

TEXTO PARA DISCUSSÃO/Nº 250

DIRETORIA DE PESQUISA / RJ

**Previsão de Séries
Econômicas com Modelos
Bayesianos Univariados**

Ana Beatriz S. Monteiro

Hélio S. Migon

Ajax R. B. Moreira

ABRIL DE 1992

Instituto de Pesquisa Econômica Aplicada

O Instituto de Pesquisa Econômica Aplicada - IPEA
é uma Fundação vinculada ao Ministério da Economia,
Fazenda e Planejamento

PRESIDENTE

Roberto Brás Matos Macedo

DIRETOR EXECUTIVO

Líscio Fábio de Brasil Camargo

DIRETOR DE ADMINISTRAÇÃO E DESENVOLVIMENTO INSTITUCIONAL

Antônio Emílio Sendim Marques

DIRETOR DE PESQUISA

Ricardo Varsano

DIRETOR DE POLÍTICAS PÚBLICAS

Antonio Carlos da Ressurreição Xavier

TEXTO PARA DISCUSSÃO tem o objetivo de divulgar
resultados de estudos desenvolvidos no IPEA, informando
profissionais especializados e recolhendo sugestões.

Tiragem: 150 exemplares

SERVIÇO EDITORIAL

Brasília - DF:

SBS. Q. 1, Bl. J, Ed. BNDES - 10º andar

CEP 70.076

Rio de Janeiro - RJ:

Av. Presidente Antônio Carlos, 51 - 14º andar

CEP 20.020

SUMÁRIO

1. INTRODUÇÃO
 2. ASPECTOS METODOLÓGICOS
 - 2.1 - Uma Classe de Modelos Lineares Dinâmicos
 - 2.2 - Principais Passos da Modelagem Dinâmica Bayesiana
 - 2.3 - Modelos Auto-regressivos de Segunda Ordem
 - 2.4 - Mudanças Estruturais Indicadas pelo Monitoramento dos Erros de Previsão
 - 2.5 - Reversão Cíclica como um Problema de Decisão
 3. MODELOS PARA O ÍNDICE DE PRODUÇÃO DA INDÚSTRIA GERAL
 - 3.1 - Características dos Modelos e Critérios de Comparação
 - 3.2 - Comparação nas Décadas de 70 e 80
 - 3.3 - Comparação na Década de 80
 - 3.4 - Determinação de Pontos de Reversão Cíclica
 4. PROJEÇÕES DAS SÉRIES ECONÔMICAS
 - 4.1 - Índices de Produção da Indústria
 - 4.2 - Componentes do PIB Real Trimestral
 - 4.3 - Balança Comercial
 - 4.4 - Indicadores de Consumo
 5. CONCLUSÃO
- BIBLIOGRAFIA
-

**PREVISÃO DE SÉRIES ECONÔMICAS COM
MODELOS BAYESIANOS UNIVARIADOS**

Ana Beatriz S. Monteiro*
Hélio S. Migon**
Ajax R. B. Moreira***

*Do IPEA-Rio e do Instituto de Matemática da UFF.

**Do Instituto de Matemática da UFRJ.

***Do IPEA-Rio.

Este trabalho contou com o apoio financeiro do Projeto
IPEA/PNUD/BRA 89/008/GESEP 2721/BR.

SINOPSE

Este artigo descreve a utilização de modelos bayesianos univariados no acompanhamento da conjuntura econômica, explicitando as capacidades descritiva e preditiva destes modelos. Adicionalmente, utilizando a sua estrutura probabilística e seqüencial, foram implementados recursos que monitoram a trajetória das séries identificando mudanças estruturais e indicam a ocorrência de reversão na atividade econômica.



1. INTRODUÇÃO

O acompanhamento da conjuntura econômica pretende interpretar a situação atual e fazer algum tipo de prognóstico para o futuro. Este acompanhamento pode ser realizado, do ponto de vista quantitativo, de uma forma apenas descritiva, observando-se a tendência de um conjunto selecionado de séries, ou utilizando-se um modelo formal que relacione estatística e dinamicamente as séries relevantes. Estes procedimentos não se excluem. O uso de um modelo formal explicita as hipóteses do analista e utiliza a interligação entre as séries econômicas para que se possa fazer uma melhor previsão. Todavia, tal uso pode ser um constrangimento, uma vez que limita o analista a um certo corpo de hipóteses quando, de fato, a explicação dos acontecimentos exige que sejam enfatizados aspectos diferentes ao longo do tempo. Neste artigo vamos nos ater apenas à discussão dos instrumentos estatísticos necessários à abordagem descritiva.

O procedimento descritivo utiliza uma estrutura conceitual para identificar as séries relevantes e interpretar as tendências verificadas em cada série. Uma das vantagens do mesmo é sua capacidade de acompanhar a volatilidade do fato econômico com a agilidade necessária. Isto envolve o uso da informação com a maior atualidade possível e na mesma periodicidade com que as séries são geradas (a maioria mensalmente). Como tais séries freqüentemente têm um padrão sazonal bastante significativo torna-se necessário que este seja isolado para extrair a tendência da série.

Os modelos estruturais foram propostos por Harvey (1981) numa abordagem clássica e por Harrison e Stevens (1976) numa abordagem bayesiana. Ambos postulam que as séries podem ser descritas por componentes como a tendência ou a sazonalidade, e podem substituir outros métodos de decomposição com vantagens. Isto porque, tendo fundamento probabilístico, são capazes de extrapolar a tendência e a própria série e medir a imprecisão correspondente de uma forma rigorosa. No Brasil, Migon e Brasil (1986) e Pereira (1988) utilizaram modelos estruturais respectivamente dentro de uma abordagem bayesiana e clássica sobre a série do produto da indústria geral. Neste artigo propomos o uso da abordagem bayesiana para descrever e prever um conjunto de séries mensais e reais relevantes para a análise da conjuntura. Procuramos, também, mostrar as vantagens da abordagem bayesiana através da comparação do desempenho de alguns modelos ajustados para diversas séries econômicas brasileiras, especialmente a do produto da indústria.

Os modelos estruturais admitem suas componentes como estocásticas permitindo que elas flutuem suavemente ao longo do tempo o que possibilita o ajustamento local às modificações no processo gerador dos dados. A abordagem bayesiana dispõe também de um processo de monitoramento que identifica os períodos em que houve uma mudança estrutural - modificação brusca na trajetória da série. Nestes períodos o algoritmo é alterado no sentido de

torná-lo mais adaptativo às novas condições.

Um tema relevante na literatura econômica é a questão do reconhecimento da reversão da tendência. Supõe-se que a economia tem dois estados, o de recessão e o de crescimento, e a questão consiste em identificar tão cedo quanto possível o momento em que a economia muda de estado. O modelo bayesiano proposto, além de descrever a série e efetuar previsões também será utilizado para derivar uma medida da probabilidade de ocorrência de reversão.

Na próxima Seção é apresentada a abordagem bayesiana, incluindo a modelagem de processos auto-regressivos de segunda ordem, discutindo critérios de monitoramento para determinação de mudanças estruturais e conceituando uma medida de reversão de tendência. Na Seção 3, modelos bayesianos, clássicos e de função de transferência para o índice de produção da indústria geral são comparados. Na Seção 4, mostra-se o desempenho preditivo do modelo dinâmico bayesiano para um conjunto razoavelmente amplo de séries - indicadores do produto da indústria, componentes do produto interno bruto, componentes da balança comercial, indicadores de consumo e emprego.

2. ASPECTOS METODOLÓGICOS

Os modelos dinâmicos bayesianos [West e Harrison (1989)] são descritos por um sistema de equações que representam a distribuição das observações e a evolução dos parâmetros. Nesta seção descreveremos as componentes do modelo básico e os principais aspectos envolvidos nesta abordagem. Apresentaremos também uma representação para a componente cíclica através de um processo auto-regressivo de segunda ordem e discutiremos um procedimento de identificação de mudanças estruturais na série denominado monitoramento bayesiano dos erros de previsão. Finalmente estabeleceremos um critério de decisão para a identificação da ocorrência de reversões, momentos em que ocorrem pontos de máximo ou de mínimo na trajetória da série.

2.1. Uma Classe de Modelos Lineares Dinâmicos

Apresentamos um subconjunto de modelos lineares dinâmicos que representam séries temporais univariadas - observadas em pontos equiespaçados no tempo - e que admitem a inclusão de regressores com parâmetros variando no tempo. Denotando-se por (y_t) a série temporal de interesse em cada instante de tempo, podemos representá-la por duas componentes: nível subjacente (μ_t) e erro observacional (ε_t) :

$$y_t = \mu_t + \varepsilon_t$$

O nível compreende, em geral, três componentes que aproximam localmente a tendência, a sazonalidade e os efeitos de

regressores. A tendência é função polinomial do tempo, enquanto a sazonalidade pode ser descrita por fatores sazonais - forma livre - ou através de harmônicos. Nos modelos aditivos essas componentes se superpõem, ocasião em que a série se decompõe em tendência, sazonalidade, efeito dos regressores e erro observacional. A seguir, descreveremos com detalhes essas componentes.

Tendência

A tendência subjacente à série temporal pode ser representada por uma função estocástica do tempo. Uma aproximação local para a tendência envolvendo apenas termos de primeira ordem da expansão de série de Taylor é apresentada abaixo, onde o termo (μ') representa a primeira derivada da tendência e (ε) a componente estocástica. Esta aproximação será tão melhor quanto for mais suave a função:

$$\mu(t + \delta t) \approx \mu(t) + \delta t \mu'(t) + \varepsilon(t)$$

A função de tendência considerada em momentos equiespaçados no tempo pode ser representada em cada período pelos parâmetros (θ_{1t} , θ_{2t}), satisfazendo as equações recursivas abaixo, onde $(w_1, w_2) \sim N(0, W_t)$. Esses parâmetros correspondem, respectivamente, ao nível presente (μ) e à derivada (μ'):

$$\theta_{1t} = \theta_{1t-1} + \theta_{2t-1} + w_{1t}$$

$$\theta_{2t} = \theta_{2t-1} + w_{2t}$$

A trajetória esperada da tendência pode ser obtida por substituições sucessivas, resultando em:

$$\mu(t+k) \approx \theta_{1t+k} = \theta_{1t} + k \theta_{2t}$$

Sazonalidade

A componente sazonal, de forma análoga, será de natureza estocástica para adaptar-se suavemente às modificações dos dados ao longo do tempo. Se o período sazonal for (s) podemos descrever a sazonalidade através de (s) fatores sazonais. As equações abaixo definem, no caso em que s for igual a 3, a dinâmica dos parâmetros onde θ_{it} é o efeito sazonal do mês $t-i$. Estas equações implicam que os períodos $t+k$ e $t+k+3$ têm os mesmos fatores sazonais caracterizando uma função periódica:

$$\theta_{1t} = \theta_{3t-1}, \quad \theta_{2t} = \theta_{1t-1}, \quad \theta_{3t} = \theta_{2t-1}$$

Quando este fator é composto com uma tendência devemos impor a condição de que sua soma seja nula e interpretá-lo como variação em torno do nível. Vale mencionar que combinações de senos e co-senos podem substituir os efeitos acima quando se desejam formas restritas e mais parcimoniosas de representar a sazonalidade. Uma ampla discussão desses modelos pode ser encontrada em West e Harrison (1989), Seção 8.4.

Regressores

Para cada variável independente (x) teremos um parâmetro que pode variar suavemente no tempo. Este procedimento é útil para se considerarem omissões de variáveis, mudanças estruturais suaves, representação de relações não-lineares através de aproximações locais lineares, etc. O caso de regressão múltipla é obtido pela superposição dos diversos regressores.

Componente Residual

A componente residual, erro observacional, será representada por uma variável aleatória com distribuição normal de média zero e variância descrita pela lei do tipo potência abaixo, onde (c) é o fator de escala e (α) relaciona a variância residual com a tendência da série. Esta lei de variância só se aplica, é claro, a séries com tendência não-negativa. A constante (c) será estimada seqüencialmente a partir dos dados. O caso em que o parâmetro (α) é nulo corresponde à variância constante:

$$\text{Var}(\varepsilon_t) = c \mu_t^\alpha$$

2.2. Principais Passos da Modelagem Dinâmica Bayesiana

As principais operações envolvidas na análise bayesiana são a evolução e a atualização. Na primeira evoluem-se temporalmente as distribuições de probabilidade que descrevem os parâmetros do modelo de forma a se obterem as distribuições *a priori* dos parâmetros, antes da observação dos dados do período correspondente. Após observarmos uma nova realização da série temporal a distribuição *a priori* é combinada com a verossimilhança gerando a distribuição *a posteriori* dos parâmetros. A natureza seqüencial - define-se para cada período - deste método está de acordo com o fato de a publicação dos dados e o uso do modelo ocorrerem também de forma seqüencial ao longo do tempo.

Serão descritas, também, as operações de previsão e suavização. O procedimento de previsão consiste em descrever a trajetória dos parâmetros condicionado a um certo momento passado, e em utilizar a equação de observação que relaciona os parâmetros da série de interesse. A suavização, por outro lado, descreve a evolução dos parâmetros considerando-se todos os dados observados. Um dos usos da suavização diz respeito a análises do tipo "o que teria acontecido se" como em Harrison (1989).

Modelo Dinâmico Geral

Neste texto suporemos que as observações se distribuem segundo uma distribuição de probabilidade normal e que os modelos envolvidos são lineares, ainda que o método suporte generalizações nos dois sentidos.¹ As observações, condicionadas aos parâmetros, serão descritas pela seguinte distribuição de probabilidade:

$$(y_t | \mu_t, \phi_t) \sim N(\mu_t, \phi_t^{-1}), \quad \text{onde } \phi = [\text{Var}(\varepsilon_t)]^{-1}$$

O nível do processo está relacionado com o vetor de parâmetros (θ) através da expressão (1) - equação da observação - onde (F_t) é um vetor de dimensão adequada e os parâmetros de estado do sistema evoluem segundo a equação (2):

$$\mu_t = F_t \theta_t \quad (1)$$

$$\theta_t = G_t \theta_{t-1} + w_t, \quad w_t \sim N(0, W_t / \phi_t) \quad (2)$$

A matriz G_t é composta por blocos que descrevem as diversas componentes. Por exemplo, nos blocos que representam o crescimento linear e a sazonalidade faz-se respectivamente:

$$G_t = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad G_t = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

As informações necessárias, no tempo ($t-1$), para implementar a análise bayesiana estão descritas, respectivamente, pelas distribuições *a posteriori* normal-gama:

$$(\theta_{t-1} | D_{t-1}, \phi_{t-1}) \sim N(m_{t-1}, C_{t-1} / \phi_{t-1})$$

$$(\phi_{t-1} | D_{t-1}) \sim \Gamma(n_{t-1}/2, s_{t-1}/2)$$

Os valores $(m_0; C_0, n_0, s_0)$ que caracterizam as distribuições *a posteriori* dos parâmetros no instante inicial podem ser determinados subjetivamente ou a partir de "priors de referência" como descrito em West e Harrison (1989). O próximo passo da análise será descrever a evolução temporal usando (2).

¹As generalizações para modelos não-normais e não-lineares estão em Migon e Gamerman (1989). Veja as referências ali encontradas.

Evolução Temporal

A expressão (2) descreve, através de G_t , a evolução determinística do modelo e também o acréscimo de incerteza traduzido pela variância da componente aleatória (w). Assim:

$$(\theta_t | D_{t-1}, \phi_t) \sim N(a_t, R_t / \phi_t)$$

onde $a_t = G_t m_{t-1}$ e $R_t = G_t' C_{t-1} G_t + W_t$

Esta distribuição é denominada distribuição *a priori* para θ_t . Por integração é possível obter-se a distribuição *a priori* marginal de $(\theta_t | D_{t-1})$ a qual será uma t-Student.

Atualização

Após observarmos y no período t , combinamos a distribuição *a priori* conjunta de θ e ϕ com a verossimilhança para se obterem as distribuições *a posteriori* de θ e ϕ que são, também, distribuições normal-gama. Estas expressões são obtidas como função do erro de previsão um passo à frente e dos parâmetros da distribuição *a priori* e podem ser encontradas em West e Harrison (1989, Seção 4.5, p. 118-119)]:

$$(\theta_t | \phi_t, D_t) \sim N(m_t, C_t / \phi_t)$$

$$(\phi_t | D_t) \sim \Gamma(n_t / 2, s_t / 2)$$

Previsão

O valor futuro da série é descrito por uma distribuição condicional aos dados passados, denominada distribuição preditiva, a qual, como indicado abaixo, será uma t-Student. Observe que, dispondo-se da distribuição conjunta das previsões futuras, $(y_{t+1}, \dots, y_{t+k})$, pode-se obter a distribuição de funções desses valores futuros. Por exemplo, poderíamos estar interessados na soma das previsões da série ao longo dos próximos trimestres, baseada em dados mensais até o último mês observado:

$$(y_{t+k} | D_t) \sim t\text{-St}(f_t ; q_t)$$

Intervenção e Observações Perdidas

A natureza subjetiva do método bayesiano permite que se combinem informações objetivas com outras de natureza subjetiva. Um exemplo seria a consideração de alterações estruturais esperadas por especialistas para um certo momento futuro. Neste caso, a tendência do processo poderia ser acrescida da variável aleatória δ com média $m(\delta)$ e variância $v(\delta)$ fornecida pelos especialistas, o

que altera o nível e a variabilidade desta componente do processo estocástico. As observações perdidas são consideradas de forma direta, evoluindo o processo - passagem da *priori* para a *posteriori* - sem a incorporação de informações novas. Isto equivale a se manter inalterado o valor esperado dos parâmetros e a se ampliarem as incertezas.

Suavização

Os parâmetros do modelo estrutural dinâmico admitem interpretação e variam ao longo do tempo. Por exemplo, num modelo de crescimento linear com dados mensais, um dos parâmetros representa o crescimento mensal do processo em estudo, outro a componente sazonal, e outro poderia ser o efeito de uma variável exógena. Permitir que estes parâmetros variem no tempo é compreender que os aspectos qualitativos dos modelos são mais estáveis do que os quantitativos. Assim, os parâmetros estarão aptos a acomodar mudanças estruturais que venham a ocorrer. Prosseguindo com este raciocínio, não é difícil conceber a importância da revisão dos parâmetros no instante (t) após observar os novos dados em (t+1). A distribuição dos parâmetros do período t condicionada a dados posteriores a t será chamada de "distribuição suavizada" e denotada pela expressão abaixo, onde m^s é o vetor das médias e C^s a matriz de variância-covariância suavizadas.² Este procedimento permite interpretar as componentes ao longo do tempo e descrever a magnitude dos efeitos de intervenções:

$$(\theta_t | D_T, \phi_t) \sim N(m_t^s, C_t^s / \phi_t)$$

2.3. Modelos Auto-regressivos de Segunda Ordem

Algumas séries econômicas, em princípio, podem ter uma componente cíclica que dê conta de um eventual ciclo, não-sazonal e estocástico, inerente ao processo gerador da série. Um exemplo seria o nível de atividade econômica que teria subjacente um ciclo de negócios. Na literatura, autores como Pereira (1988) incorporaram num modelo estrutural esta componente através de um processo restrito auto-regressivo de segunda ordem, representado por uma combinação de senos e co-senos amortecidos no tempo, o qual corresponde a um processo auto-regressivo de segunda ordem AR(2) com raízes complexas conjugadas.

No âmbito dos modelos bayesianos em que todos os parâmetros são estimados seqüencialmente, a inclusão de uma componente deste tipo resultaria num modelo não-linear, em que seriam estimados o coeficiente de amortecimento e o período do ciclo. O modelo AR(2)

²As expressões de m^s e C^s encontram-se em West e Harrison (1989, p. 121-124).

irrestrito será escrito na estrutura do Modelo Dinâmico Bayesiano, após transformações, e estimado seqüencialmente. Desta forma, os dados poderão indicar a inexistência do ciclo ao longo do tempo, ou a sua flutuação. Um processo AR(2) é definido por qualquer uma das expressões abaixo, onde $\varepsilon \sim N(0, \sigma^2)$ são independentes temporalmente e (ϕ_1, ϕ_2) os parâmetros que satisfazem certas restrições e variam no tempo:

$$y_t = a_t, \quad a_t = \phi_1 a_{t-1} + \phi_2 a_{t-2} + \varepsilon_t \quad (3)$$

Por substituição, obtém-se:

$$y_t = (\phi_1 + \phi_2 B) y_{t-1} + \varepsilon_t$$

O processo será estacionário se as raízes subjacentes aos parâmetros ϕ estiverem fora do círculo unitário. Neste caso, fica caracterizado um processo do tipo AR(2). Incluindo o nível na equação (3) obtemos a expressão abaixo:

$$y_t = \mu_t + (\phi_1 + \phi_2 B) (y_{t-1} - \mu_{t-1}) + \varepsilon_t$$

Denote-se por $a_{(t-k/t-1)} = y_{t-k} - \mu_{t-k}$ para $k=1,2$ as quantidades aleatórias com média $\bar{a}_{t-k} = E(y_{t-k} - \mu_{t-k} | D_{t-1})$ que são, respectivamente, a média *a posteriori* em $t-1$ e a média da distribuição suavizada, a um passo, em $t-2$. Verifica-se que:

$$y_t = \mu_t + \phi_{1t} \bar{a}_{t-1} + \phi_{2t} \bar{a}_{t-2} + \varepsilon_t, \quad \varepsilon_t \sim N(0, \sigma^2)$$

$$\mu_t = \mu_{t-1} + w_{1t}$$

$$\phi_{kt} = \phi_{kt-1} + \omega_{kt}, \quad k = 1, 2 \text{ e } (\omega_1, \omega_2)_t \sim N(0, \tilde{W}_t)$$

O modelo acima é a composição de uma "tendência" e uma regressão nas variáveis $(\bar{a}_{t-1}, \bar{a}_{t-2})$. As principais vantagens dessa representação são, de um lado, a maneira irrestrita de considerar o processo AR(2) e, de outro, a possibilidade de a variância residual ser estimada seqüencialmente. Formas alternativas de representar componentes ARMA na estrutura dos modelos dinâmicos bayesianos podem ser encontradas em West e Harrison (1989).

Apresentamos a seguir quatro exemplos de processo AR(2) gerados numa amostra de 240 observações e com raízes escolhidas convenientemente na região de admissibilidade, correspondendo a: (0.1, 0.9) raízes reais próximas da fronteira de inversibilidade; (1, 0.89) raízes complexas conjugadas; (0.8, 0.16) raiz real com multiplicidade 2 e, finalmente, (-0.9, -0.1) raízes no interior do espaço de parâmetros. Como se pode observar abaixo, a localização

das raízes não parece prejudicar a estimação.

Comparação entre o Parâmetro Original e o Estimado					
Parâmetro ϕ_1		Parâmetro ϕ_2		Parâmetro σ	
Original	Estimada	Original	Estimada	Original	Estimada
0.1	0.0919	0.9	0.8992	5.00	5.02
1.0	1.0202	0.89	0.9202	5.00	4.85
0.8	0.7645	0.16	0.1844	5.00	4.58
-0.9	-0.8889	-0.10	-0.09742	5.00	5.39

2.4. Mudanças Estruturais Indicadas pelo Monitoramento dos Erros de Previsão

Um problema relevante em séries econômicas é a determinação dos momentos de mudanças estruturais e o conseqüente uso da intervenção subjetiva. Em modelos dinâmicos bayesianos é possível introduzir-se um mecanismo de acompanhamento dos erros de previsão para a determinação dos momentos de mudanças estruturais. Este mecanismo de monitoramento é baseado no fator de Bayes e se inspira na piora da capacidade preditiva do modelo em análise. Estas mudanças estruturais certamente prejudicam a descrição dos fenômenos em estudo e, eventualmente, comprometem a capacidade de resposta dos modelos. A forma mais simples de se proceder à intervenção consiste em aumentar a variância da evolução dos parâmetros, o que, intuitivamente, corresponde a desprezar mais fortemente as observações passadas. Formas mais elaboradas envolveriam a modelagem do efeito da mudança estrutural incluindo a quantificação subjetiva desses eventos como, por exemplo, queda no nível de certa magnitude ou mudança de inclinação.

Podemos, intuitivamente, examinar o procedimento de se utilizarem os erros de previsão para diagnosticar a perda de aderência entre o modelo e os dados. Por exemplo, ocorrendo no instante t um erro de previsão padronizado - dividido pelo seu desvio-padrão - superior a dois, é provável que alguma excepcionalidade esteja ocorrendo. Esta pode ser uma observação espúria ou indicação do início de uma mudança estrutural. Julgar o modelo considerando o erro de previsão acumulado até o instante (t) - procedimento habitual de avaliação da capacidade preditiva de um modelo - torna o procedimento menos sensível do que considerar os erros acumulados localmente.

A densidade preditiva $p(y_t | D_{t-1}, M)$ obtida segundo o modelo M serve para avaliar se uma observação futura está de acordo com o modelo especificado. Valores grandes da densidade preditiva indicam a coerência entre o modelo M e os dados. Esta função considera a previsão pontual e também sua precisão. Agregando-se esta função sobre as (s) últimas observações temos:

$$p(y_t \dots y_{t-s} | D_{t-s-1}, M) = p(y_t | D_{t-1}) p(y_{t-1} | D_{t-2}) \dots p(y_{t-s} | D_{t-s-1})$$

a qual é denominada densidade preditiva conjunta para uma janela de (s) períodos. Seu valor é interpretado como a verossimilhança do modelo M e presta-se a comparar modelos. A razão entre as verossimilhanças dos modelos M e M' é chamada de fator de Bayes de janela (s) do modelo M com respeito ao modelo M'. Valores grandes deste fator (> 3, na escala do log) indicam que o modelo M é preferível ao modelo M' :

$$H(t,s) = p(y_t \dots y_{t-s} | D_{t-s-1}, M) / p(y_t \dots y_{t-s} | D_{t-s-1}, M')$$

Se o modelo M' for semelhante ao M, exceto por admitir maior variabilidade dos seus parâmetros, o fator de Bayes pode ser utilizado para identificar a ocorrência de uma mudança estrutural, sinalizada pela preferência pelo modelo M'. Se H(t,1) sinalizar a preferência pelo modelo M' é possível que esteja ocorrendo apenas uma observação espúria. No entanto, se H(t,s) para s>1 for significativo, então o procedimento está indicando a ocorrência de uma mudança estrutural, pois está sendo constatada a preferência sistemática pelo modelo M'.³

2.5. Reversão Cíclica como um Problema de Decisão

A probabilidade de ocorrência de uma reversão pode ser calculada a partir de uma definição clara de reversão e de um modelo probabilístico para as observações passadas,⁴ isto é, utilizando a distribuição preditiva para o cálculo das probabilidades de reversão e uma função de perda para representar as conseqüências de decisões erradas sobre o reconhecimento de uma reversão. Inicialmente definiremos reversão sobre os dados observados e a seguir estenderemos o conceito para a tendência, componente não observável.

Sejam $D_t = (y_1 \dots y_t)$ as informações disponíveis até o período t e consideremos a observação futura Y_{t+1} . Definiremos uma reversão para baixo baseada em (y_{t-1}, y_t, Y_{t+1}) por:

RB : ocorreu uma reversão para baixo : se $y_{t-1} < y_t > Y_{t+1}$

$\overline{\text{RB}}$: não ocorreu reversão para baixo : se $y_{t-1} < y_t < Y_{t+1}$

³O detalhamento do monitoramento que envolve a especificação do tamanho da janela (s) e do modelo alternativo consta em West e Harrison (1989, Seções 2.4 e 2.5).

⁴Ver Zellner (1987) e referências ali mencionadas.

A reversão para cima (RC) pode ser definida de forma análoga e estas definições podem ser estendidas para incluir várias observações passadas e futuras. Para fins de tomada de decisão sob incerteza necessitamos calcular as probabilidades desses eventos:

$$\text{Prob}[\text{RB}_t | D_t] = \begin{cases} \text{Prob}(Y_{t+1} < y_t | D_t) & \text{se } y_{t-1} < y_t \\ 0 & \text{em caso contrário} \end{cases}$$

As séries econômicas podem ter estruturas mais elaboradas, incluindo componentes sazonais, cíclicas e erro observacional, que perturbam a identificação da trajetória de mais longo prazo. Por isso a reversão deve ser caracterizada a partir da tendência ou nível corrente da série. Assim:

RB : ocorreu uma reversão para baixo : se $\mu_{t-1} < \mu_t > \mu_{t+1}$

$\bar{\text{RB}}$: não ocorreu reversão para baixo : se $\mu_{t-1} < \mu_t < \mu_{t+1}$

$$\text{Prob}[\text{RB}_t | D_t] = \text{Prob}[(\mu_{t-1} < \mu_t) \text{ e } (\mu_{t+1} < \mu_t) | D_t] = \\ \text{Prob}(\mu_{t-1} < \mu_{t+1} < \mu_t | D_t) + \text{Prob}(\mu_{t+1} < \mu_{t-1} < \mu_t | D_t)$$

As quantidades denotadas por μ são não observáveis e portanto teremos de utilizar a distribuição conjunta de $(\mu_{t-1}, \mu_t, \mu_{t+1} | D_t)$ para calcular a probabilidade de ocorrência de uma reversão. As parcelas envolvidas nesta expressão são o parâmetro μ suavizado a um passo, isto é $\mu_{t-1} | D_t$, sua *posteriori* em t e a sua *priori* em $t+1$. Os resultados que apresentamos utilizam uma aproximação para aquela probabilidade que consiste em substituir μ_{t-1} e μ_t por seus valores esperados. No caso anterior a expressão era exata pois envolvia apenas quantidades observáveis exceto o termo preditivo que está sendo considerado corretamente nos dois casos. Assim:

$$\text{Prob}[\text{RB}_t | D_t] = \begin{cases} \text{Prob}(\mu_{t+1} < m_t | D_t) & \text{se } m_{t-1} < m_t \\ 0 & \text{caso contrário} \end{cases}$$

A decisão pela ocorrência ou não de uma reversão envolve a consideração dos custos incorridos numa decisão errada. Ou seja, é necessário incluir uma função de perda (L) descrita na tabela abaixo. Embora tenhamos salientado o caráter preditivo do problema, nas aplicações nos limitamos a cálculos baseados na *posteriori*.

Ações	Eventos	
	RB	Não(RB)
a1: Identifica a ocorrência de RB	0	c
a2: Identifica a não ocorrência de RB	d	0

A perda esperada para cada uma das ações será:

$$E(L(a1)) = c * (1 - p[RB]) , \quad E(L(a2)) = d P[RB]$$

A ação (a1) será escolhida se $E(L(a1)) < E(L(a2))$, ou seja, se $p[RB]/(1-P[RB]) < c/d$. De forma análoga, a identificação de uma reversão para cima é dada por $p[RC]/(1-P[RC]) < c'/d'$. A escolha dos custos é subjetiva e retrata a gravidade dos alarmes falsos. Assim, c muito maior que d é uma indicação de que identificar (t-1) como um RB, quando efetivamente não é, é mais custoso do que não considerá-lo quando de fato ele é um RB. Nesta aplicação, por simplicidade utilizaremos $c=d$, o que implica considerarmos a ação (a1) quando $P[RB] > 0.5$.

3. MODELOS PARA O ÍNDICE DE PRODUÇÃO DA INDÚSTRIA GERAL

Nesta seção apresentaremos os resultados de diversos modelos aplicados ao índice base fixa da produção da indústria geral, que permitem comparar o desempenho destes modelos quanto à sua capacidade preditiva e aderência. Os modelos considerados são variantes do modelo dinâmico bayesiano (MDB), um modelo estrutural clássico (MEC) descrito em Pereira (1988), e um modelo de função de transferência (MFT) descrito em Coelho e Tenenblat (1989).

A economia brasileira atravessou nos últimos 20 anos dois períodos marcadamente diferentes. A década de 70 apresentando crescimento e estabilidade e a de 80 instabilidade com períodos de crescimento seguidos de outros de estagnação. Por isso, os modelos terão seu desempenho avaliado a partir da década de 70 ou, alternativamente, usando apenas dados da de 80.

3.1. Características dos Modelos e Critérios de Comparação

O (MDB) além da sua versão básica pode considerar como em Pereira (1988) uma componente de flutuação cíclica, supondo que o produto da indústria contenha um padrão cíclico não-sazonal na sua trajetória, e outra que incorpora o número de dias trabalhados como regressor, supondo que parte da flutuação sazonal decorre da estabilidade da quantidade produzida em cada dia. A vantagem da incorporação destas componentes deve ser avaliada empiricamente pois ao tornar o modelo menos parcimonioso pode piorar a sua capacidade preditiva e descritiva.

Do ponto de vista conceitual nada garante a existência do ciclo e, quanto ao efeito do número de dias, vale lembrar que as indústrias de processo funcionam continuamente, nas demais a intensidade do trabalho pode flutuar ao longo dos meses. De outro lado, o monitoramento aumenta a capacidade descritiva mas não necessariamente a preditiva. Para comparar esta multiplicidade de alternativas serão feitas medidas de aderência, que incluem o

logaritmo da verossimilhança preditiva (LVP), o erro-padrão (EP),⁵ o Theil-U e projeções a h passos dentro da amostra para medir a capacidade preditiva através do desvio absoluto médio (DAM).

Apesar da característica preditiva dos modelos estruturais, o modelo que apresenta o menor erro de previsão a um período não é necessariamente o melhor para as previsões para um número maior de períodos. Isto contrasta com a análise de conjuntura que se interessa por aquilo que vai ocorrer nos próximos meses, ou nos próximos trimestres, e não no próximo mês, tornando ambíguo o critério de seleção dos modelos. Por isso, serão apresentadas medidas do (DAM) para erros mensais, com antecipação de três e seis meses (M3 e M6), e trimestrais para o próximo trimestre (T1) e o seguinte (T2).

As equações abaixo definem os erros mensais e trimestrais, onde (y) é a série em questão:

$$e_{t+h} = y_{t+h} - E(y_{t+h} / D_t)$$

$$a_{t+h} = (1/3) * \left(\sum_{j=0}^2 (y_{t+h-j} - E(y_{t+h-j} / D_t) \right)$$

3.2. Comparação nas Décadas de 70 e 80

O (MDB), além dos blocos de tendência linear e sazonalidade, incluirá, em alguns casos, o regressor dias úteis e a componente auto-regressiva de segunda ordem para representar eventuais ciclos. Também será testado o funcionamento do monitoramento bayesiano das modificações estruturais ocorridas na série. O modelo (MFT), originalmente identificado para a década de 80, foi reestimado para as duas décadas e corresponde à equação abaixo onde (x) é o número de dias úteis. Este modelo, estimado com dados mensais de 1979 a 1988, forneceu um valor de α significativamente igual a 1, o que implica simplificações Bell(1987) utilizadas em Coelho (1989):

$$(1 - B)(1 - B^{12})y_t = \beta (1 - B) (1 - B^{12})x_t + (1 - \alpha B) \varepsilon_t$$

A Tabela 1 foi obtida utilizando-se dados do período (janeiro de 1971 a maio de 1991), sendo que as 13 primeiras observações não foram consideradas no cálculo das estatísticas. Nela são comparados o desempenho do modelo (MFT) com três alternativas do

⁵ A variância observacional, na abordagem bayesiana, é estimada seqüencialmente e tem como distribuição a posteriori uma Gama $\Gamma(\eta_t, s_t)$. A sua estimativa pontual será $\sigma^2 = s_t / \eta_t$ e o erro-padrão, simplesmente a raiz quadrada desta variância.

modelo (MDB). Um modelo básico(MDB), um modelo básico com intervenção automática (MDB+MIA) indicada pelo monitoramento bayesiano e o modelo básico acrescido de uma componente auto-regressiva de segunda ordem (MDB+AR) para representar variações cíclicas eventualmente existentes nos dados.

O efeito do monitoramento bayesiano e das conseqüentes intervenções descritas pelo aumento da incerteza é significativo. O T-U decresce cerca de 7% (0.05 unidades), o logaritmo do fator de Bayes é 681, o desvio absoluto médio decresce de 10% (3,7 para 3,3). O efeito da componente AR é menos expressivo. A verossimilhança piora, o T-U é praticamente igual ao do modelo básico e o DAM se comporta como no caso anterior. Do ponto de vista da capacidade preditiva os modelos bayesianos são equivalentes, porém superam expressivamente o (MFT).

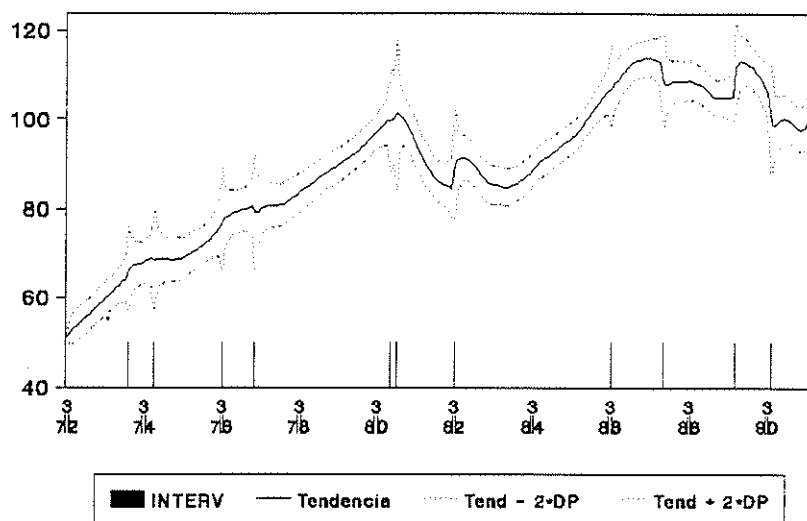
TABELA 1 : Comparação dos Modelos para Indústria Geral a partir da Década de 70

	LVP	T-U	EP	DAM*				
				M1	M3	M6	T1	T2
MFT	-	-	-	-	5.7	7.8	7.4	8.9
MDB (0.9)	-681	0.69	3.8	3.7	4.6	5.5	3.9	4.9
MDB+MIA	-613	0.64	5.4	3.3	4.6	5.9	3.6	5.2
MDB+AR	-943	0.68	2.7	3.3	5.2	6.5	4.2	5.8

*(T)Mi : DAM do erro de previsão a i (trimestres) meses

O gráfico a seguir mostra a evolução do desvio absoluto médio do erro de previsão a seis períodos ao longo da amostra. Neste gráfico observa-se que a vantagem do modelo MDB na década de 70 reduz-se gradualmente na década de 80. Este fato é consistente com as características dos modelos. O modelo (MFT) tem tendência de crescimento nula, mas, de outro lado, tem a vantagem de estimar apenas um único parâmetro. Estas características o prejudicam na década de 70 e o favorecem num período instável como os anos 80.

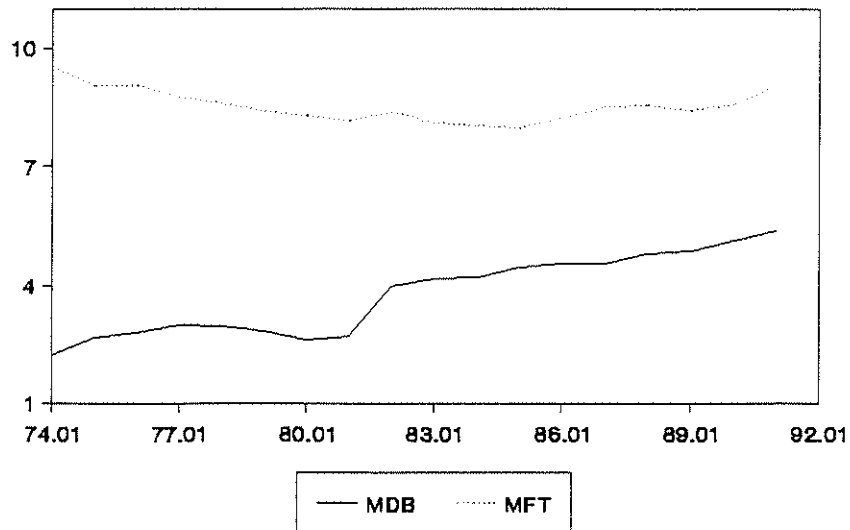
PROD. INDUSTRIAL GERAL
Tendencia Suavizada e Intervencao



O funcionamento do monitoramento⁶ é mostrado no gráfico a seguir no qual é apresentada a tendência da série suavizada com o seu intervalo de confiança e os momentos de mudança estrutural selecionados, adequadamente, pelo monitoramento.

⁶O monitoramento adotado é o descrito em West e Harrison (1989), utilizado com a *a priori* de referência definida pelos autores. Vale dizer que foi necessário reduzir o fator de desconto do resíduo de 0.98 para 0.95.

**MFT x MLD
DAM - TRIMESTRE 2**

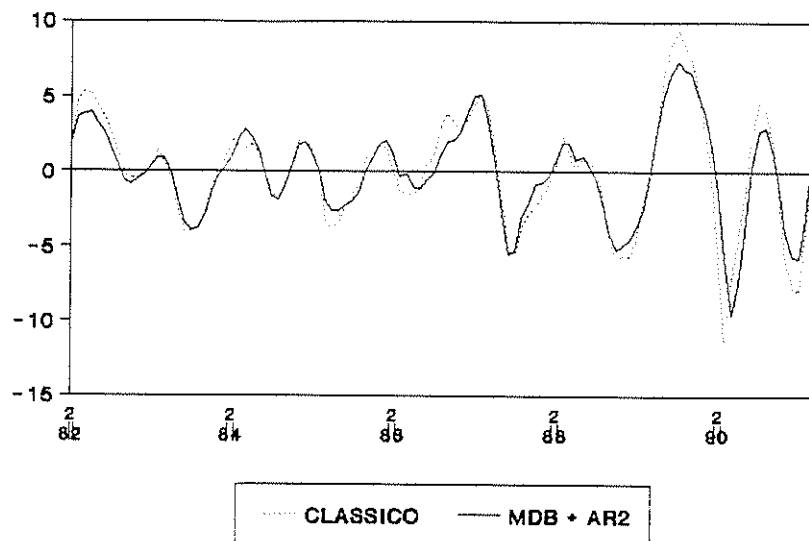


O modelo MDB com componente AR(2) só apresentou raízes complexas conjugadas na década de 70. Curiosamente, o termo auto-regressivo do modelo MDB (equação 4) tem, como mostra o gráfico a seguir, uma trajetória semelhante ao da componente cíclica do modelo MEC, cujo ciclo é gerado por processo AR(2) com raízes complexas:

$$b_t = \phi_{1t} \bar{a}_{t-1} + \phi_{2t} \bar{a}_{t-2} \quad (4)$$

Esta semelhança, possivelmente, é consequência do termo estocástico da componente AR(2) no modelo MEC, pois a trajetória gerada pela componente determinística dos dois modelos é distinta, e o modelo MDB não tem um termo estocástico associado a esta componente. De qualquer forma este termo reduziu a capacidade preditiva do modelo, não sendo portanto indicado o seu uso.

**COMPARAÇÃO CICLO
MDB x MEC**



3.3. Comparação na Década de 80

Na década de 80 foi utilizado também o modelo (MEC), e versões do modelo (MDB) que dão conta da maior instabilidade da série no período através da ampliação da incerteza indicada por menores fatores de desconto para o bloco da tendência (0.85), o que corresponde à redução da "meia vida" deste bloco para quatro meses.

O modelo MEC é constituído das mesmas componentes do (MDB), diferindo apenas por considerar os choques aleatórios sobre os parâmetros de forma aditiva e estimar os hiperparâmetros, variância destes choques, utilizando toda a amostra. Este modelo é estimado num programa específico, que infelizmente não publica os parâmetros necessários para realizar previsão a h passos dentro da amostra. Esta dificuldade foi contornada incorporando a estrutura deste modelo no programa que projeta o (MDB). Para avaliar a exatidão desta adaptação foi comparado o erro de previsão a um passo gerado no programa original (f) com o gerado na adaptação (e), obtendo-se os seguintes resultados, que foram considerados

⁷Foi utilizado o programa STAMP sobre a série para dois períodos, janeiro de 1980 a dezembro de 1989, e para testar o efeito da mudança da amostra para o período de janeiro de 1981 a maio de 1991.

satisfatórios:

$$\sum f_t^2 = 17.1 \quad \sum e_t^2 = 17.5 \quad \sum (f_t - e_t)^2 = 1$$

A variância do choque aleatório dos parâmetros nos modelos (MDB) e (MEC) só podem ser comparados de forma aproximada porque, no primeiro, a variância da distribuição *a priori* é ampliada proporcionalmente à variância da *posteriori* do parâmetro, e no segundo ela é ampliada acrescentando uma quantidade fixa que é o hiperparâmetro estimado. Tomando apenas o último período é possível calcular qual o fator de desconto implícito no modelo (MEC) relacionando a variância do parâmetro e o hiperparâmetro correspondente. Como o modelo (MDB) básico não tem a componente cíclica foi definida uma variável aleatória (U) que corresponde à soma do nível (u) e da componente cíclica(c). Lembrando que os choques são supostos independentes.

$$V(U) = V(u) + V(c) + 2*Cov(u,c)$$

$$V(\varepsilon_U) = v(\varepsilon_u) + v(\varepsilon_c)$$

O fator de desconto é então a razão entre as variâncias do parâmetro e a do choque aleatório acrescida da do parâmetro. A Tabela 2 a seguir apresenta o resultado da estimação dos hiperparâmetros, da variância dos parâmetros e os fatores de desconto implícitos no modelo MEC e os utilizados no MDB. Vale comentar a semelhança dos resultados, exceto para a componente do nível cuja "meia vida" é de cerca de dois meses indicando "sobreajustamento". O decaimento do ciclo foi de (0.88), a frequência foi de (.346) e a correlação entre o nível e o ciclo foi de (-0.54).

TABELA 2 : Relação entre Hiperparâmetros do Modelo Clássico e os Fatores de Desconto do Modelo Bayesiano

	Hiperparâmetros	Var(parâmetros)	Fator de Desconto	
			MEC	MDB
Nível	0	8.48	-	
Ciclo	3.46	7.63	-	
Nível+Ciclo	3.46	7.42	.68	.90
Inclinação	0.042	0.28	.87	.90
Sazonalidade	0.006	0.15	.96	.98

A Tabela 3 trata especificamente do período janeiro de 1980 a maio de 1991, o qual corresponde à maior instabilidade econômica. Nela são feitas comparações do modelo bayesiano com diferentes fatores de desconto, do MEC estimado com dois conjuntos de hiperparâmetros, e com o MFT. O fator de Bayes da comparação dos modelos com desconto de .9 e .8 é 17, o que indica a preferência pelo segundo. Isto caracteriza a volatilidade dos dados neste período. O modelo MEC apresentou um comportamento preditivo a um passo melhor que os modelos MDB embora não difiram quanto à capacidade preditiva para horizontes maiores. Ressalta-se, para esta amostra, que o comportamento do MFT é ligeiramente (6,3 para 6,5, ou seja, 3%) melhor que os demais. Isto se deve, possivelmente, ao comportamento dos dados nesta década caracterizado pela instabilidade e a ausência de crescimento. Ainda assim este modelo não é adequado para se prever o agregado de um trimestre e dois trimestres, quando a variação relativa é de 40% (9,9 para 6,0). O modelo (MEC) escolhendo de forma ótima as variâncias de evolução dos parâmetros "superajustam" os dados como é sugerido pela não melhoria da capacidade preditiva.

TABELA 3 : Comparação de Modelos para a Indústria Geral a partir da Década de 80

	LVP	T-U	EP	DAM*				
				M1	M3	M6	T1	T2
MFT	-	-	-	3.1	5.2	6.3	7.6	9.9
MEC - 1	-346	.58	2.0	3.2	5.3	6.6	4.3	6.0
MEC - 2	-349	.60	2.6	3.3	5.2	6.5	4.2	6.0
MDB-D(.90)	-393	.56	4.5	4.9	5.8	6.8	4.9	6.2
MDB-D(.80)	-376	.49	3.5	4.0	5.5	6.7	4.6	6.4

*(T)Mi : DAM do erro de previsão a i (trimestres) meses

Na Tabela 4 diferentes versões do modelo MDB são comparadas e mostra-se que a aderência do modelo aos dados está relacionada com a piora na capacidade preditiva, especialmente no caso do uso do monitoramento. Coerente com o cenário de instabilidade e estagnação da década de 80, o melhor modelo é o estático bayesiano (MEB), que supõe a ausência da tendência de crescimento e mudanças estruturais indicadas pelo monitoramento. Este é o que apresenta a maior aderência expressa pela (LVP) dentre os modelos com a mesma capacidade preditiva. O modelo (MDB) foi aplicado à média móvel da série, pois este filtro retira a componente sazonal, reduz a instabilidade da série e facilita a identificação dos momentos reversão. De outro lado implica a piora da capacidade preditiva

devido à perda de informação indicada pela função de previsão.⁸

TABELA 4: Comparação de Modelos Dinâmicos Bayesianos Aplicados ao Índice da Produção Industrial Geral

	LVP	T-U	Sem Monitoramento DAM					Com Monitoramento DAM			
			M1	M3	M6	T1	T2	LVP	M1	M6	T2
MDB (.90)	-401	.56	5.0	5.9	6.9	4.7	6.1	-391	4.8	7.6	6.6
MDB-D(.90)	-393	.56	4.9	5.8	6.8	4.9	6.2	-382	4.8	7.3	6.8
MDB-D(.85)	-382	.52	4.3	5.5	6.5	4.5	6.0	-368	4.3	7.3	6.9
MEB-D(.90)	-422	.68	6.1	6.4	7.1	5.9	6.7	-391	5.0	6.3	5.9
MDB+MM		1.16	7.6	12.3	15.8	13.5	14.2		7.2	15.9	14.5

*(T)M_i : DAM do erro de previsão a i (trimestres) meses

3.4. Determinação de Pontos de Reversão Cíclica

O procedimento de indicação de reversão foi aplicado sobre as componentes estimadas da série da produção da indústria geral. Foram utilizados critérios aproximados para o cálculo da probabilidade de reversão e para a identificação da situação de queda (subida) da série:

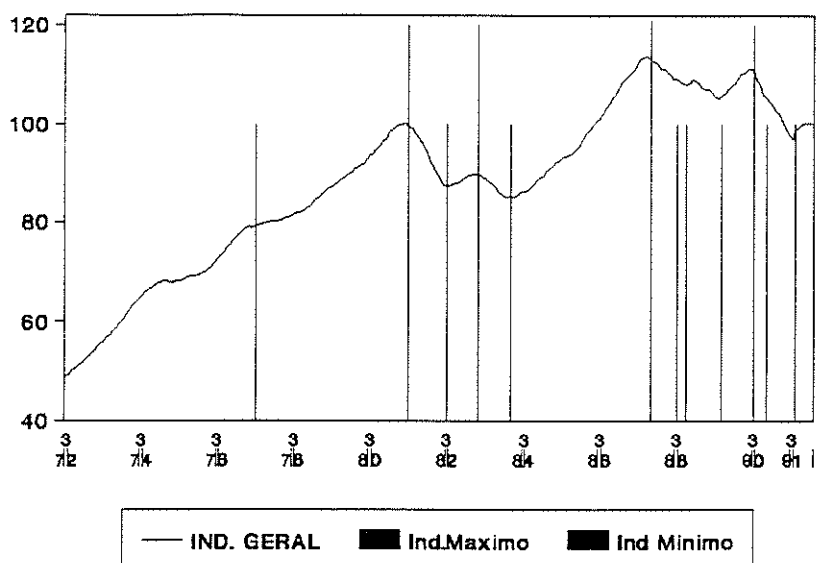
$$\text{Prob}[RB_t | D_t] = \text{Prob}(\mu_{t+1} < m_t | D_t) \quad \text{se } M_{t-1} < M_t$$

e o caso contrário, onde M_t = média móvel da série e m_t = média da posteriori.

⁸O nível da série pode ser projetado pela expressão abaixo, onde (β) é a estimativa da inclinação da trajetória da média móvel:

$$E(y_t | D_{t-1}) = y_{t-12} + 12 \beta_t$$

Indicador de Reversao - Indústria Geral



Os resultados mostram a capacidade do método em indicar as reversões adequadamente. Das oito reversões ocorridas este procedimento indicou adequadamente todas. Entretanto indicou, erradamente, como reversão quatro momentos.

4. PROJEÇÕES DAS SÉRIES ECONÔMICAS

A conjuntura econômica é descrita por variáveis nominais e reais. No Brasil a extrema volatilidade dos preços enfraquece, no caso de séries nominais, o conceito de tendência e esconde uma eventual sazonalidade. Portanto, este tipo de instrumental não se aplica a séries nominais, que não foram incorporadas neste estudo. As séries reais selecionadas - veja Boletim do Grupo de Acompanhamento de Conjuntura - podem ser agrupadas nas que: a) descrevem o produto da indústria, indicador geral e desagregado por uso; b) determinam o produto interno bruto segundo a metodologia do PIB trimestral (...) que inclui indicadores de atividade dos setores mais importantes da economia; c) contabilizam o desempenho da balança comercial; e d) indicam o nível de consumo real em São Paulo medindo o estado de "aquecimento" da economia em alguns segmentos mais importantes.

As séries que têm grande alteração do seu nível deixam em aberto a possibilidade da componente sazonal ser proporcional ao nível da série, o que exigiria o produto da componente sazonal pelo nível (modelo multiplicativo). De outro lado, séries que não tenham marcada trajetória de crescimento - e são muitas na década de 80 no Brasil - poderiam admitir modelos estacionários. Para cada série foram testadas diferentes opções do modelo linear de crescimento com componentes sazonais aditivas.

Os resultados seguem a notação: modelo dinâmico bayesiano (MDB) com componentes sazonais aditivas; com alteração do desconto do nível [MDB(.85)]; com efeito dos dias úteis (MDB-D); com componentes sazonais multiplicativas (MDB-M); modelo linear estacionário (MEB); componente de ciclo (MDB+AR). Sempre que foi testado apenas um tipo de modelo esta indicação foi omitida, e então foi utilizado o (MDB).

Os fatores de desconto⁹ utilizados foram os recomendados na literatura West (1989). Estes fatores controlam a adaptabilidade dos parâmetros. Um modo mais intuitivo de entender o funcionamento destes fatores é relacioná-los com a "meia vida" dos parâmetros, ou seja, o número de períodos decorridos para que o efeito de uma observação sobre o parâmetro decaia para a metade. Assim, o bloco de tendência tem uma "meia vida" de seis meses, o sazonal de 30 meses. No caso de intervenção esta "meia vida" cai para um e seis meses, respectivamente.

4.1. Índices de Produção da Indústria

O IBGE apura, a partir de uma amostra do tipo painel, um índice de quantidade do produto da indústria. Este índice, pela sua abrangência e cuidado metodológico, é amplamente utilizado como indicador nacional do nível de atividade. Além deste índice são apurados índices do produto industrial agregado por gênero de indústria e por uso do bem produzido. Foram selecionados para a previsão o índice da indústria geral e os agregados por uso. Estes últimos, devido a sua definição, não são uma desagregação do índice geral que não pode ser derivado deterministicamente daqueles agregados por uso. Esta situação é adequada para o uso da técnica da previsão hierárquica, que utiliza a interdependência entre as séries para melhorar a capacidade preditiva. Entretanto, no caso de cinco séries com componentes sazonais, este procedimento torna-se muito oneroso computacionalmente - envolve uma matriz de covariância com cerca de 75 componentes -, além de apresentar dificuldades conceituais no caso do uso do monitoramento.

O procedimento adotado consistiu em utilizar modelos para cada uma das séries isoladamente, exclusive uma - a de bens de consumo intermediário, que será calculada a partir das previsões das demais para garantir a coerência entre as previsões dos índices.

⁹Fator de desconto de 0.9 para o bloco de tendência e de 0.98 para o bloco sazonal. No caso de intervenção indicada pelo monitoramento os fatores de desconto são 0.1 para o bloco de tendência e 0.9 para o sazonal. Também foi considerado para a estimação da variância do resíduo o fator de desconto de 0.95 e no caso de intervenção, de 0.63.

Foi ajustado um modelo que relaciona a série de forma irrestrita (dinamicamente) - o índice de bens de consumo intermediário com os demais índices, ou seja, com os valores contemporâneos e defasados de todos os outros índices, e valores defasados dela mesma.

Os modelos utilizados nos índices de produção de bens de consumo durável e não-durável foram o uso ou não dos dias úteis como variável explicativa - que ainda que melhore a aderência, piora a capacidade preditiva - e o uso do monitoramento que só apresentou resultados marginalmente significativos para a série dos bens de consumo duráveis.

TABELA 5 : Comparação de Modelos para Índices de Bens de Consumo

	Sem Monitoramento						Com Monitoramento		
	T-U	M1	M3	M6	T1	T2	M1	M6	T2
Durável									
MDB	.81	9.7	10.2	11.0	7.7	9.3	9.2	10.5	9.0
MDB-D	.80	9.2	9.8	10.6	8.5	9.9	8.5	11.1	10.2
Não-Durável									
MDB	.54	5.3	6.9	7.6	5.5	6.8	5.1	8.2	7.3
MDB-D	.50	4.9	6.7	7.4	5.8	7.1	4.5	8.5	8.4

A previsão do índice de produção da indústria de transformação pode ser derivada dos índices da indústria geral e da extrativa mineral, pois ponderadores fixos os relacionam entre si. Conforme comentado, os índices agregados por uso não estão relacionados de forma determinística com o índice geral da produção industrial. O procedimento, já descrito, que realiza projeções compatíveis para as duas desagregações - por uso e gênero - resulta num erro absoluto médio para previsões da série de Bens de Consumo Intermediário, realizadas com três períodos de antecipação, de 6.4%.

Na Tabela 6 apresentamos o resultado para o índice de produção da indústria extrativa mineral e produtora de bens de capital. O caso da indústria extrativa mineral apresenta uma tendência e sazonalidade marcante, cabendo então cogitar do efeito da sazonalidade proporcional. Será testada também a contribuição dos dias úteis. A inclusão dos dias úteis ou o uso de componentes sazonais multiplicativas não apresenta vantagens significativas. Já o monitoramento melhora a aderência e em cerca de 20% a capacidade preditiva do modelo.

O caso do índice de produção de bens de capital foi testado com o efeito dos dias úteis, e da inclusão de uma componente auto-regressiva que dá conta de eventuais ciclos de produção. Ambos não apresentam resultados significativos. Já o monitoramento melhora marginalmente os resultados.

TABELA 6 : Comparação de Modelos para Extrativa Mineral e Bens de Capital

	Sem Monitoramento						Com Monitoramento		
	T-U	M1	M3	M6	T1	T2	M1	M6	T2
Extrativa Mineral									
MDB	.62	6.3	8.1	10.6	6.7	9.3	5.0	8.8	7.3
MDB-D	.66	6.5	8.2	10.6	7.0	9.5	5.1	8.8	7.9
MDB-M	.79	4.3	7.5	9.8	5.9	8.0	5.1	10.0	7.7
Bens de Capital									
MDB	.79	6.6	7.7	9.7	6.3	8.6	6.1	9.5	8.4
MDB-D	.82	6.7	7.6	9.6	6.8	9.0	6.1	9.5	8.7
MDB+AR		1.10	6.2	8.3	9.9		6.7	8.5	

4.2. Componentes do PIB Real Trimestral

O PIB trimestral é um índice de Laspeyres encadeado que compõe índices de produção (de oferta) dos principais grupos de atividade da economia. As séries apresentadas nas Tabelas 6 e 7 são as principais componentes deste indicador, exceto os índices de produção da indústria de transformação e extrativa mineral - já tratados anteriormente -, e dos índices de produção do setor agrícola que não são geradas mensalmente. São construídos a partir de medidas de produção anual dos principais produtos e do padrão sazonal das correspondentes safras.

O consumo de energia elétrica mede a produção dos serviços de utilidade pública. Apresenta uma tendência e padrão sazonal bem definido. Foi então avaliado o efeito de componentes sazonais multiplicativas, que não resultou em melhoria. Também a inclusão do monitoramento não é vantajosa. O consumo de óleo diesel - proxy - para a produção do setor de transportes tem uma trajetória semelhante à série anterior. Também de forma semelhante o uso de componentes multiplicativas e do monitoramento são prejudiciais.

TABELA 7 : Comparação de Modelos para os Índices do
Produto Interno Bruto

	Sem Monitoramento						Com Monitoramento		
	T-U	M1	M3	M6	T1	T2	M1	M6	T2
Consumo de Elétrica									
MDB	.40	3.5	4.0	4.7	3.1	3.8	3.4	5.0	4.1
MDB-M	.85	3.6	4.1	4.6	3.0	3.8	4.1	6.4	5.1
Consumo de Óleo Diesel									
MDB	.56	6.8	7.3	7.7	5.3	5.8	7.1	8.9	6.9
MDB-M	.78	7.8	7.9	8.3	5.9	6.1	8.7	12.9	9.2
Insumos da C.Civil									
MDB	.71	6.0	8.3	10.3	6.9	9.4	5.0	10.1	9.2
Produto Pecuária									
MDB	.50	5.5	6.8	7.7	5.6	7.1	5.2	8.3	7.4
MDB+AR	.86	5.6	8.2	10.3	6.5	8.8			

O índice de produção de insumos da construção civil é uma *proxy* do produto dos serviços deste setor. O uso do monitoramento apresenta melhoras apenas marginais. A produção pecuária é um índice composto por diversas componentes, das quais a produção de carne bovina é a que tem o maior peso. A literatura menciona a existência de um ciclo de cerca de quatro anos para esta produção, que está relacionado com o período de maturação dos animais para o abate. Para captar este ciclo foi incluída a componente cíclica. De fato, o monitoramento e a inclusão desta componente não melhoraram as previsões, indicando a já habitual "sobreparametrização".

A atividade dos setores de serviços não-comércio e das instituições financeiras foi representada por séries de emprego setoriais - na hipótese de produtividade constante da mão-de-obra. Por semelhança de natureza foi incluída também a série do emprego na indústria. Todas as séries se caracterizam por um comportamento menos instável e sem componentes sazonais. Em todos os casos o monitoramento apresentou um desempenho significativo, melhorando em até 40% a capacidade preditiva dos modelos.

TABELA 8 : Comparação de Modelos para os Índices de
Empregos Setoriais

	Sem Monitoramento						Com Monitoramento		
	T-U	M1	M3	M6	T1	T2	M1	M6	T2
Industrial	.28	2.0	3.1	4.4	2.6	3.9	1.2	3.4	3.0
Financeiro	.34	1.7	3.1	4.4	2.6	3.9	0.8	3.3	2.8
Demais Serv.	.25	1.8	2.4	3.3	2.0	2.9	1.3	3.0	2.7

4.3. Balança Comercial

A desagregação habitual das componentes da balança comercial que foi adotada permite um acompanhamento mais atualizado das informações, pois é a publicada com agilidade pelo Departamento de Comércio Exterior, além de conter algum conteúdo econômico. Separa a pauta de exportação nos produtos básicos, manufaturados e semimanufaturados, grupos que têm apresentado um comportamento específico, e especialmente os básicos, que estão mais sujeitos às condições do mercado de *commodities*. De outro lado, as importações de petróleo por sua natureza administrada torna mais previsível o comportamento do agregado de importações que a exclui. Vale lembrar o objetivo de ajustamento de um modelo descritivo que permita extrapolar a tendência da série, o que exclui o uso da taxa de câmbio que seria a variável-chave para explicar boa parte do comportamento destas séries. A exportação de manufaturados e a importação, exclusive petróleo, têm um padrão de crescimento que, embora instável, justifica a tentativa de uso de modelos multiplicativos.

TABELA 9 : Comparação de Modelos para os Índices da
Balança Comercial

	Sem Monitoramento						Com Monitoramento		
	T-U	M1	M3	M6	T1	T2	M1	M6	T2
Exportação									
Manufaturados	.9	21.9	26.8	32.7	20.3	27.3	20.3	33.5	28.3
Semimanufatur.	1.0	25.3	30.8	36.1	23.6	30.7	25.0	40.5	32.6
MDB-M	1.1	26.2	32.3	37.9	23.8	31.0	25.9	40.6	30.0
Básicos	.86	12.7	15.3	17.3	12.3	15.4	12.3	19.9	17.8
Importação									
Exc. Petróleo	.78	9.5	11.7	14.3	8.8	12.0	9.0	14.0	11.8
MDB-M	.90	9.3	12.1	14.1	9.0	11.4	9.8	18.0	13.6

Os resultados apresentados na Tabela 9 mostram que o uso de modelos multiplicativos ou do monitoramento não melhora o péssimo desempenho preditivo dos modelos. Chama a atenção para o fato de que o desvio médio absoluto é três vezes maior que os casos anteriormente expostos, tornando praticamente inútil o uso destes modelos para extrapolação. Vale dizer que este resultado retrata a instabilidade intrínseca das séries, tornando igualmente imprecisos quaisquer procedimentos de extrapolação.

4.4. Indicadores de Consumo

Os indicadores de consumo são agregados nominais coletados pela Federação de Comércio de São Paulo e posteriormente deflacionados por algum índice de preços. A coleta realizada por uma instituição não-especializada e o deflacionamento de valores nominais deixam - alguma dúvida quanto à qualidade das séries. Além disso, cobrem apenas a cidade de São Paulo, reduzindo a sua abrangência. Apesar destes pontos negativos, são as únicas medidas do lado da demanda.

Os resultados apresentados na Tabela 9 mostram também que o efeito do monitoramento é marginal, melhorando a capacidade descritiva, mas melhorando muito pouco e, em alguns casos piorando, a capacidade preditiva. A qualidade preditiva medida pelo desvio absoluto médio do erro de previsão no segundo trimestre é da mesma magnitude - cerca de 10% - dos casos anteriores, exceto os componentes da balança comercial.

TABELA 10 : Comparação de Modelos para os Índices de Consumo em São Paulo

	Sem Monitoramento						Com Monitoramento		
	T-U	M1	M3	M6	T1	T2	M1	M6	T2
Geral	.53	7.1	8.7	11.0	7.2	9.7	6.2	10.9	9.3
Bens de Consumo									
Não-Durável	.67	10.5	14.4	18.6	12.0	16.6	7.8	16.1	13.8
Semidurável	.32	10.3	12.5	13.7	9.9	12.2	9.8	16.1	14.0
Durável	.52	8.2	9.9	11.4	8.0	10.3	6.7	11.9	10.2
Bens de Investimento									
Veículos	1.0	8.9	10.3	13.0	8.3	11.5	8.6	13.4	11.7
Mat. Construção	.93	7.0	9.0	10.5	7.5	9.5	6.3	11.1	9.9

5. CONCLUSÃO

Este artigo comparou modelos alternativos aplicados à série do produto da indústria geral. Os resultados sugerem que os modelos estruturais têm capacidade preditiva tão boa ou melhor do que o não-estrutural tomado como exemplo, e que a abordagem bayesiana também apresenta resultados equivalentes aos da clássica. Conceitualmente os modelos estruturais têm a característica de descreverem qualitativamente os atributos das séries - tendência, sazonalidade - e a abordagem bayesiana dos modelos estruturais é, adicionalmente, mais simples no uso, além de dispor de recursos de monitoramento e intervenção - que aumentam a capacidade de descrever os fatos econômicos.

O procedimento de intervenção utilizado se limitou às indicações do monitoramento, apesar de potencialmente ser capaz de incorporar a opinião dos especialistas. Isto se torna possível pelo fato de os parâmetros dos modelos terem interpretação econômica e serem conceitualmente variáveis aleatórias, cuja distribuição de probabilidade pode ser imputada subjetivamente. Vale dizer que, apesar de não terem sido apresentados os gráficos dos resultados, o monitoramento indicou corretamente, para a maioria das séries, os seus momentos de mudança de inclinação.

O desempenho do modelo nas séries foi heterogêneo - erros médios absolutos variando de 2 a 12% -, indicando a diversidade das características do conjunto de séries analisado. Particularmente, as séries relativas ao comércio exterior foram as que apresentaram o pior desempenho. As características probabilísticas e de flexibilidade dos modelos estruturais bayesianos sugerem que métodos alternativos que se limitam a extrapolar a trajetória da série terão também um desempenho igualmente ruim no conjunto de séries mais voláteis.

Finalmente, os resultados aproximados obtidos para o indicador de reversão encorajam o desenvolvimento de procedimentos numéricos mais precisos.

BIBLIOGRAFIA

- BELL, W. A note on overdifferencing and the equivalence of seasonal time series with monthly means and models with $(0,1,1)_{12}$ Seasonal parts when $(\theta = 1)$. Journal of Business and Statistics, 5, 3, 1987.
- COELHO, C.H.M e Tenenblat, M. Forecasting with trading day variation, a seasonal common factor and variance change (Relatório Interno, IPEA), 1991.
- GAMERMAN, D. e MIGON, H. S. - Hierarquical Dynamic Models - To Appear in JASA-B.
- HARRISON, P. J. Bayesian forecasting in O.R. Tech. Report-Warwick University, 1989.
- HARRISON, P. J. e STEVENS, C. F. Bayesian forecasting. Journal of the Royal Statistical Society, Series B, 38, p. 205-247, 1976.
- HARVEY, A. C. Time Series Models - Philip Allan Oxford, 1981.
- MIGON, H. S. e BRASIL, G. H. Modelos de Previsão para séries de produção e preços: métodos bayesianos e Box-Jenkins. (IPEA-Rio TDI n° 94), 1986.
- MIGON, H. S. e GAMERMAN, D. - Generalized exponential growth models: a Bayesian approach (Relatório Técnico n°. 41). Laboratório de Estatística, Universidade Federal do Rio de Janeiro, 1989.
- PEREIRA, P.V. e HARVEY, A.C. Trend, seasonality and seasonal adjustment. TDI n° 154, IPEA-Rio, 1988.
- WEST, M., e HARRISON P. J. Bayesian forecasting and dynamic Models, Springer Verlag, 1989.
- ZELLNER A., HONG, C. e GULATI G.H. Turning points in economic time series, loss structures and Bayesian forecasting: Graduate School os Business University of Chicago, Discussion Paper, 1987.
- STAMP - Structural time series packge, Harvey A.C., s. d.
- PRODUTO Interno Bruto Real Trimestral - Metodologia, Fundação IBGE, junho de 1989, mimeo.

