4
A Modelagem usando Memória Longa

No capítulo anterior, a modelagem da série de carga considerou a existência de componentes determinísticos para a tendência, ciclos e efeitos de dias da semana. Nesse capítulo, de maneira diferente, todos os componentes são considerados estocásticos. A abordagem usa processos de memória longa para modelar determinadas características das séries temporais de carga.

A memória longa em processos estacionários tem duas definições tradicionais alternativas, uma no domínio da freqüência e a outra no domínio do tempo. Estas definições não são necessariamente equivalentes (ver Hosking (1996)), embora os processos de memória longa clássicos relatados na literatura respeitem ambas.

4.1
Definição

4.1.1
Memória Longa Ordinária

Os processos estacionários de memória longa são definidos pelo comportamento do espectro nas freqüências próximas de zero, conforme é mostrado a seguir.

Seja $f(\lambda)$, a função densidade espectral de um processo estacionário X_t. Se existir uma função positiva, $c_f(\lambda)$ que varie lentamente enquanto λ tende a zero, tal que $d \in (0; 0,5)$ e,

$$f(\lambda) \sim c_f(\lambda)|\lambda|^{-2d} \text{ quando } \lambda \longrightarrow \infty, \tag{4-1}$$

então X_t é um processo com memória longa e d é o parâmetro de memória (longa).

Alternativamente, seja ρ_k, a autocorrelação de defasagem k da série X_t. Se existir um número real $d \in (0; 0,5)$ e uma função positiva $c_\rho(k)$, que varie

PUC-Rio - Certificação Digital Nº 9824881/CA

lentamente quando k tende para o infinito, tal que:

$$\rho_k \sim c_\rho(k)k^{2d-1} \quad \text{quando} \quad k \longrightarrow \infty \tag{4-2}$$

então X_t tem uma dependência de memória longa.

Os modelos estocáticos de séries temporais constituem a classe de modelos dinâmicos de previsão e na literatura existem diversos nomes para esta classe de modelo, por exemplo, modelos *ARMA* (*AutoRegressive Moving Average*), modelos *ARIMA* (*AutoRegressive Integrated Moving Average*), método Box & Jenkins, modelos lineares de séries temporais, etc. Um tratamento geral desse tipo de modelo pode ser encontrado, por exemplo, em Box et al. (1994). O modelo *ARIMA* básico pode não ser o mais apropriado para descrever uma série de carga, uma vez que a estas séries incorporam-se variações de tendência, nível e da própria variância que podem não ser capturadas por diferenciações inteiras. Conseqüentemente, pode ser necessária uma diferenciação fracionária e neste caso o modelo obtido é denominado *ARFIMA (AutoRegressive Fractionaly Integrated Moving Average).*

X_t segue um modelo $ARFIMA(p,d,q)$ se $\Phi(B)(1-B)^d X_t = \Theta(B)\epsilon_t$, onde ϵ_t é um ruído branco com média zero e variância constante, B é o operador de retardo tal que $BX_t = X_{t-1}$, d não está restrito a valores inteiros como na especificação do modelo ARIMA. Os termos $\Phi(B) = 1-\phi_1 B-\cdots-\phi_p B$ e $\Theta(B) = 1+\theta_1 B+\cdots+\theta_q B^q$ são, respectivamente, os polinômios auto-regressivo e de médias-móveis. Os processos $ARFIMA$ são estacionários e apresentam memória longa se as raízes de $\Phi(B)$ estão fora do círculo unitário e $d \in (0;0,5)$. Se $d<0$ o processo tem memória curta e é considerado "antipersistente"(Mandelbrot 1977) e a equação (4-1) garante que o espectro tenha um zero na freqüência zero. Se $d=0$ o processo $ARFIMA$ reduz-se a um processo $ARMA$. Se as raízes de $\Theta(B)$ estiverem fora do círculo unitário e $d>-0.5$, o processo é inversível. As autocorrelações de um processo ARFIMA seguem $\rho(k) \sim ck^{2d-1}$ quando k tende à infinito, satisfazendo desta forma a equação (4-2). Uma diferença não-inteira pode ser expandida em um polinômio infinito auto-regressivo ou médias-móveis usando o teorema binomial:

$$(1-B)^d = \sum_{k=0}^{\infty} \binom{d}{k} (-B)^k,$$

onde $\binom{d}{k} = \frac{\Gamma(d+1)}{\Gamma(k+1)\Gamma(d-k+1)}$. A função $\Gamma(.)$, denominada função gama, é definida para todo $t>0$ por:

$$\int_0^\infty x^{t-1}e^{-x}dx \tag{4-3}$$

PUC-Rio - Certificação Digital Nº 9824881/CA

e que por um processo assintótico de integração por partes conduz a uma generalização da função fatorial para n inteiro:

$$\Gamma(n+1) = n!.$$

4.1.2
Memória Longa Generalizada

A memória longa sazonal não é geralmente definida na literatura, porque a definição em (4-1) guarda analogias com comportamentos espectrais diferentes dentro de um contexto sazonal. Em geral, são definidos os processos com propriedades análogas e posteriormente são explorados seus comportamentos espectrais e o comportamento da autocovariância. Neste trabalho são usados os processos de Gegenbauer como em Gray, Zhang and Woodward (1989) e em Chung (1996), mas modelos alternativos foram usados, por exemplo por: Porter-Hudak (1990), Ray (1993), Arteche & Robinson (2000) e Arteche (2002). Os processos de Gegenbauer foram sugeridos por Hosking (1981) e mais tarde formalizados por Gray et al. (1989) e são definidos por:

$$(1 - 2\gamma B + B^2)^d X_t = \epsilon_t, \tag{4-4}$$

onde $|\gamma| \leq 1$ e ϵ_t é um ruído branco.

Este processo é denominado processo de Gegenbauer porque a representação de Wold, que é encontrada através da expansão pelo teorema binomial em um $MA(\infty)$, da equação (4-4) admite como coeficientes a importante classe de polinômios ortogonais chamados de polinômios de Gegenbauer (Gray et al. 1989). Se $|\gamma| < 1$ e $0 < d < 0,5$, as autocorrelações do processo definido em (4-4) podem ser aproximadas por:

$$\rho(k) = C\ cos(k\nu)k^{2d-1} \quad \text{quando} \quad k \longrightarrow \infty, \tag{4-5}$$

onde C é uma constante não dependente de k (mas que pode depender de d e ν) e $\nu = arccos(\gamma)$. Isso significa que as autocorrelações se comportam como uma senóide amortecida hiperbolicamente na freqüência ν e que γ determina o tamanho do ciclo (ou sazonalidade). Além disso, a função densidade espectral obedece

$$f(\lambda) = \frac{\sigma^2}{2\pi} 2|cos\lambda - cos\nu|^{-2d},$$

comportando-se quando $\lambda \longrightarrow \nu$ como

$$f(\lambda) \sim |\lambda^2 - \nu^2|^{-2d}, \quad para \quad 0 < \lambda < \pi. \tag{4-6}$$

PUC-Rio - Certificação Digital Nº 9824881/CA

Os processos de Gegenbauer são uma generalização de uma integração fracionária, dando origem ao nome memória longa generalizada. Para o caso em que $\gamma = 1$, o processo (4-4) é definido como um modelo de integração fracionária ($FI(2d)$), ou equivalentemente, um modelo $ARFIMA(0, 2d, 0)$. Além disso, a analogia (com relação à memória longa) dos processos de Gegenbauer com os processos FI vem do fato de que os últimos têm um pólo ou um zero na freqüência zero, enquanto os primeiros terão um pólo ou um zero na freqüência ν, dependendo se d é positivo ou negativo, respectivamente. Observa-se que os processos de Gegenbauer podem ser generalizados em processos de k-fatores de Gegenbauer como em Gray et al. (1989) e em Ferrara & Guégan (2001), permitindo que diferentes filtros de Gegenbauer $(1 - 2\gamma_i B + B^2)^{di}$; $i = 1, \ldots, k$ sejam aplicados a X_t. O modelo desenvolvido combina processos de Gegenbauer e $ARFIMA$ para modelar a carga sem tendência, depois de removidos os efeitos de calendário.

Ensaios realizados com os dados utilizados para prever carga com memória longa generalizada (sazonal), mostraram que, mesmo quando existe uma quebra estrutural, algumas propriedades estatísticas não são perdidas, de maneira que as previsões não perdem precisão drasticamente.

PUC-Rio - Certificação Digital Nº 9824881/CA

4.1.3
Os Dados

A idéia é identificar as dependências da carga diária na carga passada, sem considerar a existência de componentes determinísticos e sem utilizar variáveis externas explicativas, como temperatura por exemplo, embora reconhecendo a importância de variáveis climáticas neste caso. A razão, como já explicitado na seção 3.5, é a qualidade destes dados. A base de dados utilizada é a mesma descrita e definida no capítulo 2. Cada série é tratada como uma série única e todas são modeladas por uma mesma especificação geral, embora sejam estimadas independentemente uma da outra (modelagem univariada).

Uma janela de estimação de quatro anos é usada, acreditando-se ser suficiente para uma boa estimação. Algumas experiências com janelas de estimação maiores e menores produziram resultados pouco diferentes dos obtidos. Como o foco concentra-se nas previsões com horizontes múltiplos de 24 horas à frente, as influências das defasagens de ordem até 23 podem ser desconsideradas sem afetar a preditividade, reforçando a justificativa do uso dos dados seccionais. A abordagem de modelagem hora-por-hora foi também usada por Ramanathan et al. (1997), que venceram uma competição de previsão de carga da *Puget Sound Power and Light Company*, embora a modelagem siga uma abordagem diferente da adotada neste trabalho. Ao contrário deles, não são usados componentes determinísticos na modelagem da carga e nem variáveis externas como as relacionadas à temperatura. Incorporar algumas medidas de temperatura (máximos, médias

e outras) à modelagem poderia melhorar substancialmente a previsão, particularmente no período do verão, da região geográfica compreendida pelo estudo, em que os aparelhos de condicionamento de ar constituem uma grande parte da carga. Os erros de previsão são em geral maiores neste período.

4.1.4
Formulação do Modelo

Uma grande variedade de modelos e de métodos tem sido desenvolvida para prever a demanda de energia, e a maior parte dos esforços é dedicada às técnicas de inteligência artificial, em particular a modelagem com redes neurais [1]. Diante desta avalanche de aplicações neurais em previsão de carga, aqui são adotados métodos estatísticos lineares, que parecem explicar os dados em um nível razoável, além de iluminar o que está sendo modelado.

Seja a carga horária representada por $X_{h,d}$ e $L_{h,d} = log(X_{h,d})$. O modelo proposto, denominado modelo de memória longa generalizada ajustado por variáveis binárias (*DAGLM - Dummy-Adjusted Generalized Long Memory*) é definido por:

$$L_{h,d} = L^N_{h,d} + WD_{h,d} + \alpha_h H_d + Z_{h,d} \tag{4-7}$$

onde: $L^N_{h,d}$ é o nível estocástico (que segue alguma tendência definida, possivelmente, por fatores macroeconômicos e demográficos); $WD_{h,d}$ é o efeito do dia da semana; H_d é o efeito dos feriados (ponderado pelo parâmetro α_h); e $Z_{h,d}$ é o processo estocástico:

$$(1 - 2\gamma B + B^2)^{d1_h}(1 - B)^{d2_h}(1 - \phi_{h,d}B)Z_{h,d} = \epsilon_{h,d}, \tag{4-8}$$

onde $\epsilon_{h,d}$ é um ruído branco.

A razão para trabalhar com a série logaritmizada é permitir que a sazonalidade semanal e o efeito dos feriados sejam modelados aditivamente quando eles são multiplicativos na série original $X_{h,d}$. Acredita-se que estes efeitos sejam multiplicativos ao relacionar-se aos hábitos de consumo, variando proporcionalmente ao número de consumidores. Entretanto, a mesma modelagem definida em (4-7) e (4-8) aplicada à $X_{h,d}$, em vez de $L_{h,d}$, produziu resultados muito similares. A escolha da transformação logarítmica destina-se apenas a manter a coerência de opinião e não é orientada à resultado.

O primeiro termo multiplicativo no lado esquerdo da equação (4-8) refere-se à sazonalidade anual (memória longa), onde γ é tal que o período seja de 365

[1] Hippert, Pedreira & Souza (2001) produziram uma revisão de trabalhos que utilizam técnicas variadas de redes neurais em previsão de demanda de carga.

PUC-Rio - Certificação Digital Nº 9824881/CA

dias e $d1_h$ é o grau de diferenciação. O segundo refere-se à memória longa pura, onde $d2_h$ é o grau de diferenciação. O terceiro refere-se ao termo auto-regressivo, com valores diferentes de $\phi_{h,d}$ para diferentes dias da semana.

Nenhuma suposição é feita sobre a tendência que influencia o nível estocástico. $L^N_{h,d}$ é estimado calculando-se a média dos valores correspondentes ao período compreendido entre meio ano antes até meio ano depois do dia d, de tal maneira que a sazonalidade anual não interfira na estimação do nível.

A previsão é feita através da extrapolação da tendência linear estimada pela regressão local de $\widehat{L}^N_{h,d}$. O efeito de calendário é modelado através de $WD_{h,d}$ que é codificado como variável binária e estimado por técnicas de MQO e H_d que assume valor 1 para os feriados; $0,5$ para os meio-feriados e 0 para os outros casos. Os dias modelados como meio-feriados são aqueles em que ocorrem meio-expediente, ou é feriado em apenas uma parte da região, ou é feriado para uma parcela da população, como ocorre com os dias das categorias profissionais, de tal forma que o efeito do feriado seja suavizado. Dois exemplos de meio-feriado são: quarta-feira de cinzas após o carnaval, onde o expediente ocorre apenas à tarde; e dia de São Sebastião, quando ocorre a festa do santo padroeiro de apenas uma das cidades integrantes da região em estudo. H poderia ter um refinamento maior, por exemplo, uma medida para "intensidade do feriado"para diferentes tipos de feriado, mas há um compromisso entre a melhoria do ajuste e o erro da estimação. O exame dos erros de previsão quando não são considerados os efeitos dos feriados mostra erros grandes que desaparecem quando são incluídos estes efeitos.

A memória longa sazonal e a ordinária são estimadas usando a abordagem de Whittle (1951) e detalhes, como consistência e normalidade assintótica, podem ser encontrados em Fox & Taqqu (1986) e em Ferrara & Guégan (1999), respectivamente, para a memória longa ordinária e generalizada. Chung (1996) observa que, quando o parâmetro γ é conhecido, como aqui ocorre, a estimação da memória longa sazonal é virtualmente idêntica àquela da memória longa ordinária. Entretanto, existem diferentes estimadores de memória longa na literatura e aqui eles foram escolhidos pela simplicidade e eficiência quando comparados entre si. Os termos auto-regressivos semanais são estimados sobre os resíduos resultantes, usando técnicas de MQO.

A Figura 4.1 mostra a função de autocorrelação (FAC) da série de 26 de dezembro de 1990 a 25 de dezembro de 1998 (oito anos), para horas representativas dos diferentes períodos do dia, após a remoção da tendência e dos efeitos de calendário. O formato da FAC assemelha-se a uma senóide amortecida como em (4-5), com periodicidade anual o que justifica plenamente o uso de processos de Gegenbauer para modelar $Z_{h,d}$. Uma exceção ocorre na hora 20 em que ou existe um componente senoidal de periodicidade semestral ou existem

PUC-Rio - Certificação Digital Nº 9824881/CA

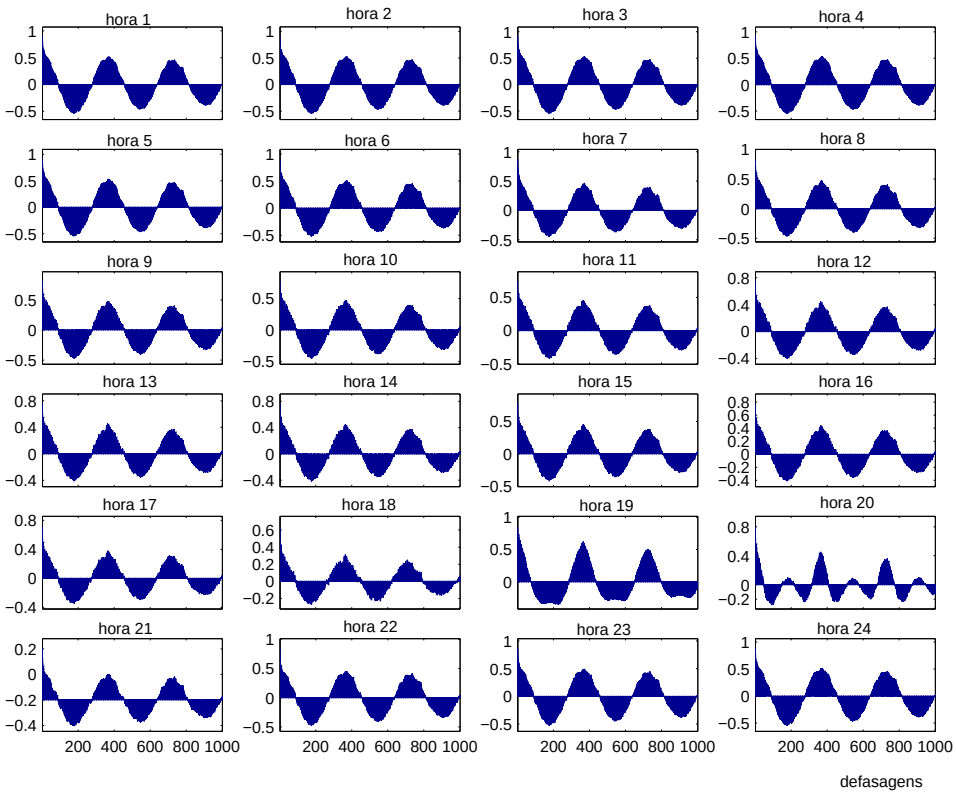

Figura 4.1: função de autocorrelação, após a remoção da tendência e dos efeitos de calendário, até a defasagem de ordem 1000. Observe a semelhança com uma senóide amortecida

dois componentes senoidais, um com periodicidade anual e o outro com periodicidade semestral. Este comportamento não ocorre em nenhuma outra hora, sendo específico da hora 20. Entretanto, o modelo proposto não usa processos de Gegenbauer de k-fatores (k seria igual a 2 neste caso) como proposto por Gray et al. (1989) e estudado por ((Ferrara & Guégan 2000), (Ferrara & Guégan 2001)) para modelar a hora 20, obtendo entretanto bons resultados. De fato, na hora 20 ocorrem os erros mais baixos, exceto para 1 e 2 passos à frente, quando a hora 21 apresenta erros menores.

A Figura 4.2 mostra os periodogramas (normal e suavizado por uma janela de Parzen com grau de amortecimento igual a 0,9, representados respectivamente por pontos e por uma linha contínua) para a mesma série, em escala logarítmica. Observa-se a semelhança com o comportamento descrito por (4-6). A oitava freqüência de Fourier, que corresponde a um período anual, pois são usados oito anos de dados, mostra um pico mais visível no periodograma normal, e suas vizinhanças mostram uma suavização (na escala log-log) para ambos os lados. O gráfico correspondente à hora 20 na figura 4.2 é consistente com a FAC mostrada na figura 4.1, pois existem picos na oitava e na décima-sexta freqüências de Fourier, significando que pode haver um único fator(semestral) ou dois fatores (anual e semestral) de Gegenbauer.

PUC-Rio - Certificação Digital Nº 9824881/CA

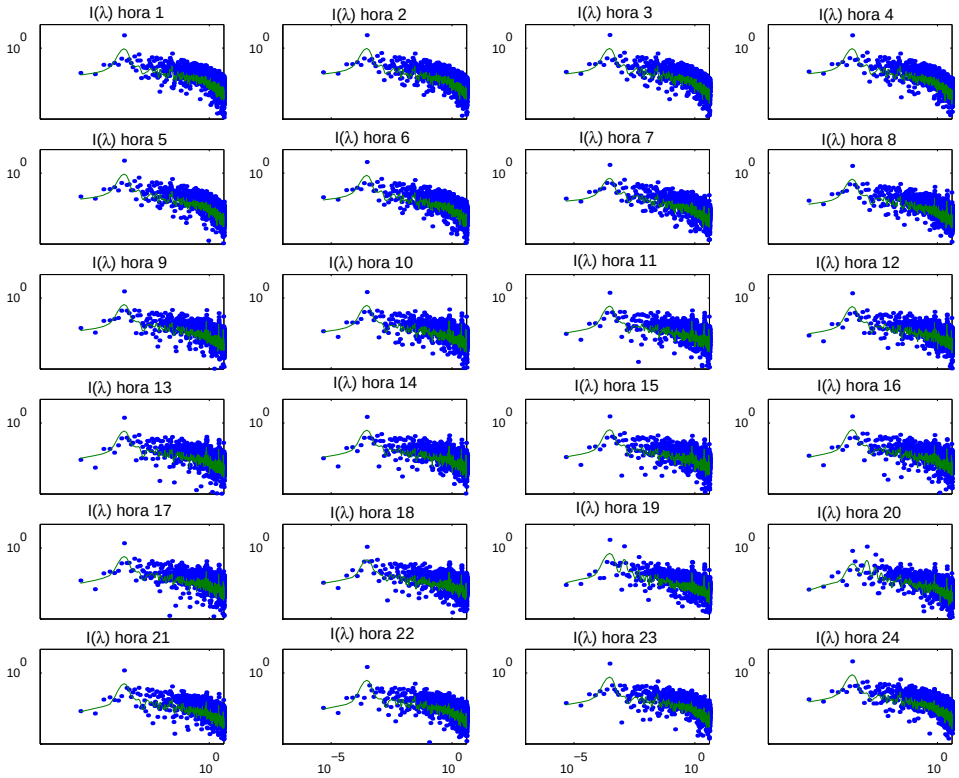

Figura 4.2: periodograma normal (pontos) e suavisado (linha), após a remoção da tendência e dos efeitos de calendário (dias da semana e feriados).

Os parâmetros são estimados com dados históricos e espera-se que a dependência desconhecida observada no passado esteja presente no período de previsão. A tabela 4.1 apresenta os valores dos parâmetros estimados com o modelo formulado em (4-7).

PUC-Rio - Certificação Digital Nº 9824881/CA

PUC-Rio - Certificação Digital Nº 9824881/CA

Tabela 4.1: Parâmetros estimados do modelo *DAGLM*

hora	$\widehat{d1}_h$	$\widehat{d2}_h$	$\widehat{WD}_{dom}$	$\widehat{WD}_{seg}$	$\widehat{WD}_{ter}$	$\widehat{WD}_{qua}$	$\widehat{WD}_{qui}$	$\widehat{WD}_{sex}$	$\widehat{WD}_{sab}$	$\widehat{\alpha}$	$\widehat{\phi}_{dom}$	$\widehat{\phi}_{seg}$	$\widehat{\phi}_{ter}$	$\widehat{\phi}_{qua}$	$\widehat{\phi}_{qui}$	$\widehat{\phi}_{sex}$	$\widehat{\phi}_{sab}$
1	0.361	-0.370	-0.013	-0.047	0.001	0.006	0.010	0.013	0.031	0.006	0.407	0.695	0.495	0.552	0.289	0.574	0.541
2	0.374	-0.395	-0.013	-0.044	-0.001	0.007	0.010	0.013	0.030	0.015	0.434	0.687	0.513	0.560	0.301	0.558	0.550
3	0.385	-0.407	-0.019	-0.041	0.001	0.010	0.011	0.013	0.025	0.016	0.473	0.665	0.485	0.525	0.335	0.607	0.606
4	0.394	-0.411	-0.026	-0.039	0.003	0.012	0.012	0.015	0.022	0.010	0.450	0.646	0.480	0.582	0.316	0.624	0.639
5	0.398	-0.408	-0.036	-0.034	0.006	0.015	0.015	0.018	0.017	-0.004	0.471	0.617	0.485	0.559	0.331	0.679	0.604
6	0.396	-0.410	-0.063	-0.020	0.015	0.024	0.022	0.023	-0.002	-0.042	0.493	0.537	0.439	0.538	0.338	0.713	0.673
7	0.350	-0.340	-0.121	0.006	0.035	0.042	0.038	0.039	-0.039	-0.126	0.537	0.291	0.400	0.406	0.363	0.700	0.751
8	0.315	-0.280	-0.175	0.028	0.048	0.057	0.052	0.055	-0.066	-0.194	0.613	0.295	0.305	0.340	0.244	0.714	0.620
9	0.280	-0.192	-0.220	0.046	0.059	0.068	0.065	0.068	-0.084	-0.252	0.494	0.309	0.286	0.304	0.096	0.789	0.401
10	0.251	-0.112	-0.246	0.057	0.065	0.072	0.070	0.073	-0.091	-0.285	0.367	0.163	0.278	0.345	-0.002	0.739	0.171
11	0.230	-0.072	-0.257	0.064	0.069	0.076	0.073	0.074	-0.099	-0.300	0.298	0.158	0.389	0.378	0.040	0.527	-0.109
12	0.238	-0.126	-0.257	0.066	0.070	0.078	0.074	0.073	-0.105	-0.306	0.398	0.151	0.379	0.383	0.109	0.631	0.088
13	0.239	-0.140	-0.249	0.066	0.069	0.076	0.072	0.073	-0.105	-0.297	0.439	0.183	0.379	0.354	0.125	0.599	0.089
14	0.244	-0.160	-0.255	0.068	0.072	0.079	0.075	0.075	-0.115	-0.303	0.473	0.338	0.468	0.414	0.116	0.543	0.113
15	0.248	-0.177	-0.261	0.072	0.077	0.085	0.078	0.076	-0.126	-0.311	0.481	0.409	0.511	0.289	0.126	0.547	0.147
16	0.253	-0.201	-0.263	0.073	0.078	0.086	0.080	0.074	-0.128	-0.308	0.508	0.462	0.600	0.212	0.156	0.597	0.208
17	0.219	-0.128	-0.253	0.072	0.075	0.086	0.079	0.066	-0.126	-0.297	0.321	0.336	0.648	-0.029	0.127	0.537	0.128
18	0.198	-0.149	-0.192	0.053	0.054	0.066	0.061	0.048	-0.090	-0.257	0.441	0.165	0.567	-0.024	0.114	0.523	0.203
19	0.271	-0.062	-0.131	0.032	0.039	0.044	0.041	0.030	-0.055	-0.207	0.409	0.086	0.472	-0.125	-0.040	0.296	0.112
20	0.292	-0.121	-0.096	0.021	0.026	0.032	0.028	0.018	-0.029	-0.156	0.526	0.238	0.449	0.127	0.068	0.374	0.203
21	0.279	-0.273	-0.081	0.019	0.023	0.029	0.025	0.015	-0.030	-0.114	0.755	0.457	0.695	0.329	0.439	0.508	0.312
22	0.311	-0.355	-0.072	0.017	0.025	0.030	0.025	0.014	-0.039	-0.097	0.909	0.462	0.608	0.360	0.484	0.523	0.422
23	0.330	-0.346	-0.058	0.010	0.020	0.025	0.023	0.017	-0.038	-0.063	0.776	0.481	0.639	0.350	0.468	0.493	0.461
24	0.343	-0.349	-0.050	0.006	0.010	0.018	0.017	0.024	-0.024	-0.040	0.772	0.477	0.618	0.336	0.509	0.523	0.400

4.2
Resultados da previsão

São realizados dois ensaios de previsão com o modelo *DAGLM*, um utilizando dados até 31 de dezembro de 2000, onde não ocorre nenhuma quebra estrutural significativa e outro incluindo dados até setembro de 2002 para avaliar o desempenho do modelo diante da ocorrência de uma quebra estrutural, neste caso devido ao racionamento de energia ocorrido entre 2001 e 2002.

Os dados especificados no capítulo 2, compreendendo o período de 1 de janeiro de 1990 a 30 de setembro de 2002, foram utilizados para testar o modelo. No primeiro ensaio, o período usado para estimar os parâmetros do modelo compreende o intervalo entre 26 de dezembro de 1994 e 25 de dezembro de 1998. O modelo é estimado uma vez usando uma janela de estimação de quatro anos que parece ser suficiente. Os resultados obtidos foram comparados com os resultados do modelo *DASARIMA* de forma idêntica à realizada no capítulo anterior com o modelo *TLSAR*.

No segundo ensaio, diferentemente do que é feito nos ensaios anteriores, o modelo é reestimado recursivamente, isto é, para cada nova observação incorporada, e considera o período que vai de 1 de janeiro de 2000 a 30 de setembro de 2002 para avaliar o desempenho das previsões. Neste período ocorre uma quebra estrutural (resposta à campanha de racionamento) e o objetivo principal é avaliar o comportamento do modelo diante deste evento.

4.2.1
Um ensaio antes do Racionamento (sem quebra estrutural)

Neste ensaio é estimado um modelo *DAGLM* para cada hora do dia. Os modelos são estimados uma única vez usando uma janela de estimação de quatro anos e os parâmetros estimados para cada hora encontram-se na tabela 4.1. Os graus de integração fracionária, sazonal e ordinária, variam entre as séries correspondentes à cada hora, embora $d1$ seja sempre positivo (memória longa) e $d2$ sempre negativo (memória curta). A memória longa (sazonal) é mais forte para as horas iniciais e finais do dia, coincidindo com as horas onde a carga apresenta menor intervalo de variação.

O efeito de calendário, modelado por $WD_{h,d}$ e por $\alpha_h H_d$, também 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. O uso de um parâmetro auto-regressivo para cada dia é inteiramente justificado conforme pode ser comprovado pela análise da tabela 4.1 (por exemplo, $\phi_{h,d}$ estimado para a hora 19, variando de -0.125 na quata-feira a 0.472 na terça-

PUC-Rio - Certificação Digital Nº 9824881/CA

feira). Observa-se que o ϕ_{ter} corresponde ao efeito de segunda-feira no resíduo de terça-feira, e assim por diante. Outras janelas de estimação variando de três a oito anos foram testadas, mas os resultados foram praticamente os mesmos. O modelo *DASARIMA* comporta-se muito bem nas primeiras horas, quando a carga varia em um intervalo menor.

análise e comentários 1

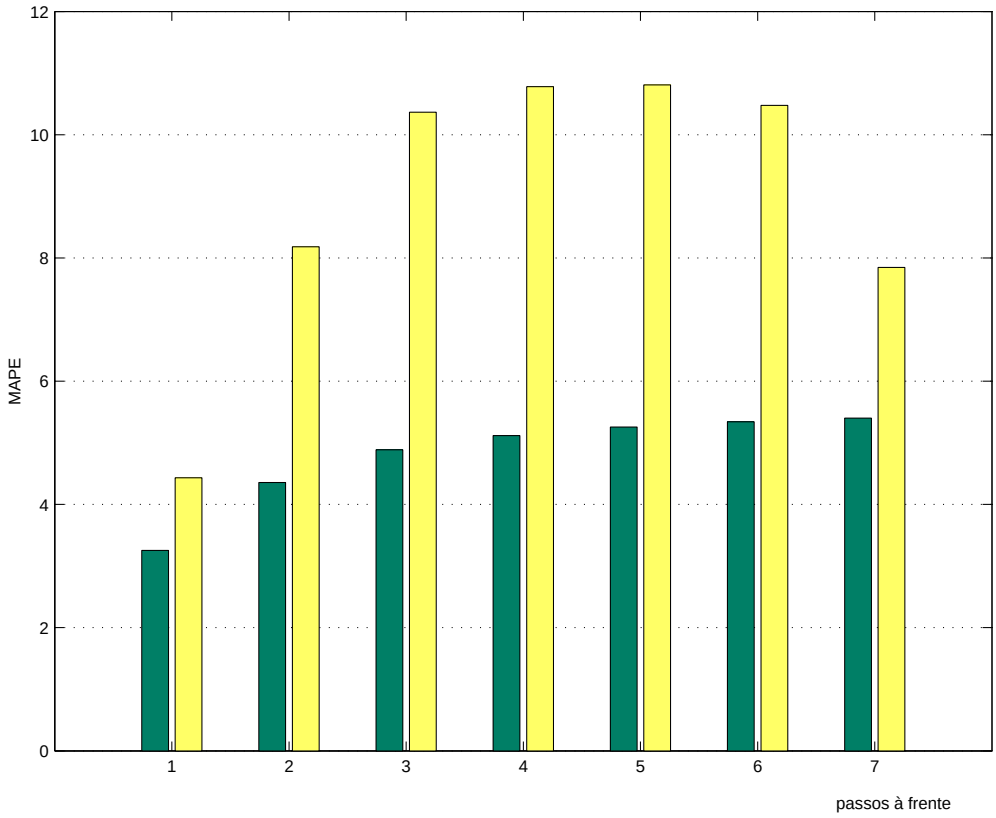

Figura 4.3: MAPE médio (1999 – 2000) - *DAGLM* (escuro), *DASARIMA* (claro)

Uma análise dos erros do período de previsão completo (1999 – 2000) do modelo *DAGLM* mostra valores entre 2.43% (1 passo à frente - hora 21) e 7.62% (7 passos à frente - hora 2). O modelo *DASARIMA* apresenta valores de erros entre 3.01% (1 passo à frente - hora 21) e 17.95% (5 passos à frente - hora 16). O melhor MAPE para cada hora considerando o número de passos à frente é mostrado em negrito nas tabelas 4.2 a 4.4.

análise e comentários 2

A tabela 4.3 mostra o MAPE de um a sete dias à frente do ano de 1999 para o modelo *DAGLM* e o *DASARIMA*. O erro médio para o ano de 1999 do modelo *DAGLM* ficou entre 3.03% e 4.83% e para o modelo *DASARIMA* entre 4.08% e 10.40%, conforme mostra a figura 4.4.

PUC-Rio - Certificação Digital Nº 9824881/CA

Tabela 4.2: MAPE - total (1999 – 2000) dos modelos *DAGLM* e *DASARIMA*

hora	*DAGLM* passos à frente							*DASARIMA* passos à frente						
	1	2	3	4	5	6	7	1	2	3	4	5	6	7
1	**4.07**	5.78	6.60	7.00	7.24	7.35	7.44	4.23	**5.35**	**6.09**	**6.27**	**6.28**	**6.14**	**5.21**
2	**4.06**	5.82	6.73	7.13	7.38	7.52	7.62	4.19	**5.26**	**5.90**	**6.03**	**5.99**	**5.95**	**5.16**
3	**3.96**	5.71	6.68	7.07	7.33	7.46	7.58	4.10	**5.12**	**5.79**	**5.90**	**5.82**	**5.77**	**4.99**
4	**3.79**	5.46	6.47	6.83	7.09	7.23	7.33	4.04	**4.97**	**5.68**	**5.80**	**5.77**	**5.65**	**4.83**
5	**3.62**	5.19	6.16	6.54	6.79	6.95	7.07	3.85	**4.66**	**5.42**	**5.56**	**5.57**	**5.38**	**4.59**
6	**3.36**	4.84	**5.64**	**5.99**	**6.20**	6.38	6.50	3.74	**4.74**	5.77	6.21	6.21	**5.79**	**4.75**
7	**3.22**	**4.38**	**5.02**	**5.25**	**5.44**	**5.58**	**5.66**	4.08	6.59	8.47	9.12	9.18	8.57	6.45
8	**2.98**	**3.99**	**4.51**	**4.71**	**4.89**	**5.02**	**5.10**	4.39	8.76	11.23	11.94	12.06	11.48	8.50
9	**2.89**	**3.86**	**4.26**	**4.45**	**4.60**	**4.71**	**4.78**	4.78	10.51	13.33	14.02	14.16	13.74	10.15
10	**2.89**	**3.76**	**4.14**	**4.27**	**4.37**	**4.43**	**4.50**	5.08	11.70	14.72	15.35	15.46	15.18	11.27
11	**2.96**	**3.78**	**4.12**	**4.27**	**4.35**	**4.40**	**4.43**	5.17	12.25	15.65	16.15	16.23	16.05	11.70
12	**2.97**	**3.80**	**4.14**	**4.26**	**4.32**	**4.35**	**4.37**	5.31	12.42	16.04	16.47	16.53	16.40	11.84
13	**2.99**	**3.90**	**4.25**	**4.38**	**4.45**	**4.47**	**4.51**	5.22	12.12	15.79	16.29	16.33	16.00	11.45
14	**3.16**	**4.18**	**4.56**	**4.72**	**4.77**	**4.79**	**4.81**	5.33	12.43	16.36	16.90	16.90	16.56	11.78
15	**3.30**	**4.33**	**4.75**	**4.91**	**4.99**	**5.01**	**5.04**	5.45	12.87	17.24	17.74	17.80	17.39	12.10
16	**3.32**	**4.33**	**4.73**	**4.91**	**5.02**	**5.08**	**5.10**	5.44	12.78	17.35	17.90	17.95	17.42	11.99
17	**3.43**	**4.14**	**4.47**	**4.58**	**4.67**	**4.72**	**4.72**	5.10	11.88	16.40	16.92	17.01	16.37	11.11
18	**3.20**	**3.84**	**4.07**	**4.15**	**4.20**	**4.21**	**4.25**	4.53	9.11	12.07	12.56	12.62	12.14	8.66
19	**3.01**	**3.57**	**3.82**	**4.01**	**4.15**	**4.22**	**4.26**	3.79	6.23	7.68	8.07	8.07	7.84	6.15
20	**2.53**	**3.08**	**3.26**	**3.39**	**3.46**	**3.50**	**3.53**	3.27	5.03	5.69	5.96	5.99	5.78	4.93
21	**2.43**	**3.14**	**3.48**	**3.66**	**3.72**	**3.79**	**3.84**	3.01	4.68	5.56	5.84	5.89	5.56	4.49
22	**2.77**	**3.75**	**4.18**	**4.41**	**4.52**	**4.60**	**4.65**	3.22	5.08	6.41	6.82	6.80	6.33	4.84
23	**3.31**	**4.57**	**5.14**	**5.44**	**5.57**	**5.65**	5.71	4.09	5.65	6.95	7.36	7.29	6.82	**5.39**
24	**3.84**	**5.36**	**6.07**	**6.47**	**6.64**	**6.76**	6.83	4.96	6.19	7.17	7.51	7.46	7.12	**6.01**
min	2.43	3.08	3.26	3.39	3.46	3.50	3.53	3.01	4.66	5.42	5.56	5.57	5.38	4.49
med	3.25	4.36	4.89	5.12	5.26	5.34	5.40	4.43	8.18	10.37	10.78	10.81	10.48	7.85
max	4.07	5.82	6.73	7.13	7.38	7.52	7.62	5.45	12.87	17.35	17.90	17.95	17.42	12.10

O modelo *DAGLM* é melhor que o modelo *DASARIMA* para todas as horas 1 passo à frente. O modelo *DASARIMA* é melhor principalmente nas cinco horas do início do dia, mais de 1 passo à frente. As horas do meio do dia apresentam uma grande variação na precisão da previsão, com o modelo *DAGLM* apresentando bons resultados, alcançando MAPE de 3% para 1 passo à frente a 4% para 7 passos à frente. Neste período, o modelo *DASARIMA* alcança MAPE de dois dígitos para mais de 1 passo à frente. As horas de pico entre 19:00/21:00 apresentam os melhores MAPE, em parte, porque a carga é mais elevada, mas a previsibilidade proporcional é mais precisa para ambos os métodos.

Os MAPE do ano de 2000 são apresentados na tabela 4.4, no mesmo formato usado na tabela 4.3. Os resultados são ligeiramente piores, principalmente porque a tendência linear não é reestimada, mas parece sofrer uma ruptura em 2000 e assemelha-se a uma linearidade por partes. Certamente são, como pode ser analisado comparando as tabelas 4.3 e 4.4. A razão para esta aparente quebra pode estar relacionada à fatores macroeconômicos, que estão fora do escopo deste trabalho.

Mesmo assim, os resultados são bons e são qualitativamente iguais aos do ano de 1999, com uma diferença na sexta hora para mais de 1 passo à frente

PUC-Rio - Certificação Digital Nº 9824881/CA

Tabela 4.3: MAPE – ano 1999 dos modelos *DAGLM* e *DASARIMA*

hora	*DAGLM* passos à frente							*DASARIMA* passos à frente						
	1	2	3	4	5	6	7	1	2	3	4	5	6	7
1	**3.76**	5.14	5.82	6.08	6.30	6.49	6.71	3.93	**4.93**	**5.63**	**5.84**	**5.83**	**5.63**	**4.72**
2	**3.68**	5.09	5.84	6.12	6.33	6.59	6.81	3.85	**4.79**	**5.37**	**5.51**	**5.46**	**5.40**	**4.67**
3	**3.54**	4.97	5.75	6.00	6.18	6.47	6.70	3.67	**4.55**	**5.21**	**5.36**	**5.26**	**5.14**	**4.40**
4	**3.35**	4.67	5.54	5.73	5.92	6.18	6.40	3.46	**4.39**	**5.07**	**5.24**	**5.20**	**5.02**	**4.17**
5	**3.21**	4.39	5.20	5.39	5.56	5.83	6.04	3.37	**4.16**	**4.95**	**5.18**	**5.16**	**4.90**	**4.01**
6	**2.91**	**4.09**	**4.70**	**4.93**	**5.07**	**5.35**	5.54	3.26	4.31	5.29	5.90	5.95	5.43	**4.30**
7	**2.85**	**3.69**	**4.16**	**4.30**	**4.43**	**4.63**	**4.72**	3.62	6.21	8.24	9.08	9.16	8.42	6.27
8	**2.74**	**3.37**	**3.75**	**3.88**	**4.04**	**4.22**	**4.31**	4.09	8.43	11.04	11.80	11.99	11.38	8.38
9	**2.75**	**3.45**	**3.76**	**3.86**	**3.98**	**4.12**	**4.21**	4.63	10.33	13.18	13.76	13.98	13.62	10.05
10	**2.82**	**3.46**	**3.78**	**3.86**	**3.93**	**4.00**	**4.07**	4.91	11.38	14.42	14.89	15.07	14.84	11.05
11	**2.92**	**3.59**	**3.86**	**3.97**	**4.01**	**4.04**	**4.05**	4.95	11.81	15.19	15.48	15.67	15.57	11.40
12	**2.89**	**3.61**	**3.88**	**3.96**	**3.97**	**3.98**	**4.00**	5.02	12.06	15.61	15.81	15.93	15.96	11.58
13	**2.99**	**3.79**	**4.07**	**4.17**	**4.19**	**4.19**	**4.24**	5.04	11.69	15.32	15.64	15.72	15.50	11.19
14	**3.15**	**4.06**	**4.40**	**4.51**	**4.49**	**4.49**	**4.52**	5.10	12.01	15.91	16.14	16.20	16.04	11.45
15	**3.22**	**4.16**	**4.52**	**4.63**	**4.67**	**4.67**	**4.73**	5.16	12.45	16.78	17.08	17.16	16.89	11.79
16	**3.21**	**4.11**	**4.48**	**4.60**	**4.70**	**4.74**	**4.77**	5.11	12.40	16.92	17.33	17.34	16.93	11.70
17	**3.26**	**3.84**	**4.17**	**4.27**	**4.35**	**4.40**	**4.43**	4.82	11.57	16.16	16.57	16.64	16.11	10.97
18	**2.95**	**3.49**	**3.71**	**3.79**	**3.81**	**3.85**	**3.90**	4.27	8.97	11.95	12.46	12.56	12.19	8.65
19	**2.82**	**3.38**	**3.57**	**3.73**	**3.87**	**3.94**	**3.99**	3.54	6.04	7.55	8.02	8.01	7.89	6.12
20	**2.35**	**2.87**	**2.93**	**3.04**	**3.05**	**3.07**	**3.12**	3.07	4.90	5.63	5.98	5.94	5.77	4.88
21	**2.23**	**2.83**	**3.10**	**3.21**	**3.27**	**3.35**	**3.47**	2.85	4.47	5.44	5.84	5.82	5.44	4.42
22	**2.56**	**3.39**	**3.75**	**3.88**	**3.93**	**4.02**	**4.13**	3.08	4.85	6.21	6.73	6.70	6.09	4.76
23	**3.04**	**3.99**	**4.45**	**4.66**	**4.72**	**4.87**	5.01	3.35	4.83	6.18	6.67	6.55	6.07	**4.63**
24	**3.54**	**4.75**	**5.33**	**5.60**	**5.72**	**5.92**	6.10	3.76	4.97	6.00	6.38	6.34	5.93	**4.76**
min	2.23	2.83	2.93	3.04	3.05	3.07	3.12	2.85	4.16	4.95	5.18	5.16	4.90	4.01
med	3.03	3.92	4.36	4.51	4.60	4.73	4.83	4.08	7.77	9.97	10.36	10.40	10.09	7.51
max	3.76	5.14	5.84	6.12	6.33	6.59	6.81	5.16	12.45	16.92	17.33	17.34	16.93	11.79

em que o *DASARIMA* é melhor. O erro médio para o ano de 2000 do modelo *DAGLM* ficou entre 3.47% e 5.97% e para o modelo *DASARIMA* entre 4.65% a 11.10%, conforme mostra a figura 4.5.

análise e comentários 3

Os piores MAPE do modelo *DAGLM* são referentes às primeiras horas do dia 7 passos à frente, crescendo de menos de 7% em 1999 a mais de 8% em 2000. Sobre os resultados de 1 passo à frente, o melhor MAPE encontra-se entre 2.23% em 1999 a 2.63% em 2000 (hora 21) e o pior entre 3.76% (hora 1) a 4.43% (hora 2). A hora 20 é a melhor hora prevista para 7 passos à frente usando o MAPE como critério, apresentando MAPE de 3.12% em 1999 e de 3.93% em 2000. É importante observar que quando é feita referência a k passos à frente, considerando os dados seccionais, a previsão refere-se a $24k$ horas à frente, porque os dados originais são de hora em hora. Isto leva a interpretar como $24k$ passos à frente, de modo que $(1, 2, \ldots, 7)$ passos à frente correspondem realmente a $(7, 14, \ldots, 168)$ passos à frente de hora em hora. Para permitir uma boa visualização sobre o desempenho do modelo, foram analisadas quatro curvas de carga diária, uma de cada estação do ano. Uma comparação dos modelos conforme as estações do ano é mostrada na figura 4.6.

PUC-Rio - Certificação Digital Nº 9824881/CA

Tabela 4.4: MAPE – ano 2000 dos modelos *DAGLM* e *DASARIMA*

hora	*DAGLM* passos à frente							*DASARIMA* passos à frente						
	1	2	3	4	5	6	7	1	2	3	4	5	6	7
1	**4.37**	6.42	7.38	7.92	8.18	8.20	8.16	4.53	**5.77**	**6.55**	**6.71**	**6.73**	**6.65**	**5.70**
2	**4.43**	6.54	7.62	8.15	8.43	8.45	8.42	4.53	**5.73**	**6.42**	**6.55**	**6.54**	**6.50**	**5.65**
3	**4.37**	6.45	7.61	8.15	8.46	8.44	8.45	4.44	**5.58**	**6.27**	**6.36**	**6.35**	**6.33**	**5.48**
4	**4.23**	6.24	7.40	7.93	8.26	8.28	8.26	4.35	**5.30**	**6.05**	**6.15**	**6.15**	**6.06**	**5.24**
5	**4.03**	5.99	7.12	7.69	8.01	8.06	8.09	4.24	**5.07**	**5.82**	**5.92**	**5.97**	**5.87**	**5.11**
6	**3.81**	5.60	6.58	7.05	7.32	7.40	7.46	4.14	**5.10**	**6.21**	**6.48**	**6.46**	**6.16**	**5.14**
7	**3.60**	**5.08**	**5.88**	**6.21**	**6.43**	**6.53**	6.59	4.47	6.93	8.68	9.12	9.16	8.72	**6.57**
8	**3.22**	**4.61**	**5.26**	**5.54**	**5.73**	**5.81**	**5.88**	4.70	9.08	11.41	12.08	12.13	11.58	8.61
9	**3.03**	**4.28**	**4.75**	**5.03**	**5.22**	**5.30**	**5.35**	4.92	10.69	13.48	14.28	14.33	13.86	10.25
10	**2.97**	**4.06**	**4.49**	**4.67**	**4.80**	**4.86**	**4.92**	5.24	12.01	15.02	15.80	15.85	15.52	11.48
11	**3.00**	**3.97**	**4.39**	**4.56**	**4.69**	**4.76**	**4.81**	5.39	12.69	16.11	16.81	16.79	16.53	11.99
12	**3.05**	**3.99**	**4.41**	**4.56**	**4.68**	**4.73**	**4.75**	5.60	12.78	16.47	17.14	17.12	16.84	12.10
13	**3.00**	**4.00**	**4.44**	**4.59**	**4.71**	**4.76**	**4.78**	5.40	12.56	16.27	16.94	16.93	16.49	11.71
14	**3.17**	**4.29**	**4.72**	**4.93**	**5.05**	**5.10**	**5.10**	5.55	12.85	16.80	17.65	17.60	17.09	12.11
15	**3.38**	**4.51**	**4.98**	**5.19**	**5.31**	**5.35**	**5.36**	5.74	13.29	17.70	18.40	18.45	17.89	12.41
16	**3.43**	**4.55**	**4.98**	**5.23**	**5.33**	**5.41**	**5.43**	5.77	13.17	17.77	18.47	18.57	17.90	12.29
17	**3.60**	**4.43**	**4.77**	**4.89**	**4.98**	**5.04**	**5.02**	5.39	12.19	16.65	17.26	17.39	16.63	11.25
18	**3.44**	**4.19**	**4.43**	**4.52**	**4.58**	**4.58**	**4.59**	4.79	9.24	12.18	12.66	12.68	12.09	8.67
19	**3.21**	**3.77**	**4.07**	**4.30**	**4.44**	**4.50**	**4.53**	4.03	6.43	7.81	8.13	8.14	7.80	6.18
20	**2.71**	**3.30**	**3.60**	**3.74**	**3.87**	**3.93**	**3.93**	3.47	5.15	5.74	5.95	6.05	5.80	4.99
21	**2.63**	**3.45**	**3.85**	**4.11**	**4.16**	**4.23**	**4.22**	3.17	4.89	5.67	5.84	5.95	5.69	4.56
22	**2.98**	**4.12**	**4.62**	**4.94**	**5.11**	**5.17**	5.16	3.36	5.30	6.61	6.92	6.90	6.56	**4.92**
23	**3.58**	**5.15**	**5.82**	**6.22**	**6.41**	**6.43**	6.41	3.99	5.64	6.95	7.28	7.24	6.79	**5.30**
24	**4.14**	5.96	6.81	7.34	7.56	7.60	7.56	4.50	**5.79**	**6.78**	**7.09**	**6.98**	**6.70**	**5.59**
min	2.63	3.30	3.60	3.74	3.87	3.93	3.93	3.17	4.89	5.67	5.84	5.95	5.69	4.56
med	3.47	4.79	5.42	5.73	5.91	5.96	5.97	4.65	8.47	10.64	11.08	11.10	10.75	8.05
max	4.43	6.54	7.62	8.15	8.46	8.45	8.45	5.77	13.29	17.77	18.47	18.57	17.90	12.41

Tabela 4.5: MAPE de datas selecionadas dos modelos *DAGLM* e *DASARIMA*

data	dia da semana	estação do ano	*MAPE*	
			DASARIMA	*DAGLM*
12/04/1999	segunda	outono	2.52	0.89
11/07/1999	domingo	inverno	1.83	1.26
06/10/1999	quarta	primavera	4.03	2.89
02/02/2000	quarta	verão	2.58	2.11

A avaliação do desempenho do modelo para um período de 7 dias entre 9 e 15 de maio de 1999 mostra que o modelo *DAGLM* consegue MAPE de 2.17% e o modelo *DASARIMA* de 3.05%. A figura 4.7 apresenta as curvas de cargas real e prevista do período.

4.2.2
Um ensaio incluindo o Racionamento

Este ensaio abrange o período de racionamento ocorrido no Brasil entre 4 de junho de 2001 e 28 de fevereiro de 2002. De maneira diferente dos ensaios anteriores, os parâmetros são estimados recursivamente e isto melhora a capacidade preditiva do modelo e sua capacidade de incorporar mudanças, como por exemplo, a quebra estrutural causada pelo racionamento. Para contextualizar a análise, a crise energética brasileira pode ser resumida como:

PUC-Rio - Certificação Digital Nº 9824881/CA

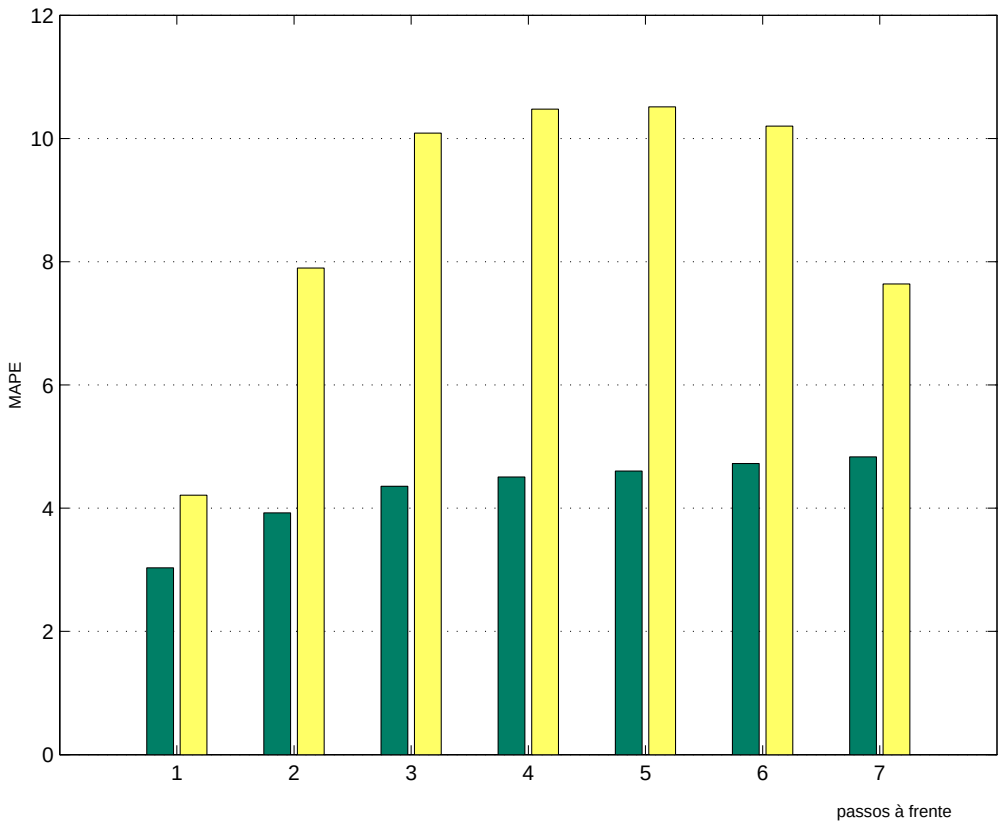

Figura 4.4: MAPE médio 1999 – *DAGLM* (escuro), *DASARIMA* (claro)

A Crise Energética Brasileira

Na maioria dos países, os investimentos em energia são balanceados entre diferentes tipos de fontes energéticas, com o objetivo de evitar que problemas específicos de um tipo de geração afetem grande parte da produção de eletricidade. A diversificação diminui o risco quando os impactos negativos não são fortemente correlacionados entre as diversas alternativas. Entretanto, em grande parte devido aos generosos volumes de água que existem no país, a matriz energética brasileira é altamente dependente das plantas de hidroelétricas, que constituem 87% do total de recursos de eletricidade. A geração de eletricidade através de hidroelétricas é, por sua vez, altamente dependente das condições climáticas, como os períodos chuvosos. Outras fontes de eletricidade no Brasil são usinas termoelétricas (10%) e nucleares (2%).

A crise energética no Brasil foi percebida no início de 2001, quando foi anunciado que os reservatórios estavam com nível médio em torno de somente 34% de sua capacidade, quando precisavam estar com pelo menos metade da capacidade a fim de suportar a estação seca em operação normal. Este quadro era verdadeiro apenas para região sudeste e centro-oeste, enquanto a região norte estava em uma situação ligeiramente menos desfavorável (o racionamento foi retardado em dois meses para esta região) e a região sul, por outro lado, descartava excesso de água de seus reservatórios. Entretanto, a dimensão continental

PUC-Rio - Certificação Digital Nº 9824881/CA

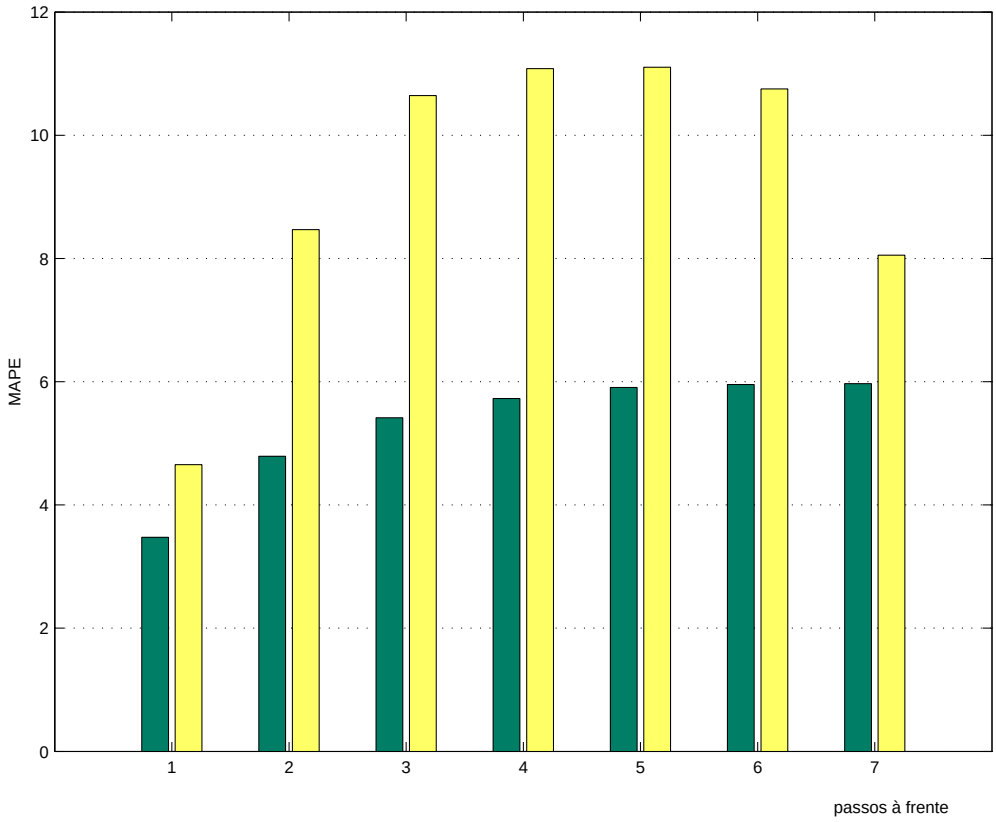

Figura 4.5: MAPE médio 2000 - *DAGLM* (escuro), *DASARIMA* (claro)

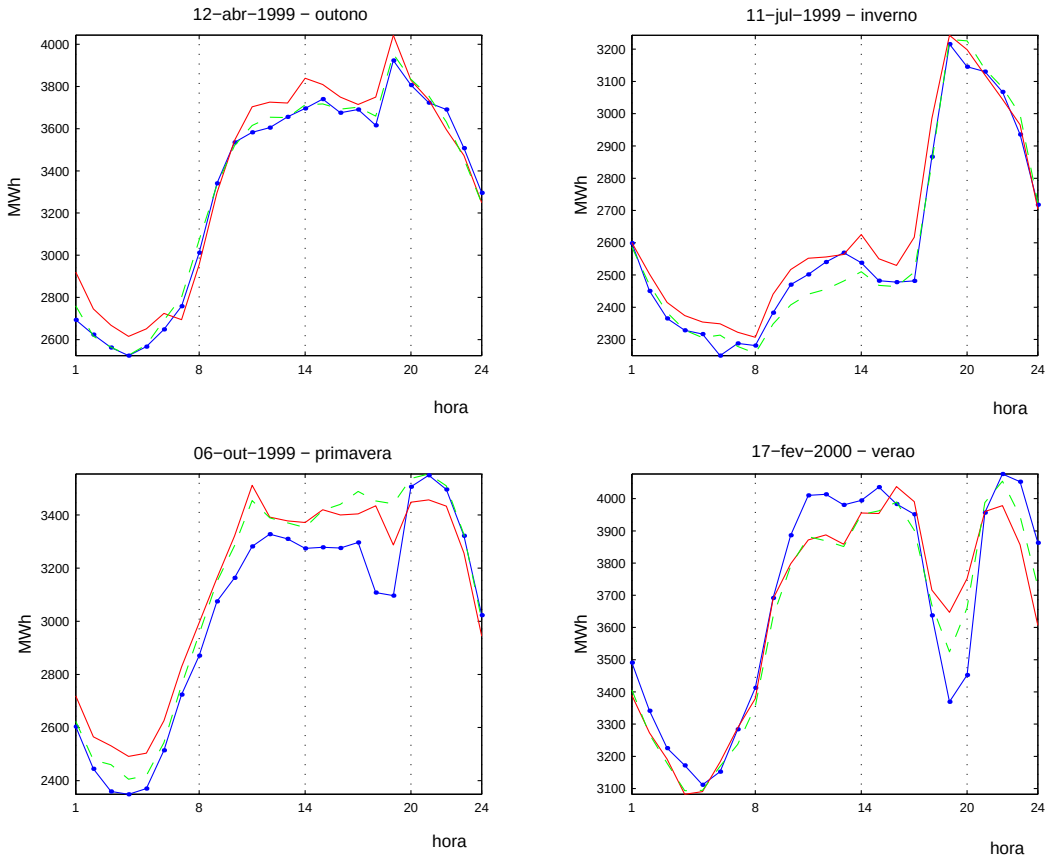

Figura 4.6: curva de carga diária das 4 estações do ano - real (ponto), *DASARIMA* (linha), DAGLM (tracejado)

PUC-Rio - Certificação Digital Nº 9824881/CA

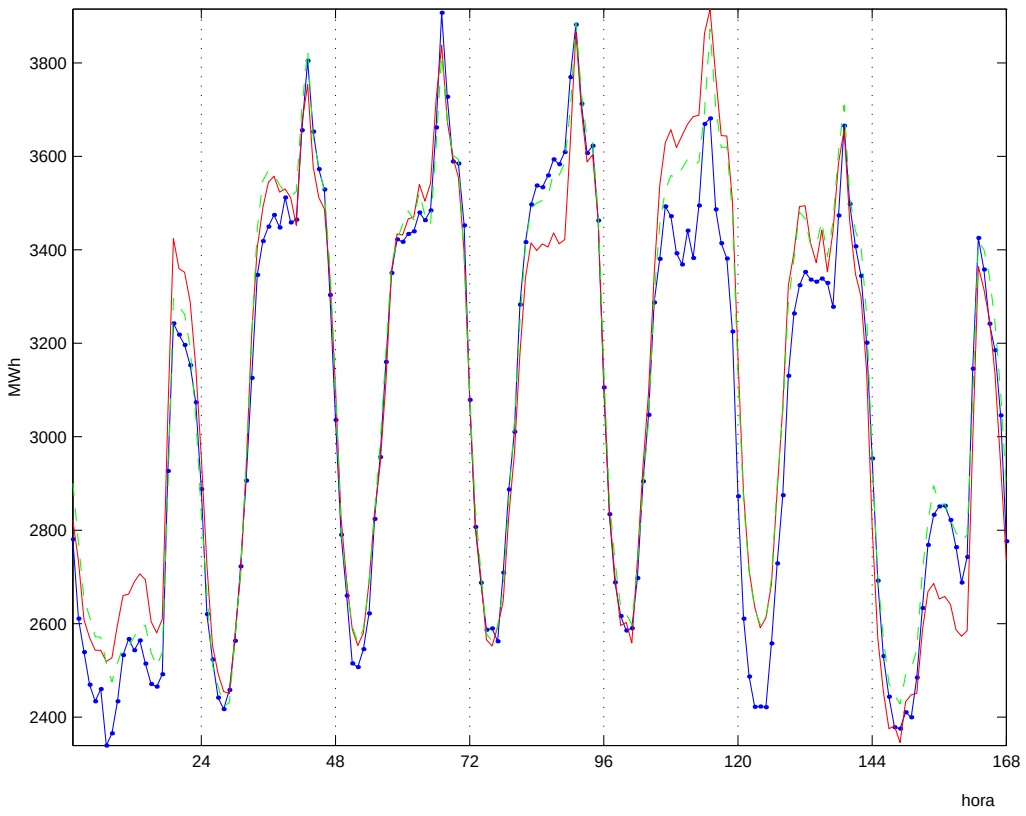

Figura 4.7: curva de carga de 09 a 15/05/1999 - real (ponto), *DASARIMA* (linha), DAGLM (tracejado)

do país, aliada à falta de investimentos em linhas de transmissão, provocou a existência de regiões geográficas "eletricamente isoladas".

Esta situação ocorreu principalmente devido à falta de chuvas e de investimentos em geração e transmissão de eletricidade. A primeira causa resultou no menor volume hídrico estocado nos últimos 20 anos, enquanto a segunda foi reduzida à metade nos anos de 1990. Como se não fosse suficiente, o consumo de eletricidade estava crescendo acentuadamente à uma taxa de aproximadamente 3% ao ano desde o início da década de 1990. Para administrar estes problemas, uma campanha de racionamento foi anunciada, tendo início no dia 4 de junho de 2001 e terminando no dia 28 de fevereiro de 2002, quando as condições técnicas permitiram o retorno do consumo à um patamar normal. A necessidade de redução no consumo da eletricidade, entretanto, foi anunciada aproximadamente um mês antes do início do racionamento e as pessoas adotaram mais cedo medidas de uso racional de energia a fim de diminuir o risco de interrupção no fornecimento de energia. A Figura 4.10 compara a carga média diária dos anos 2000 e 2001, onde se pode observar pela inspeção visual uma queda no nível da carga em 2001, ocorrendo suavemente a partir do início de maio (em torno do 125° dia de 2001) a meados do mês de julho (em torno do 195° dia de 2001). Após este período, a carga de 2001 parece ter incorporado esta queda de nível, seguindo seu padrão sazonal característico.

PUC-Rio - Certificação Digital Nº 9824881/CA

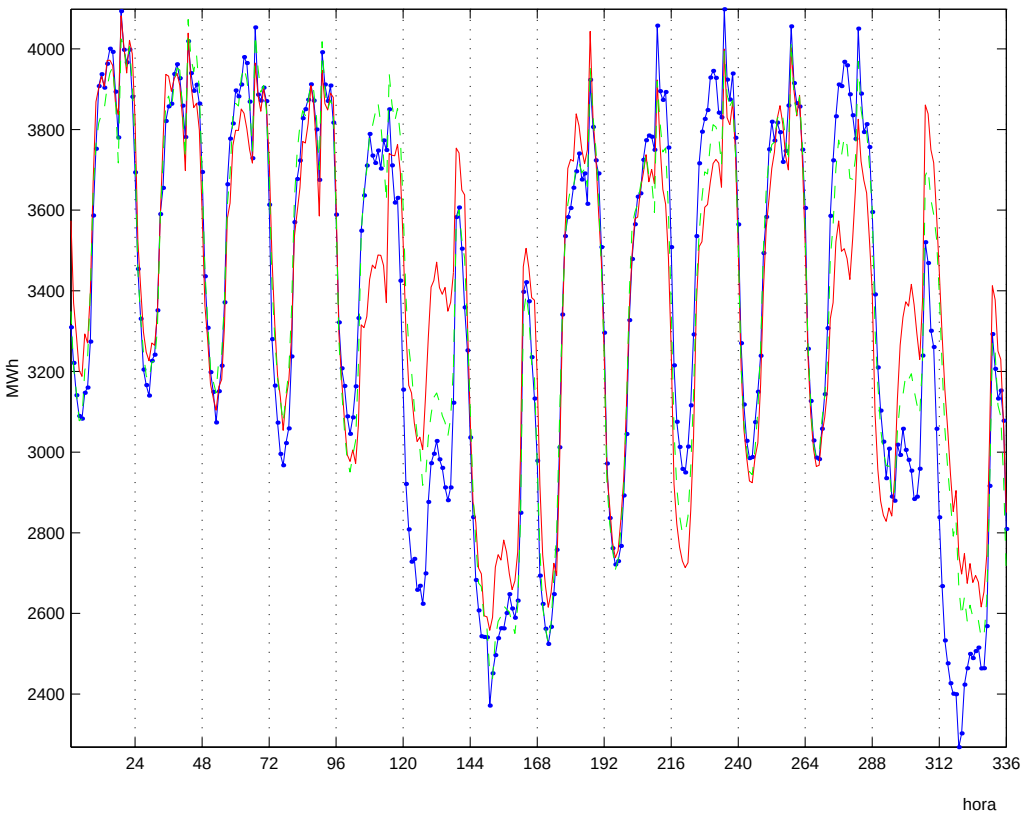

Figura 4.8: curva de carga de 05 a 18/04/1999 - real (ponto), *DASARIMA* (linha), DAGLM (tracejado)

O racionamento consistiu basicamente em estabelecer uma meta de consumo (20% inferior ao consumo médio observado no mesmo período do ano anterior) para cada unidade consumidora (casa, loja, indústria). Os consumidores que consumissem abaixo da meta estabelecida seriam recompensados com um bônus na conta de eletricidade e aqueles que consumissem acima da meta seriam sobretaxados e correriam o risco de ter o fornecimento interrompido por três dias. A campanha foi bem sucedida e o consumo teve grande diminuição, graças à conscientização popular e ao trabalho técnico. Embora fosse esperado o crescimento do consumo após o término do racionamento, isto não ocorreu. Medidas como substituições de lâmpadas, mudanças nos hábitos e aumento da eficiência no uso da eletricidade, reduziram o desperdício. Também, o aumento no preço da energia, imposto pelo governo para compensar as perdas de receita das empresas concessionárias, contribuiu para que a demanda permanescesse em nível baixo. O processo de previsão foi conturbado por uma quebra estrutural na série de demanda. Deste modo, um dos pressupostos de um modelo de previsão, passado e futuro com características semelhantes, não é atendido.

O modelo *DAGLM* se readapta rapidamente aos novos dados e consegue erros de previsão quase tão baixos como os observados no ensaio anterior (antes do racionamento). Assim, mesmo com uma quebra estrutural no nível da série,

PUC-Rio - Certificação Digital Nº 9824881/CA

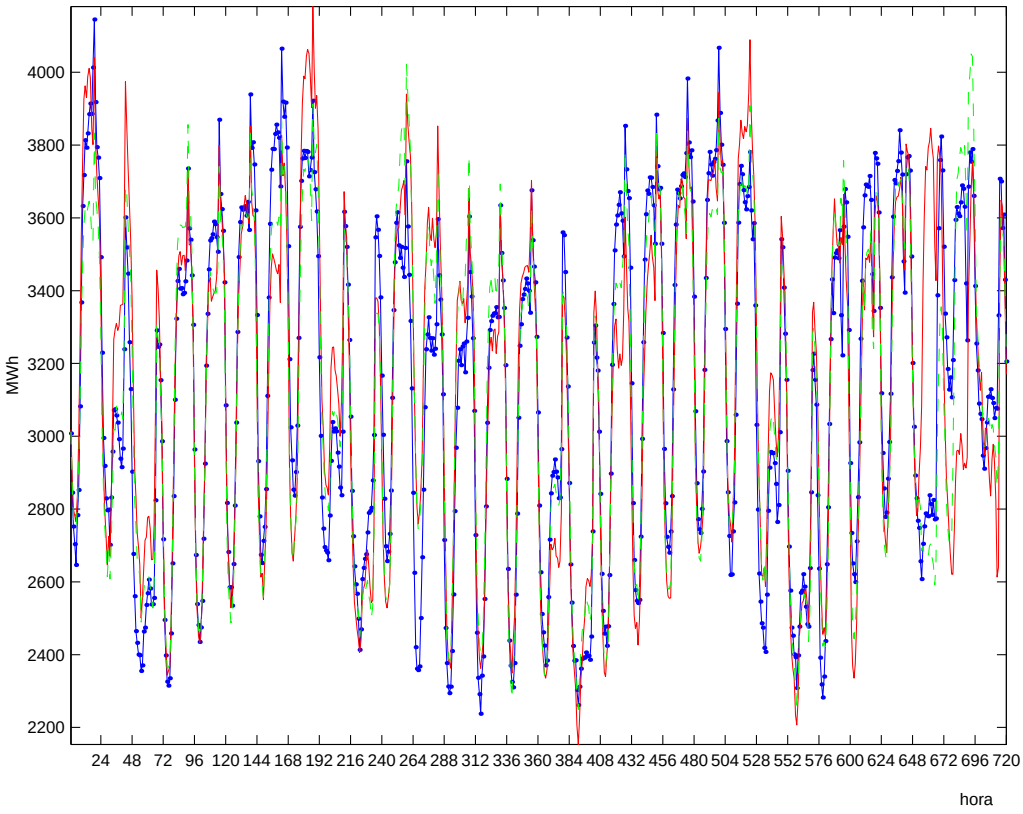

Figura 4.9: curva de carga de 15/09/2000 a 14/10/2000 - real(ponto), *DASARIMA*(linha), DAGLM(tracejado)

algumas propriedades estatísticas não são perdidas e são capturadas pelo modelo.

análise e comentários 4

A tabela 4.6 mostra os erros para cada hora, dividido em três partes: o ano de 2000 (antes da campanha de racionamento), o ano de 2001 (incluindo a campanha de racionamento), e o período compreendido entre 1 de outubro de 2001 e 30 de setembro de 2002 (último ano de dados, parte durante a campanha, parte depois do término da campanha), que é identificado como 2002 [2].

[2] de 1 de outubro de 2001 até 30 de setembro de 2002

PUC-Rio - Certificação Digital Nº 9824881/CA

PUC-Rio - Certificação Digital Nº 9824881/CA

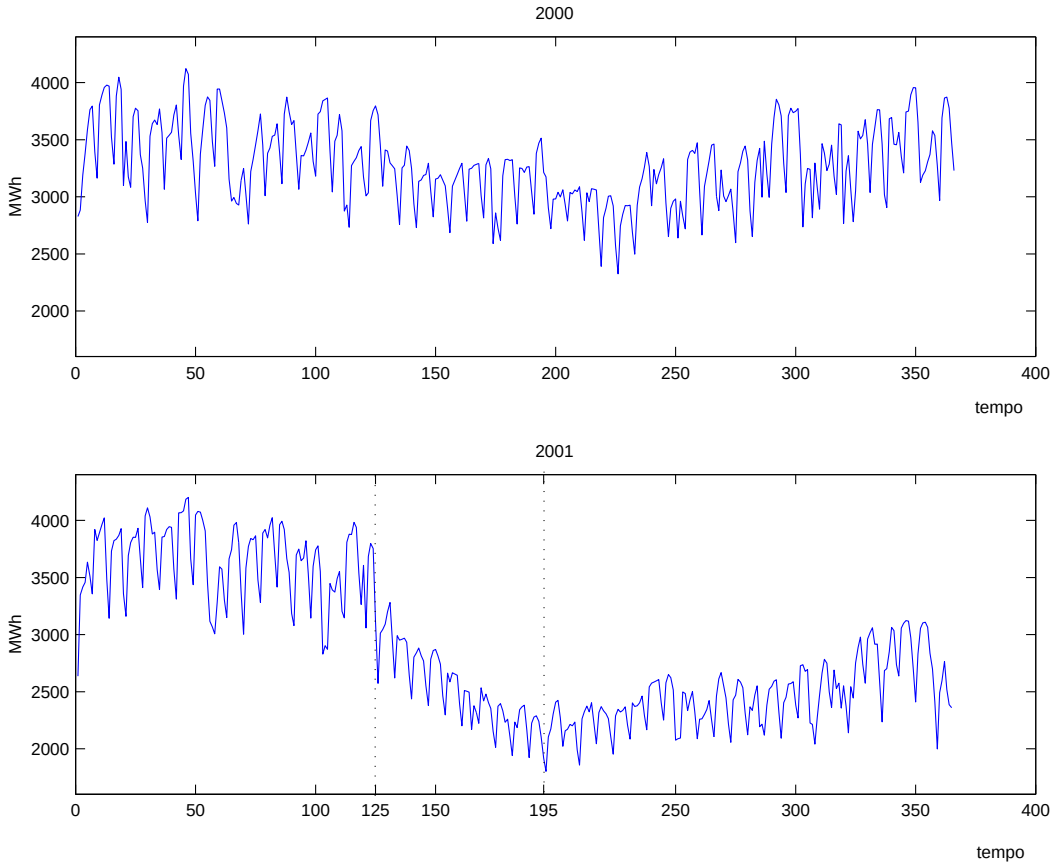

Figura 4.10: Carga diária dos anos 2000 e 2001

PUC-Rio - Certificação Digital Nº 9824881/CA

Tabela 4.6: MAPE dos anos de 2000 (antes do racionamento), 2001 (incluindo o racionamento no final) e 2002* (incluindo o racionamento no início).

hora	2000 passos à frente							2001 passos à frente							2002* passos à frente						
	1	2	3	4	5	6	7	1	2	3	4	5	6	7	1	2	3	4	5	6	7
1	4.34	6.41	7.42	7.92	8.12	8.11	8.06	4.11	6.21	7.28	8.18	8.96	9.48	9.88	4.60	6.31	7.35	8.16	8.68	9.00	9.22
2	4.35	6.54	7.68	8.20	8.40	8.40	8.37	4.12	6.29	7.41	8.30	9.02	9.58	10.0	4.50	6.31	7.45	8.28	8.76	9.14	9.42
3	4.30	6.48	7.64	8.15	8.38	8.36	8.36	4.02	6.18	7.28	8.13	8.86	9.40	9.84	4.36	6.34	7.34	8.09	8.61	8.98	9.34
4	4.15	6.27	7.43	7.91	8.18	8.18	8.18	3.85	5.98	7.08	7.97	8.70	9.23	9.59	4.36	6.27	7.28	8.10	8.58	8.96	9.29
5	4.00	6.02	7.13	7.66	7.91	7.93	7.98	3.74	5.74	6.87	7.73	8.44	8.98	9.43	4.16	6.02	7.10	7.84	8.27	8.67	9.01
6	3.79	5.58	6.54	7.00	7.21	7.28	7.35	3.70	5.58	6.59	7.35	7.98	8.50	8.89	3.86	5.56	6.57	7.21	7.66	8.05	8.34
7	3.59	5.05	5.80	6.14	6.34	6.44	6.48	3.50	5.05	5.98	6.65	7.25	7.69	8.08	3.81	5.01	5.83	6.31	6.67	7.04	7.30
8	3.23	4.51	5.13	5.43	5.65	5.76	5.85	3.19	4.56	5.40	6.02	6.59	7.01	7.36	3.55	4.62	5.31	5.75	6.14	6.48	6.70
9	3.10	4.16	4.64	4.95	5.14	5.26	5.34	3.14	4.24	5.03	5.57	6.05	6.46	6.78	3.23	4.27	4.99	5.43	5.81	6.12	6.34
10	2.99	3.98	4.39	4.64	4.78	4.86	4.93	3.09	4.22	4.94	5.46	5.96	6.29	6.64	3.28	4.22	4.93	5.31	5.70	5.93	6.19
11	2.96	3.94	4.31	4.57	4.70	4.78	4.82	3.13	4.36	5.12	5.58	6.04	6.38	6.67	3.17	4.05	4.79	5.16	5.51	5.74	5.92
12	3.00	3.89	4.29	4.52	4.61	4.69	4.72	3.10	4.26	5.01	5.52	5.91	6.25	6.53	2.99	3.92	4.55	4.99	5.30	5.53	5.73
13	2.96	3.95	4.35	4.58	4.67	4.74	4.75	3.17	4.45	5.18	5.68	6.11	6.47	6.75	2.98	3.99	4.63	5.02	5.38	5.64	5.82
14	3.20	4.25	4.66	4.92	5.02	5.07	5.06	3.34	4.76	5.58	6.15	6.60	7.01	7.30	3.57	4.67	5.27	5.69	6.04	6.32	6.49
15	3.32	4.38	4.88	5.15	5.27	5.31	5.30	3.47	4.92	5.85	6.42	6.87	7.30	7.57	3.70	4.86	5.53	5.98	6.34	6.65	6.76
16	3.36	4.40	4.87	5.14	5.27	5.35	5.35	3.45	4.92	5.87	6.47	6.92	7.32	7.62	3.66	4.83	5.50	5.95	6.28	6.58	6.74
17	3.30	4.26	4.64	4.77	4.87	4.94	4.94	3.30	4.69	5.54	6.11	6.54	6.90	7.21	3.38	4.45	5.11	5.50	5.78	6.08	6.27
18	3.22	4.02	4.28	4.41	4.49	4.49	4.52	3.28	4.57	5.23	5.67	6.07	6.44	6.75	3.09	4.09	4.75	5.10	5.48	5.85	6.13
19	2.97	3.59	3.97	4.20	4.35	4.45	4.52	3.36	4.33	4.81	5.10	5.47	5.80	6.17	3.01	3.92	4.52	4.98	5.40	5.85	6.21
20	2.70	3.25	3.56	3.70	3.83	3.89	3.92	2.98	3.85	4.40	4.80	5.16	5.47	5.82	2.89	3.66	4.18	4.52	4.91	5.25	5.55
21	2.57	3.44	3.84	4.11	4.17	4.22	4.21	2.74	3.80	4.52	5.06	5.45	5.75	6.07	2.73	3.67	4.23	4.58	4.89	5.21	5.50
22	2.93	4.11	4.67	4.98	5.15	5.20	5.15	2.99	4.36	5.11	5.71	6.22	6.57	6.88	3.05	4.25	4.91	5.38	5.75	6.05	6.33
23	3.60	5.16	5.88	6.26	6.43	6.42	6.37	3.47	5.15	6.05	6.72	7.32	7.69	8.04	3.68	5.04	5.91	6.47	6.85	7.12	7.33
24	4.13	6.02	6.93	7.39	7.57	7.55	7.48	3.89	5.90	6.96	7.71	8.43	8.88	9.29	4.27	6.00	6.95	7.58	8.04	8.26	8.52

Está claro na tabela 4.6 que o desempenho das previsões se deteriora um pouco, de 2000 para 2001, mas não tanto quanto se poderia esperar considerando a quebra estrutural no nível da série de carga. Os resultados para 1 a 3 passos à frente apresentam uma redução do MAPE na primeira e nas últimas horas (onde o pior caso do MAPE é encontrado), sendo a redução maior para 1 passo à frente. As outras ocorrências mostram de forma geral um aumento no MAPE, da ordem de $0,2\%$ na média para 1 passo à frente nas horas do meio do dia e um aumento em torno de 2% para 7 passos à frente. De 2001 para 2002*, ocorre um aumento do MAPE 1 passo à frente em certas horas, mas para mais de 1 passo à frente os MAPE sempre diminuem. Por exemplo, o MAPE mínimo e máximo 1 passo à frente em 2000 são respectivamente $2,57\%$ (hora 21) e $4,35\%$ (hora 2); em 2001, $2,74\%$ (hora 21) e $4,12\%$ (hora 2); e em 2002 *, $2,73\%$ (hora 21) e $4,60\%$ (hora 1). Para 7 passos à frente os valores são: em 2000, $3,92\%$ (hora 20) e $8,37\%$ (hora 2); em 2001, $5,82\%$ (hora 20) e $10,0\%$ (hora 2); e em 2002*, $5,50\%$ (hora 21) e $9,42\%$ (hora 2).

análise e comentários 5

A figura 4.11 mostra os erros de previsão 1 passo à frente para o período de 21 de março de 2001 a 7 de setembro 2001, para horas representativas dos diferentes períodos do dia. Observa-se que os erros se tornam consistentemente negativos aproximadamente quinze dias antes e consistentemente positivos depois que a necessidade de redução do consumo foi anunciada (marcado com uma linha tracejada horizontal, em torno do 125° dia de 2001, no início de maio), mostrando que as previsões se tornam enviesadas para cima. Entretanto, este viés diminui lentamente, demorando aproximadamente cem dias para alcançar um nível em torno de zero. O intervalo dos erros, por outro lado, parece diminuir após esta data para certas horas. Isto não significa necessariamente que o MAPE deve diminuir porque, como pode ser observado na figura 4.10, a carga também diminui (20% aproximadamente) a partir desta data. De fato, comparando os erros dos anos (tabela 4.6), fica claro que os MAPE em geral aumentam.

análise e comentários 6

A Figura 4.12 mostra a evolução do parâmetro $d1$ para as horas representativas das diferentes partes do dia, referindo-se à intensidade da memória longa sazonal, no decorrer do ano de 2001. Em geral, este parâmetro aumenta durante a queda no nível, mas posteriormente, quando esta queda é incorporada à série, tende a retornar para sua escala de valores anterior. A exceção é a hora 1, representando as primeiras e últimas horas do dia quando se espera uma diminuição das atividades sociais, que parece incorporar a alteração em $d1$. Nota-se, entretanto, que mesmo as mudanças no parâmetro $d1$ estimado para diferentes horas não são drásticas, sendo percebidas apenas porque a figura 4.12 está em escala

PUC-Rio - Certificação Digital Nº 9824881/CA

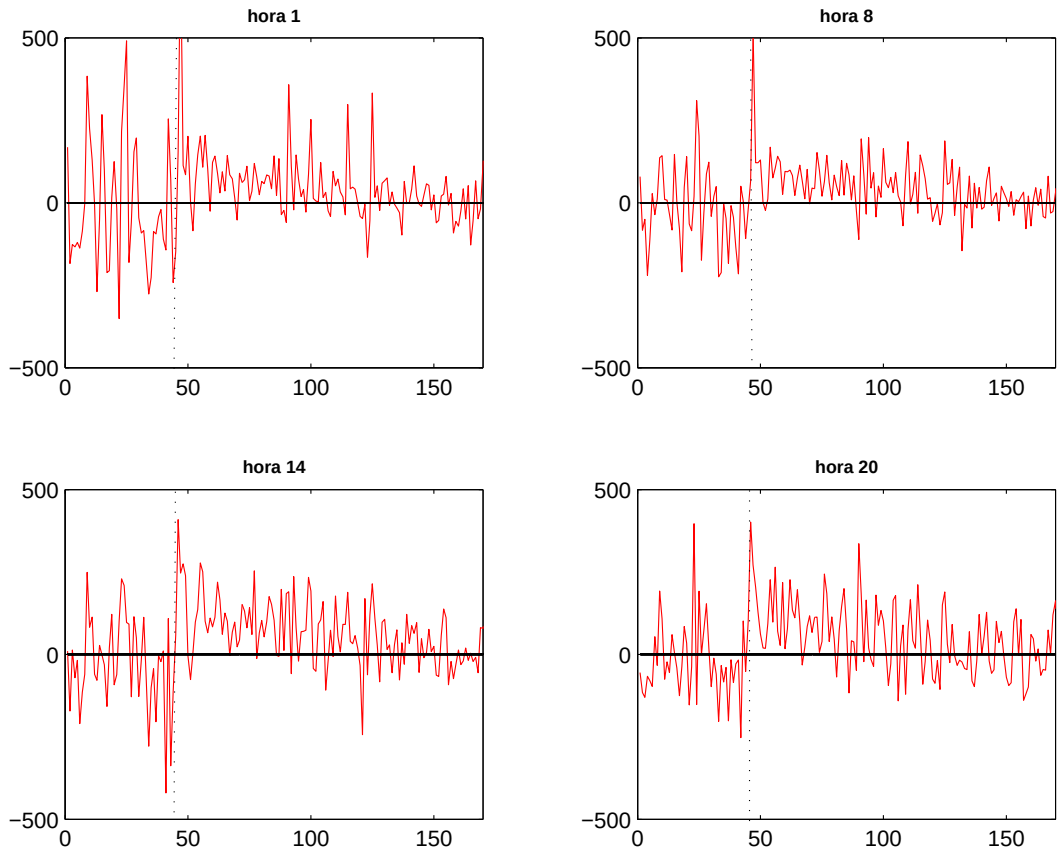

Figura 4.11: erro de previsão 1 passo à frente no período de 21 de março de 2001 a 7 de setembro de 2001.

reduzida. Todas estas evidências conduzem a concluir que mesmo apresentando uma quebra estrutural, algumas propriedades estatísticas da série de demanda de eletricidade não são perdidas, e estas propriedades são capturadas pelo menos parcialmente pelo modelo *DAGLM*.

4.3 Resumo

Prever o perfil da carga diária foi o primeiro objetivo dos trabalhos em previsão de carga de curto prazo. O perfil da curva de carga é previsto de forma simples, segura e robusta, pela previsão 1 passo à frente das 24 séries correspondentes à cada hora do dia. Além disso, são feitas previsões para até 7 dias à frente, que corresponde na abordagem de previsão longitudinal ao horizonte de 168 passos à frente.

A razão para o grande intervalo de variação de resultados, relatados na literatura, é ocasionada pelas peculiaridades do conjunto de dados e pelo número de unidades consumidoras, considerando que o nível do desvio da carga agregada diminui quando o número de unidades consumidoras aumenta, fazendo com que a diferença nos resultados pareça natural.

PUC-Rio - Certificação Digital Nº 9824881/CA

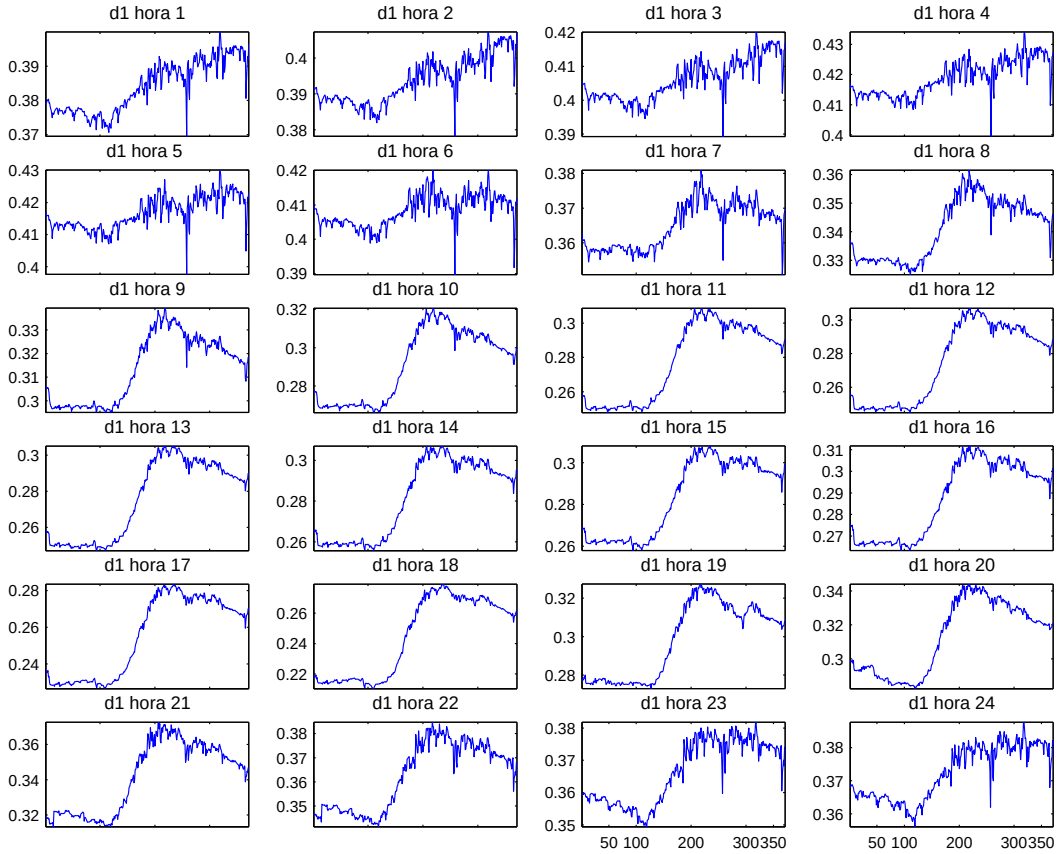

Figura 4.12: evolução do parâmetro $d1$ (relativo à memória longa sazonal) durante o ano de 2001 (escalas diferentes)

Uma parte do conjunto de dados é utilizada para estimar os parâmetros dos modelos utilizando uma janela de 4 anos de dados. O desempenho dos modelos foi testado com os dados correspondentes aos anos de 1999 e 2002. Foram obtidos erros médios de aproximadamente 2%. Os resultados foram, entretanto, melhores no outono.

É proposto um modelo estocástico empregando técnicas lineares sofisticadas, como a memória longa generalizada induzida por processos de Gegenbauer, para modelar o comportamento sazonal da carga. São realizados dois ensaios com o modelo, o primeiro com um período de previsão sem quebra estrutural e o segundo incluindo o período de racionamento. O desempenho do modelo é comparado com o modelo *DASARIMA* durante o período de previsão. O modelo é claramente superior ao modelo de referência.

O primeiro ensaio, produziu para o ano 1999, previsões para um dia (24 horas) à frente com erro médio percentual (MAPE) variando entre $2,23\%$ (21ª hora) e $3,76\%$ (1ª hora). Para dois dias à frente, a variação dos erros está entre $2,83\%$(21ª hora) e $5,14\%$(1ª hora). Para sete dias à frente, por sua vez, o MAPE não aumenta demasiadamente, permanecendo entre $3,12\%$(20ª hora) e $6,81\%$(2ª hora).

PUC-Rio - Certificação Digital Nº 9824881/CA

Para o ano 2000, as previsões para um dia (24 horas) à frente apresentaram erro médio percentual MAPE) variando entre $2,5\%$ (21ª hora) e $4,4\%$ (2ª hora). Para dois dias à frente, a variação dos erros está entre $3,3\%$(20ª hora) e $6,5\%$(2ª hora). Para sete dias à frente, por sua vez, o MAPE não aumenta demasiadamente, permanecendo entre $3,9\%$(20ª hora) e $8,4\%$(2ª hora).

O segundo ensaio, estuda o comportamento da demanda de carga elétrica durante o racionamento ocorrido em 2001/2002 . A campanha de racionamento, quase em âmbito nacional, foi provocada por uma combinação de eventos que ocorreram no Brasil, tais como um período extremamente seco e falta de investimentos na geração e na transmissão, aliados ao fato de ser o Brasil altamente dependente de energia hidroelétrica. O período de racionamento é estudado e mostra uma quebra estrutural na série, com redução da carga em aproximadamente 20%, se comparada com o mesmo período do ano anterior. Mesmo assim, a precisão da previsão diminui somente marginalmente.

Os efeitos de calendário são modelados da mesma forma apresentada no ensaio que estuda o comportamento do modelo sem quebra estrutural. A precisão da previsão é avaliada antes, durante e depois do período de racionamento. As características de memória longa generalizada na série e também algumas propriedades estatísticas não são perdidas após a quebra estrutural.

O modelo desenvolvido aplica-se à dados seccionais, isto é, a carga de cada hora do dia é tratada separadamente como uma série. No caso estudado, mesmo na presença de uma quebra estrutural, a precisão da previsão não se deteriora drasticamente, e o modelo readapta-se rapidamente à nova situação. Conclui-se que existe a presença de características de memória longa generalizada nos dados estudados e que mesmo com a ocorrência de uma grande quebra estrutural, como a imposta pela campanha de racionamento, não destruiu essas propriedades estatísticas.

PUC-Rio - Certificação Digital Nº 9824881/CA