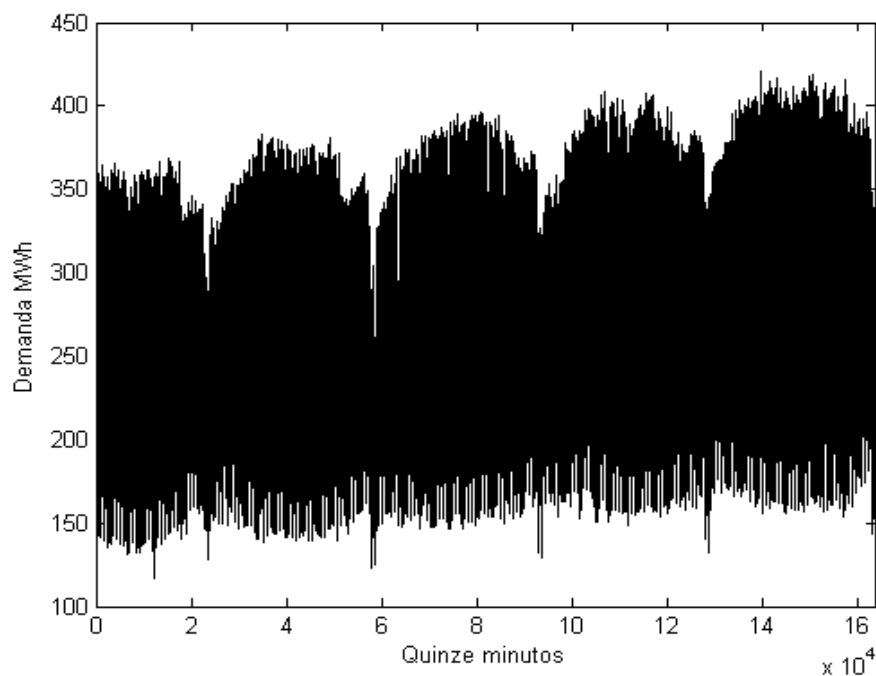
É possível encontrar na literatura diversas aplicações que utilizam técnicas estatísticas e de computação intensiva para realizar previsões de curto prazo na área de energia elétrica. Entre elas podem-se citar: Esteves (2003), que modela dados com múltiplos ciclos através do Método de Amortecimento Direto e Holt-Winters com múltiplos ciclos e Lourenço (1998) que faz uso de Redes Neurais e Lógica Fuzzy para modelar dados horários. Miranda (2007) apresenta uma extensão do método Holt-Winters com múltiplos ciclos que considera a temperatura como variável exógena, enquanto Rizzo (2001) aplica Amortecimento Direto e o Modelo de Gupta a dados horários. Sobral (1999) combina Redes Neurais, Lógica Fuzzy, modelos estatísticos e dados de temperatura para modelar dados de alta frequência.

Uma gama variada de outras técnicas e modelos pode ser empregada no estudo de previsão a curto prazo. O HPA é um desses métodos inovadores que tem obtidos bons resultados (Ivaha *et al.*, 2007).

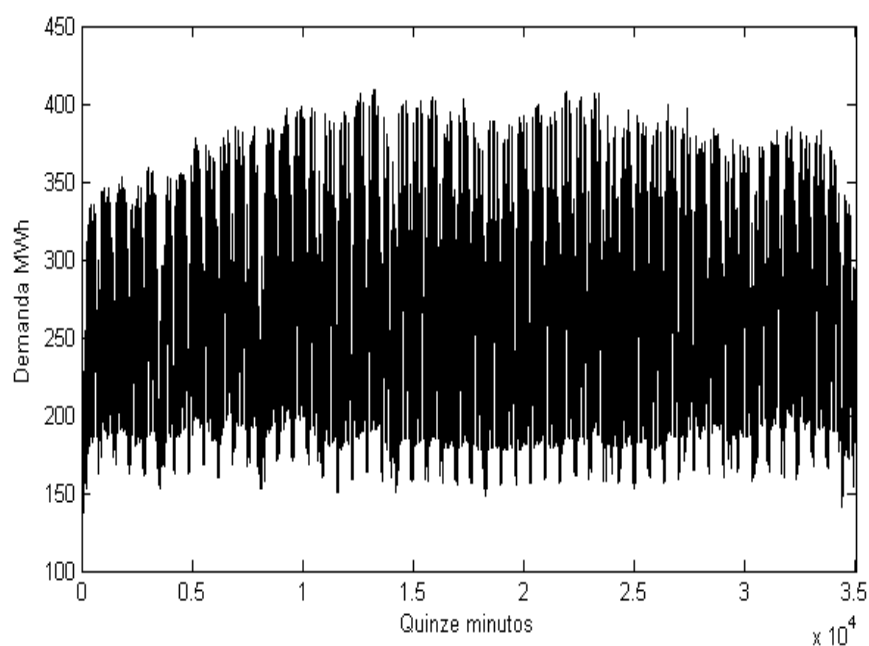
Neste capítulo são apresentados os resultados obtidos pela aplicação da metodologia *Hierarchical Profiling Approach* a dados de demanda brasileiros. O objetivo do modelo é realizar previsões a cada 15 minutos para sete dias à frente, ou seja, prever 672 passos.

Para tanto, o modelo foi aplicado à base de dados de uma grande concessionária de energia elétrica do Brasil, localizada na região Sudeste. Os dados referem-se à demanda de energia elétrica em MWh no período compreendido entre maio de 2002 e dezembro de 2006. As informações foram verificadas a cada 15 minutos, fornecendo 96 observações diárias e 163.676 no total.

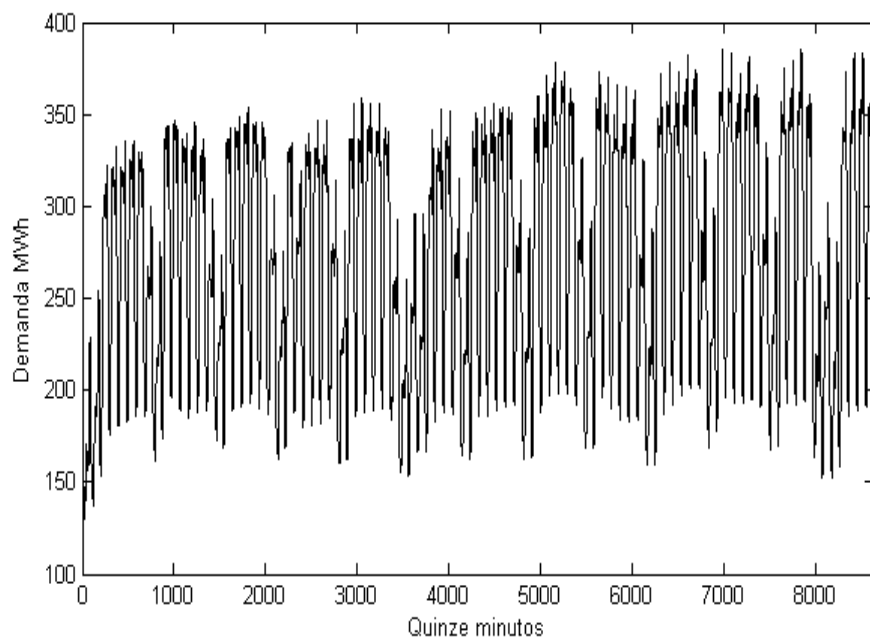
As características da série de carga estão ilustradas nas Figuras 4.1 a 4.6, nas quais se podem observar as sazonalidades existentes.



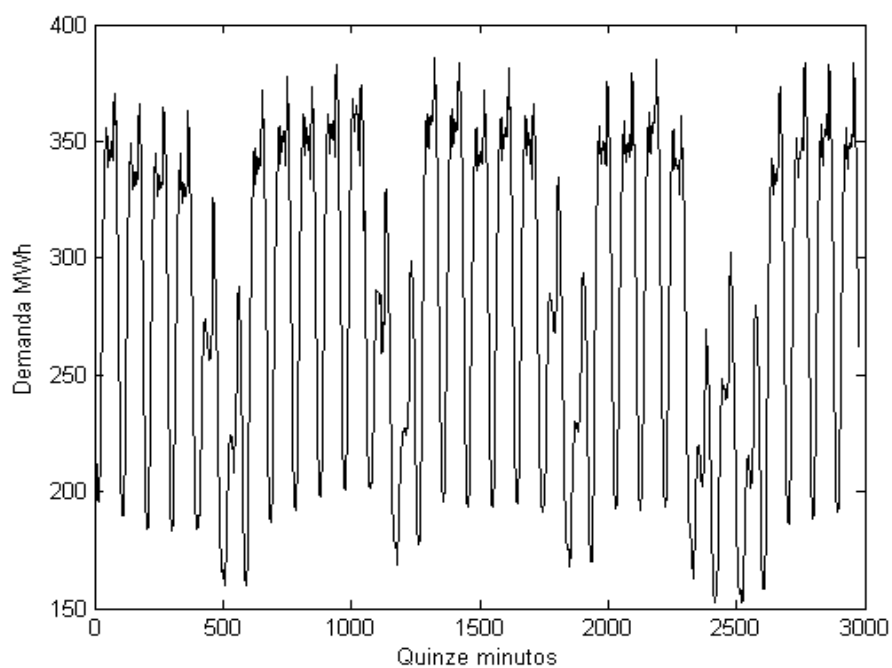
**Figura 4.1** Série de demanda de energia em todo o histórico



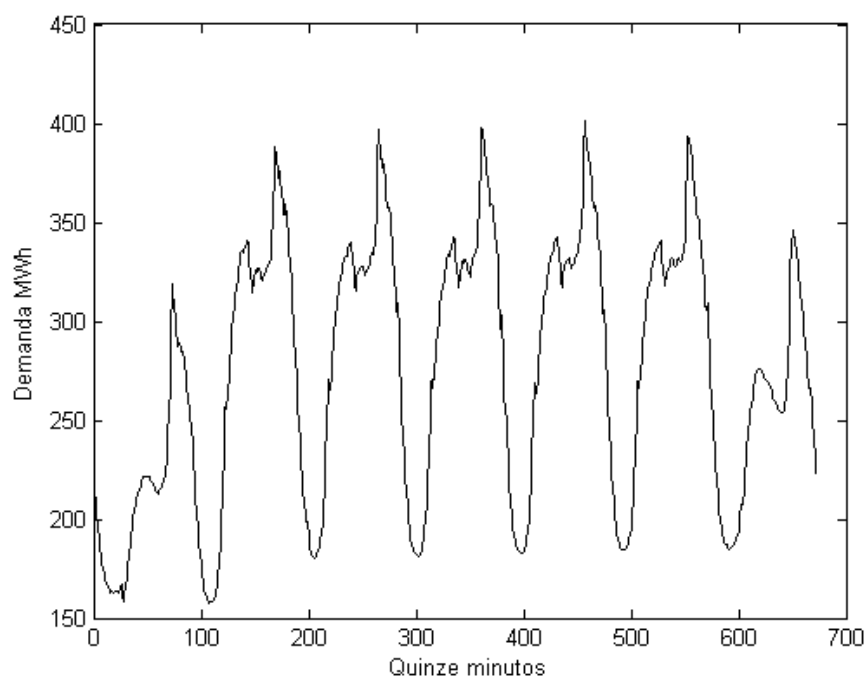
**Figura 4.2** Demanda de energia a cada quinze minutos de um ano



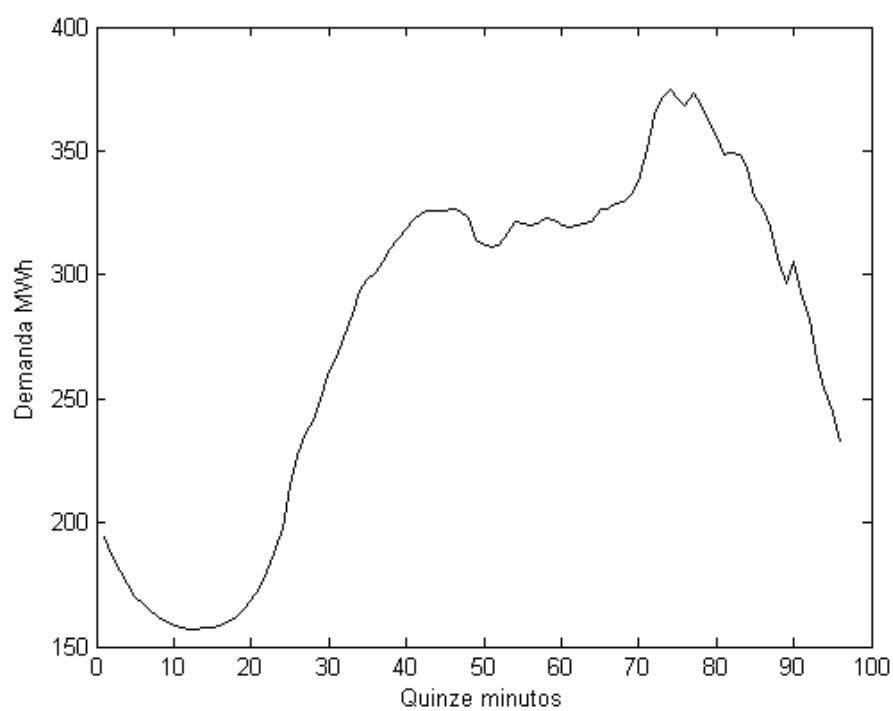
**Figura 4.3** Demanda de energia a cada quinze minutos de um trimestre



**Figura 4.4** Demanda de energia a cada quinze minutos de um mês



**Figura 4.5** Demanda de energia a cada quinze minutos de uma semana



**Figura 4.6** Demanda de energia a cada quinze minutos de um dia

Na Figura 4.1 nota-se que a periodicidade anual é bastante clara, visto que é possível distinguir facilmente cada ano observado. Nota-se, ainda, que há uma tendência crescente para os quatro anos de dados. Quando o período analisado é restrito a um ano de dados, tal tendência não é percebida.

Quando a análise é feita no âmbito mensal, Figura 4.4, pode-se perceber claramente cada divisão equivalente a uma semana, o que é um indício da presença de sazonalidade semanal. Esta análise pode ser estendida para a Figura 4.5, que mostra uma semana de dados: os sete dias aparecem bastante destacados e pode-se notar que a demanda para o sábado e o domingo apresenta comportamento diferente dos demais dias da semana.

Analizando a Figura 4.6, pode-se notar que a demanda é bastante baixa nas horas iniciais do dia e aumenta significativamente no fim do mesmo. Este comportamento permanece praticamente inalterado para os dias úteis sofrendo alterações nos fins de semana e feriados.

Maiores informações sobre os dados serão apresentadas a cada fase da especificação do HPA. A cada perfil considerado será possível obter informações mais detalhadas sobre a dinâmica subjacente aos dados, além de modelar cada comportamento particular que influencia a previsão.

#### **4.1.**

##### **Perfil de Nível 1 – Padrões para os dias da semana**

O primeiro nível do modelo HPA trata da periodicidade mais geral observada nos dados, ou seja, o perfil de nível 1 modela a sazonalidade dentro do ano. Para isto, analisa-se, em primeiro plano, o comportamento da demanda segundo o dia da semana analisado.

Dado que podem existir até setes padrões diferentes para este perfil, foram identificados os dias da semana para o conjunto de dados em estudo. As

observações correspondentes a cada dia da semana foram agrupadas e calculou-se uma média para cada dia da semana. Assim, foram obtidos sete padrões de demanda que correspondem ao comportamento médio durante cada dia da semana para os dados observados. Cada dia da semana padrão é uma série composta por 96 observações, ou seja, uma série específica de um dia da semana onde cada observação corresponde à média de todos os dados observados para um determinado horário correspondente.

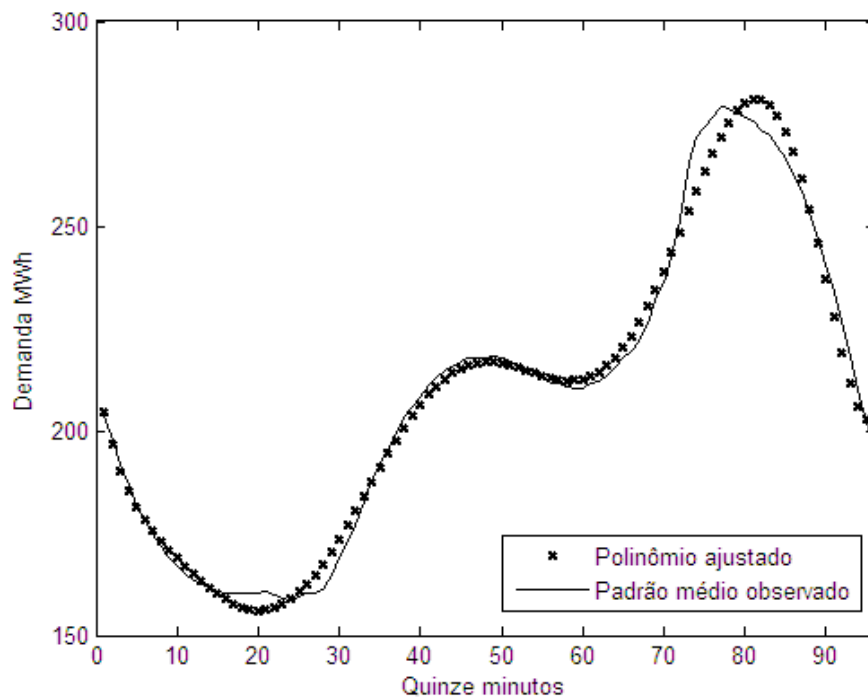
Uma análise mais cuidadosa revelou que não há diferenças significativas entre os padrões obtidos para as terças, quartas e quintas-feiras. Assim, optou-se por criar um único padrão para estes dias da semana. Para o nível 1, então, foram identificados cinco padrões de comportamento diários que são responsáveis por modelar a sazonalidade dentro do dia.

Segundo Al-Madfai (2005), há diversas abordagens possíveis para se modelarem os padrões obtidos. Optou-se, neste trabalho, por realizar um ajuste polinomial de curvas para estimar a equação que representa o comportamento de cada dia da semana. As curvas foram estimadas via Mínimos Quadrados Ordinários. Foram testados polinômios com graus variando de 1 a 20. A melhor curva para cada perfil foi escolhida com base em dois critérios: maximização do Coeficiente de Correlação de Pearson (equação 2.48) e minimização do Erro Médio Percentual Absoluto, MAPE (equação 2.11).

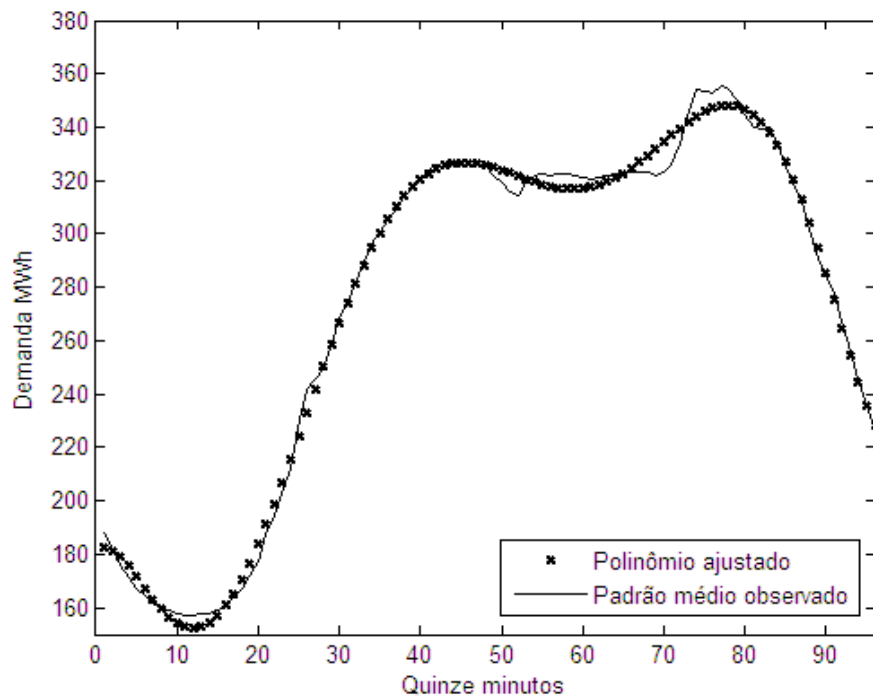
As Figuras 4.7 a 4.12 apresentam a comparação entre o perfil padrão observado e o perfil estimado para cada dia da semana, seguidos pelos respectivos polinômios estimados.

A equação do polinômio estimado para o padrão do domingo é:

$$\begin{aligned} \text{dompad} = & 215,13 - 11,95t + 1,56t^2 - 1,3t^3 + 6,02 \cdot 10^{-3}t^4 - 1,51 \cdot 10^{-4}t^5 + \\ & + 2,05 \cdot 10^{-6}t^6 - 1,42 \cdot 10^{-8}t^7 + 3,93 \cdot 10^{-11}t^8 \end{aligned} \quad (4.1)$$



**Figura 4.7** Padrão médio observado e polinômio estimado para o perfil de domingo



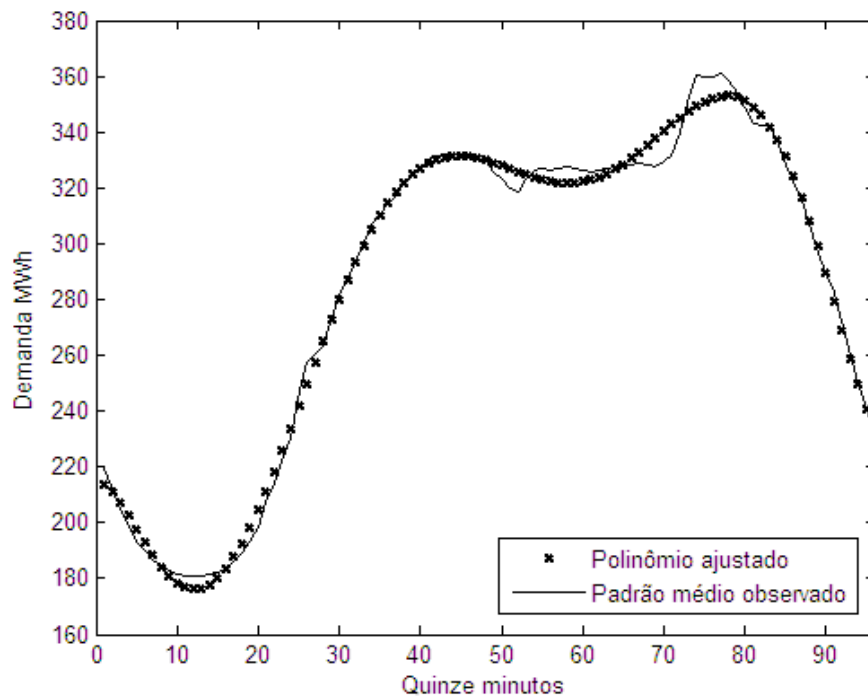
**Figura 4.8** Padrão observado e polinômio estimado para o perfil de segunda-feira

Para o padrão das segundas-feiras a equação do polinômio estimado é:

$$\text{segpad} = 181,06 + 2,68t - 1,46t^2 + 0,13t^3 - 4,27 \cdot 10^{-3}t^4 + 6,72 \cdot 10^{-5}t^5 + 5,08 \cdot 10^{-7}t^6 + 1,48 \cdot 10^{-9}t^7 \quad (4.2)$$

O polinômio estimado para o perfil que combina as terças, quartas e quintas-feiras é mostrado a seguir. A comparação gráfica pode ser observada na Figura 4.9.

$$\text{tqqpad} = 214,96 + 8,35 \cdot 10^{-2}t - 1,18t^2 + 0,11t^3 - 3,87 \cdot 10^{-3}t^4 + 6,19 \cdot 10^{-5}t^5 - 4,73 \cdot 10^{-7}t^6 + 1,39 \cdot 10^{-9}t^7 \quad (4.3)$$

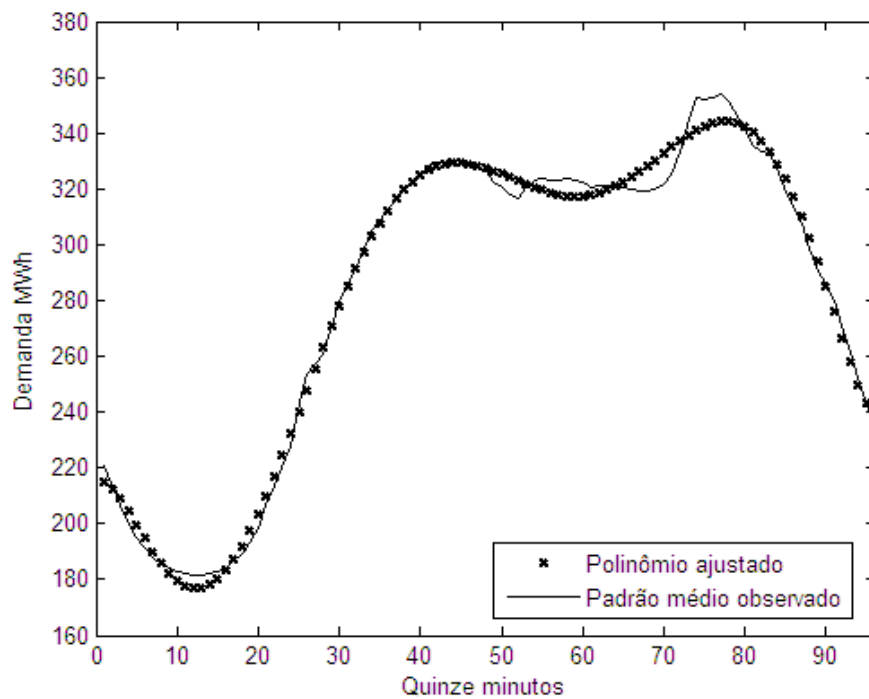


**Figura 4.9** Padrão observado e polinômio estimado para o perfil médio de terça, quarta e quinta-feira

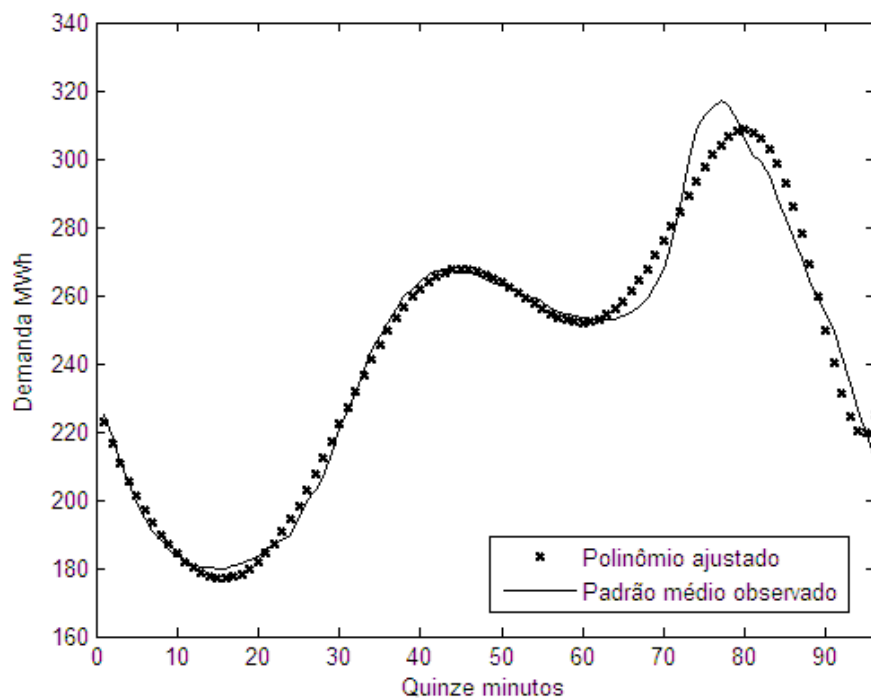
Para o padrão das sextas-feiras a equação do polinômio estimado é:

$$\text{sexpad} = 215,37 + 0,73t - 1,27t^2 + 0,12t^3 - 3,96 \cdot 10^{-3}t^4 + 6,29 \cdot 10^{-5}t^5 - 4,79 \cdot 10^{-7}t^6 + 1,40 \cdot 10^{-9}t^7 \quad (4.4)$$





**Figura 4.10** Padrão observado e polinômio estimado para o perfil de sexta-feira



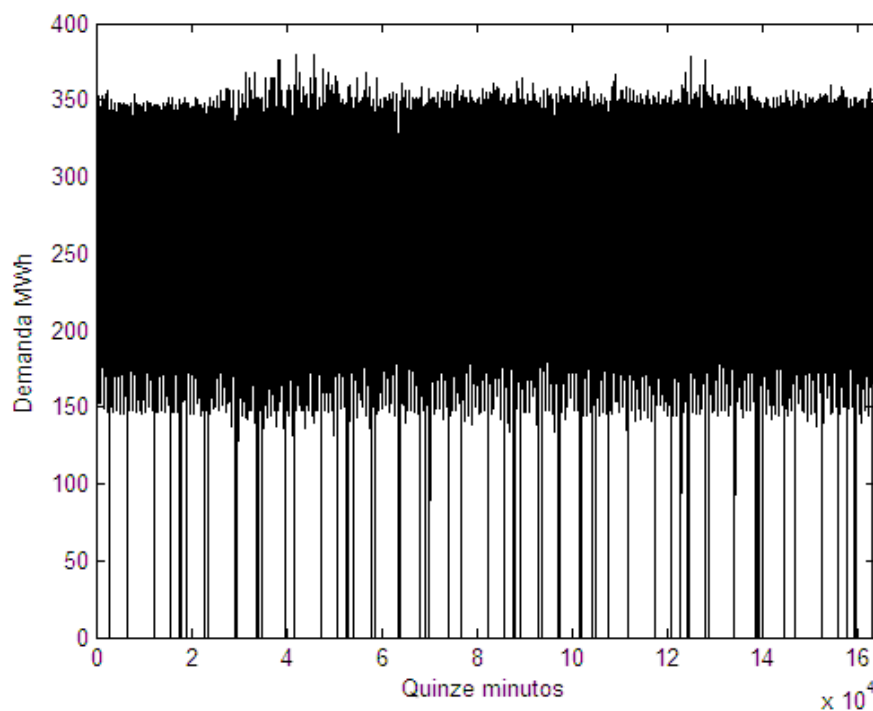
**Figura 4.11** Padrão observado e polinômio estimado para o perfil de sábado

O polinômio estimado para o perfil dos sábados é mostrado a seguir.

$$\begin{aligned} \text{sabpad} = 230,6 - 8,18t + 0,67t^2 - 0,06t^3 + 3,31 \cdot 10^{-5}t^4 - 9,95 \cdot 10^{-5}t^5 + \\ + 1,52 \cdot 10^{-6}t^6 - 1,14 \cdot 10^{-8}t^7 + 3,35 \cdot 10^{-11}t^8 \end{aligned} \quad (4.5)$$

Definidos os cinco padrões referentes ao perfil de nível 1, o próximo passo foi criar uma série temporal do mesmo tamanho que a série original, cujos valores são os padrões encontrados para os dias da semana.

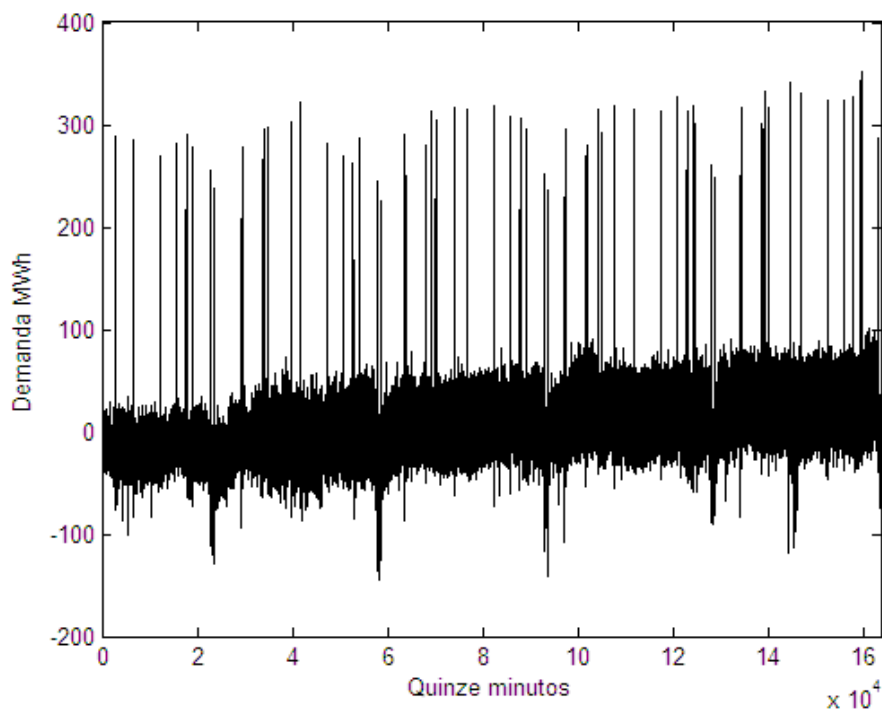
Os dias de domingo da nova série, por exemplo, são compostos pelos valores previstos após a aplicação do polinômio estimado para o padrão de domingo; o mesmo foi feito para os quatro padrões restantes encontrados. Cabe salientar que este procedimento foi realizado apenas para os dias em que não há ocorrência de feriados. Os dias de feriado receberam valor zero. Esta série, que pode ser vista na Figura 4.12, representa o perfil obtido no nível 1.



**Figura 4.12** Série obtida no nível 1

Para prosseguir a análise para os próximos níveis é necessário criar uma série ajustada sem o efeito da sazonalidade dentro do dia, modelada no nível 1. Para isso, utiliza-se o resíduo produzido pela diferença entre a série observada e a série obtida no nível 1.

A série remanescente após o ajuste de nível 1 é exposta na Figura 4.13. Nota-se que, desconsiderando os picos que correspondem aos feriados observados nos quatro anos de dados, a série apresenta-se dispersa em torno de zero. Este comportamento é esperado visto que, após a modelagem do último nível, a série remanescente deve ter o comportamento de um ruído aleatório, sem o efeito da sazonalidade, tendência ou outra função determinística.



**Figura 4.13** Série residual proveniente do ajuste de nível 1

#### 4.2.

#### Perfil de Nível 2 – Tratamento de feriados

Retirado o efeito da sazonalidade dentro do dia da semana para os dias úteis, a próxima etapa consiste em realizar o mesmo procedimento para os dias de feriado.

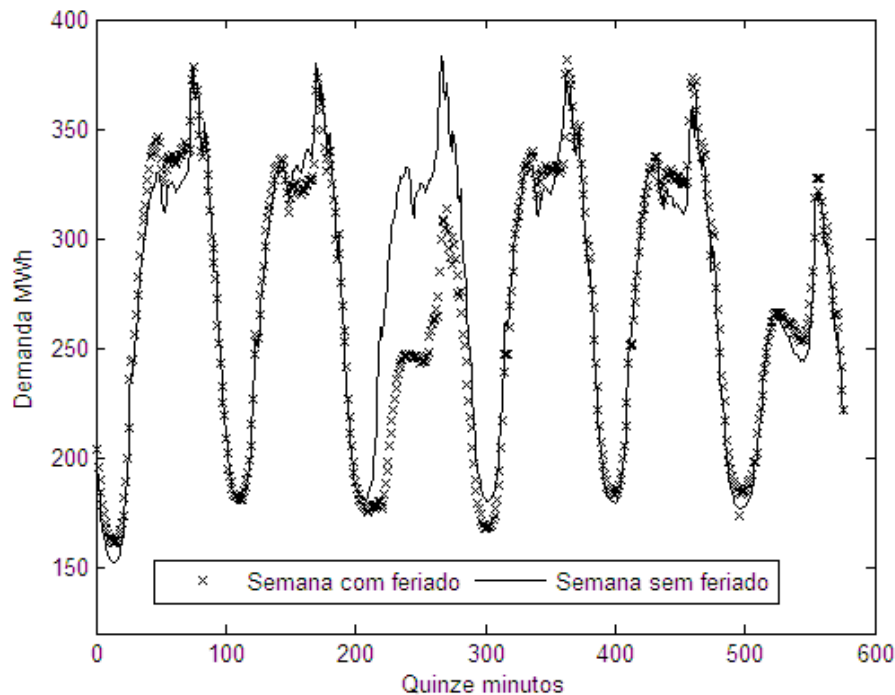
O nível 2 do HPA consiste na modelagem de observações atípicas, como os feriados, e outros eventos discrepantes que possam prejudicar as previsões. Cabe lembrar que a série a ser tratada neste nível não é mais a série original e sim, a série de resíduos gerada após a modelagem dos perfis de nível 1.

Com a finalidade de retirar o efeito dos dias atípicos, a primeira etapa consiste em identificar o dia da semana de ocorrência do feriado e agrupar os dados segundo este critério. Feito isso, o procedimento de modelagem é idêntico ao adotado no tratamento dos dias úteis do nível 1 do modelo.

Assim como em outros países, o Brasil possui uma gama de feriados que geram um perfil de demanda de carga elétrica diferente do comportamento dos dias úteis. Além de causar impacto na previsão, os feriados também distorcem a identificação e estimação do modelo (Vaage, 2000).

No meio acadêmico é possível ignorar os feriados ao verificar o poder preditivo de um modelo. Tal possibilidade não é verificada no mercado, visto que as concessionárias de energia elétrica necessitam de previsões acuradas para estes dias a fim de evitar sobrecargas no sistema ou mesmo subutilização da carga gerada.

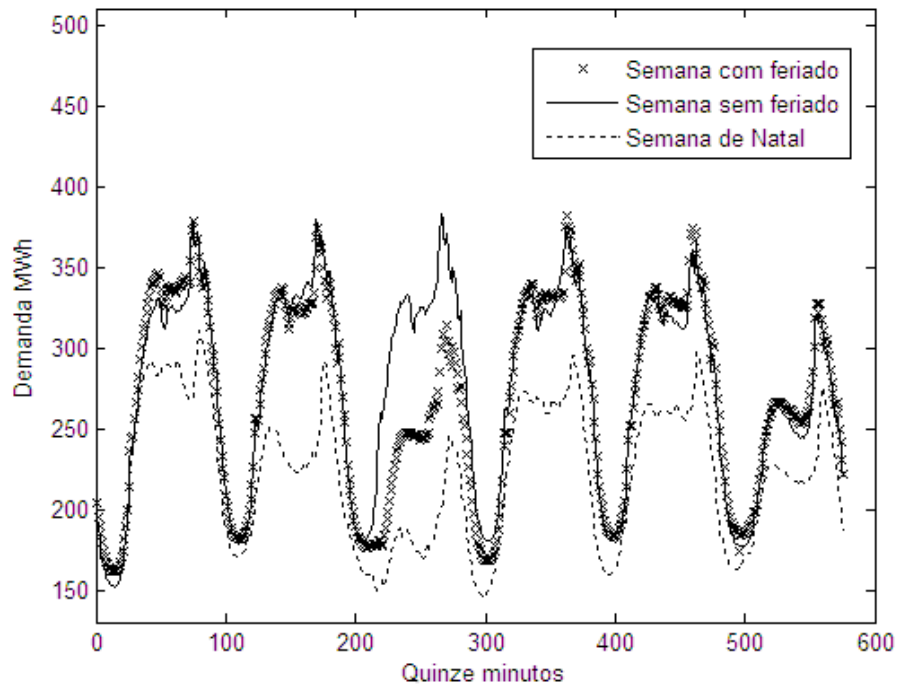
A diferença entre uma semana comum e uma semana onde há um feriado pode ser observada claramente na Figura 4.14. A ocorrência do feriado não parece influenciar significativamente a demanda dos dias próximos a este evento.



**Figura 4.14** Demanda de carga elétrica em uma semana que ocorreu um feriado e em uma semana onde não houve feriado

Miranda (2007) observa o problema da influência de feriados na demanda de energia elétrica e utiliza fatores de correção baseados no comportamento das semanas anterior e posterior ao feriado para minimizar este efeito. Tendo em vista que o feriado gera comportamentos distintos em dias da semana diferentes, este problema foi solucionado pela autora com a criação de regras para cada dia da semana. Tais regras foram ajustadas no mesmo período para o qual a previsão foi realizada, a fim de proporcionar uma melhor correção.

Os fatores de correção não serão aplicados nos feriados comuns deste trabalho, porém os feriados de Natal e Ano Novo receberão tratamento especial. Observou-se, para este conjunto de dados, que os feriados de Natal e Ano Novo comportam-se de maneira peculiar se comparados aos dias úteis e demais feriados. A demanda de carga elétrica para estes dias é bem inferior à dos demais. Este comportamento pode ser observado na Figura 4.15.



**Figura 4.15** Demanda em uma semana de feriado comum, uma semana de ocorrência de Natal e em uma semana onde não houve feriado

Tendo em vista estas discrepâncias, optou-se pelo desenvolvimento de um banco de regras que foi incorporado ao modelo de previsão de forma exógena. Este banco de regras contém fatores de redução/aumento que serão incorporados à previsão da energia, sendo assim capaz de ajustar melhor o modelo, reduzindo o erro de previsão (Bunn & Farmer, 1985).

No caso estudado nessa dissertação, havia disponibilidade de quatro anos de dados, o que não é suficiente para análise de feriados, porém, é possível realizar algumas análises capazes de reduzir os erros do modelo de previsão gerados para esses dias.

O primeiro passo desta análise consiste na identificação dos dias de ocorrência de feriados. Com isso, pode-se calcular a diferença percentual entre o dia do feriado e o mesmo dia da semana, observado nas semanas anterior e posterior (visto que, caso não ocorra nenhum fato extraordinário, o montante de carga do dia observado será semelhante ao das semanas antecedente e subsequente).

O objetivo deste procedimento é analisar a variação de carga elétrica de um dia normal e de um feriado. Se na semana anterior/posterior a um feriado que está sendo estudado, houve ocorrência de outro feriado, este dia será descartado da análise. Neste caso, o procedimento é retroceder/avançar uma semana a mais para realizar a análise.

Após o cálculo destes fatores, os feriados de Natal e Ano Novo podem ser corrigidos. A aplicação das regras faz com que estes feriados “especiais” apresentem, agora, comportamento semelhante aos demais feriados. Assim, as análises posteriores referentes a feriados poderão incluir mais estas duas datas. Os feriados tratados neste trabalho são mostrados nas Tabelas 4.1 e 4.2.

**Tabela 4.1** Lista de feriados fixos

<b>Feriados Fixos</b>
Confraternização Universal
Tiradentes
Dia do Trabalho
Independência do Brasil
Nossa Senhora Aparecida
Finados
Proclamação da República
Consciência Negra
Natal

**Tabela 4.2** Lista de feriados móveis

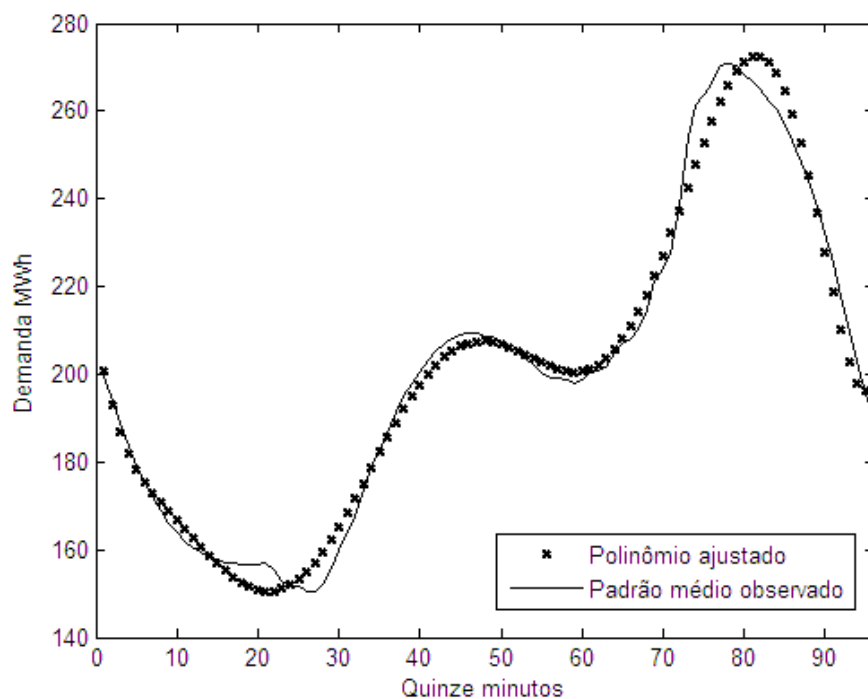
<b>Feriados Móveis</b>
Carnaval
Sexta-feira da Paixão
Páscoa
Corpus Christi

Cabe ressaltar que o feriado da Consciência Negra só é considerado a partir do ano 2006, quando foi instituído.

Cada dia da semana com feriado teve seu comportamento médio representado por uma série composta por 96 observações. Cada instante é obtido pela média do período correspondente aos feriados ocorridos no mesmo dia da semana considerado.

Foram, então, estimados sete polinômios correspondentes a cada dia da semana de ocorrência de feriado, que são responsáveis por definir o comportamento padrão da demanda de carga elétrica nos dias de feriado.

Os ajustes e equações para os sete padrões são apresentados a seguir. Os critérios de seleção dos polinômios foram os mesmos utilizados para os dias úteis.

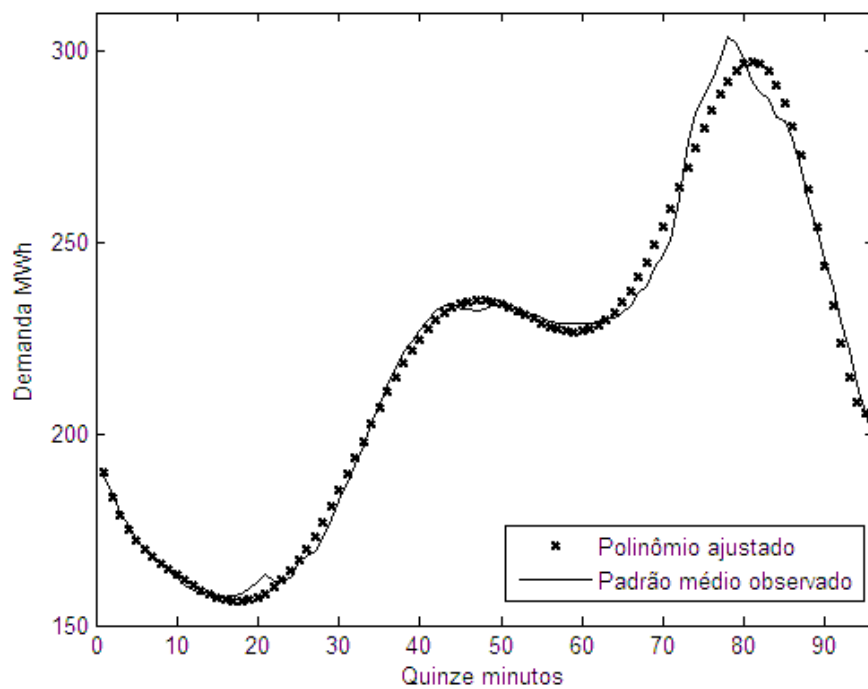


**Figura 4.16** Padrão observado e polinômio estimado para o perfil de domingo de feriado



A equação que define o polinômio ajustado para o padrão médio observado para os feriados que ocorrem aos domingos é:

$$\begin{aligned} \text{dompadfer} = & 211,67 - 12,3t + 1,72t^2 - 0,15t^3 + 6,81 \cdot 10^{-3}t^4 - 1,70 \cdot 10^{-4}t^5 + \\ & + 2,30 \cdot 10^{-6}t^6 - 1,58 \cdot 10^{-8}t^7 + 4,36 \cdot 10^{-11}t^8 \end{aligned} \quad (4.6)$$



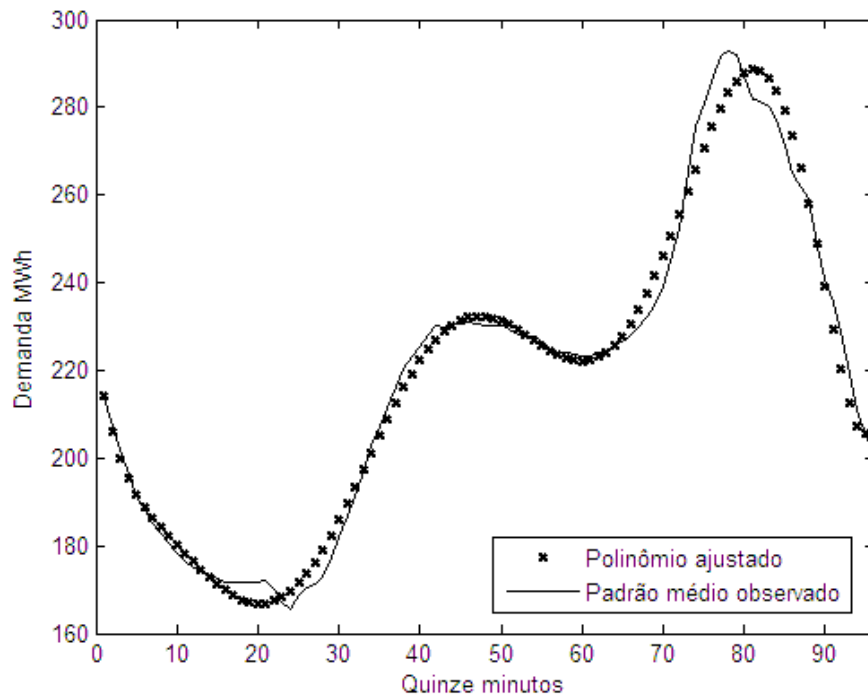
**Figura 4.17** Padrão observado e polinômio estimado para o perfil dos feriados às segundas-feiras

A equação que define o comportamento médio da demanda em um feriado que ocorre numa segunda-feira é:

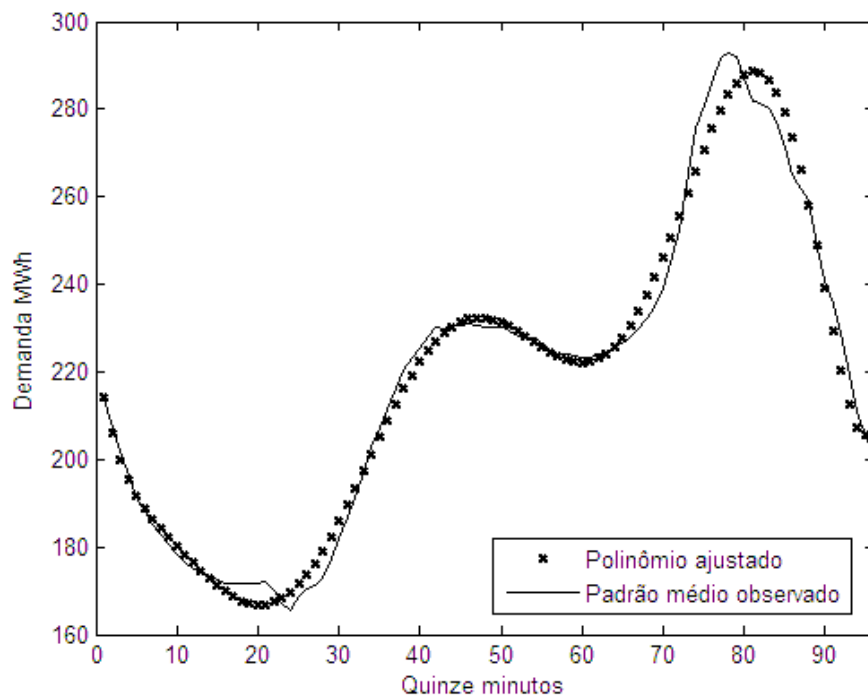
$$\begin{aligned} \text{segpadfer} = & 198,12 - 9,54t + 1,34t^2 - 0,12t^3 + 5,82 \cdot 10^{-3}t^4 - 0,02 \cdot 10^{-3}t^5 + \\ & + 2,11 \cdot 10^{-6}t^6 - 14,88 \cdot 10^{-9}t^7 + 41,64 \cdot 10^{-12}t^8 \end{aligned} \quad (4.7)$$

De maneira análoga, o perfil padrão para um feriado que ocorre numa terça-feira é definido pela equação:

$$\begin{aligned} \text{terpadfer} = & 216,38 - 13,27t + 1,88t^2 - 0,15t^3 + 7,04 \cdot 10^{-3}t^4 - 173,91 \cdot 10^{-6}t^5 + \\ & + 2,34 \cdot 10^{-6}t^6 - 16,09 \cdot 10^{-9}t^7 + 44,27 \cdot 10^{-12}t^8 \end{aligned} \quad (4.8)$$



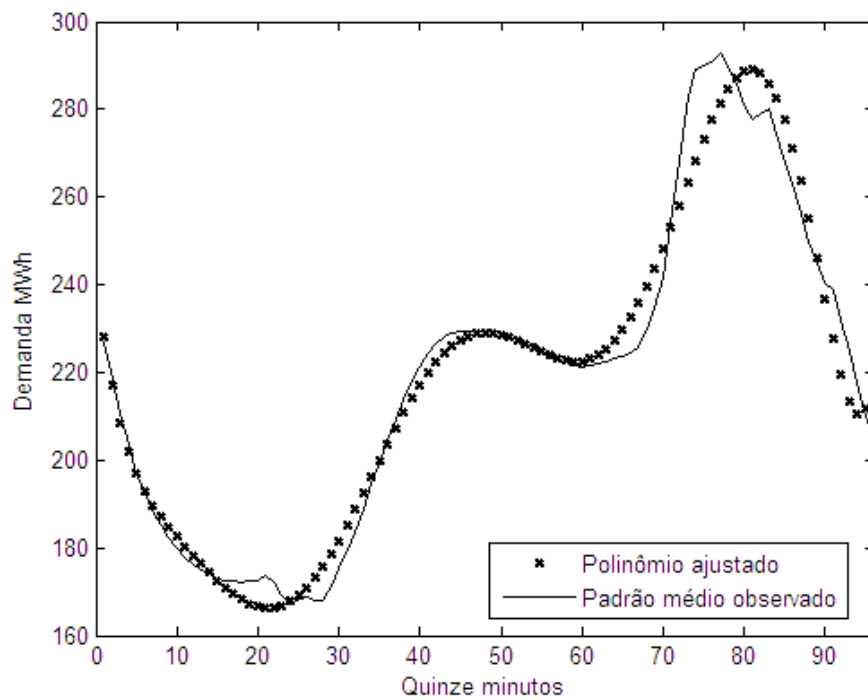
**Figura 4.18** Padrão observado e polinômio estimado para o perfil dos feriados às terças-feiras



**Figura 4.19** Padrão observado e polinômio estimado para o perfil dos feriados às quartas-feiras

O polinômio que define o padrão para um feriado que ocorre numa quarta-feira é:

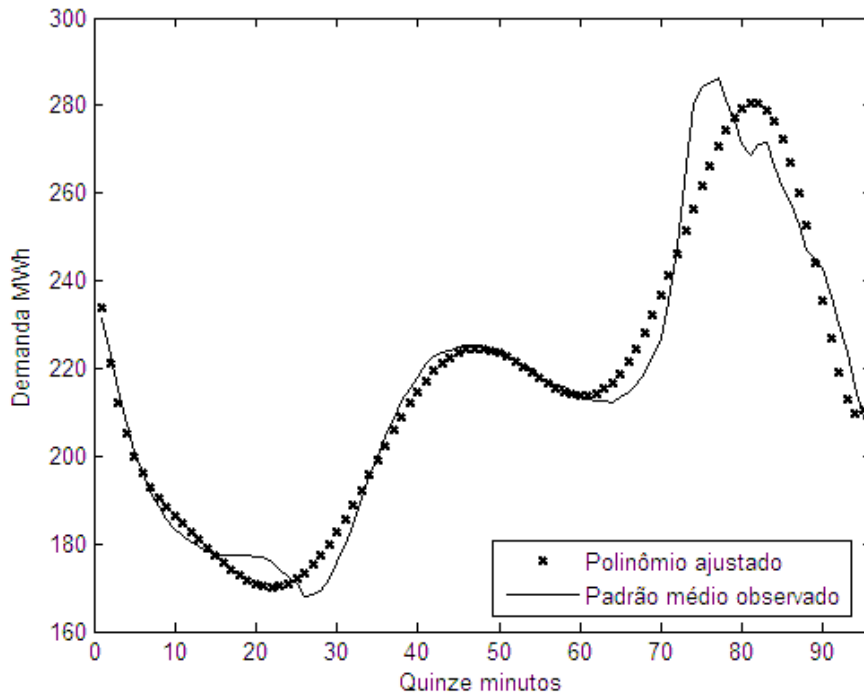
$$\begin{aligned} \text{quapadfer} = & 225,16 - 12,42t + 1,74t^2 - 0,15t^3 + 6,97 \cdot 10^{-3}t^4 - 175,01 \cdot 10^{-6}t^5 + \\ & + 2,37 \cdot 10^{-6}t^6 - 16,44 \cdot 10^{-9}t^7 + 45,41 \cdot 10^{-12}t^8 \end{aligned} \quad (4.9)$$



**Figura 4.20** Padrão observado e polinômio estimado para o perfil dos feriados às quintas-feiras

O padrão que representa o comportamento médio que ocorre às quintas-feiras pode ser definido por:

$$\begin{aligned} \text{quipadfer} = & 243,28 - 17,11t + 2,31t^2 - 0,18t^3 + 6,97 \cdot 10^{-3}t^4 - 199,25 \cdot 10^{-6}t^5 + \\ & + 2,65 \cdot 10^{-6}t^6 - 18,15 \cdot 10^{-9}t^7 + 49,80 \cdot 10^{-12}t^8 \end{aligned} \quad (4.10)$$



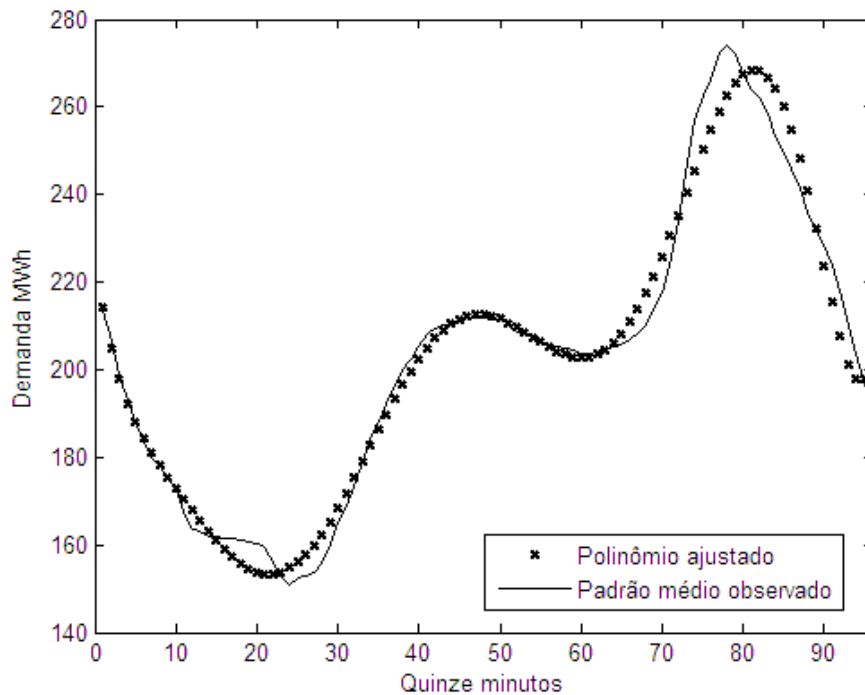
**Figura 4.21** Padrão observado e polinômio estimado para o perfil dos feriados às sextas-feiras

O polinômio que representa o padrão para os feriados às sextas-feiras é:

$$\begin{aligned} \text{sexpadfer} = & 250,13 - 18,80t + 2,58t^2 - 0,20t^3 + 8,79 \cdot 10^{-3}t^4 - 210,01 \cdot 10^{-6}t^5 + \\ & + 2,76 \cdot 10^{-6}t^6 - 18,68 \cdot 10^{-9}t^7 + 50,85 \cdot 10^{-12}t^8 \end{aligned} \quad (4.11)$$

Por fim, definiu-se o padrão que representa o comportamento dos feriados que ocorrem aos sábados. A equação que define este padrão é:

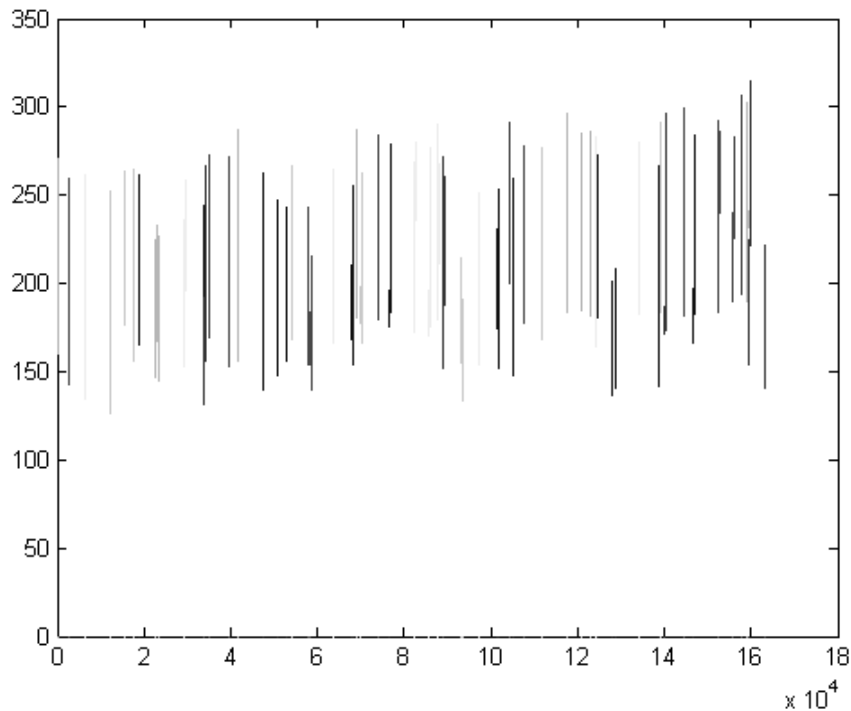
$$\begin{aligned} \text{sabpadfer} = & 226,26 - 13,68t + 1,81t^2 - 0,15t^3 + 7,09 \cdot 10^{-3}t^4 - 176,93 \cdot 10^{-6}t^5 + \\ & + 2,39 \cdot 10^{-6}t^6 - 16,50 \cdot 10^{-9}t^7 + 45,47 \cdot 10^{-12}t^8 \end{aligned} \quad (4.12)$$



**Figura 4.22** Padrão observado e polinômio estimado para o perfil dos feriados aos sábados

Após a estimação dos sete padrões que compõem o perfil de nível 2, o passo seguinte consiste em criar uma série temporal composta pelos padrões encontrados para os dias de feriado. A série apresenta valor nulo para as observações correspondentes aos dias úteis.

Os polinômios estimados foram alocados nos dias de feriado em toda a extensão da série. Assim, é possível determinar matematicamente o comportamento da demanda de carga de energia elétrica em qualquer dia de feriado deste conjunto de dados, sabendo apenas o dia da semana de ocorrência. A série criada para o nível 2 pode ser vista na Figura 4.23.

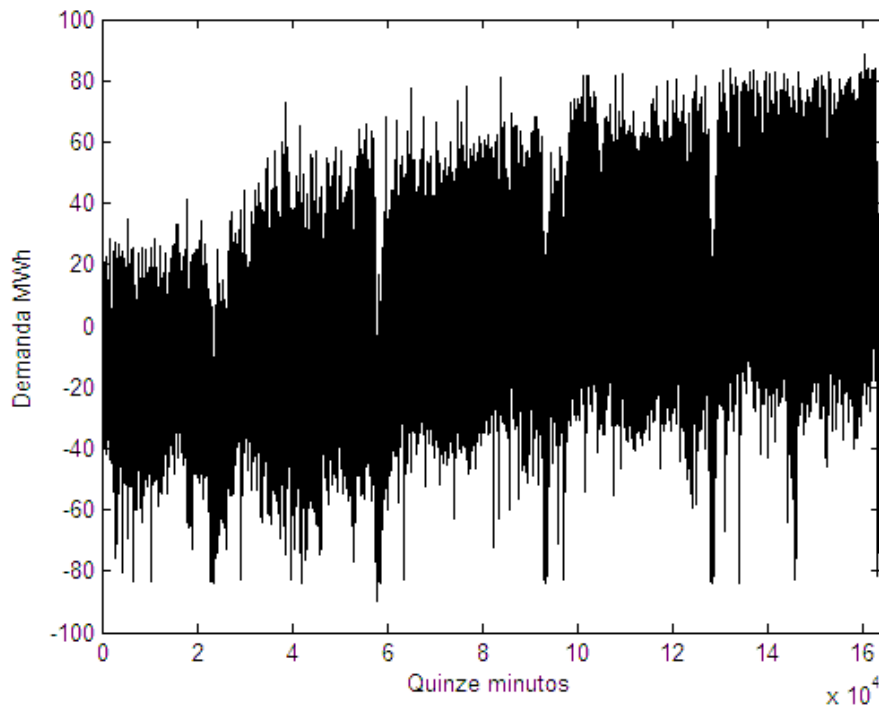


**Figura 4.23** Série de polinômios para os feriados obtida no nível 2

Para a modelagem dos próximos níveis, é necessário criar uma série ajustada que esteja sem o efeito da sazonalidade dentro do dia (tratada no nível 1) e, também, sem o efeito provocado pela ocorrência dos feriados (nível 2). Para isso, calculam-se os resíduos produzidos pela diferença entre a série observada, a série obtida no nível 1 e a série de ajuste de polinômios estimados para os feriados, obtida no nível 2.

A série remanescente após o ajuste dos níveis 1 e 2 é exposta na Figura 4.24. Pode-se observar, numa comparação com a série de dados original, que a tendência de crescimento é mais explícita na série ajustada. Nesta última, a média da série é aproximadamente nula.

O objetivo para os próximos níveis é retirar o efeito da componente determinística remanescente, a fim de obter um resíduo semelhante a um ruído branco, que será modelado com vistas a obter as previsões para os dados. Para isso, é necessário modelar a componente sazonal dos dados. O próximo nível tem a função de identificar e tratar tal componente.



**Figura 4.24** Série de resíduos obtida após o ajuste dos níveis 1 e 2

#### 4.3.

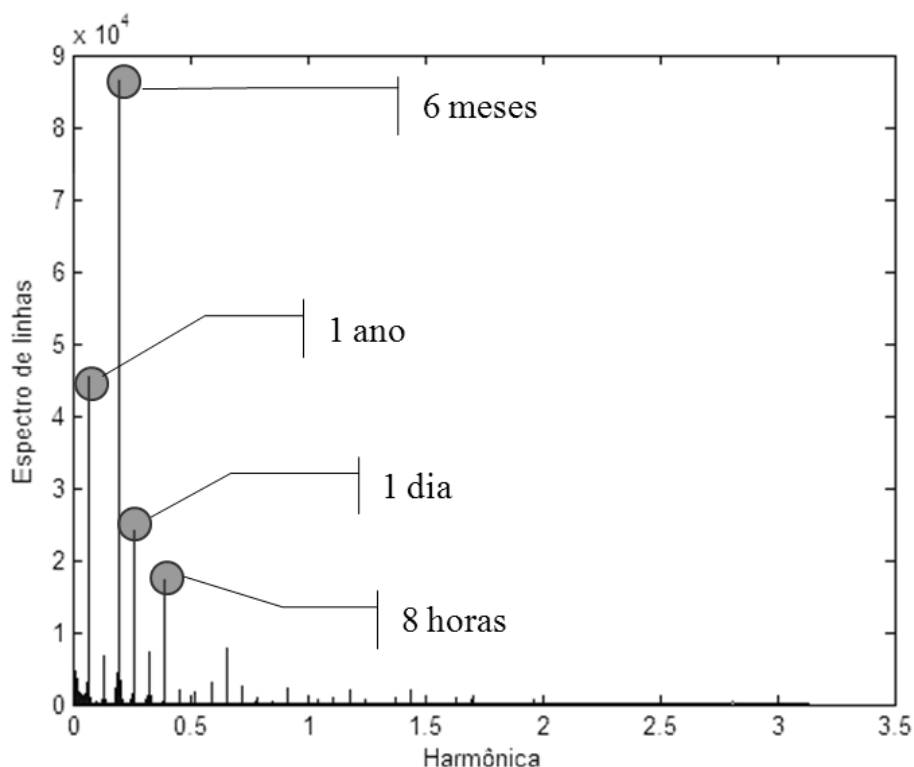
##### **Perfil de Nível 3 – Identificação e modelagem do padrão sazonal**

Segundo Al-Madfai *et al.* (2007), o perfil de nível 3 tem a função de modelar periodicidades anuais ou maiores, ou seja, modela a forma geral da série em torno do ano, incluindo tendência, sazonalidade e distúrbios periódicos.

Para isso é necessário investigar a série temporal com a finalidade de identificar a periodicidade presente nos dados. Tendo em vista que a sazonalidade dentro dia e o efeito proveniente da ocorrência de feriados foram retirados da série, a periodicidade não é mais notada de forma intuitiva como na série original.

Uma ferramenta útil para identificar periodicidades desconhecidas em um conjunto de dados é o Periodograma (seção 2.6.2). Aplicou-se esta técnica à série resultante do ajuste dos níveis 1 e 2 e identificaram-se periodicidades correspondentes a oito horas, um dia, seis meses e um ano.

O Periodograma pode ser observado na Figura 4.25. Os picos maiores representam as periodicidades mais significativas contidas nos dados.



**Figura 4.25** Periodograma da série de demanda de carga de energia elétrica

Os quatro maiores picos encontrados foram submetidos aos testes de Fisher e Whittle e as periodicidades ( $\omega$ ) foram consideradas estatisticamente significativas para este conjunto de dados. Outras periodicidades foram testadas, mas concluiu-se que elas são estatisticamente irrelevantes.

Conhecidas as frequências presentes nos dados é possível modelar a parcela final da componente determinística: a sazonalidade. Tendo em vista que tratam-se de periodicidades múltiplas há algumas abordagens específicas para tratar estes dados. Neste trabalho, optou-se por fazer uso da Regressão Harmônica com frequências conhecidas.

Como o objetivo é obter a componente aleatória após este ajuste, é necessário que os resíduos se apresentem dispersos aleatoriamente em torno de zero. Para isso, além de simplesmente modelar a componente sazonal, é preciso



tratar, também, a tendência dos dados. Assim, a técnica correta a ser empregada é a Regressão Harmônica com tendência polinomial para frequências conhecidas.

A equação da regressão é:

$$\begin{aligned} X_t = & \mu + at + bt^2 + \dots + c^n + A_1 \cos(\omega_1 t) + B_1 \text{sen}(\omega_1 t) + \\ & + A_2 \cos(\omega_2 t) + B_2 \text{sen}(\omega_2 t) + A_3 \cos(\omega_3 t) + B_3 \text{sen}(\omega_3 t) + \\ & + A_4 \cos(\omega_4 t) + B_4 \text{sen}(\omega_4 t) + \varepsilon_t \end{aligned} \quad (4.13)$$

onde:

$$\omega_1 = 0,0001793146$$

$$\omega_2 = 0,0003586293$$

$$\omega_3 = 0,0656291616$$

$$\omega_4 = 0,1963495408$$

e  $\mu, a, b, \dots, c, A_1, B_1, A_2, B_2, A_3, B_3, A_4$  e  $B_4$  são os parâmetros a estimar.

O grau do polinômio que define a tendência foi testado iterativamente até chegar ao modelo com menor MAPE e todos os parâmetros significativos. Foram estimados inúmeros modelos com variadas configurações até se chegar ao modelo mais adequado sob o ponto de vista estatístico.

Assim, a tendência é definida por um polinômio de grau 3 e os parâmetros estimados para a regressão que modela a sazonalidade são:

$$\mu = -24,3784$$

$$a = 335,35 \cdot 10^{-6}$$

$$b = 337,31 \cdot 10^{-12}$$

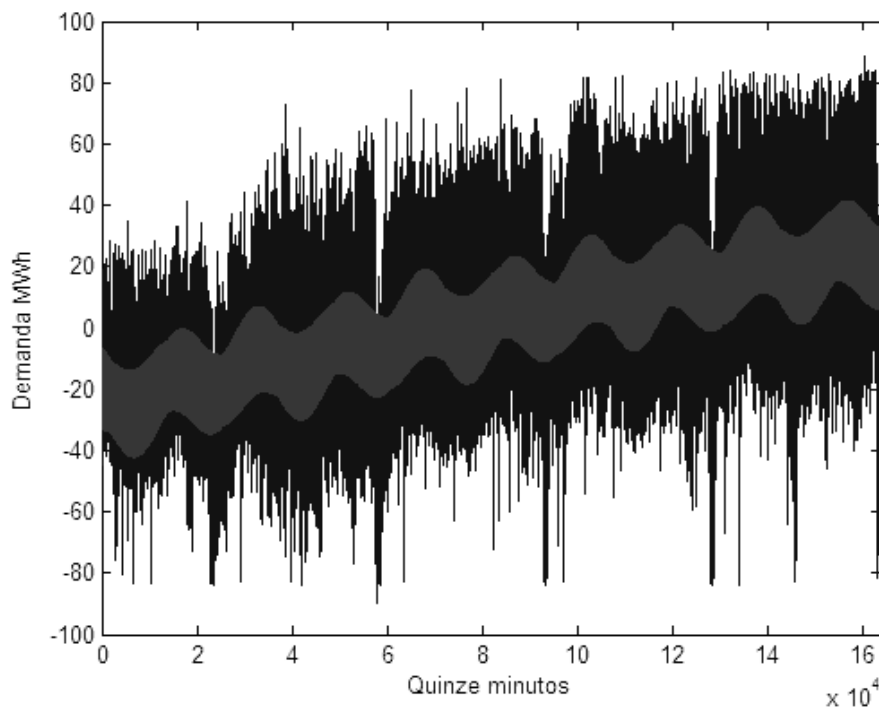
$$c = -3,44 \cdot 10^{-15}$$

$$A_1 = 0,3149$$

$$B_1 = -0,4667$$

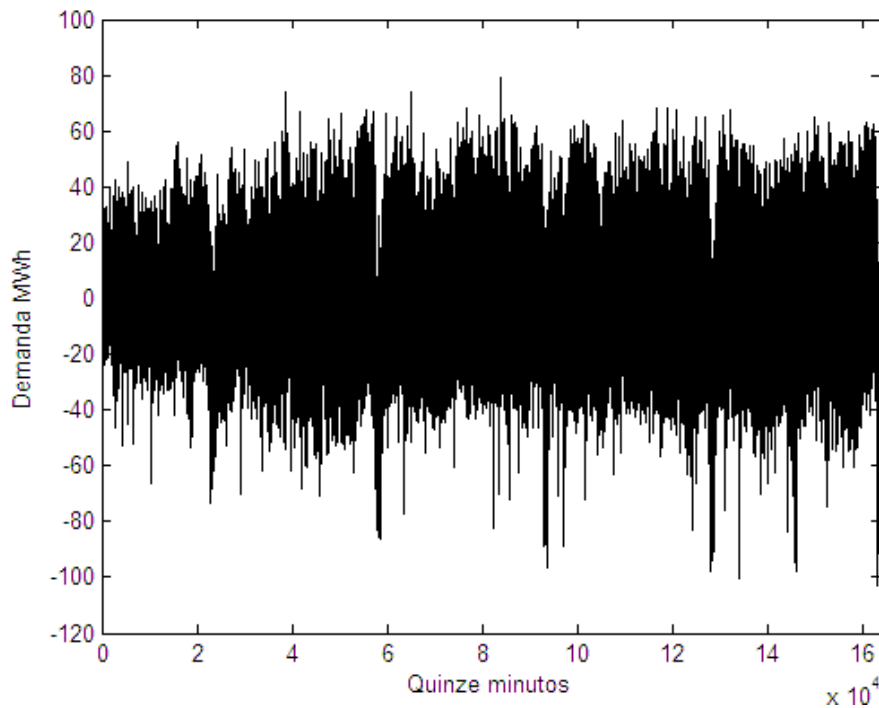
$$\begin{aligned}A_2 &= 4,2393 \\B_2 &= -3,5282 \\A_3 &= 5,9748 \\B_3 &= -0,1392 \\A_4 &= -2,8741 \\B_4 &= 7,8447\end{aligned}$$

A Figura 4.26 mostra a comparação entre valores observados e ajustados segundo a Regressão Harmônica. Nota-se que a dinâmica dos dados não é completamente entendida pelo modelo, porém espera-se que os resíduos apresentem o comportamento aleatório para prosseguir às próximas etapas do HPA.



**Figura 4.26** Ajuste da série resultantes dos níveis 1 e 2 via Regressão Harmônica

Pode-se perceber, no ajuste realizado, que o modelo conseguiu captar a tendência contida nos dados. Dado que a regressão também modelou a sazonalidade, a diferença entre a série inicial (sem o efeito dos dois primeiros níveis) e série resultante (sem o efeito dos três níveis) é um ruído aleatório, conforme mostra a Figura 4.27.



**Figura 4.27** Série de resíduos obtidos após o ajuste dos níveis 1, 2 e 3

#### 4.4 Criação da Função Determinística

Após a identificação dos perfis e o processo inicial de modelagem, o “modelo fundamental” da série foi obtido.

Os dados foram examinados para perfis de níveis superiores a 3 e nada foi encontrado. Isto significa que toda a componente determinística foi capturada pelos ajustes dos níveis discutidos anteriormente.

Ajustados os perfis, a função determinística  $f(t)$  pode ser obtida através da soma dos perfis obtidos nos níveis 1, 2 e 3. Cabe ressaltar que não há uma equação simples que caracterize esta função.

Pode-se considerar que, para este conjunto de dados, cada dia da semana observado tem uma equação definida que depende do ajuste nos perfis iniciais.

Para uma segunda-feira, por exemplo, a função determinística que define o comportamento da demanda de carga de energia elétrica é:

$$f_{seg}(t) = (1-i)secpad(t) + i.secpadfer(t) + X_t(t) \quad (4.14)$$

onde  $secpad(t)$  é o polinômio definido para o padrão de uma segunda-feira de não feriado,  $secpadfer(t)$  é o polinômio que define o padrão para uma segunda-feira de feriado,  $X_t$  é a equação estimada pela Regressão Harmônica que define o padrão sazonal e  $i$  é um indicador da forma:

$$i = \begin{cases} 1 & \text{se é feriado} \\ 0 & \text{caso contrário} \end{cases}$$

que pode ser reescrita da forma:

$$\begin{aligned} f_{seg}(t) = & (1-i).181,06 + 2,68t - 1,46t^2 + 0,13t^3 - 4,27.10^{-3}t^4 + 6,72.10^{-5}t^5 + \\ & + 5,08.10^{-7}t^6 + 1,48.10^{-9}t^7 + i.[226,26 - 13,68t + 1,81t^2 - 0,15t^3 + \\ & + 7,09.10^{-3}t^4 - 176,93.10^{-6}t^5 + 2,39.10^{-6}t^6 - 16,50.10^{-9}t^7 + \\ & + 45,47.10^{-12}t^8] - 24,4 + 335,35.10^{-6}t + 337,31.10^{-12}t^2 - 3,44.10^{-15}t^3 + \\ & + 0,31\cos(1,79.10^{-4}t) - 0,47\sin(1,79.10^{-4}t) + 4,24\cos(2,15.10^{-3}t) - \\ & + 3,53\sin\cos(2,15.10^{-3}t) + 5,97\cos(9,32.10^{-3}t) - 0,14\sin(9,32.10^{-3}t) + \\ & - 2,87\cos(0,07t) + 7,84\sin(0,07t) \end{aligned}$$

As equações que definem a componente determinística para os demais dias da semana são criadas de forma análoga à equação para as segundas-feiras.

## 4.5

### Modelagem da Componente Estocástica

Conforme exposto na equação (3.1), a série pode ser decomposta em duas componentes: estocástica e determinística. Dado que a última foi modelada, a série resultante da diferença entre a série original e os ajustes dos níveis contém apenas a componente estocástica.

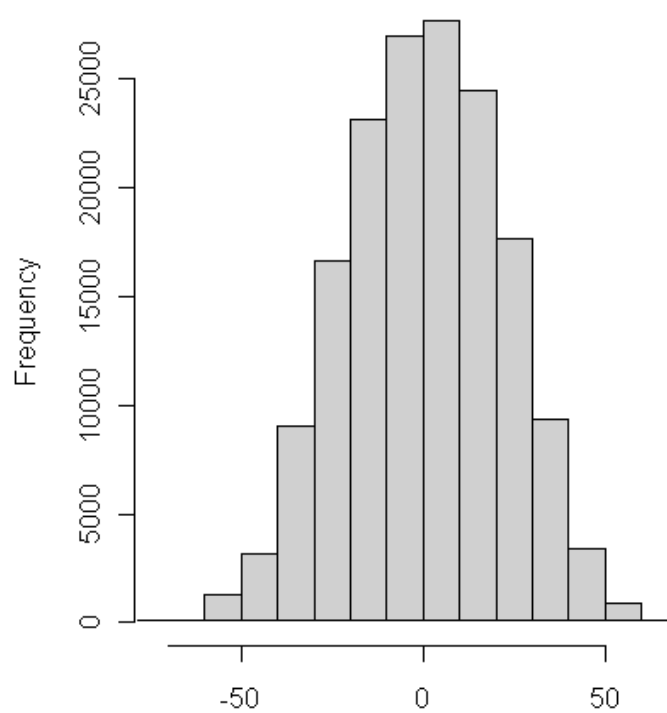
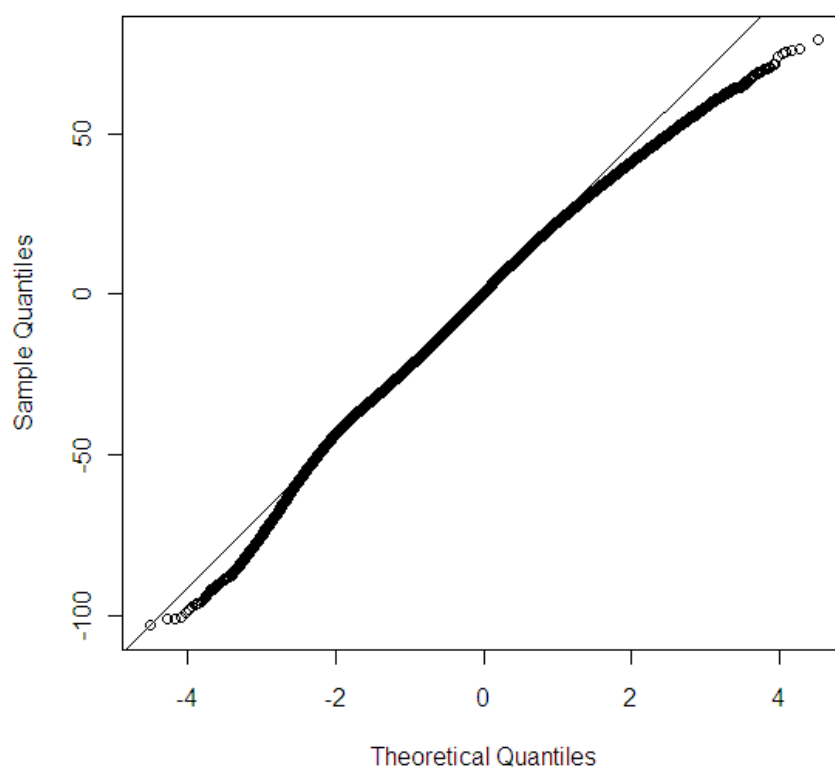
O objetivo final da modelagem antes de fazer a previsão é, então, modelar a componente estocástica. Visto que se trata de um ruído sem presença de tendência ou componente sazonal, qualquer abordagem simples pode ser utilizada para realizar tal tarefa.

Neste trabalho, optou-se por fazer uso da conhecida metodologia de Box & Jenkins. Dadas as características da série, será utilizado um modelo do tipo ARMA(p,q).

Antes de proceder à modelagem é necessário verificar os pressupostos de normalidade, estacionariedade e homocedasticidade.

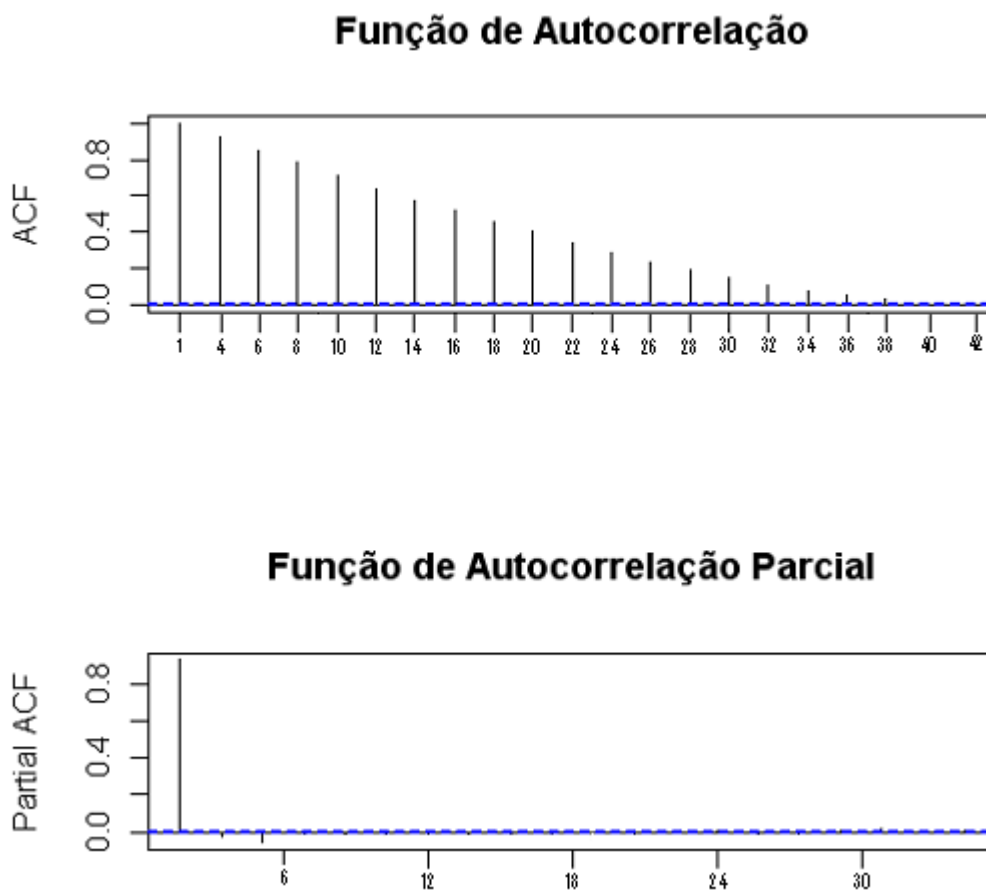
A análise da série resultante do ajuste dos três níveis considerados (Figura 4.27) revela que os dados são estacionários, visto que não é possível notar a presença da componente de tendência na série.

Para verificar a hipótese de Normalidade, podemos fazer uso de alguns testes estatísticos, como o Jarque-Bera, além das análises do histograma e do gráfico Normal Q-Q Plot (Figuras 4.28 e 4.29).

**Figura 4.28** Histograma dos dados**Figura 4.29** Normal Q-Q Plot dos dados de demanda de carga elétrica

As figuras anteriores revelam que os dados se distribuem normalmente. O histograma apresenta simetria e a forma de sino característica da distribuição Normal. No gráfico Normal Q-Q não há desvios significantes da reta Normal. Esta análise é comprovada pela aplicação do teste de Jarque-Bera, que indica normalidade ao nível de 5% de significância.

Cumpridos os pressupostos, deve-se recorrer à análise do correlograma e autocorrelograma para identificar as ordens  $p$  e  $q$  do modelo. Estes gráficos podem ser observados na Figura 4.30.



**Figura 4.30** FAC e FACP da série de demanda de carga elétrica

Pode-se afirmar que a FAC apresenta decaimento exponencial lento e a FACP apresenta alguns *lags* significantes. Este é um indicativo de que trata-se de um modelo ARMA.

Partiu-se da estimação e análise residual de um modelo ARMA(1,1) e, via Critério de Sobrefixação, chegou-se ao modelo ARMA(2,2), cuja equação é:

$$Z_t = -0,0001 + 0,8084Z_{t-1} - 0,8240Z_{t-2} + a_t + 0,8647a_{t-1} - 0,0242a_{t-2} \quad (4.15)$$

onde  $Z_{t-1}$  e  $Z_{t-2}$  são as observações defasadas em 1 e 2 *lags*, respectivamente, e  $a_t$ ,  $a_{t-1}$  e  $a_{t-2}$ , o ruído.

Os resíduos gerados pelo modelo foram analisados segundo os testes de Portmanteau e Jarque-Bera, onde concluiu-se que os mesmos são descorrelatados e normais. Pode-se afirmar, então, que o modelo é bem especificado e permite realizar previsões.

Finalizada a modelagem da componente estocástica, pode-se concluir que a equação do modelo final é dada pela soma da componente determinística (que contém os perfis) com a componente estocástica.

O modelo HPA para a demanda de carga elétrica de 15 em 15 minutos é:

$$y_t = f(t) + Z_t \quad (4.16)$$

onde  $y_t$  são os dados observados,  $f(t)$  a componente determinística especificada para cada dia da semana e  $Z_t$  é a componente estocástica, dada pela equação (4.15).

A seguir, podem ser obtidas as previsões 672 passos à frente, objetivo do trabalho.