

3

A Modelagem em dois níveis

Neste e no próximo capítulo são propostos, testados e avaliados três modelos. Muitos trabalhos afirmam que o relacionamento entre a carga e os fatores exógenos são complexos e não-lineares, tornando a previsão uma tarefa difícil para modelos que utilizem técnicas estatísticas convencionais. Ressaltam ainda que a maioria das técnicas tradicionais não oferecem a precisão desejada e não são suficientemente robustas, pois têm dificuldade para incorporar variáveis de mudanças bruscas como as climáticas, por exemplo. Como uma ilustração da importância da precisão das previsões, foi estimado que um incremento de 1% no erro de previsão pode aumentar em 10 milhões de libras o custo de operação por ano para uma empresa de eletricidade no Reino Unido (Bunn & Farmer 1985b). A literatura clássica de séries temporais, ao longo dos anos, tem apresentado um grande número de técnicas de previsão desenvolvidas para modelar a demanda de carga elétrica (ver Bunn & Farmer (1985a) para uma revisão completa dos métodos clássicos) e outras técnicas têm sido apresentadas pela comunidade de inteligência computacional.

3.1

Introdução

Como exemplo de técnicas clássicas, Ramanathan, Engle, Granger, Vahid-Araghi & Brace (1997) adotaram uma estratégia de modelagem hora por hora, onde o problema de previsão para cada hora é tratado separadamente. Eles propuseram um modelo de regressão múltipla para cada hora do dia e com dias da semana e de final de semana tratados separadamente, resultando num total de 48 modelos. Ramanathan, Granger & Engle (1985) desenvolveram uma metodologia em duas etapas. Primeiramente é formulado um modelo que traduz a carga horária por variáveis que refletem mudanças bruscas como temperatura, hora do dia, dia da semana, mês correspondente. Numa segunda etapa são definidos fatores que substituem as variáveis de influência lenta, como preço de energia elétrica, crescimento demográfico e industrial.

Schneider, Takenawa & Schiffman (1985) formularam um sistema de previsão da carga em termos de 4 componentes: a carga básica, a carga sensível ao clima, um componente de erro e um fator de correção. A carga básica é modelada

como a soma de funções senoidais. Os efeitos climáticos foram capturados por uma função não-linear. O viés na previsão destes termos é usado para construir o fator de correção que é adicionado à previsão para obter uma previsão atualizada. Gupta (1985) propôs uma adaptação deste modelo incorporando dados históricos e informações climáticas.

Ackerman (1985) comparou três modelos alternativos: (a) uma diferenciação de ordem 24 dos dados históricos modelados como um processo AR de primeira ordem; (b) diferenças consecutivas modeladas como um AR(24) combinado com um MA de primeira ordem e (c) um modelo AR simples em um conjunto de dados diferentes que eliminou a entrada de carga despachada.

Harvey & Koopman (1993) propuseram um modelo usando *splines* variantes no tempo. Os resultados prévios da literatura (Ramanathan et al. 1997, Connor & Martin 1994), mostram que o ciclo diário é forte, ratificando a decisão de desenvolver um modelo para cada hora.

Dentre as propostas de inteligência computacional, Connor & Martin (1994) propuseram um modelo de rede neural recorrente para cada hora. Choeiki et al. (1997) propuseram um procedimento de mínimos quadrados ponderados para treinar uma rede neural simples que produz como saída a previsão da carga horária de um dia inteiro. Lu et al. (1993) compararam duas arquiteturas de redes neurais: (a) a primeira é um modelo de previsão hora por hora e (b) uma arquitetura para prever um dia inteiro. Outras abordagens podem ser vistas em Park, Park & Lee (1991a), Yuan & Fine (1993), Park, El-Sharkawi, Marks, Atlas & Damborg (1991b).

3.2

O Modelo linear em dois níveis

A abordagem proposta é baseada na construção de um modelo, estimado individualmente para cada hora do dia, ou seja, os parâmetros são estimados 24 vezes, uma para cada hora do dia. O uso desta abordagem evita complexidade na modelagem do perfil diário da carga, que varia para diferentes tipos de dias. Ramanathan et al. (1997) também utilizam uma abordagem que propõem um modelo de previsão para cada hora, mas incluem nos modelos a dinâmica dos erros de horas anteriores, o que torna as abordagens totalmente diferentes.

Seja $L_{h,d}$, a série temporal representando a carga da hora h do dia d , em que $h = 1, \dots, 24$, $d = 1, \dots, D$ onde D é o total de dias observados.

3.2.1

Formulação Matemática

Considere-se a série formada pelos valores de carga de cada hora do dia, modelada pela soma de duas componentes. A primeira, determinística, representando a tendência, o ciclo anual e os efeitos dos diferentes tipos de dia. A outra componente é descrita por um modelo linear auto-regressivo (AR) de baixa ordem. O modelo **auto-regressivo sazonal em dois níveis**, denominado aqui *TLRAR* (sigla para *Two-Level Seasonal AutoRegressive* por coerência com a literatura corrente) é definido como:

$$L_{h,d} = L_{h,d}^R + L_{h,d}^I, \quad (3-1)$$

onde

$$L_{h,d}^R = \alpha_0 + \rho d + \sum_{r=1}^H (\alpha_r \cos(\omega r d) + \beta_r \sin(\omega r d)) + \sum_{i=1}^D \mu_i \delta_i \quad (3-2)$$

é a “carga regular”, e

$$L_{h,d}^I = \phi' \mathbf{z}_{h,d} + \varepsilon_{h,d} \quad (3-3)$$

é a “carga irregular”, onde: $\omega = 2\pi/365 \text{ rd/s}$ é a frequência fundamental; $\alpha_r \cos(\omega r) + \beta_r \sin(\omega r)$ é a harmônica de ordem r ; δ_i , $i = 1, \dots, D$ são variáveis binárias identificando os dias da semana, feriados, dias especiais e outros dias atípicos (feriados de categorias profissionais, meio-expediente, ponto facultativo, dias após feriados, entre outros), totalizando 15 tipos diferentes de dias (a tabela 3.1 mostra a classificação completa dos tipos de dias); α_0 , ρ , α_r , β_r , $r = 1, \dots, H$ e μ_i , $i = 1, \dots, D$ são parâmetros a serem estimados; $\mathbf{z}_{h,d}$ é formado por uma constante e um subconjunto de p defasagens de $L_{h,d}^I$; $\phi \in \mathbb{R}^{p+1}$ é um vetor de parâmetros; e $\varepsilon_{h,d}$ é um ruído branco com variância σ_ϵ^2 .

Vários aspectos do modelo *TLRAR* merecem comentários. Primeiramente, a tendência é modelada como uma função determinística do tempo. Modelar-se o ciclo anual como uma soma de senos e cossenos. A motivação para modelagem com funções senoidais pode ser facilmente visualizada pela inspeção da figura 2.1. O efeito dos dias da semana e dias especiais, como os feriados, é modelado com variáveis binárias. Diversos autores consideram que terças, quartas, quintas e sextas-feiras podem ser modeladas como um mesmo tipo de dia, entretanto, devido a disponibilidade de grande massa de dados, cada tipo de dia é aqui modelado por uma variável binária.

3.2.2

Estratégia de Modelagem

A metodologia desenvolvida é dividida em duas etapas: na primeira são estimadas as componentes determinísticas relativas à tendência, à sazonalidade anual, ao efeito dos dias da semana, dos dias especiais, dos feriados, dos meio-expedientes e dos dias intercalados. Na segunda etapa ajusta-se um modelo linear auto-regressivo aos resíduos da primeira.

Classificação dos dias

Na literatura, o perfil da carga é previsto pela curvas de cargas de dias similares no histórico da carga. Por esta razão, os dias são agrupados dentro de classes de diferentes tipos de dias. Informar o modelo sobre o tipo de dia é importante, porque nos dias úteis as cargas são mais elevadas que nos outros dias. A classificação de dias utilizada é mostrada na tabela 3.1.

Tabela 3.1: Classificação dos tipos de dias para efeito de previsão

código	descrição
1	domingo
2	segunda-feira
3	terça-feira
4	quarta-feira
5	quinta-feira
6	sexta-feira
7	sábado
8	feriado oficial ou religioso
9	dia após feriado
10	dia anterior a feriado
11	dia útil entre feriado e final de semana
12	sábado após feriado
13	meio-expediente manhã
14	meio-expediente tarde
15	feriado profissional

Estimação dos parâmetros

O número de harmônicas é determinado minimizando-se o Critério da Informação Bayesiana de Schwarz-*SBIC* (Schwarz 1978), definido por:

$$SBIC = \ln(\hat{\sigma}^2) + p \frac{\ln(n)}{n} \quad (3-4)$$

onde: $\hat{\sigma}^2$ é o estimador de máxima verossimilhança da variância dos resíduos, σ_ϵ^2 ; p é o número de parâmetros estimados no modelo, incluindo o termo constante; e n é o número de observações da série.

Um outro critério de seleção, o Critério da Informação de Akaike-*AIC* (Akaike 1974), também foi calculado, mas apresentou pouca diferença no proces-

so de seleção. O *SBIC* foi escolhido por ser mais parcimonioso.

As harmônicas capturam os efeitos sazonais e outros efeitos relativos às estações do ano e comportamentos sociais como, por exemplo, períodos de férias de verão. A figura 3.1 mostra o perfil das harmônicas selecionadas para o ciclo anual. A fundamental captura o efeito do ciclo sazonal anual e as demais harmônicas, uma combinação de outros efeitos contidos nas séries, sendo as variações climáticas das estações do ano um efeito importante. Além disso, os pontos de mínimos e máximos se alteram ao longo do ano de acordo com as variações da temperatura.

A seleção dos parâmetros dos modelos lineares auto-regressivos também é baseada no *SBIC*, com o objetivo de identificar a combinação de defasagens que propicia o menor *SBIC* com o conjunto de dados utilizados na estimação dos parâmetros.

Resumidamente, o procedimento de estimação pode ser descrito como:

- Para cada hora estimar os parâmetros: $\alpha_0, \rho, \alpha_r, \beta_r$ e μ_i , com $r = 1, \dots, H$ e $i = 1, \dots, D$ em (3-2) por Mínimos Quadrados Ordinários (MQO). O número de harmônicas, H , é determinado pela minimização do *SBIC*. As variáveis binárias representando os diferentes tipos de dia são fixadas em 15, como descrito na tabela 3.1.
- Após estimar a “carga regular”, calcular os resíduos $\hat{L}_{h,d}^S = L_{h,d} - \hat{L}_{h,d}^P$.
- Utilizando o *SBIC*, outra vez, selecionar a melhor combinação de defasagens de $\mathbf{z}_{h,d}$ em (3-3) entre as sete primeiras defasagens da série.
- Após estimar o modelo, testar a correlação dos resíduos. Se houver evidências de correlação incluir outras defasagens e voltar ao passo anterior.

Tabela 3.2: Classificação dos tipos de dias para o modelo *DASARIMA*

código	descrição
1	dias da semana (dom, seg, ter, qua, qui, sex e sab)
2	feriado oficial ou religioso
3	dia após feriado
4	dia anterior a feriado
5	dia útil entre feriado e final de semana
6	sábado após feriado
7	meio-expediente manhã
8	meio-expediente tarde
9	feriado profissional

3.3

Modelo de Referência

3.3.1

Formulação Matemática

Muitos autores tem utilizado como referência um modelo *SARIMA* (*e.g.*, Darbellay & Slama (2000)). A referência aqui adotada também será um modelo (*SARIMA*) porém, ajustado por variáveis binárias para capturar os efeitos dos feriados e dias especiais e será denominado *DASARIMA* (sigla para *Dummy-Adjusted Seasonal Integrated AutoRegressive Moving Average* por coerência com a literatura corrente).

Seja $\overline{\Delta_7 \Delta_1 L_{h,d}}$ as séries resultantes da primeira e sétima diferenças aplicadas às séries originais extraídos os efeitos dos tipos de dias através de variáveis binárias:

$$\overline{\Delta_7 \Delta_1 L_{h,d}} = \Delta_7 \Delta_1 L_{h,d} - \sum_{i=1}^D \alpha_i \delta_{i,d} - \sum_{i=1}^D \lambda_i \delta_{i,d-1} - \sum_{i=1}^D \gamma_i \delta_{i,d-7} \quad (3-5)$$

Então o modelo *DASARIMA* é definido por:

$$(1 - \phi_1 B) \overline{\Delta_7 \Delta_1 L_{h,d}} = (1 + \theta_1 B + \theta_2 B^2 + \theta_3 B^3)(1 + \beta B^7) \epsilon_{h,d} \quad (3-6)$$

onde: $\Delta_j = (1 - B^j)$, $j = 1, 7$; B é o operador de retardo; θ_k com $k = 1, 2, 3$, α_i , β , γ_i e λ_i com $i = 1, \dots, D$ são parâmetros; δ_i com $i = 1, \dots, D$ são variáveis binárias identificando feriados (públicos, religiosos e profissionais) e dias especiais; e $\epsilon_{h,d}$ é o erro da hora h do dia d , um ruído branco.

Para o modelo *DASARIMA* foi adotada a classificação de dias mostrada na tabela 3.2. Esta classificação considera que a aplicação de diferenças de primeira e sétima ordem capturam os efeitos do dia anterior e o ciclo semanal, respectivamente, permitindo classificar os dias da semana em um único tipo.

3.3.2

Estratégia de Modelagem

Estima-se um modelo *DASARIMA* para cada hora do dia. A estrutura de cada modelo foi definida pelo exame das funções de autocorrelação (FAC) e de autocorrelação parcial (FACP) dos dados, e pela análise da estatística t . Os termos correspondentes aos parâmetros não significativos foram desprezados. A tabela 3.3 mostra os parâmetros selecionados para cada um dos modelos com os desvios-padrão assinalados abaixo do valor de cada parâmetro.

Tabela 3.3: Parâmetros significativos dos modelos *DASARIMA*

hora	$\hat{\mu}$	$\hat{\phi}_1$	$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\theta}_3$	$\hat{\beta}$
1	0.018 (0.315)	0.299 (0.080)	-0.318 (0.079)	-0.164 (0.018)	-0.076 (0.023)	-0.817 (0.011)
2	-0.008 (0.280)	0.307 (0.081)	-0.324 (0.080)	-0.165 (0.018)	-0.087 (0.024)	-0.823 (0.010)
3	-0.012 (0.271)	0.305 (0.082)	-0.284 (0.081)	-0.173 (0.018)	-0.098 (0.024)	-0.825 (0.010)
4	-0.008 (0.269)	0.282 (0.087)	-0.252 (0.086)	-0.174 (0.018)	-0.112 (0.025)	-0.819 (0.010)
5	-0.012 (0.275)	0.262 (0.089)	-0.219 (0.088)	-0.177 (0.018)	-0.114 (0.025)	-0.813 (0.011)
6	-0.007 (0.311)	0.261 (0.098)	-0.230 (0.097)	-0.149 (0.018)	-0.111 (0.025)	-0.787 (0.011)
7	0.013 (0.426)	-	-0.043 (0.147)	-0.094 (0.019)	-0.105 (0.024)	-0.726 (0.012)
8	0.034 (0.689)	-	-0.076 (0.286)	-0.041 (0.028)	-0.049 (0.022)	-0.623 (0.014)
9	0.080 (0.924)	-	-0.114 (0.363)	-	-0.033 (0.018)	-0.576 (0.014)
10	0.074 (1.097)	-	-0.143 (0.382)	-	-	-0.530 (0.015)
11	0.070 (1.098)	-	-0.157 (0.503)	-	-	-0.542 (0.015)
12	0.057 (1.073)	-	-0.133 (0.460)	-	-	-0.544 (0.015)
13	0.059 (0.989)	-	-0.135 (0.384)	-0.042 (0.054)	-	-0.567 (0.014)
14	0.053 (1.035)	-	-0.129 (0.292)	-0.064 (0.041)	-	-0.563 (0.015)
15	0.045 (1.064)	-	-0.122 (0.299)	-0.063 (0.041)	-0.035 (0.027)	-0.570 (0.014)
16	0.034 (1.019)	-	-0.125 (0.274)	-0.056 (0.039)	-0.053 (0.025)	-0.576 (0.014)
17	0.057 (0.889)	-	-0.179 (0.242)	-0.041 (0.047)	-0.055 (0.020)	-0.607 (0.014)
18	0.069 (0.719)	-0.548 (0.165)	0.301 (0.164)	-0.165 (0.046)	-0.080 (0.017)	-0.629 (0.014)
19	0.031 (0.617)	-0.476 (0.218)	0.264 (0.217)	-0.158 (0.050)	-0.070 (0.019)	-0.633 (0.014)
20	0.032 (0.480)	-	-0.167 (0.242)	-0.052 (0.044)	-0.065 (0.024)	-0.669 (0.013)
21	0.017 (0.423)	-	-0.131 (0.137)	-0.108 (0.026)	-0.062 (0.024)	-0.693 (0.013)
22	0.001 (0.387)	0.223 (0.104)	-0.300 (0.103)	-0.116 (0.020)	-0.076 (0.024)	-0.733 (0.012)
23	0.003 (0.342)	0.236 (0.084)	-0.292 (0.083)	-0.158 (0.019)	-0.091 (0.023)	-0.775 (0.011)
24	0.025 (0.369)	0.280 (0.081)	-0.316 (0.080)	-0.169 (0.019)	-0.080 (0.023)	-0.777 (0.011)

3.4

O Experimento

A abordagem de modelagem utilizada considera que o método estatístico baseia-se em princípios onde elementos subjetivos não têm dominância. Um dos problemas de especificação que requer atenção especial é a seleção de variáveis (Rech, Teräsvirta & Tschernig 2001). O ensaio realizado para mostrar a utilidade prática do modelo consiste em calcular as previsões da carga horária de 1 a 168 passos à frente com os modelos *TLSAR* e *DASARIMA*. A seção 3.4.1 mostra a especificação e os resultados da estimação dos modelos. Os resultados e as comparações da previsão são apresentados na seção 3.4.2.

3.4.1

Especificação e Estimação

As condições de carga mudam relativamente muito devagar e fatores básicos fazem as características de carga de um dia, por alguma razão, poder serem vistas nas cargas dos dias anteriores no caso dos dias úteis. Isto porém, não é verificado

de um domingo para uma segunda-feira, por exemplo. O efeito do calendário, modelado por δ_i , não é fixo para todas as horas, pois é evidente que os dias úteis afetam mais a carga das horas do meio do dia (horas de trabalho) do que as horas iniciais e finais do dia. As séries de carga relativas à cada hora mostram comportamentos diferentes entre si, porém as séries relativas as horas da madrugada (22:00h até 6:00h) apresentam mais regularidade entre si. O perfil das harmônicas da componente determinística, conforme mostra a figura 3.1, apresenta um comportamento associado às estações do ano e os pontos de mínimo acompanham o comportamento da sociedade, aparecendo mais claramente nas horas do início e final do dia.

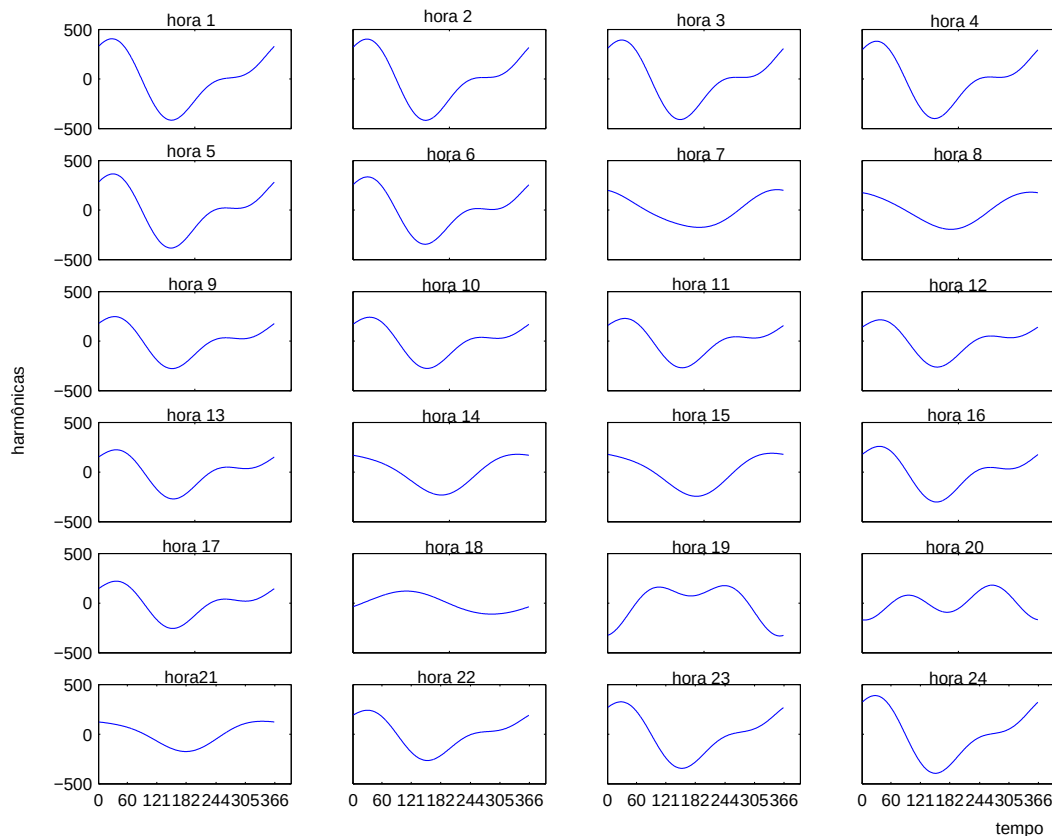


Figura 3.1: perfil das harmônicas da carga regular.

Na tabela 3.4 são apresentadas os valores dos parâmetros selecionados estimados conforme definido em (3-1-3-3). A figura 3.2 mostra as funções de autocorrelação (FAC) e a figura 3.3 as funções de autocorrelações parciais (FACP) das componentes irregulares dos modelos, indicando a possibilidade de ajuste de um modelo $AR(p)$. A figura 3.4 mostra os valores observados das séries relativas a cada hora e a componente determinística estimada para o período de estimação dos parâmetros. A análise visual mostra que os parâmetros estimados capturam bem os ciclos existentes nas séries.

Tabela 3.4: Parâmetros estimados do modelo *TLRAR*

hora	harmônicas #	$\hat{\phi}_0$	$\hat{\phi}_1$	$\hat{\phi}_2$	$\hat{\phi}_3$	$\hat{\phi}_4$	$\hat{\phi}_5$	$\hat{\phi}_6$	$\hat{\phi}_7$	$\hat{\sigma}^2$
1	2	-0.052	0.855	-0.137	—	0.046	—	—	0.057	1.000
2	2	-0.082	0.875	-0.148	—	0.044	—	—	0.058	0.902
3	2	-0.067	0.912	-0.181	—	0.051	—	—	0.052	0.792
4	2	-0.081	0.924	-0.191	—	0.056	—	—	0.049	0.725
5	2	-0.083	0.933	-0.199	—	0.060	—	—	0.047	0.664
6	2	-0.103	0.916	-0.175	—	0.054	—	—	0.053	0.596
7	5	-0.164	0.814	-0.090	—	0.047	—	—	0.068	0.526
8	3	-0.136	0.721	—	—	—	—	—	0.086	0.525
9	2	-0.140	0.660	—	—	—	—	—	0.115	0.653
10	2	-0.164	0.610	—	—	—	—	—	0.143	0.735
11	2	-0.211	0.577	—	—	—	—	—	0.156	0.790
12	2	-0.219	0.583	—	—	—	—	—	0.159	0.765
13	2	-0.224	0.592	—	—	—	—	—	0.154	0.787
14	3	-0.201	0.582	-0.004	0.042	—	—	—	0.138	0.888
15	3	-0.219	0.588	-0.007	0.045	—	—	—	0.130	0.967
16	2	-0.229	0.611	—	—	—	—	—	0.135	0.937
17	2	-0.267	0.555	—	—	0.047	—	—	0.136	0.939
18	1	-0.283	0.519	—	—	0.084	—	—	0.127	0.787
19	5	-0.318	0.597	—	—	0.072	—	0.043	0.092	0.596
20	6	-0.291	0.641	—	—	0.063	—	—	0.110	0.468
21	3	-0.215	0.706	-0.043	0.049	—	—	0.056	0.077	0.458
22	2	-0.210	0.784	-0.080	—	0.045	—	0.031	0.058	0.599
23	2	-0.185	0.825	-0.121	—	0.054	—	—	0.058	0.811
24	2	-0.163	0.858	-0.144	—	0.046	—	—	0.059	0.944

3.4.2

Resultados da Previsão

Um ensaio de previsão utilizando o modelo *TLRAR* foi realizado com a base de dados descrita no capítulo 2, dividida em dois sub-conjuntos. O sub-conjunto formado pelos dados que compreendem o período entre 01 de janeiro de 1990 a 31 de dezembro de 1998 é usado para estimar os parâmetros do modelo, definido pelas equações (3-1-3-2) uma única vez, o que parece ser suficiente. O sub-conjunto restante, compreendendo o período entre o primeiro dia de 1999 e o último dia de 2000, é separado em dois períodos de um ano para permitir uma análise mais detalhada.

A previsão é feita de 1 a 7 passos à frente para cada uma das horas, considerando o conjunto das 24 previsões 1 passo à frente, o que produz desde a previsão do dia inteiro de uma vez (perfil de carga) até a curva prevista de carga relativa a uma semana, considerando todos os 168 valores previstos. A idéia pode ser adaptada para uma parte do dia de acordo com os patamares (médio, leve e pesado) de carga e então, a carga de poucas horas poderia ser prevista com base em poucas horas anteriores, de forma similar, como parte da curva de carga diária. Este raciocínio conduz à criação de um modelo robusto. Os resultados obtidos com o modelo *TLRAR* são comparados com os resultados do modelo *DASARIMA* através da medida de erro definida a seguir.

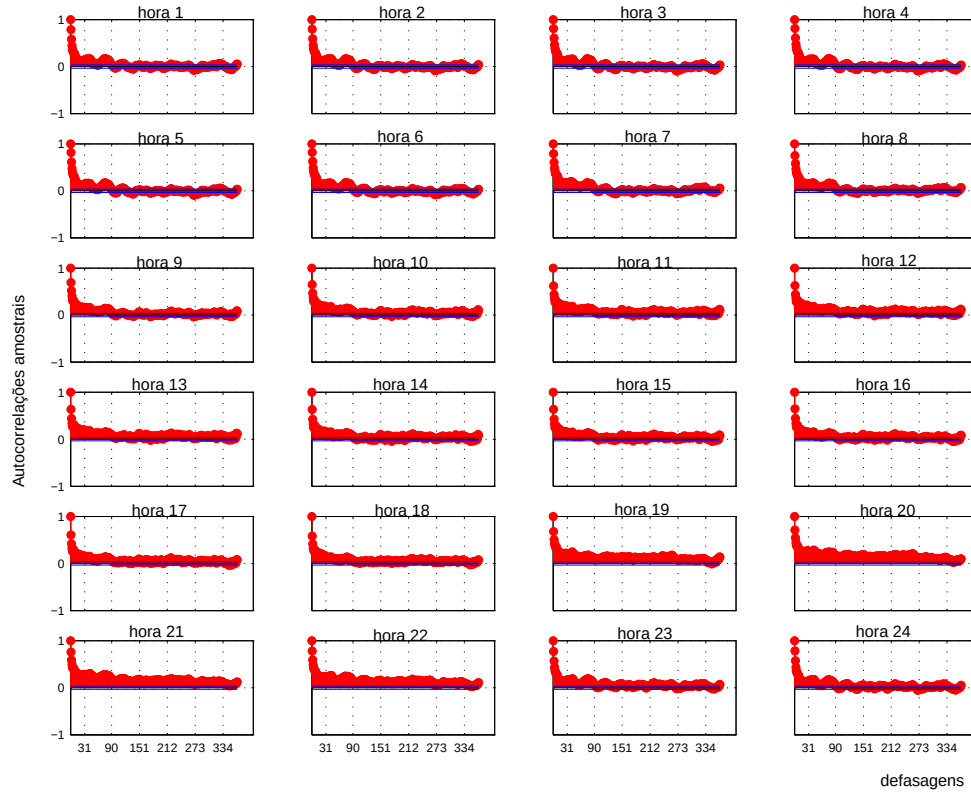


Figura 3.2: autocorrelações das componentes irregulares

Medida do erro

A medida de desempenho mais utilizada na literatura de previsão de carga (por exemplo: Lourenço et al. (1999), Darbellay & Slama (2000)) é o erro absoluto médio percentual (MAPE), que mede a proporcionalidade entre o erro de previsão e o valor observado da carga, e é definido como:

$$MAPE = \frac{1}{N} \sum_{i=1}^N \left[\left| \frac{\hat{L}_i - L_i}{L_i} \right| \right] * 100\%, \quad (3-7)$$

onde: N é o número de valores previstos; L_i é o valor da i -ésima carga; $\hat{L}_i$ é o valor da i -ésima carga prevista.

Outras medidas poderiam ser usadas e alguns trabalhos sugerem que estas medidas devem penalizar os maiores erros. Outros sugerem que deve ser uma medida de fácil compreensão e relacionada às necessidades dos tomadores de decisões. Como exemplo da primeira forma, Armstrong & Collopy (1992) sugerem a raiz do erro quadrático médio (RMSE), enquanto para a outra forma, Bakirtzis et al. (1996) usam o erro absoluto médio (MAE). Entretanto, falta comparabilidade à essas medidas para séries diferentes e, talvez por isso, o MAPE permaneça como a medida padrão. O MAPE permite algumas considerações comparativas (embora os resultados não sejam diretamente comparáveis em situações diferentes).

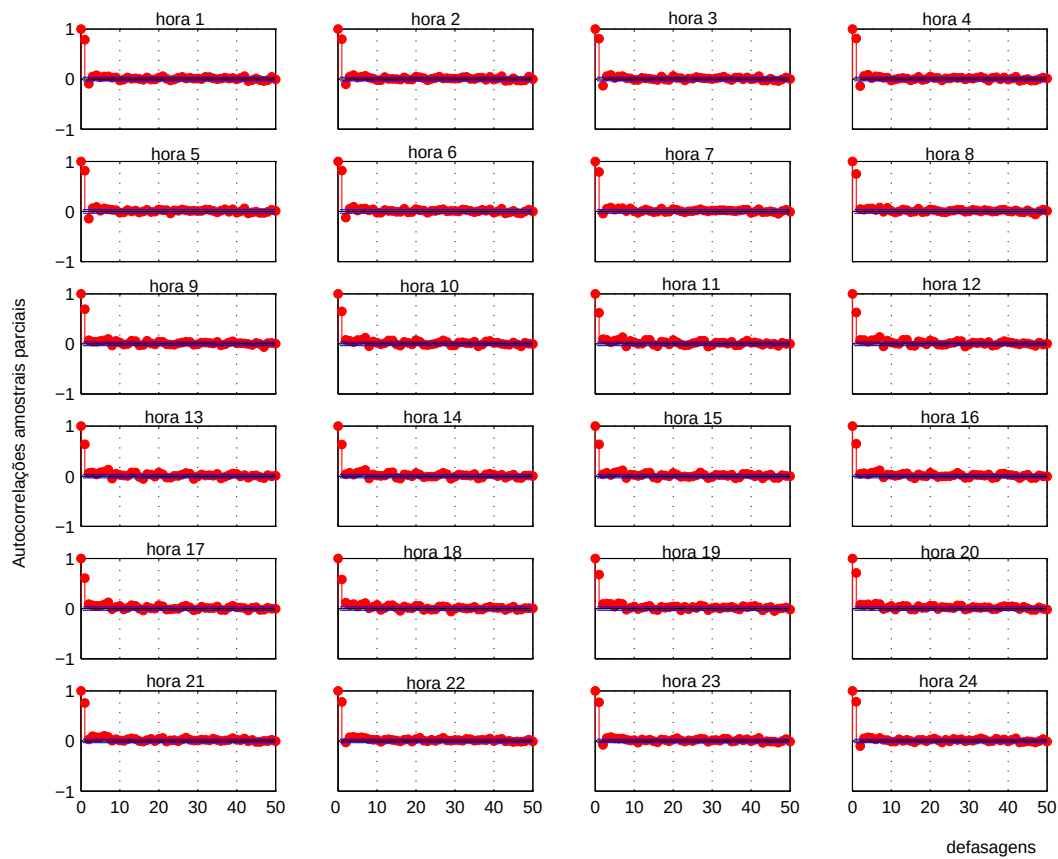


Figura 3.3: autocorrelações parciais das componentes irregulares

análise e comentários 1

A análise dos erros de previsão do período completo (1999–2000), mostra valores médios entre 2.51% (1 passo à frente) e 7.47% (7 passos à frente) para o *TLSAR* e entre 3.01% (1 passo à frente) e 17.95% (5 passos à frente) para o modelo *DASARIMA*. O *TLSAR* é melhor 1 passo à frente para todas as horas, variando entre 2.51% e 4.48%. O modelo *DASARIMA* é melhor nas cinco primeiras horas, a maioria das vezes, com pequena diferença, mas apresenta nas horas do meio do dia, erros da ordem de dois dígitos. A tabela 3.5 mostra os erros de todas as horas de 1 a 7 passos à frente com o melhor MAPE em negrito. A figura 3.5 mostra os MAPE médios para o período completo de previsão.

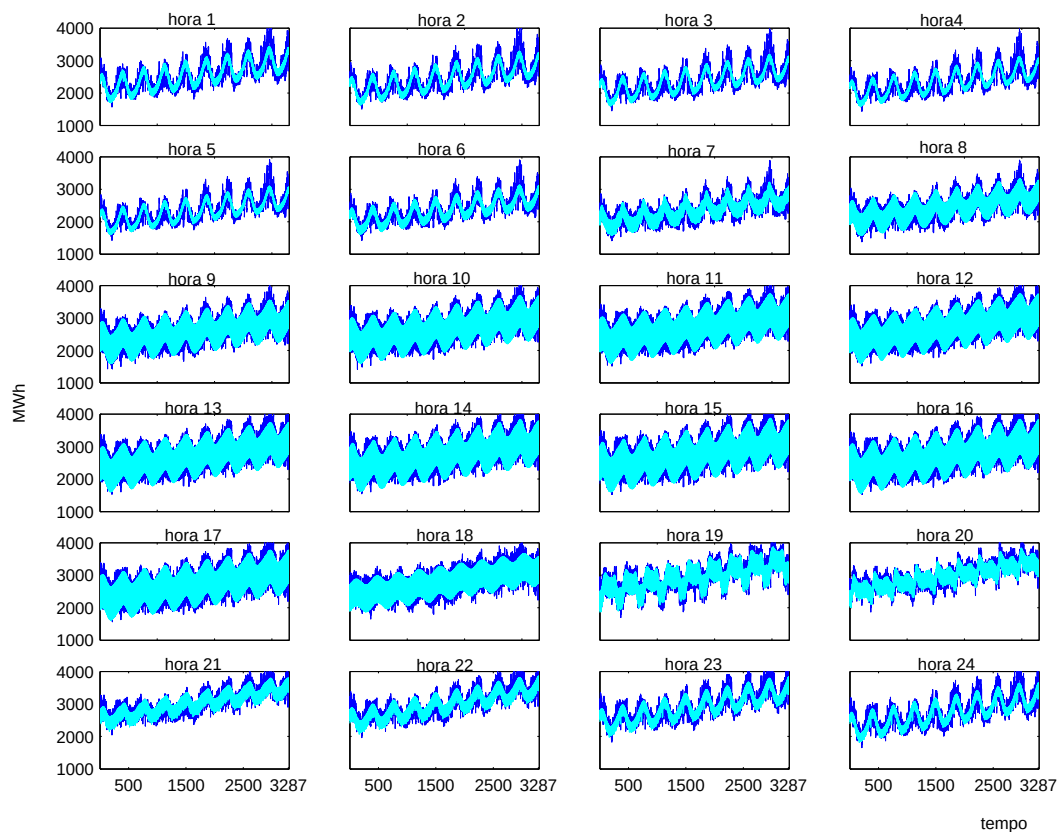


Figura 3.4: carga real (escuro), componente determinística(claro)

análise e comentários 2

A análise dos erros do ano de 1999, mostra valores médios entre 2.31% (1 passo à frente) e 6.56% (7 passos à frente) para o *TLSAR* e entre 2.85% (1 passo à frente) e 17.34% (5 passos à frente) para o modelo *DASARIMA*. O *TLSAR* é melhor 1 passo à frente para todas as horas, apresentando erro máximo de 4.71%. O *DASARIMA* é melhor nas cinco primeiras horas, a maioria das vezes, com pequena diferença, mas apresenta nas horas do meio do dia, erros da ordem de dois dígitos. A tabela 3.6 mostra os erros de todas as horas de 1 a 7 passos à frente com o melhor MAPE em negrito. A figura 3.6 mostra os MAPE médios para o período completo de previsão.

análise e comentários 3

A análise dos erros do ano de 2000, mostra valores médios entre 2.72% (1 passo à frente) e 6.56% (7 passos à frente) para o *TLSAR* e entre 3.17% (1 passo à frente) e 18.57% (5 passos à frente) para o modelo *DASARIMA*. O *TLSAR* é melhor 1 passo à frente para todas as horas, apresentando erro máximo de 8.63%. O *DASARIMA* é melhor nas cinco primeiras horas, a maioria das vezes, com pequena diferença, mas apresenta nas horas do meio do dia, erros da ordem de dois dígitos. A tabela 3.7 mostra os erros de todas as horas de 1 a 7 passos à frente com o melhor MAPE em negrito. A figura 3.7 mostra os MAPE médios para o período completo de previsão. Para este ano os resultados

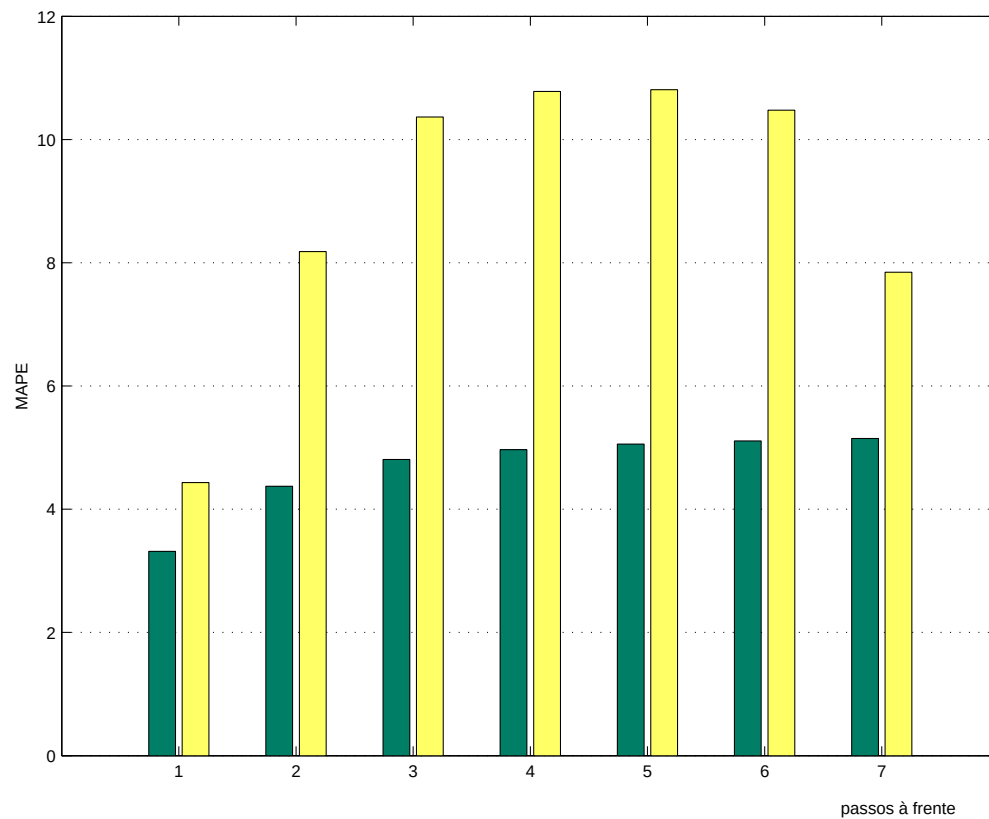


Figura 3.5: MAPE médio (1999 – 2000) - *TLSAR* (escuro) e *DASARIMA* (claro)

de ambos os modelos são piores e este fato pode ter influência da mudança de nível ocorrida. A figura 3.8 mostra as curvas de carga diárias, real e previstas pelos modelos *TLSAR* e *DASARIMA* de dias úteis escolhidos arbitrariamente, nas quatro estações do ano. Os modelos apresentam melhor ajuste nos meses de clima mais ameno do que naqueles que apresentam temperaturas elevadas. Os resultados do modelo *DASARIMA* são melhores nas primeiras cinco horas do dia, período menos importante quanto a acurácia dos modelos e para horizontes de previsão mais distantes são claramente imprecisos, com erros crescendo bastante rápido e inaceitavelmente grandes.

Tabela 3.5: MAPE - (1999 – 2000) dos modelos *TLSAR* e *DASARIMA*

hora	<i>TLSAR</i>							<i>DASARIMA</i>						
	passos à frente							passos à frente						
	1	2	3	4	5	6	7	1	2	3	4	5	6	7
1	4.05	5.85	6.59	6.89	7.09	7.19	7.26	4.23	5.35	6.09	6.27	6.28	6.14	5.21
2	4.07	5.93	6.72	7.07	7.28	7.40	7.47	4.19	5.26	5.90	6.03	5.99	5.95	5.16
3	3.95	5.81	6.67	7.04	7.23	7.35	7.44	4.10	5.12	5.79	5.90	5.82	5.77	4.99
4	3.91	5.64	6.51	6.87	7.09	7.21	7.30	4.04	4.97	5.68	5.80	5.77	5.65	4.83
5	3.69	5.37	6.19	6.54	6.75	6.88	6.96	3.85	4.66	5.42	5.56	5.57	5.38	4.59
6	3.41	4.91	5.62	5.95	6.12	6.24	6.33	3.74	4.74	5.77	6.21	6.21	5.79	4.75
7	3.19	4.33	4.84	5.10	5.24	5.33	5.39	4.08	6.59	8.47	9.12	9.18	8.57	6.45
8	2.99	3.92	4.34	4.51	4.62	4.69	4.75	4.39	8.76	11.23	11.94	12.06	11.48	8.50
9	2.99	3.82	4.14	4.25	4.33	4.37	4.39	4.78	10.51	13.33	14.02	14.16	13.74	10.15
10	2.99	3.74	4.01	4.06	4.10	4.12	4.13	5.08	11.70	14.72	15.35	15.46	15.18	11.27
11	2.98	3.70	3.94	3.98	4.01	4.03	4.03	5.17	12.25	15.65	16.15	16.23	16.05	11.70
12	2.99	3.64	3.86	3.90	3.92	3.93	3.94	5.31	12.42	16.04	16.47	16.53	16.40	11.84
13	3.07	3.78	4.00	4.05	4.06	4.07	4.08	5.22	12.12	15.79	16.29	16.33	16.00	11.45
14	3.32	4.11	4.36	4.41	4.42	4.42	4.42	5.33	12.43	16.36	16.90	16.90	16.56	11.78
15	3.42	4.26	4.55	4.62	4.63	4.62	4.63	5.45	12.87	17.24	17.74	17.80	17.39	12.10
16	3.39	4.21	4.54	4.61	4.63	4.64	4.64	5.44	12.78	17.35	17.90	17.95	17.42	11.99
17	3.27	3.96	4.21	4.26	4.29	4.29	4.30	5.10	11.88	16.40	16.92	17.01	16.37	11.11
18	3.05	3.64	3.88	3.96	4.00	4.01	4.02	4.53	9.11	12.07	12.56	12.62	12.14	8.66
19	2.94	3.52	3.77	3.86	3.91	3.93	3.95	3.79	6.23	7.68	8.07	8.07	7.84	6.15
20	2.59	3.12	3.28	3.32	3.39	3.41	3.44	3.27	5.03	5.69	5.96	5.99	5.78	4.93
21	2.51	3.18	3.41	3.51	3.55	3.59	3.62	3.01	4.68	5.56	5.84	5.89	5.56	4.49
22	2.80	3.70	4.10	4.23	4.28	4.33	4.36	3.22	5.08	6.41	6.82	6.80	6.33	4.84
23	3.59	4.80	5.28	5.46	5.54	5.58	5.63	4.09	5.65	6.95	7.36	7.29	6.82	5.39
24	4.48	5.97	6.62	6.87	6.98	7.05	7.13	4.96	6.19	7.17	7.51	7.46	7.12	6.01
min	2.51	3.12	3.28	3.32	3.39	3.41	3.44	3.01	4.66	5.42	5.56	5.57	5.38	4.49
med	3.32	4.37	4.81	4.97	5.06	5.11	5.15	4.43	8.18	10.37	10.78	10.81	10.48	7.85
max	4.48	5.97	6.72	7.07	7.28	7.40	7.47	5.45	12.87	17.35	17.90	17.95	17.42	12.10

3.5

Resumo

Prever o perfil da carga diária foi o objetivo dos primeiros estudos relatados na literatura. O perfil da curva de carga foi previsto pela previsão das séries resultantes do seccionamento da série de dados históricos em 24 séries referentes às 24 horas do dia.

A idéia básica desenvolvida para modelar a série, pode ser aplicada à séries com características similares. O objetivo é propor um modelo para realizar previsões a curto prazo, um dia inteiro à frente. Este horizonte de previsão é importante para o processo de planejamento da operação do sistema. Os modelos univariados são utilizados para modelar as série de cargas, apesar de ser possível produzir melhores previsões através de um modelo estrutural que inclua informações sobre temperatura, umidade do ar, ventos, etc... Este é um fato que merece maior atenção, pois dados climáticos poderiam melhorar substancialmente a previsão, particularmente no período do verão em que os aparelhos de condicionamento de ar são responsáveis por grande parte da carga consumida. Entretanto, o efeito da temperatura na carga é não-linear, como atestado por grande número de artigos (veja Hippert, Pedreira & Souza (2001)), e incluí-la adicionaria ao modelo linear, provavelmente, um relacionamento não-linear.

Uma parte do conjunto de dados foi utilizada para estimar os parâmetros representativos da sazonalidade e tendência, utilizando uma abordagem estática

Tabela 3.6: MAPE - ano 1999 dos modelos *TLSAR* e *DASARIMA*

hora	<i>TLSAR</i> passos à frente							<i>DASARIMA</i> passos à frente						
	1	2	3	4	5	6	7	1	2	3	4	5	6	7
1	3.76	5.12	5.67	5.82	5.95	6.07	6.18	3.93	4.93	5.63	5.84	5.83	5.63	4.72
2	3.68	5.17	5.74	5.93	6.06	6.19	6.33	3.85	4.79	5.37	5.50	5.45	5.39	4.66
3	3.60	5.06	5.67	5.89	5.96	6.10	6.25	3.77	4.65	5.30	5.43	5.28	5.21	4.49
4	3.63	4.93	5.54	5.76	5.87	6.01	6.15	3.72	4.65	5.31	5.46	5.38	5.23	4.43
5	3.38	4.55	5.14	5.37	5.47	5.62	5.73	3.47	4.24	5.02	5.21	5.18	4.89	4.08
6	3.08	4.15	4.63	4.84	4.93	5.07	5.17	3.34	4.38	5.34	5.93	5.97	5.42	4.36
7	2.83	3.60	3.94	4.06	4.13	4.20	4.27	3.69	6.26	8.26	9.13	9.20	8.42	6.32
8	2.69	3.28	3.53	3.61	3.67	3.74	3.80	4.09	8.43	11.04	11.8	11.99	11.38	8.38
9	2.74	3.29	3.47	3.51	3.56	3.59	3.60	4.63	10.33	13.18	13.76	13.98	13.62	10.05
10	2.76	3.28	3.46	3.46	3.49	3.50	3.51	4.91	11.38	14.42	14.89	15.07	14.84	11.05
11	2.76	3.28	3.45	3.44	3.44	3.46	3.47	4.95	11.81	15.19	15.48	15.67	15.57	11.40
12	2.71	3.21	3.32	3.33	3.33	3.33	3.34	5.02	12.06	15.61	15.81	15.93	15.96	11.58
13	2.86	3.38	3.52	3.55	3.55	3.56	3.56	5.04	11.69	15.32	15.64	15.72	15.50	11.19
14	3.11	3.71	3.88	3.92	3.90	3.90	3.89	5.10	12.01	15.91	16.14	16.20	16.04	11.45
15	3.23	3.84	4.02	4.07	4.08	4.07	4.07	5.16	12.45	16.78	17.08	17.16	16.89	11.79
16	3.18	3.70	3.97	3.99	4.01	4.01	4.02	5.11	12.40	16.92	17.33	17.34	16.93	11.70
17	3.04	3.50	3.72	3.76	3.78	3.79	3.80	4.82	11.57	16.16	16.57	16.64	16.11	10.97
18	2.81	3.25	3.43	3.51	3.53	3.54	3.55	4.27	8.97	11.95	12.46	12.56	12.19	8.65
19	2.73	3.27	3.41	3.46	3.51	3.54	3.56	3.54	6.04	7.55	8.02	8.01	7.89	6.12
20	2.33	2.76	2.80	2.82	2.87	2.89	2.93	3.07	4.90	5.63	5.98	5.94	5.77	4.88
21	2.31	2.78	2.89	2.94	2.92	2.95	3.02	2.85	4.47	5.44	5.84	5.82	5.44	4.42
22	2.56	3.18	3.45	3.46	3.44	3.47	3.53	3.08	4.85	6.21	6.73	6.70	6.09	4.76
23	3.50	4.35	4.68	4.69	4.67	4.73	4.82	4.19	5.66	6.96	7.43	7.33	6.85	5.47
24	4.71	5.85	6.28	6.36	6.35	6.43	6.56	5.42	6.60	7.57	7.93	7.95	7.54	6.43
min	2.31	2.76	2.8	2.82	2.87	2.89	2.93	2.85	4.24	5.02	5.21	5.18	4.89	4.08
med	3.08	3.85	4.15	4.23	4.27	4.32	4.38	4.21	7.9	10.09	10.48	10.51	10.20	7.64
max	4.71	5.85	6.28	6.36	6.35	6.43	6.56	5.42	12.45	16.92	17.33	17.34	16.93	11.79

onde os parâmetros dos modelos são estimados uma única vez. O desempenho dos modelos foi testado sobre os dados de carga do período referente à 01/01/1999 a 31/12/2000.

Foram obtidos erros médios de aproximadamente 2.5%. Os resultados foram, entretanto, melhores no período entre os meses de junho e agosto. Estes erros percentuais são menores do que muitos daqueles relatados na literatura. Alguns autores têm relatado erros de previsão para carga máxima e para carga diária total em torno de 2% (por exemplo Park et al. (1991a), Peng et al. (1992)). A razão para a escolha da medida de erro depende do conjunto de dados utilizado e dos objetivos do sistema de previsão. Resultados de métodos diferentes não devem ser comparados se se utilizam conjuntos de dados diferentes, considerando que alguns dados são mais ruidosos do que outros. A comparação entre os métodos é válida se o mesmo conjunto de dados é usado com diferentes métodos.

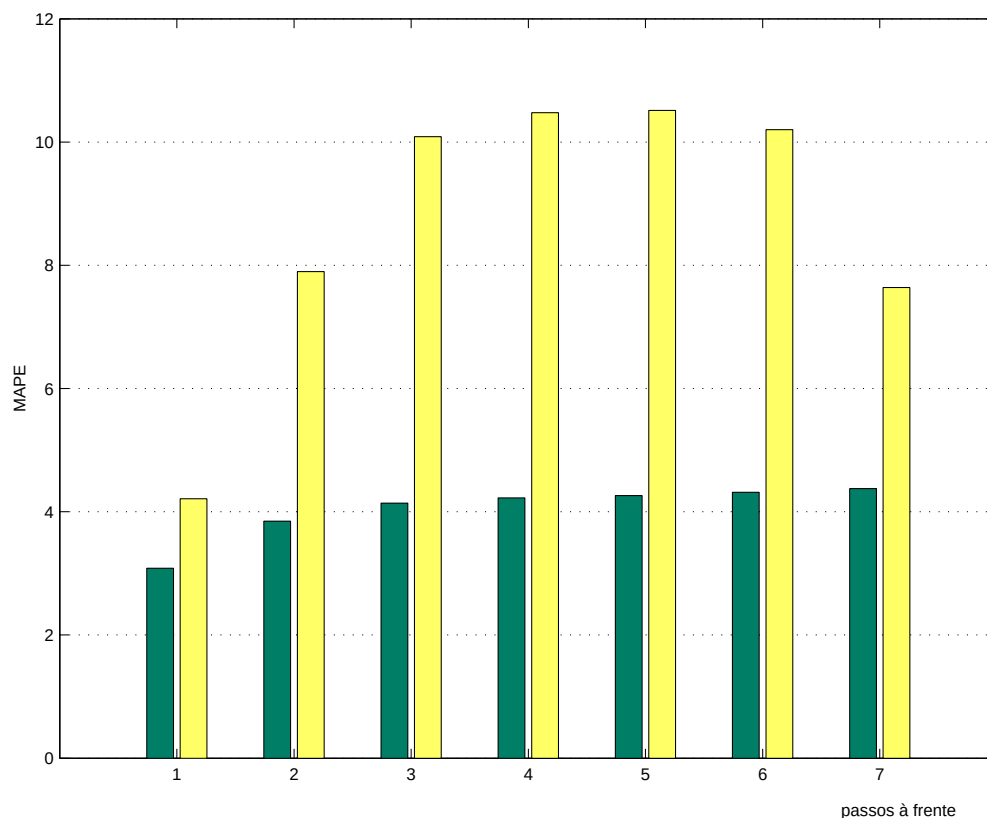


Figura 3.6: MAPE médio 1999 - *TLSAR* (escuro) e *DASARIMA* (claro)

Tabela 3.7: MAPE - ano 2000 dos modelos *TLSAR* e *DASARIMA*

hora	<i>TLSAR</i>							<i>DASARIMA</i>						
	passos à frente							passos à frente						
	1	2	3	4	5	6	7	1	2	3	4	5	6	7
1	4.34	6.58	7.51	7.95	8.22	8.30	8.33	4.53	5.77	6.55	6.71	6.73	6.65	5.70
2	4.45	6.70	7.70	8.20	8.50	8.60	8.62	4.53	5.73	6.42	6.55	6.54	6.50	5.65
3	4.30	6.56	7.67	8.20	8.50	8.60	8.63	4.44	5.58	6.27	6.36	6.35	6.33	5.48
4	4.18	6.34	7.49	7.98	8.30	8.40	8.45	4.35	5.30	6.05	6.15	6.15	6.06	5.24
5	3.99	6.18	7.24	7.72	8.03	8.14	8.18	4.24	5.07	5.82	5.92	5.97	5.87	5.11
6	3.74	5.66	6.61	7.06	7.31	7.41	7.48	4.14	5.10	6.21	6.48	6.46	6.16	5.14
7	3.55	5.06	5.74	6.13	6.35	6.45	6.51	4.47	6.93	8.68	9.12	9.16	8.72	6.57
8	3.29	4.55	5.15	5.42	5.58	5.64	5.69	4.70	9.08	11.41	12.08	12.13	11.58	8.61
9	3.23	4.36	4.82	4.99	5.11	5.15	5.17	4.92	10.69	13.48	14.28	14.33	13.86	10.25
10	3.21	4.20	4.56	4.65	4.70	4.73	4.75	5.24	12.01	15.02	15.80	15.85	15.52	11.48
11	3.20	4.12	4.44	4.52	4.57	4.59	4.60	5.39	12.69	16.11	16.81	16.79	16.53	11.99
12	3.27	4.06	4.41	4.47	4.51	4.52	4.53	5.60	12.78	16.47	17.14	17.12	16.84	12.10
13	3.28	4.17	4.48	4.54	4.57	4.59	4.59	5.40	12.56	16.27	16.94	16.93	16.49	11.71
14	3.52	4.50	4.83	4.90	4.93	4.93	4.94	5.55	12.85	16.80	17.65	17.60	17.09	12.11
15	3.61	4.68	5.08	5.15	5.17	5.18	5.18	5.74	13.29	17.70	18.40	18.45	17.89	12.41
16	3.61	4.72	5.11	5.22	5.26	5.26	5.26	5.77	13.17	17.77	18.47	18.57	17.90	12.29
17	3.50	4.42	4.71	4.75	4.79	4.78	4.79	5.39	12.19	16.65	17.26	17.39	16.63	11.25
18	3.29	4.04	4.32	4.42	4.47	4.48	4.48	4.79	9.24	12.18	12.66	12.68	12.09	8.67
19	3.15	3.77	4.12	4.26	4.31	4.32	4.33	4.03	6.43	7.81	8.13	8.14	7.80	6.18
20	2.85	3.48	3.75	3.82	3.91	3.93	3.95	3.47	5.15	5.74	5.95	6.05	5.80	4.99
21	2.72	3.58	3.93	4.09	4.18	4.22	4.23	3.17	4.89	5.67	5.84	5.95	5.69	4.56
22	3.04	4.22	4.76	5.00	5.13	5.18	5.19	3.36	5.30	6.61	6.92	6.90	6.56	4.92
23	3.68	5.26	5.88	6.21	6.39	6.43	6.43	3.99	5.64	6.95	7.28	7.24	6.79	5.30
24	4.26	6.08	6.96	7.38	7.62	7.67	7.70	4.50	5.79	6.78	7.09	6.98	6.70	5.59
min	2.72	3.48	3.75	3.82	3.91	3.93	3.95	3.17	4.89	5.67	5.84	5.95	5.69	4.56
med	3.55	4.89	5.47	5.71	5.85	5.90	5.92	4.65	8.47	10.64	11.08	11.10	10.75	8.05
max	4.45	6.70	7.70	8.20	8.50	8.60	8.63	5.77	13.29	17.77	18.47	18.57	17.90	12.41

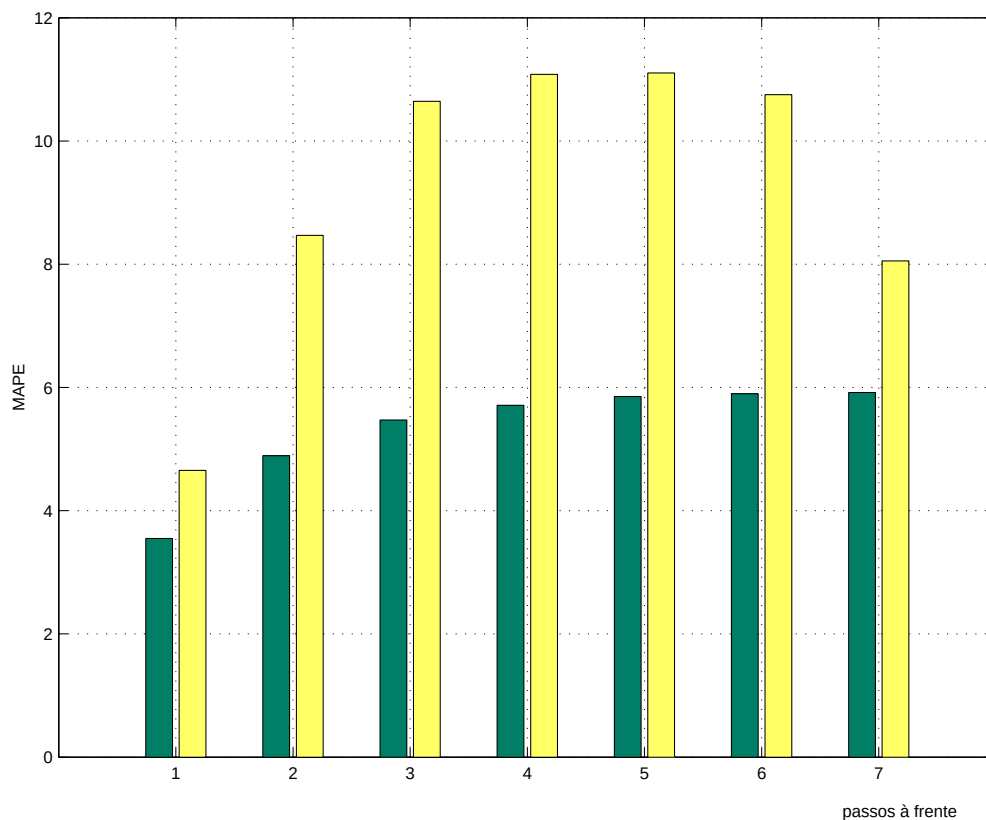


Figura 3.7: MAPE médio 2000 - *TLSAR* (escuro) e *DASARIMA* (claro)

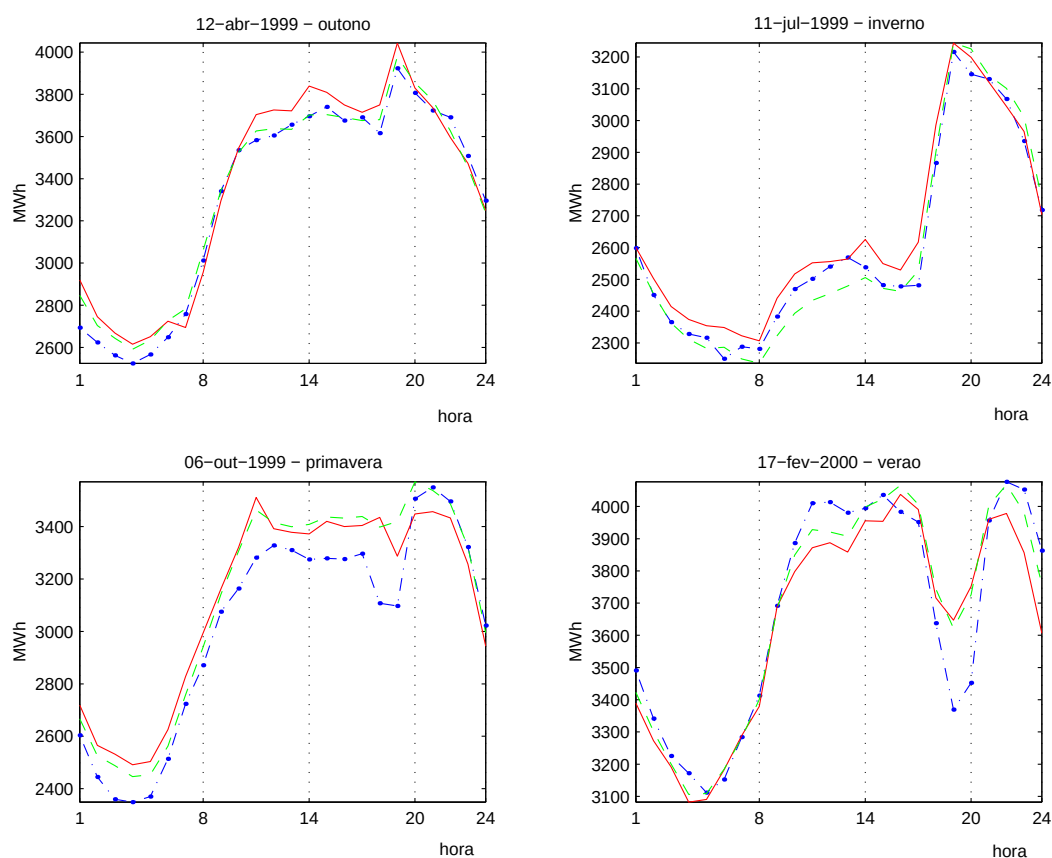


Figura 3.8: carga real (ponto), prevista *DASARIMA* (linha) e *TLSAR* (tracejado), 1 passo à frente.