



Modulo: Fundamentos de Economia de Energia

Sessão: Previsão de Demanda de Energia Elétrica
29/04/2006; 13:30 às 17:30

Prof. Reinaldo Castro Souza

DEE & IEPUC; PUC-Rio

PROGRAMAÇÃO



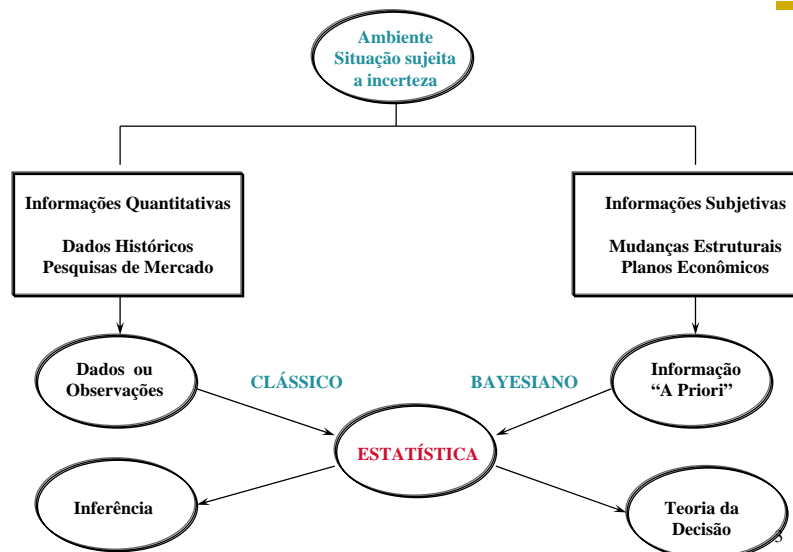
Sessão 1: Estado da Arte

Sessão 2: Previsões no Setor Elétrico
(Curto x Longo Prazo ; Ativo x Reativo)

Sessão 3: Como Gerar Previsões Estatísticas de Mercado

Sessão 4: Ilustração e Demonstração de Sistemas de Previsão

MODELAGEM ESTATÍSTICA



SÉRIE TEMPORAL



- “É um conjunto de dados discretos observados em intervalos de tempo equidistantes e que apresentam uma **DEPENDÊNCIA SERIAL** entre eles”
- Representação: $Z_t ; t = 1,2,\dots$

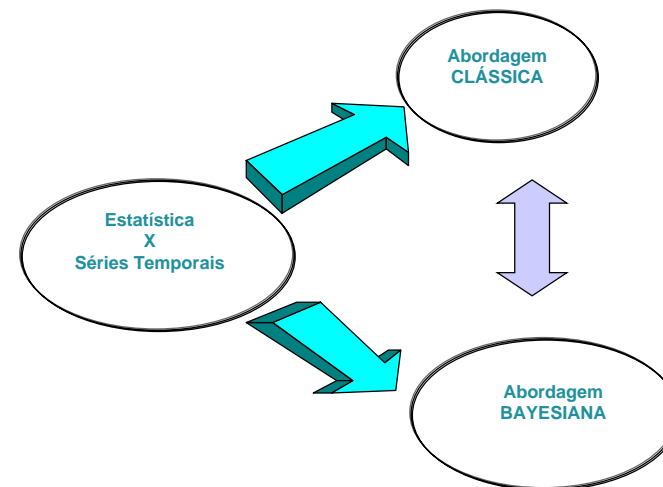
SÉRIE HISTÓRICA (SÉRIE TEMPORAL) x AMOSTRA ALEATÓRIA (ESTATÍSTICA)



- Peso de um mesmo indivíduo ao longo do tempo (SÉRIE)
- Peso de indivíduos selecionados aleatoriamente em uma dada população (AMOSTRA ALEATÓRIA)
- Exemplos de Séries Temporais:
 - Market Share; Distribuição; Promoção; Índices e/ou Indicadores; Preço; Propaganda; Consumo; Demanda; Produção; Precipitações; etc...

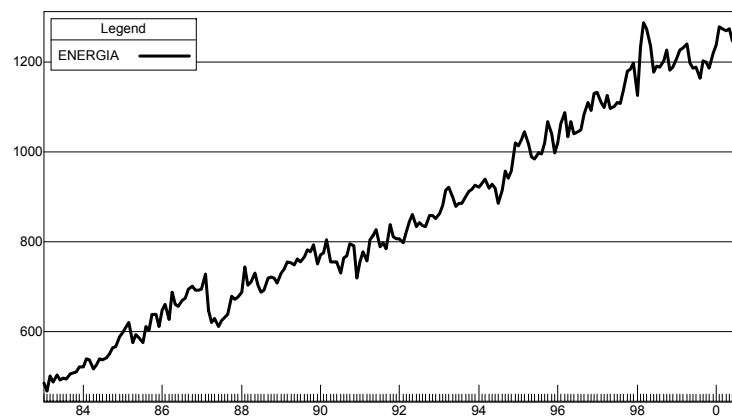
5

CONCLUSÃO



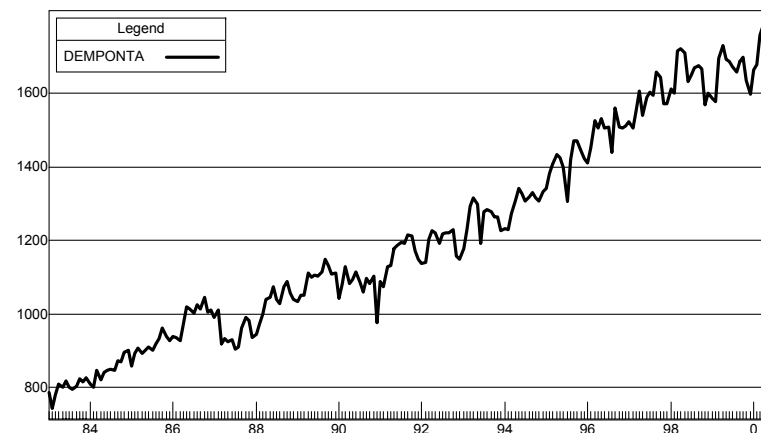
6

EXEMPLO: COELBA - ENERGIA



7

EXEMPLO: COELBA DEMANDA NA PONTA



8

PREVISÃO DE UMA SÉRIE TEMPORAL



- Dado o conjunto de observações passadas (Série Histórica):

$$Z_t = (Z_1, Z_2, \dots, Z_t)$$

- Determinar, com base nas propriedades estatísticas desta Série Histórica (e possivelmente de outras), quais os valores prováveis:

$$Z_{t+1}, Z_{t+2}, \dots, Z_{t+k} \quad (k \text{ é o Horizonte de Previsão})$$

- Notação: Z_{t+i} ; $i = 1, 2, \dots, k$

9

HORIZONTE DE PREVISÃO



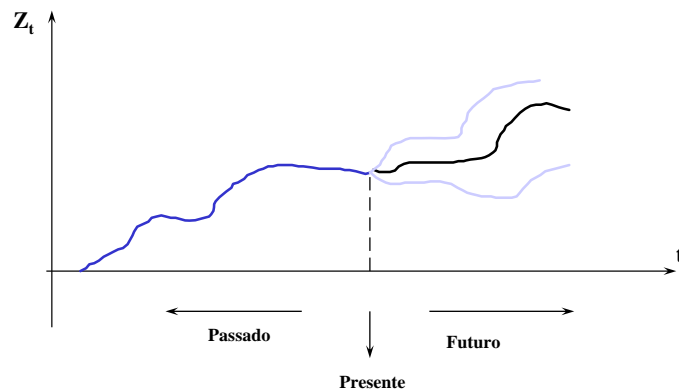
- Previsão a curto prazo x futurologia, onde “CURTO PRAZO” depende do grau de previsibilidade da série (Aspecto crítico: os métodos garantem previsão “ótima” 1-passo à frente).
- Precisão diminui à medida em que aumenta o horizonte de previsão e o erro da previsão 1-passo à frente servirá como indicador do horizonte.

10

DISTRIBUIÇÃO DE PROBABILIDADES



- Previsão sob a forma de uma Distribuição de Probabilidades ao invés de Previsão Pontual e Intervalo de Previsão com uma dada Probabilidade.



11

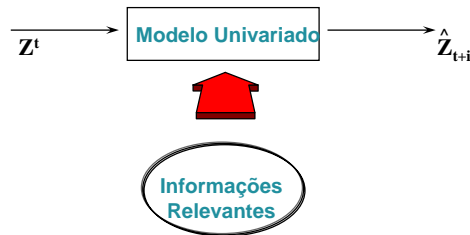
MACRO-CLASSIFICAÇÃO DOS MODELOS DE PREVISÃO



- UNIVARIADOS
- CAUSAIS
- MULTIVARIADOS

12

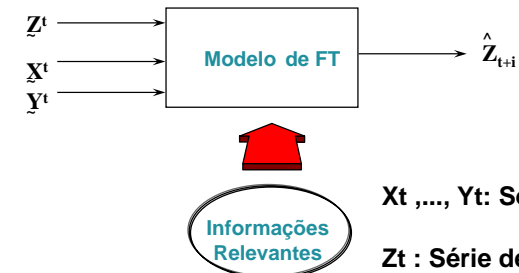
MODELOS UNIVARIADOS



Exemplo: Previsão dos valores futuros de venda de energia (Z_{t+i}) explicada somente pelos valores passados das vendas de energia (Z)

13

MODELOS CAUSAIS (OU DE FUNÇÃO DE TRANSFERÊNCIA)



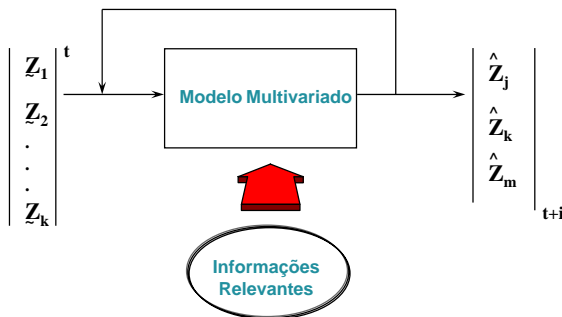
X_t, \dots, Y_t : Séries de Entrada

Z_t : Série de Saída

- Exemplo: Explicar a venda futura de energia Z_{t+i} pelo seu passado Z_t e pelas séries de temperatura (X_t) e renda (Y_t)

14

MODELOS MULTIVARIADOS



- Exemplo: Modelo para a Energia das concessionárias da região Sudeste.

15

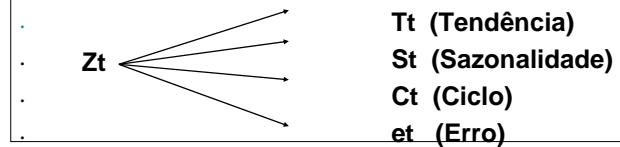
MÉTODOS (MODELOS) UNIVARIADOS



(A) Clássicos:

- Até os Anos 60:

· Decomposição



- Anos 60 (AUTOMAÇÃO)

- Métodos de Amortecimento (Alisamento) Exponencial
- Fáceis de usar porém formulação dos modelos é restritiva (Função do Tempo)

16



- Anos 70 (REVOLUÇÃO)

- Modelos BOX & JENKINS
ARIMA (p,d,q)
SARIMA (p,d,q) X (P,D,Q)s
- Baseados na Teoria Geral dos Sistemas Lineares (CONTROLES)
- Requer maior conhecimento dos usuários; formulação dos modelos é mais geral (para cada série um modelo da família (S)ARIMA)

17



- ANOS 80: COMBINAÇÃO DE PREVISÕES

- Seja $Z_t^{(i)}$ a previsão de Z_{t+1} feita no instante t pelo método M_i , $i=1,2$
 - Então: $Z_t = a \cdot Z_t^{(1)} + (1-a) \cdot Z_t^{(2)}$ é a previsão ótima combinada de Z_{t+1} feita no instante t , onde a é tal que o erro da previsão combinada é mínimo.
 - O resultado se estende facilmente para mais de 2 métodos.
- “Se duas ou mais previsões, baseadas em quaisquer métodos são combinadas, o resultado será sempre melhor que as previsões obtidas pelos métodos individuais”.
- (Granger, 1980)

18



BJ x MAE

- BJ requer conhecimento detalhado da teoria correspondente; MAE pode ser utilizado por usuários menos experientes
- BJ oferece uma classe de modelos para uma série; MAE fornece uma classe restrita de modelos para qualquer série
- Previsões com modelos BJ são, em geral, mais acuradas e robustas

“AN ATTEMPT TO FORCE SQUARE PEGS INTO ROUND HOLES”
Newbold, (1984)

19



- ANOS 90: COMPUTAÇÃO INTENSIVA

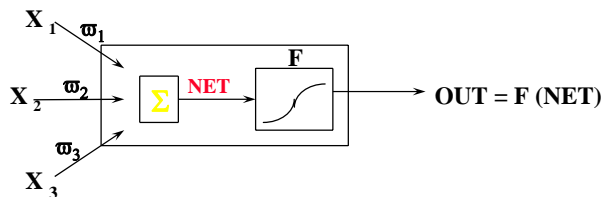
- Uso intenso de técnicas de modelagem e previsão que utilizam a computação intensiva (principalmente via paralelismo).
- Métodos mais usados:
 - BOOTSTRAP
 - NEURAL NETWORKS
 - FUZZY LOGIC

20

MODELOS DE REDES NEURAIS



- Redes Neurais são uma técnica de inteligência computacional inspirada no funcionamento dos neurônios dos seres humanos.
- Usa modelo do neurônio artificial abaixo.

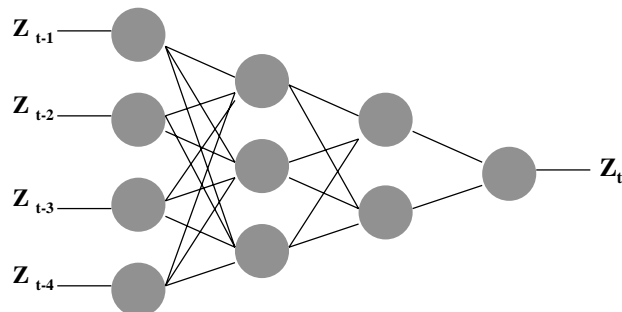


21

REDES NEURAIS



- Exemplo: Rede com 2 camadas escondidas



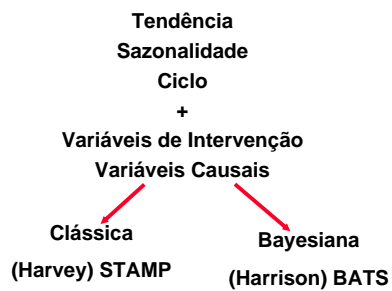
- Modela também séries sazonais. Fornece previsões pontuais (sem intervalos de confiança !!!!)

22

MODELOS CAUSAIS



- Formulação BOX & JENKINS para Modelos Causais ou de Função de Transferência.
- Modelos Estruturais



- Modelos de Regressão Dinâmica (Regressão Linear para Séries Temporais)

23

MODELOS ECONOMÉTRICOS



- Modelos desenvolvidos tomando como base uma dada teoria econômica, tendo como objetivos: controle, verificação de causalidades, teste de cenários e até mesmo previsões.
- A evidência tem mostrado (ver Nelson, 1973) que os Modelos Econométricos não podem competir com os Modelos Específicos de Séries Temporais.

“YOU MUST FALL IN LOVE WITH YOUR DATA, BUT NOT ALWAYS WITH YOUR MODEL”

G. M. Jenkins, (1979)

24

CONCLUSÕES



- Modelar Séries Temporais é uma tarefa que requer conhecimento e, acima de tudo, criatividade → **ESTADO DA ARTE**
- Previsão é um meio e não um fim → **COMPROMETIMENTO ENTRE O ANALISTA DE PREVISÃO E O TOMADOR DE DECISÃO.**

25

CONCLUSÕES



- Não existe um modelo único adequado para qualquer série. Além disto, em ambientes sujeitos a instabilidades e mudanças estruturais, um modelo em operação hoje pode não mais ser adequado amanhã → **MODELOS DINÂMICOS.**

FORECASTING IN BUSINESS IS LIKE SEX IN SOCIETY, WE HAVE TO HAVE IT, WE CANNOT GET ALONG WITHOUT IT; EVERYONE IS DOING IT, ONE WAY OR THE OTHER, BUT NOBODY IS SURE THEY'RE DOING IT THE BEST WAY.

G. W. Plossel

26

PROGRAMAÇÃO



Sessão 1: Estado da Arte

Sessão 2: Previsões no Setor Elétrico
(Curto x Longo Prazo ; Ativo x Reativo)

Sessão 3: Como Gerar Previsões Estatísticas de Mercado

Sessão 4: Ilustração e Demonstração de Sistemas de Previsão

27

PRELIMINARES



- Modelos de Previsões podem ser classificados em 3 macro-categorias

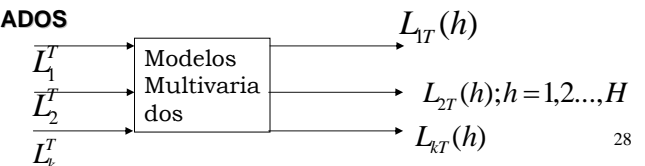
- **UNIVARIADOS**



- **CAUSAIS**



- **MULTIVARIADOS**



28

PRELIMINARES



Previsões necessárias no **SEB**, de acordo com o intervalo de observação dos dados

- **DADOS HORÁRIOS** (ou ½ hora; ¼ hora; 1/5 minutos etc...)
- Modelos de Curtíssimo/Curto prazo; previsões para horizontes de 1 dia, 1 semana até 1 mês à frente
- Variáveis meteorológicas (temperatura, Iuminosidade, vento, etc...) importantíssimas na melhora da acurácia das previsões dos modelos.

29

PRELIMINARES



- **DADOS MENSAIS** (ou semanais)
- Modelos de Médio prazo para horizontes de 1 mês a 1 ano à frente
- Variáveis Meteorológicas não são tão importantes na melhora da acurácia.

Exemplo: PREVCAR (MAPE~1%)
- **DADOS ANUAIS**
- Modelos de longo prazo para horizontes de 1 ano a 5 anos à frente (requerimento do novo SEB para a contratação nos leilões)
- Variáveis Macro-econômicas importantíssimas na construção do modelo e formulação de cenários ⇒ WHAT IF ANALYSIS? ³⁰

PRELIMINARES



PRINCIPAIS MODELOS DE PREVISÃO DE CARGAS (curtíssimo prazo)

- ESTATÍSTICOS
- Holt – Winters com multiplas sazonalidades (resultado recente!)
- Box – Jenkins com mais de 1 componente sazonal
- Modelos de Espaço de Estados (“Forecast Master Plus”; EPRI/BFS)
- Decomposição de Gupta
- Regressão Dinâmica/Pontos Cardinais (National Grid)
- “INTELIGENTES”
- Redes Neurais
- Modelos Fuzzy
- Modelos Neuro-Fuzzy
- MODELOS HÍBRIDOS
- Estatísticos & “Inteligentes”

31

IRREGULARIDADES DOS DADOS



- Dados de carga apresentam vários tipos de descontinuidades, tais como:

Outliers
Mudanças bruscas de nível
Mudanças bruscas de inclinação
Variações abruptas na dispersão
Missing Values



FILTROS DE CORREÇÃO para produzir séries “modeláveis”

32

IRREGULARIDADES DOS DADOS

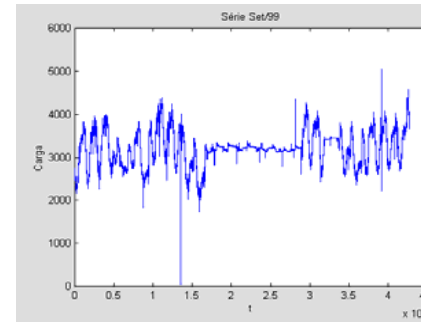


- O mesmo se aplica para os dados de temperatura (especialmete os longos “strings” de “missing values”), além da indisponibilidade dos mesmos (histórico disponível em geral se referem a medições horárias da INFRAERO nos principais aeroportos comerciais)

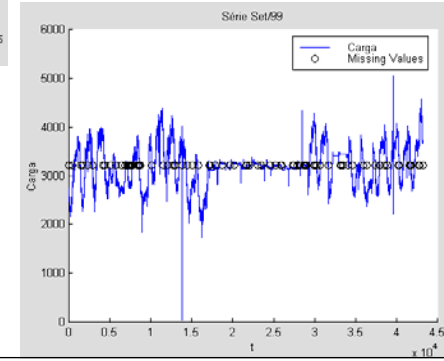


FILTROS DE CORREÇÃO para obtenção de séries “modeláveis” e **ALGORÍTMOS DE OTIMIZAÇÃO** para a criação de “proxys” de temperatura para a previsão da carga de uma dada área.

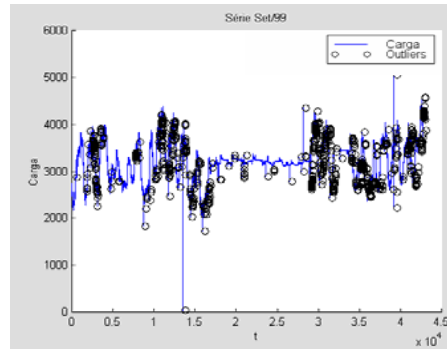
Ilustração: Dados minuto – a minuto de uma concessionária



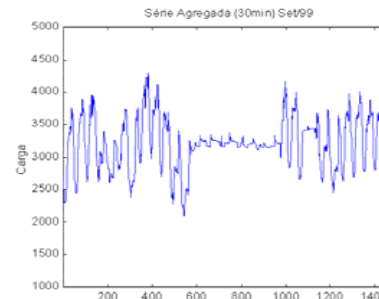
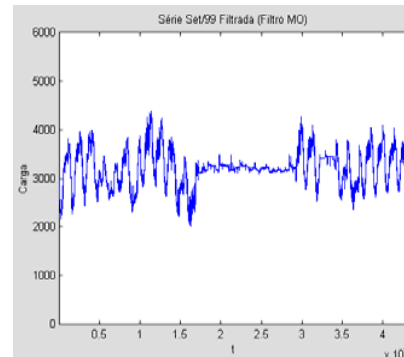
Declarando os Missing Values



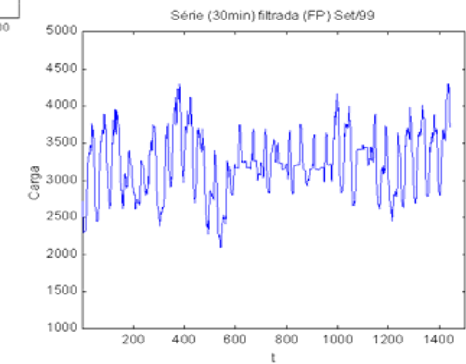
Light
Série: Março 99
Min. a Min.



Interpolando



Filtro de Padrões



Light
Série : Setembro 99
Agregada 30min. -
30min.

MODELOS ESTATÍSTICOS (Mais Usados)



1) HOLT – WINTERS MULTIPLAS SAZONALIDADES

- HW convencional (dados mensais)

$$Z_t = (a_1 + a_2 t) \cdot \rho_{m(t)} + \epsilon_t$$

$\rho_{m(t)}$ Sazonalidade (fatores) mensais, m(t): mês correspondente ao instante t

- HW para CARGA HORÁRIA

$$L_t = (a_1 + a_2 t) \cdot \rho_{1d(t)} \cdot \rho_{2s(t)} + \epsilon_t$$

$\rho_{1d(t)}$: "Sazonalidade" diária; d(t) hora do dia correspondente à hora t (1,...,24)

$\rho_{2s(t)}$: "Sazonalidade" semanal; s(t) hora no ciclo semanal correspondente ao instante t. (1,...,168) 37

MODELOS ESTATÍSTICOS (Mais Usados)



2) BJ com multiplas sazonalidades

- BJ convencional (dados mensais) SARIMA $(p, d, q) \cdot (P, D, Q)_{12}$

$$\Phi(B^{12}) \cdot \Phi(B) \nabla^d \nabla_{12}^d Z_t = \Theta(B^{12}) \cdot \Theta(B) a_t$$

- BJ para CARGA HORÁRIA:

"SARIMA" $(p, d, q) \cdot (P_1, D_1, Q_1)_{24} \cdot (P_2, D_2, Q_2)_{168}$

$$\Phi(B^{168}) \cdot \Phi(B^{24}) \cdot \Phi(B) \cdot \nabla^d \cdot \nabla_{168}^{D_1} \nabla_{24}^{D_2} L_t = \Theta(B^{168}) \cdot \Theta(B^{24}) \cdot \Theta(B) a_t$$

38

MODELOS ESTATÍSTICOS (Mais Usados)



3) DECOMPOSIÇÃO DE GUPTA

$$L(d, h) = P(d, h) + \epsilon(d, h)$$

Componente Standart Componente estocástica

$$P(d, h) = B(d, h) + S(d, h)$$

Componente Básica Componente Semanal(efeito do dia da semana)

Componente básica : obtida por um procedimento de médias móveis de tamanho 7 para cada hora

Componente semanal: obtida por um filtro de amortecimento exponencial para cada hora.

Componente estocástica: assume uma estrutura AR(1) multivariada

Obs: Permite a inclusão de variáveis meteorológicas

39

MODELOS ESTATÍSTICOS (Mais Usados)



4) MODELOS DE REGRESSÃO DINÂMICAS

- Modelos lineares que admitem "dinâmica" (estrutura de defasamentos) na série de carga e nas causais (temperatura)

Exemplo: Hora1 Granger Et alli IJF(1997)

$$\begin{aligned} \text{HOUR1} = & 1678.9 + 40.6\text{YEAR} + 32.4\text{DEC} + 0.5\text{FEB} \\ & + 71.7\text{MAR} - 193.9\text{MONDAY} \\ & + 223.9\text{DAYAFTERHOLIDAY} - 2.86\text{TEMP} \\ & + 0.39\text{TEMP}^2 \\ & - 7.68\text{SEVENDAYAVGMIDTEMP} \\ & - 3.62\text{YESTERDAYMAXTEMP} \\ & + 0.08\text{LOAD8A.M.} \\ & + 0.07\text{MONDAY*LOAD8A.M.} \\ & - 0.10\text{DAYAFTERHOLIDAY*LOAD8A.M.} \\ & + 0.52\epsilon[t-1] + 0.15\epsilon[t-2] + 0.07\epsilon[t-3] \\ & + 0.14\epsilon[t-5] \end{aligned} \quad (2)$$

- Taylor (98) propõe estrutura similar para modelar os "turning points" (pontos cardinais) da curva de carga diaria (dados 1/2 horas; NGC)

40

MODELOS ESTATÍSTICOS (Mais Usados)

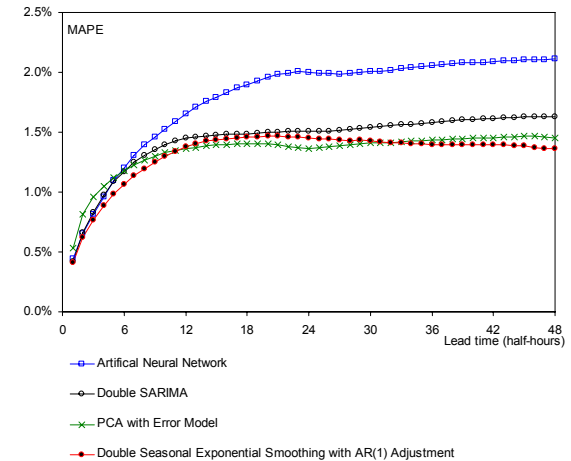


Ilustração: A Comparison of Univariate Time Series Methods for Predicting Electricity Demand up to a Day Ahead
Taylor, J, Menezes, L.M., McSharry, P. E. Annals of the ISF (2004)

- Compares and evaluates 4 univariate methods.
- Uses intra-day data (30 weeks) for the state of Rio in Brazil (hourly) and for England and Wales (half-hourly).
 - Rio: 3,360 observations for estimation and 1,680 for evaluation.
 - England and Wales: 6,720 observations for estimation and 3,360 for evaluation.

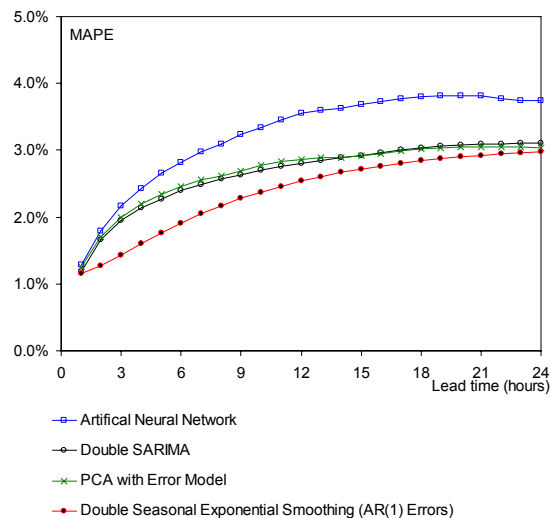
41

MODELOS ESTATÍSTICOS (Mais Usados)



42

MODELOS ESTATÍSTICOS (Mais Usados)



43

MODELOS HÍBRIDOS



- Utilização conjunta de técnicas estatísticas e “inteligentes” na formulação de modelos de previsão de carga horária de curtíssimo prazo (dados Rio)

- Ilustração Souza & Lourenço, (1998)

Redes Neurais (Classificação de padrões)



Decomposição de Gupta (Modelagem univariada dos padrões)



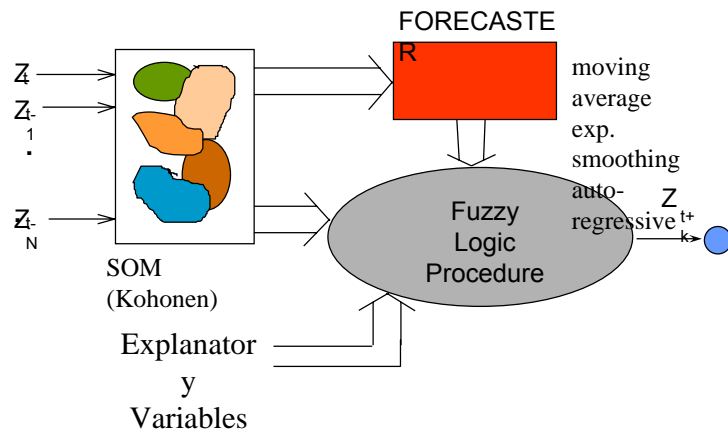
Lógica Nebulosa (Introdução do efeito da temperatura)

44

MODELOS HÍBRIDOS



- Dados de carga e temperatura horários de 96 e 99



45

MODELOS HÍBRIDOS



- **Passo 1:** Classificação dos perfis pelo algoritmo de Kohonen (50n)

1 Monday - Friday	2 Monday - Friday	3 Monday - Friday	1 SUMMER	2 SUMMER	3 SPRING
4 Saturdays	5 Monday - Friday	6 Monday - Friday	4	5 AUTUMN	6 WINTER
7 Saturdays	8 Saturdays	9 Monday - Friday	7	8	9 WINTER
10 Sundays	11 Sundays	12 Saturdays	10	11	12

- **Passo 2:** Decomposição de Gupta para cada grupo

$$L^r(d, h) = P^r(d, h) + \epsilon(d, h)$$

$$P^r(d, h) = B(d, h) + S^r(d, h)$$

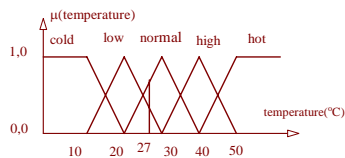
r: grupo (ou classe) identificado no passo 1.

46

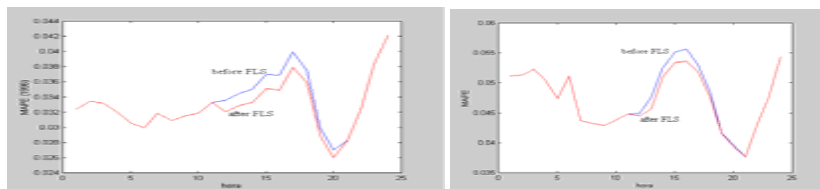
MODELOS HÍBRIDOS



- **Passo 3:** Inferência nebulosa para verificar a variação na carga produzida por variação na temperatura



Resultados



PREVISÃO CARGA REATIVA



- Dados horários de carga reativa apresentam os mesmos tipos de irregularidades apontados para o caso de carga ativa (geralmente em maior intensidade)
- Apresentam também (dependendo da composição da carga) descontinuidades causadas pela entrada/saída de compensadores.
- Possibilidades de Modelagem
- Diretamente a partir do histórico horário de carga reativa (e/ou variáveis meteorológicas) através dos modelos descritos acima.
- Utilizando a forte correlação carga ativa x carga reativa.

48

PREVISÃO CARGA REATIVA



- Modelo baseado em Cinvalar & Grainger (1988)

$$y_t = \tan(\theta)x_t - (c_0 + c_t) + u_t$$

- yt: carga reativa no instante t
- xt: carga ativa no instante t
- c0: carga reativa devido à capacitores fixos (shunt)
- ct: carga reativa devido à capacitores variantes no tempo t (switched)
- tan(θ): tangente do ângulo entre as cargas ativa e reativa
- ut: ruído aleatório.

49

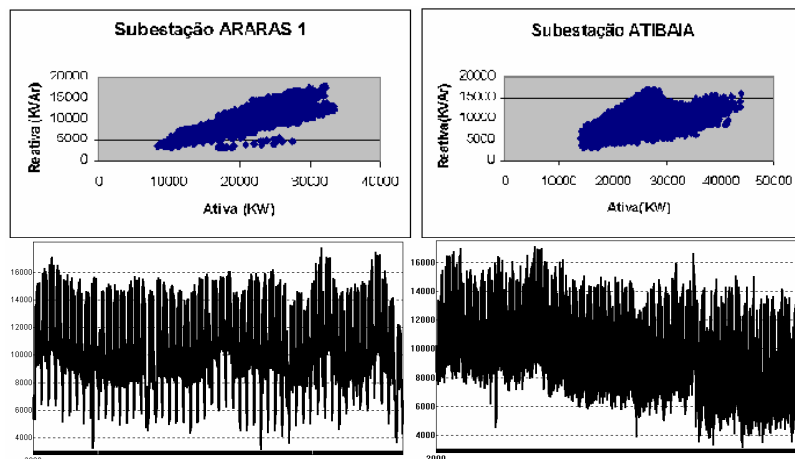
PREVISÃO CARGA REATIVA



- Aplicação: Previsão de Reativo por Subestação
Souza R. C. & Christo, E., (2004)
- Os dados em estudo são as cargas ativa e reativa das subestações ARARAS1 e ATIBAIA da concessionária ELEKTRO de Campinas – SP
- Os dados são horários e o ano escolhido para pesquisa é 2000

50

PREVISÃO CARGA REATIVA



PREVISÃO CARGA REATIVA



- “INVERNO”

$$\begin{cases} Y_t = \beta_0 + \beta_1 X_t + \beta_2 X_{t-1} + \beta_3 Y_{t-1} + \beta_4 Y_{t-168} + u_t \\ u_t - ru_{t-1} = \varepsilon_t \end{cases}$$

- Y_t : carga reativa em PU na hora t
- X_t : carga ativa em PU na hora t
- X_{t-1} : carga ativa em PU na hora $t-1$
- Y_{t-1} : carga reativa em PU na hora $t-1$
- Y_{t-168} : carga reativa em PU na hora $t-168$
- β 's: coeficientes a serem estimados no modelo
- ε_t : ruído aleatório associado ao modelo, onde se supõem que são independentes e identicamente distribuídos com densidade $N(0, \sigma^2)$.

52

PREVISÃO CARGA REATIVA



- “VERÃO”

$$\begin{cases} Y_t = \beta_0 + \beta_1 X_t + \beta_2 Y_{t-1} + \beta_3 Y_{t-168} + u_t \\ u_t - ru_t = \varepsilon_t \end{cases}$$

- Yt: carga reativa em PU na hora t
- Xt: carga ativa em PU na hora t
- Yt-1: carga reativa em PU na hora t-1
- Yt-168: carga reativa em PU na hora t-168
- β's: coeficientes a serem estimados no modelo
- εt: ruído aleatório associado ao modelo, onde se supõem que são independentes e identicamente distribuídos com densidade N(0,σ2).

53

MAPE Subestação ARARAS - 2000						
Hora	Inverno			Verão		
	Prop	RD	RN	Prop	RD	RN
1	3.22	2.29	2.01	8.95	4.72	2.71
2	1.16	1.59	1.91	2.02	14.19	2.14
3	1.45	2.38	1.65	1.54	6.60	1.57
4	1.30	1.44	1.31	1.99	6.40	2.11
5	1.81	2.21	1.72	1.68	4.60	2.31
6	1.36	1.83	2.16	2.66	11.31	2.60
7	1.70	1.29	1.60	3.31	3.66	6.39
8	0.90	0.87	1.22	1.20	4.19	1.21
9	0.54	0.71	0.98	1.16	4.94	2.12
10	0.88	0.89	0.97	1.39	2.68	1.77
11	1.27	1.41	1.94	1.67	2.99	1.99
12	1.38	1.34	2.09	2.30	3.83	1.30
13	0.51	0.44	1.55	2.82	3.74	2.22
14	1.11	1.36	1.05	1.92	2.69	2.27
15	0.82	1.19	1.12	1.74	5.03	2.60
16	0.92	0.80	1.18	2.94	2.79	2.63
17	1.52	0.83	1.21	1.90	4.71	3.21
18	1.08	1.51	0.99	3.00	2.18	2.08
19	1.35	0.85	1.29	2.00	2.39	2.47
20	0.96	1.24	1.76	3.27	3.10	2.98
21	1.14	1.13	1.24	1.14	2.67	1.16
22	1.24	1.42	1.21	1.58	2.76	1.87
23	1.06	0.95	2.24	2.09	2.14	2.73
24	1.05	1.30	1.54	2.79	4.64	4.23
Média	1.24	1.30	1.50	2.30	4.50	2.44
Carga Leve	1.55	1.69	1.68	2.90	6.70	2.81
Carga Média	0.99	0.98	1.33	2.08	3.56	2.22
Carga Pesada	1.16	1.23	1.30	2.02	2.61	2.24



54

MAPE Subestação ATIBAIA - ano 2000						
Hora	Inverno			Verão		
	Prop	RD	RN	Prop	RD	RN
1	7.89	14.91	11.21	8.73	6.28	9.48
2	3.72	5.11	9.42	3.18	10.70	8.10
3	2.14	4.30	25.71	3.18	2.43	3.32
4	2.24	5.85	5.20	2.24	1.92	2.34
5	3.67	7.83	15.58	3.63	3.88	4.35
6	1.91	9.02	18.94	4.44	7.83	5.82
7	3.20	3.83	19.51	5.23	17.89	8.58
8	3.76	5.24	13.47	3.19	5.00	3.06
9	3.89	4.35	4.48	1.97	9.49	2.47
10	1.53	2.28	5.37	2.15	7.81	19.60
11	1.96	1.59	4.31	1.18	6.40	1.75
12	1.38	1.49	3.05	1.46	7.15	2.75
13	1.06	0.88	3.03	3.41	1.76	3.34
14	2.50	3.14	4.64	1.09	3.47	2.48
15	1.64	2.40	3.02	2.05	5.25	4.30
16	1.37	1.08	2.82	1.85	3.43	2.30
17	2.59	2.37	2.98	2.08	8.68	3.76
18	2.00	2.11	11.77	2.39	3.24	3.47
19	3.22	5.18	3.50	1.92	2.41	2.53
20	2.04	3.52	6.06	6.16	3.96	5.62
21	3.30	2.55	7.65	1.96	2.74	4.36
22	1.63	1.74	4.80	2.68	3.01	2.78
23	4.44	5.09	7.64	1.84	1.86	5.29
24	3.69	4.73	8.66	3.20	3.24	4.51
Média	2.78	4.16	8.45	2.97	5.41	4.81
Carga Leve	3.65	6.74	13.54	3.96	6.22	5.76
Carga Média	2.17	2.48	4.72	1.96	5.67	4.62
Carga Pesada	2.44	3.02	6.76	2.91	2.80	4.12



55

PREVISÃO CARGA REATIVA



- Os resultados mostram que o modelo proposto superou, de maneira geral, os outros.
- O erro percentual absoluto médio (MAPE) do modelo proposto foi menor que os demais modelos em mais de 80% das horas.

56

CONCLUSÕES



- *O que usar como previsão de temperatura nos modelos causais (Temperatura-carga)? As obtidas exógenamente (Institutos Oficiais) ou geradas por modelos de séries temporais específicos?*
- *Modelos Híbridos parecem produzir formulações claras, mais precisas e com fácil entendimento por parte dos usuários.*
- *Sempre que possível, utilizar as formulações univariadas (particularmente o HW com múltiplas sazonalidades) que fornecem previsões automáticas e que servem como benchmark*

57

PROGRAMAÇÃO



Sessão 1: Estado da Arte

Sessão 2: Previsões no Setor Elétrico
(Curto x Longo Prazo ; Ativo x Reativo)

Sessão 3: Como Gerar Previsões Estatísticas de Mercado

Sessão 4: Ilustração e Demonstração de Sistemas de Previsão

58

Sumário



1) Modelos de previsão

- Introdução
- Variáveis utilizadas na modelagem
- Modelos
- Conclusões
- Modelos selecionados para implementação no SPCA

2) SPCA: Sistema de Previsão de Cargas

3) PPH (Pesquisa de Posses e Hábitos)

59

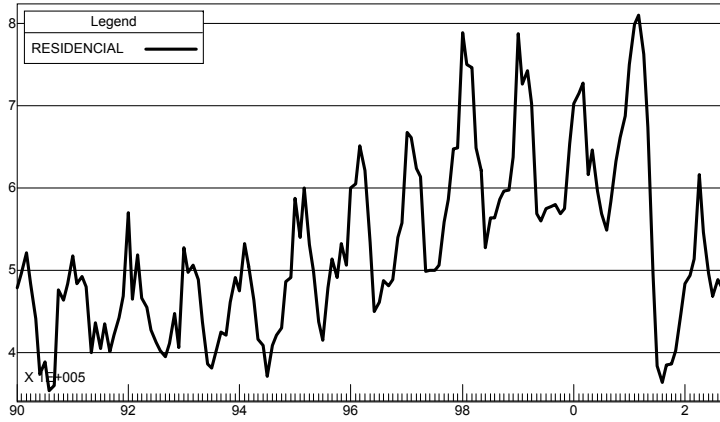
Introdução



- A elaboração de modelos de previsão causais para séries de **energia faturada da Light** abrangeu os **seguintes setores**:
 - **Residencial** Baixa Tensão (BT)
 - **Industrial** Baixa e Média Tensão (BT e MT)
 - **Comercial** Baixa e Média Tensão (BT e MT)
- A técnica de **modelagem** utilizada foi a **bottom-up** através de modelos de **Regressão Dinâmica**, isto é, a partir de um conjunto de **variáveis selecionadas previamente**, parte-se de um **modelo inicial mais simples e procede-se incrementando-o aos poucos**.

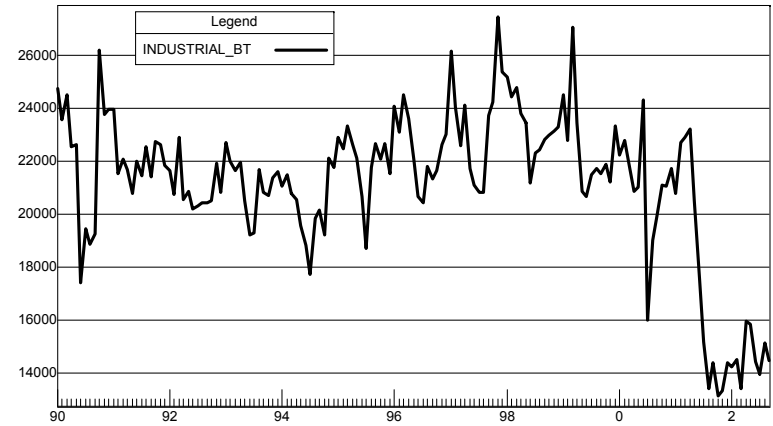
60

Residencial



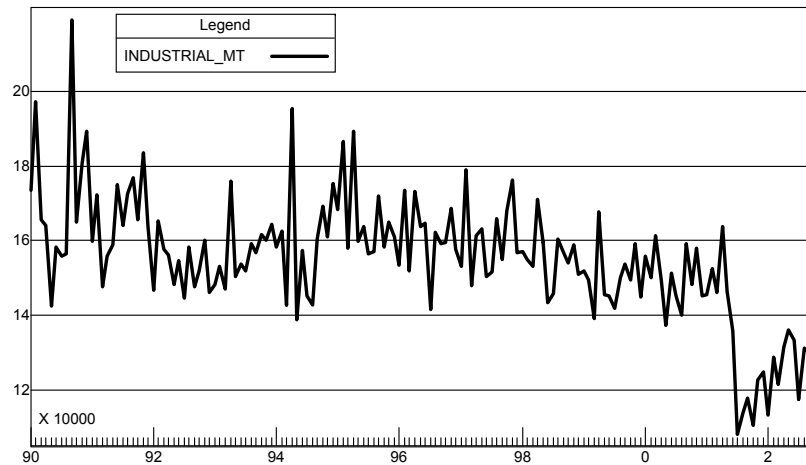
61

Industrial BT



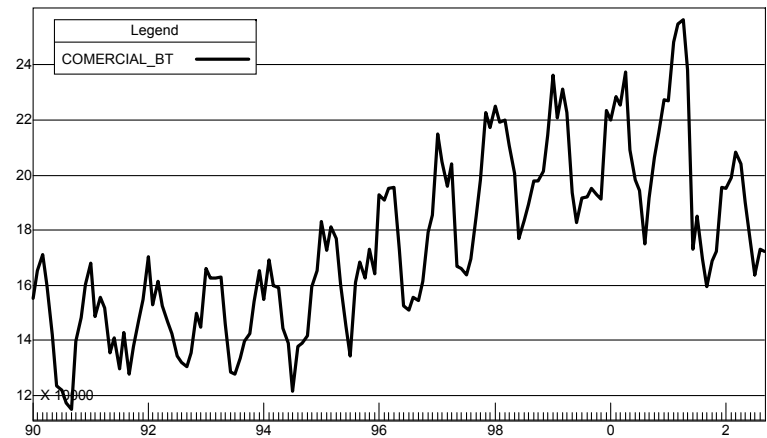
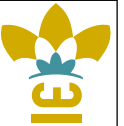
62

Industrial MT



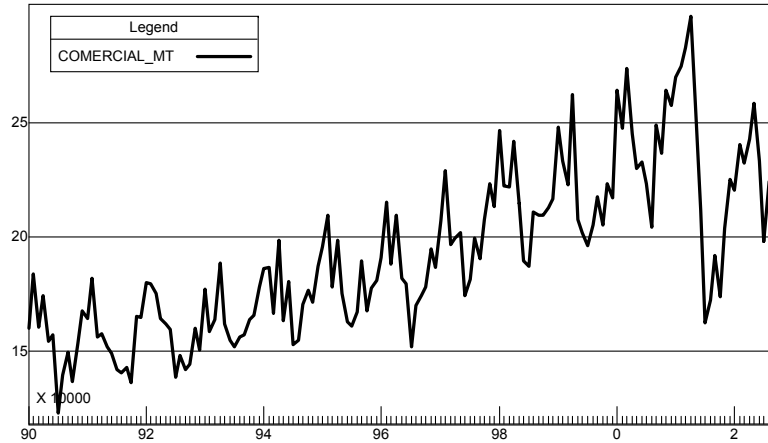
63

Comercial BT



64

Comercial MT



65

Variáveis utilizadas na modelagem



- Foram selecionadas *a priori* :
 - 22 variáveis para o setor **Residencial**
 - 21 para o setor **Industrial**
 - 18 para o setor **Comercial**
- Algumas variáveis são comuns aos três setores, entretanto existem também variáveis específicas para cada setor (e quando possível separadas por BT e MT).

66

Variáveis utilizadas na modelagem



- Para o **setor residencial**, pode-se notar que as variáveis refletem (**variáveis proxy**) “fatores” como:
 - nível de preços na economia;
 - características do faturamento da Light (como número de contas faturadas, tarifas, número de dias de faturamento, entre outras);
 - estrutura de renda dos consumidores (participação dos gastos com ar condicionado, refrigeradores, eletrodomésticos, energia elétrica no orçamento das famílias, rendimento médio real);
 - atividade econômica (população ocupada, índice de emprego, taxa de desemprego, taxa de utilização da capacidade instalada, índices de produção física);
 - variáveis de temperatura.
 - Além das variáveis listadas, foram utilizadas também variáveis de intervenção (*dummy*) para captar alguns eventos como o racionamento, mudança de nível pós-acionamento, outliers (pontos discrepantes), sazonalidade, entre outros

67

Variáveis utilizadas na modelagem



- Para os setores industrial e comercial, os mesmos fatores foram testados na modelagem.
- Entretanto, como já mencionado, utilizou-se variáveis setorializadas como, por exemplo, População Ocupada na Indústria e no Comércio, Tarifas e Número de contas por setor (separados ainda em BT e MT).

68

Setor Residencial



Variáveis	Sigla
Índice de Preços de Condicionadores de Ar	IPAC
Índice de Preços de Eletrodomésticos e Equipamentos	IPED
Índice de Preços de Refrigeradores	IPRF
Variação do Índice Nacional de Preços ao Consumidor para a Região Metropolitana	INPC
Número de Contas Faturadas da classe residencial	NCCR
Tarifa Residencial BT (Reais com correção monetária de ago-02)	TAR_RESBT
Número de dias no calendário	CALE
Número de Sábados, domingos e feriados	SDFE
Número de dias úteis	NDUT
Número de dias de faturamento BT	FATB
Participação dos Gastos com Condicionadores de Ar das famílias entre 1 e 8 M	PGAC
Participação dos Gastos com Eletrodomésticos das famílias entre 1 e 8 Mínimo	PGED
Participação dos Gastos com Energia Elétrica no Orçamento das famílias entre	PGEE
Participação dos Gastos com Refrigeradores das famílias entre 1 e 8 Mínimos	PGRF
Rendimento Médio Real do Pessoal Ocupado na RM do RJ	RMR
População ocupada de 15 anos e mais (Habitante) RM do RJ	PORJ
Índice de Emprego na Indústria Fluminense	IEIF
Taxa de Desemprego Aberto na RM do RJ	TDARJ
Taxa de Utilização da Capacidade Instalada da Indústria	TUCI
Temperatura Média dentre as Temperaturas Máximas Diárias registradas (RM)	TMMX
Temperatura Média dentre as Temperaturas Mínimas Diárias registradas (RM)	TMMN
Temperatura Média na RM do RJ	TMRM ⁶⁹

Setor Industrial



Variáveis	Sigla
Índice de Preços de Condicionadores de Ar	IPAC
Índice de Preços de Eletrodomésticos e Equipamentos	IPED
Índice de Preços de Refrigeradores	IPRF
Variação do Índice Nacional de Preços ao Consumidor para a RM do RJ	INPC
Número de Contas Faturadas da Classe Industrial: BT e MT	NCCI_BT e NCCI_MT
Tarifa Industrial (Reais com correção monetária de ago-02): BT e MT	TAR_INDBT e TAR_INDMT
Número de dias no calendário	CALE
Número de sábados, domingos e feriados	SDFE
Número de dias úteis	NDUT
Número de dias de faturamento: BT e AT	FATB e FATA
População Ocupada na Região Metropolitana do Rio de Janeiro - Indústria	POI
Rendimento Médio Real do Pessoal Ocupado na Região Metropolitana do RJ	RMR
Taxa de Desemprego Aberto na Região Metropolitana do RJ	TDARJ
Índice de Emprego na Indústria Fluminense	IEIF
Taxa de Utilização da Capacidade Instalada da Indústria	TUCI
Índice da Produção Física da Indústria Geral	PFI_IG
Índice da Produção Física da Indústria de Extrativa Mineral	PFI_IEM
Índice da Produção Física da Indústria de Transformação	PFI_IT
Temperatura Média dentre as Temperaturas Máximas Diárias registradas (RM)	TMMX
Temperatura Média dentre as Temperaturas Mínimas Diárias registradas (RM)	TMMN
Temperatura Média na Região Metropolitana do RJ	TMRM

70

Setor Comercial



Variáveis	Sigla
Índice de Preços de Condicionadores de Ar	IPAC
Índice de Preços de Eletrodomésticos e Equipamentos	IPED
Índice de Preços de Refrigeradores	IPRF
Variação do Índice Nacional de Preços ao Consumidor para a RM do RJ	INPC
Número de Contas Faturadas da Classe Comercial: BT e MT	NCCC_BT e NCCC_MT
Tarifa Comercial (Reais com correção monetária de ago-02): BT e MT	TAR_COMBT e TAR_COMMT
Número de dias no calendário	CALE
Número de sábados, domingos e feriados	SDFE
Número de dias úteis	NDUT
Número de dias de faturamento: BT e AT	FATB e FATA
População Ocupada na Região Metropolitana do RJ - Comércio	POC
Rendimento Médio Real do Pessoal Ocupado na Região Metropolitana do RJ	RMR
Taxa de Desemprego Aberto na Região Metropolitana do RJ	TDARJ
Índice de Emprego na Indústria Fluminense	IEIF
Taxa de Utilização da Capacidade Instalada da Indústria	TUCI
Temperatura Média dentre as Temperaturas Máximas Diárias registradas	TMMX
Temperatura Média dentre as Temperaturas Mínimas Diárias registradas	TMMN
Temperatura Média na Região Metropolitana do RJ	TMRM

71

Variáveis utilizadas na modelagem



- A variável a ser explicada (energia faturada) começa (para todos os três setores) em janeiro de 1990. Entretanto, uma característica importante a respeito das variáveis causais é que nem todas coincidem em relação ao período de dados.
- Isto é, algumas estão disponíveis a partir de janeiro de 1990, outras iniciam-se em janeiro de 1991, outras começam a partir de janeiro de 1992, outras a partir de janeiro de 1994 e outras ainda a partir de agosto de 1999.
- Isto significa que a metodologia *bottom-up* foi ainda incrementada com análise para samples (períodos amostrais) diferentes, ou seja, para cada setor, foram usados 5 períodos com dados iniciando-se em: 1990, 1991, 1992, 1994 e 1999 (agosto) gerando um total de 25 modelos (não necessariamente com estruturas diferentes).

72

Variáveis utilizadas na modelagem



Ano	Quantidade de variáveis		
	Residencial	Industrial	Comercial
1990	11	14	11
1991	17	17	14
1992	19	19	16
1994	20	20	17
1999	22	21	18

73

Modelos



- Atenta-se que os **modelos estimados** tiveram a **forma LOG-LOG**, ou seja trabalhou-se com as **variáveis em escala logarítmica** de forma que os **coeficientes de regressão estimados** já refletissem o **coeficiente de elasticidade**.
- O termo **elasticidade** é empregado para descrever as características de uma **relação entre duas variáveis**.
- Neste sentido, a elasticidade é **definida como o limite da razão entre a variação proporcional em uma variável em relação à variação proporcional em outra variável** (isto é, para uma variação de 1% em determinada variável X, objetiva-se saber qual a variação percentual em uma variável Y).

74

Modelos



- A seguir, são apresentadas as **estruturas dos modelos** bem como **algumas estatísticas de desempenho** como:
 - **MAPE (Mean Absolute Percentual Error)** : Calculado através da diferença entre valores estimados e reais. Equivale às previsões um passo-à-frente (no caso, para o mês seguinte).
 - **R² ajustado (coeficiente de explicação)**: indica o quanto da variação total dos dados (série dependente, neste caso, a energia faturada) é explicada pelo modelo. Calculado através da comparação do erro do modelo e a variação dos dados da série dependente em torno de sua média. Varia entre 0 e 100%.
 - **GMRAE (Geometric Mean Relative Absolute Error)**: Compara erro do modelo em questão com o erro do modelo ingênuo (aquele que usa como previsão o último dado disponível). É desejável que seja igual ou menor do que 1.

75

Modelos



- Foi feita uma **análise dos erros in-sample e out-of-sample**.
- Estratégia de análise que **consiste em guardar parte dos dados para testar o poder de generalização do modelo**.
- Para **análise in-sample** foram usados os **dados de janeiro de 1990 (1991, 1992, 1994 e agosto de 1999) a agosto de 2002**. Isto significa que, **para validar o modelo foi usado apenas o mês de setembro de 2002 (out-of-sample)**.
- Portanto, neste caso, tanto o MAPE quanto GMRAE **out-of-sample** na verdade não se referem a uma média e sim ao erro percentual (caso do MAPE) e comparação com erro do modelo ingênuo (caso do GMRAE).

76

Modelos: Residencial



Modelo	Período	Estrutura	MAPE (%) IN-SAMPLE	R ² ajustado(%)	MAPE(%) OUT-OF-SAMPLE	GMRAE OUT
1	Jan/1990 a Ago/2002	RESIDENCIAL[-1], TMRM, TMRM[-1], RMR, FATB, FATB[-1]	4,5	90,7	1,0	0,53
2	Jan/1991 a Ago/2002	RESIDENCIAL[-1], RESIDENCIAL[-12], TMRM, RMR, FATB, FATB[-1], RACIONAMENTO	3,6	93,7	0,0	0,004
3	Jan/1992 a Ago/2002	RESIDENCIAL[-1], RESIDENCIAL[-12], TMRM, RMR, FATB, FATB[-1], RACIONAMENTO	3,7	93,3	0,5	0,24
4	Jan/1994 a Ago/2002	RESIDENCIAL[-1], RESIDENCIAL[-12], TMRM, RMR, FATB, FATB[-1], RACIONAMENTO	3,7	92,1	0,6	0,33
5	Ago/1999 A Ago/2002	CTE, RESIDENCIAL[-1], TMRM, TMRM[-1], IEIF	3,9	92,8	1,7	0,86

Nota: variáveis acompanhadas de [-k] significa estrutura de defasagem de ordem k. Exemplo: RESIDENCIAL [-1] = energia faturada no mês anterior.

77

Modelos: Industrial BT



Modelo	Período	Estrutura	MAPE (%) IN-SAMPLE	R ² ajustado(%)	MAPE(%) OUT-OF-SAMPLE	GMRAE OUT
1	Jan/1990 a Ago/2002	INDUSTRIAL_BT[-1], TMRM, PFI_IT, RMR, FATB, RACIONAMENTO, POS_RACIONAMENTO, JUL_2000	3,5	87,9	3,0	0,85
2	Jan/1991 a Ago/2002	CTE, INDUSTRIAL_BT[-1], INDUSTRIAL_BT[-12], TMRM, PFI_IT, RMR, NCCCI_BT, FATB, RACIONAMENTO, JUL_2000	3,4	90,2	4,2	0,91
3	Jan/1992 a Ago/2002	CTE, INDUSTRIAL_BT[-1], INDUSTRIAL_BT[-12], TMRM, PFI_IT, RMR, NCCCI_BT, FATB, RACIONAMENTO, JUL_2000	3,4	90,7	3,7	0,80
4	Jan/1994 a Ago/2002	INDUSTRIAL_BT[-1], TMRM, RMR, FATB, SDFE, RACIONAMENTO, POS_RACIONAMENTO, JUL_2000	3,3	92,8	3,8	0,83
5	Ago/1999 A Ago/2002	CTE, INDUSTRIAL_BT[-1], TMRM, IEIF, JUL_2000	4,5	91,1	3,4	0,74

Notas: CTE = Termo constante; Variáveis como RACIONAMENTO, POS_RACIONAMENTO e JUL_2000 são variáveis de intervenção para modelar eventos (JUL_2000 por exemplo foi um pulso incluído no modelo para captar somente um outlier ocorrido em julho de 2000).

78

Modelos: Industrial MT



Modelo	Período	Estrutura	MAPE (%) IN-SAMPLE	R ² ajustado(%)	MAPE(%) OUT-OF-SAMPLE	GMRAE OUT
1	Jan/1990 a Ago/2002	CTE, INDUSTRIAL_MT[-12], TMRM, PFI_IT[-1], RACIONAMENTO, POS_RACIONAMENTO	3,9	73,5	0,1	0,20
2	Jan/1991 a Ago/2002	INDUSTRIAL_MT[-1], INDUSTRIAL_MT[-12], TMRM, FATA, RACIONAMENTO, POS_RACIONAMENTO	3,7	77,0	0,3	0,45
3	Jan/1992 a Ago/2002	INDUSTRIAL_MT[-1], INDUSTRIAL_MT[-12], TMRM, FATA, RACIONAMENTO, POS_RACIONAMENTO	3,6	79,8	0,6	0,84
4	Jan/1994 a Ago/2002	INDUSTRIAL_MT[-1], INDUSTRIAL_MT[-12], TMRM, TAR_INDMT, TAR_INDMT[-1], RACIONAMENTO[-1], INDUT[-1]	3,0	87,6	0,6	0,91
5	Ago/1999 A Ago/2002	CTE, TMRM, TAR_INDMT, RACIONAMENTO[-1]	3,9	81,2	1,1	1,6

79

Modelos: Comercial BT



Modelo	Período	Estrutura	MAPE (%) IN-SAMPLE	R ² ajustado(%)	MAPE(%) OUT-OF-SAMPLE	GMRAE OUT
1	Jan/1990 a Ago/2002	COMERCIAL_BT[-1], COMERCIAL_BT[-12], TMRM, RMR, FATB, FATB[-1], RACIONAMENTO	3,7	92,0	0,4	0,76
2	Jan/1991 a Ago/2002	COMERCIAL_BT[-1], COMERCIAL_BT[-12], TMRM, RMR, FATB, FATB[-1], RACIONAMENTO	3,6	91,9	0,3	0,55
3	Jan/1992 a Ago/2002	COMERCIAL_BT[-1], COMERCIAL_BT[-12], TMRM, RMR, FATB, FATB[-1], RACIONAMENTO	3,8	90,6	0,2	0,39
4	Jan/1994 a Ago/2002	COMERCIAL_BT[-1], COMERCIAL_BT[-12], TMRM, RMR, FATB, FATB[-1], RACIONAMENTO	3,8	85,1	0,1	0,21
5	Ago/1999 A Ago/2002	COMERCIAL_BT[-1], TMRM, IEIF, ERRO[-1]	3,5	86,9	0,5	1,04

Nota: ERRO[-1] = erro de previsão no mês anterior

80

Modelos: Comercial MT



Modelo	Período	Estrutura	MAPE (%) IN-SAMPLE	R ² ajustado(%)	MAPE(%) OUT-OF-SAMPLE	GMR AE OUT
1	Jan/1990 a Ago/2002	COMERCIAL_MT[-1], COMERCIAL_MT[-12], TMRM, RMR, NCCC_MT, FATA, FATA[-1], CALE, RACIONAMENTO	3,3	94,0	1,1	4,6
2	Jan/1991 a Ago/2002	CTE, COMERCIAL_MT[-1], COMERCIAL_MT[-12], TMRM, RMR, NCCC_MT, FATA, FATA[-1], RACIONAMENTO	3,2	93,4	1,1	4,6
3	Jan/1992 a Ago/2002	COMERCIAL_MT[-1], COMERCIAL_MT[-12], TMRM, RMR, FATA, FATA[-1], RACIONAMENTO	3,7	90,8	4,3	18,6
4	Jan/1994 a Ago/2002	CTE, COMERCIAL_MT[-1], TMRM, RMR, TUCI[-1], FATA, FATA[-1]	3,0	92,7	1,2	5,0
5	Ago/1999 A Ago/2002	COMERCIAL_MT[-1], TMRM, TMRM[-1], RMR, FATA, RACIONAMENTO, PÓS_RACIONAMENTO	2,6	92,8	5,6	24,3

81

Conclusões



- A heurística de modelagem via **Regressão Dinâmica objetiva a parcimônia nos modelos, ou seja, evita-se a superparametrização com a inclusão de muitas variáveis causais de forma desnecessária.**
- O procedimento **bottom-up** faz com que apenas as variáveis relevantes estatisticamente sejam inseridas nos modelos.

82

Conclusões



- Pode-se observar que, **em geral, os modelos apresentaram bom desempenho preditivo (MAPE baixo e R² alto).**
- Apresentaram ainda um **bom poder de generalização (erro out-of-sample próximo ao erro in-sample).**
- **Exceções** são os modelos para **Industrial (MT)** que apresentaram um **poder de explicação um pouco menor que os dos outros setores e o Comercial (MT)** que apresentou **dificuldade de generalização** (entretanto cabe lembrar que o período **out-of-sample** é constituído apenas do dado de setembro de 2002).

83

Conclusões



- Como era esperado, **samples diferentes geram estruturas de modelos diferentes.**
- Das variáveis de temperatura, a que melhor se ajustou foi a **Temperatura Média na Região Metropolitana do RJ (TMRM).**
- Outra variável freqüente nos modelos é o **Rendimento Médio Real do Pessoal Ocupado na Região Metropolitana do Rio de Janeiro (RMR).**
- Das variáveis **proxy** das características de faturamento da Light, uma variável que ajustou-se bem aos modelos foi o **Número de dias de faturamento (FATB e FATA).** O **número de contas faturadas bem como a tarifa cobrada aparecem também em alguns modelos.**

84

Conclusões



- Das **variáveis de atividade econômica**, algumas como o **Índice de Emprego na Indústria Fluminense (IEIF)**, **Índice da Produção Física da Indústria de Transformação (PFI_IT)** e **Taxa de Utilização da Capacidade Instalada na Indústria (TUCI)** ajustaram bem os modelos.
- Em todos os modelos é possível visualizar também a existência de **dinâmica (estruturas de defasagem) tanto da variável dependente (energia faturada) quanto das variáveis independentes (variáveis causais)**.
- Por fim, em quase todos os modelos é possível observar a existência de **variáveis de intervenção (como para o Racionamento e Pós-acionamento)**.

85

Conclusões



- A estratégia é partir de um modelo mais simples para um mais elaborado, observando-se a **coerência estatístico-lógica e o critério da parcimônia**. **As inclusões de variáveis não são feitas ao mesmo tempo mas passo a passo**.
- Isto significa que **pode-se chegar a diversas estruturas para modelos de regressão dinâmica, pois diversos caminhos podem se seguidos o que implica em diferentes modelos correntes sob avaliação**.
- Isto não caracteriza um problema, ao contrário, oferece várias possibilidades para se obter as previsões. **O importante é que os modelos tenham coerência estatística, lógica e erros de previsão aceitáveis**.
- Obviamente, **o modelo a ser adotado precisa ter estas características e ainda ser melhor do que os demais sob comparação**.

86

Conclusões



- Feitas estas considerações, de todos os modelos apresentados para cada setor, **selecionou-se neste momento uma estrutura apenas para realizar as previsões**.
- **Importante salientar que os modelos são dinâmicos também quanto à sua estrutura. Isto se deve ao fato de realidade também ser dinâmica**.
- Com a ocorrência de novos dados é necessário atualizar os parâmetros e verificar se a estrutura encontrada a princípio foi mantida. Obviamente espera-se que a chegada de poucos dados não deve mudar a estrutura do modelo.
- Os parâmetros precisam ser atualizados constantemente (e também não devem sofrer mudanças significativas).

87

Modelos escolhidos para implementação no SPCA



Modelo	Período	Estrutura	MAPE (%)	R ² ajustado(%)
Residencial (BT)	Jan/1991 a Ago/2002	RESIDENCIAL[-1], RESIDENCIAL[-12], TMRM, RMR, FATB, FATB[-1], RACIONAMENTO	3,6	93,7
Industrial (BT)	Jan/1991 a Ago/2002	CTE, INDUSTRIAL_BT[-1], INDUSTRIAL_BT[-12], TMRM, PFI_IT, RMR, NCCC_BT, FATB, RACIONAMENTO, JUL_2000	3,4	90,5
Industrial (MT)	Jan/1994 a Ago/2002	INDUSTRIAL_MT[-1], INDUSTRIAL_MT[-12], TMRM, TAR_INDMT, TAR_INDMT[-1], RACIONAMENTO[-1], NDUT[-1]	3,0	87,8
Comercial (BT)	Jan/1991 a Ago/2002	COMERCIAL_BT[-1], COMERCIAL_BT[-12], TMRM, RMR, FATB, FATB[-1], RACIONAMENTO	3,6	91,9
Comercial (MT)	Jan/1991 a Ago/2002	CTE, COMERCIAL_MT[-1], COMERCIAL_MT[-12], TMRM, RMR, NCCC_MT, FATA, FATA[-1], RACIONAMENTO	3,2	93,4

Nota: Estatísticas apresentadas já incorporando o dado de setembro de 2002 no período utilizado para estimar modelo.

88

REFERÊNCIAS BIBLIOGRÁFICAS



- Granger, W.J. et alli (1997). Short-run forecasts of electricity loads and peaks. IFJ, 12, 161-174.
- Bunn, D.W. & Farmer, E.D.(1985). Comparative Models for Electrical Load Forecasting Wiley, New York.
- Harvey, A. & Koopman, S.J.(1993). Forecasting Hourly Electricity Demand Using Time-Series Splines. Jasa, 88, 1228-1237.
- Souza, R.C. & Lourenço, P.M.(1999). A Neural-Fuzzy Load Forecasting Model-Application to Hourly Load Data. Anais of 19o. ISF, Washington, DC, UK.
- Souza, R.C.; Hippert, H.S.(2001). Um Previsor de Temperatura Horárias Baseado em Modelos Lineares e Redes Neurais. Publicação interna, DEE, PUC-Rio.

89

REFERÊNCIAS BIBLIOGRÁFICAS



- Taylor, J.W.(1999). Using Combined Forecasts with Changing Weights for Electricity Demand Profiling. Anais of XIV ISF, Washington, DC, UK.
- Souza, R.C.; Fariñas, M.S.; Sousa, R. L.(2004). A Methodology to Filter Time Series: Application in Minute-by-minute Electric Load Series. Revista Pesquisa Operacional, v.24, n.3,p.355-371.

Tese de Mestrado

- "Previsão da Potência Reativa da Carga". Por: Elizabeth Cardoso Bezerra, DEE, PUC/RJ (Jul/02).

Tese de Doutorado

- "Um Modelo de Previsão de Curto Prazo de Carga Elétrica Combinando Métodos Estatísticos e Inteligência Computacional". Por: Plutarcho Maravilha Lourenço, DEE, PUC-Rio (Jun/98).

90