

Previsão de séries de tempo na presença de mudança estrutural: redes neurais artificiais e modelos estruturais*

Wilson R. Corrêa§
Marcelo S. Portugal†

RESUMO

As políticas de estabilização de preços e o processo de liberalização comercial realizados nas últimas décadas aumentaram muito a dificuldade de geração de previsões precisas de séries econômicas, em função das mudanças estruturais que foram ocasionadas no processo gerador dos dados. Este artigo apresenta uma avaliação empírica da capacidade preditiva de modelos de Redes Neurais Artificiais (RNA) e modelos estruturais de série de tempo (MEST) quando existe uma mudança estrutural. Nossa preocupação básica é avaliar a capacidade relativa destes dois modelos de previsão em identificar a ocorrência da mudança estrutural e a rapidez no ajuste das previsões um passo à frente após a mudança, com e sem intervenção no modelo. Nestes exercícios empíricos usamos séries reais, tais como a taxa de inflação e o valor das importações brasileiras, e criadas por simulação. As séries geradas por simulação seguem diferentes processos ARIMA com imposição de mudanças estruturais na média e tendência. Os resultados mostram uma capacidade levemente superior das RNAs em identificar a ocorrência de quebras estruturais e realizar previsões um passo à frente nos períodos imediatamente seguintes a esta.

Palavras-chave: mudança estrutural, redes neurais, séries de tempo.

ABSTRACT

The Brazilian price stabilisation policies and trade liberalisation measures of this decade have considerably increased the difficulty in generating accurate time series forecasts due to structural changes in the data generation processes. In this paper we provide an empirical evaluation of the forecasting performance of Artificial Neural Networks (ANN) and Structural Time Series models (STS) in the presence of structural change. We are basically interested in evaluating the capability of ANN and STS models in terms of both identifying that a structural change has happened and the speed of adjustment of the one step ahead forecasts after the change. We use both real and simulated time series in these exercises. The simulated series are generated from ARIMA processes with imposed structural changes in the mean and trend. On the other hand, we also use real time series data for the Brazilian inflation rate and total imports. The results for the one step ahead forecasts show that the ANN models present a marginally better performance than the STS in the periods just after the structural change.

Key words: structural change, neural networks, time series.

* Os autores gostariam de agradecer a colaboração do bolsista de iniciação científica do CNPq Gregório Silva Caetano. Uma versão anterior deste artigo foi apresentada no XVIII International Symposium on Forecasting, em Edinburgh, Escócia.

§ Mestrando em economia na UFRGS e professor substituto na UFRGS e no Centro Regional de Economia e Administração da Universidade de Passo Fundo (rotatori@vortex.ufrgs.br).

† Professor do Curso de Pós-Graduação em Economia da UFRGS e pesquisador do CNPq (msp@vortex.ufrgs.br).

1 Introdução

A partir da década de setenta, tem sido grande a proliferação de modelos univariados de previsão com aplicação em economia. O objetivo básico destes modelos é a realização de previsões de curto prazo, utilizando apenas os dados referentes à própria série que se deseja prever. Em algumas extensões, outras séries podem ser utilizadas, mas, de forma geral, não é necessário a especificação de um modelo econométrico completo. A experiência prática na utilização de modelos univariados para a previsão de curto prazo tem se mostrado bastante satisfatória.

A renovação e o aperfeiçoamento dos modelos têm sido constante. Os modelos ARIMA, que foram largamente utilizados nas décadas de setenta e oitenta, foram rapidamente substituídos pelos chamados modelos estruturais de séries de tempo, de enfoque clássico ou bayesiano. Mais recentemente, mesmo estes modelos estruturais já começam a ser questionados quanto à sua eficiência em gerar previsões de curto prazo pelas redes neurais, desenvolvidas originalmente no campo da inteligência artificial.

O objetivo central deste artigo é avançar no estudo da comparação entre modelos de previsão univariados, de forma a avaliar a capacidade de cada modelo em lidar com a existência de quebras estruturais na série que é objeto da previsão. Nossa análise se restringirá aos modelos estruturais de séries de tempo e a modelos de redes neurais. Estes modelos foram escolhidos pela sua já demonstrada capacidade de reconhecer padrões e regularidades em séries de tempo. Nossa análise tentará avaliar a capacidade relativa de cada modelo em identificar o ponto de quebra estrutural e a rapidez do ajustamento das previsões um passo à frente após a quebra, com e sem a utilização de intervenção sobre o modelo.¹

As seções seguintes apresentam, brevemente, a metodologia a ser usada e os resultados empíricos.

2 Modelos de previsão

Nesta seção introduziremos os modelos de previsão a serem utilizados. Como mencionado acima, trataremos basicamente de dois tipos de metodologias de previsão: modelos estruturais de séries de tempo e modelos de redes neurais.

1 Comparações do desempenho preditivo destes modelos para séries de tempo sem quebra estrutural podem ser encontradas em Portugal (1995) e Fernandes, Navaux e Portugal (1996).

2.1 Modelos estruturais de séries de tempo

Nos últimos anos, especialmente após a introdução do filtro de Kalman em econometria, a literatura sobre modelos estruturais tem avançado substancialmente. O filtro de Kalman permitiu a operacionalização dos modelos estruturais, ao fornecer uma forma fácil de estimar os componentes estruturais do modelo que são variantes no tempo.²

Seguindo-se Harvey (1989), podemos dizer que a idéia básica é ver uma série de tempo como o resultado do somatório de vários componentes não diretamente observados. Em geral, assume-se que a série tenha um componente de tendência (μ_t), cíclico (ψ_t), sazonal (γ_t) e irregular (ε_t). Em termos analíticos

$$y_t = \mu_t + \gamma_t + \psi_t + \varepsilon_t.$$

Esta forma de modelagem de séries de tempo tem o grande atrativo de oferecer uma descrição das principais características da série. Ela nos possibilita a análise de mudança de tendência e alterações no padrão sazonal ou no comportamento cíclico da série. Por permitir uma direta interpretação dos componentes da série este tipo de modelo pode ser considerado mais interessante que os tradicionais modelos ARIMA. Na verdade, para cada tipo de modelo de decomposição em componentes não observados existe um modelo ARIMA correspondente.

O modelo completo é obtido após a especificação de cada um dos componentes isoladamente. Por razões de espaço vamos utilizar uma versão reduzida do modelo apresentado acima, chamado na literatura de modelo estrutural básico, que exclui o componente do ciclo. Esta simplificação não prejudica a exposição, pois o componente cíclico é modelado de forma semelhante ao componente sazonal, usando-se funções seno e cosseno. O modelo estrutural básico pode então ser especificado como

$$y_t = \mu_t + \gamma_t + \varepsilon_t \tag{1}$$

$$\mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t \tag{2}$$

2 Na verdade, os modelos estruturais de séries de tempo fazem parte de uma classe mais genérica de modelos de decomposição de uma série de tempo em seus componentes não diretamente observados. Neste artigo trataremos apenas dos modelos estruturais de inspiração clássica. Modelos muito semelhantes, de inspiração bayesiana, chamados de modelos lineares dinâmicos, podem ser encontrados em West e Harrison (1997).

$$\beta_t = \beta_{t-1} + \zeta_t \quad (3)$$

$$\gamma_t = \sum_{j=1}^{s/2} \gamma'_t$$

onde γ'_{jt} é formado por

$$\begin{bmatrix} \gamma_{jt} \\ \gamma_{jt}^* \end{bmatrix} = \begin{bmatrix} \cos \lambda_j & \text{sen } \lambda_j \\ -\text{sen } \lambda_j & \cos \lambda_j \end{bmatrix} \begin{bmatrix} \gamma_{j,t-1} \\ \gamma_{j,t-1}^* \end{bmatrix} + \begin{bmatrix} \omega_{jt} \\ \omega_{jt}^* \end{bmatrix}$$

e ε_t , η_t e ζ_t são erros aleatórios com distribuição normal e variâncias σ_ε^2 , σ_η^2 e σ_ζ^2 , ω_t e ω_t^* são também erros aleatórios normais com variância igual a σ_ω^2 , λ_j é a frequência e γ'_{jt} aparece por construção.

Para simplificar a exposição vamos supor que estamos lidando com dados trimestrais, de forma que $s = 4$. Desta forma, a frequência $\lambda_j = 2\pi j/4$ para $j = 1$ e 2 , o que produz $\lambda_j = \pi, \pi/2$. Usando relações trigonométricas básicas, podemos rescrever o sistema de equações como

$$\begin{aligned} \gamma_{1t} &= \gamma_{1,t-1}^* + \omega_{1t} \\ \gamma_{1t}^* &= -\gamma_{1,t-1} + \omega_{1t}^* \\ \gamma_{2t} &= -\gamma_{2,t-1} + \omega_{2t} \\ \gamma_{2t}^* &= -\gamma_{2,t-1}^* + \omega_{2t}^* \end{aligned} \quad (4)$$

A estimação do modelo com parâmetros variáveis pode ser feita utilizando-se o filtro de Kalman. O modelo formado pelas equações (1) a (4) pode ser facilmente colocado no formato do estado de espaço, como mostrado abaixo.

$$y_t = [1 \ 0 \ 1 \ 0 \ 1] \alpha_t + \varepsilon_t \quad (5)$$

$$\alpha_t = \begin{bmatrix} \mu_t \\ \beta_t \\ \gamma_{1t} \\ \gamma_{1t}^* \\ \gamma_{2t} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} \mu_{t-1} \\ \beta_{t-1} \\ \gamma_{1,t-1} \\ \gamma_{1,t-1}^* \\ \gamma_{2,t-1} \end{bmatrix} + \begin{bmatrix} \eta_t \\ \zeta_t \\ \omega_{1t} \\ \omega_{1t}^* \\ \omega_{2t} \end{bmatrix} \quad (6)$$

Nesta especificação temos uma variável para representar o nível da série (μ_t) e outra para a tendência propriamente dita (β_t). Nas séries que não apresentam uma tendência clara o valor de β_t será aproximadamente zero, enquanto que se houver uma taxa de crescimento constante teremos $\beta_t = \beta$.

Como mencionamos acima, a estimação dos modelos de decomposição em componentes não observáveis é feita por meio do Filtro de Kalman. O filtro é um algoritmo desenvolvido no âmbito da engenharia para resolver o problema de extração de sinal.³ Utilizando a metodologia apresentada por Atkinson, Koopman e Shephard (1997) temos os modelos representados na forma de espaço de estados, seguindo a formulação Gaussiana, dados pela equação de medida (7) e transição (8).

$$y_t = Z_t \alpha_t + \varepsilon_t \quad (7)$$

$$\alpha_t = T_t \alpha_{t-1} + \eta_t \quad (8)$$

$$\alpha_0 / Y_0 \sim N(a_0, P_0), \varepsilon_t \sim N(0, H_t), \eta_t \sim N(0, Q_t)$$

$$E(\varepsilon_t, \eta_t) = 0$$

$$E(\varepsilon_t, \alpha_0) = 0$$

$$E(\eta_t, \alpha_0) = 0$$

$$E(\varepsilon_i, \varepsilon_j) = \delta_{ij} \sigma^2 H_t$$

$$E(\eta_i, \eta_j) = \delta_{ij} \sigma^2 Q_t$$

onde δ_{ij} é o delta de Kronecker, y_t representa o vetor de medida, α_t o vetor de estado e Z_t e T_t são matrizes conhecidas.

3 Ver Kalman (1960). Diferentes interpretações do Filtro de Kalman na área da econometria e estatística clássica e Bayesiana podem ser encontradas em Duncan and Horn (1972), Meinhold e Singpurwalla (1983) Harvey (1987), Diderrich (1985).

O que o Filtro de Kalman faz é gerar a melhor previsão para o vetor de estado e sua matriz de covariância (P_t) a cada período de tempo. Uma vez que uma nova observação se torne disponível, o filtro atualiza esta previsão levando em conta o erro de previsão cometido. Neste sentido, o filtro aprende com os seus erros. A estimação das variâncias (σ^2), às vezes denominadas hiperparâmetros, é feita por máxima verossimilhança.⁴ O filtro consiste ainda no seguinte conjunto de equações

$$a_{t+1/t} = T_{t+1}a_{t/t-1} + K_t v_t \quad (9)$$

$$P_{t+1/t} = T_{t+1}P_{t/t-1}L_t' + Q_{t+1} \quad (10)$$

$$v_t = y_t - Z_t a_{t/t-1} \quad (11)$$

$$K_t = T_{t+1}P_{t/t-1}Z_t'F_t^{-1}$$

$$F_t = Z_t P_{t/t-1} Z_t' + H_t$$

$$L_t = T_{t+1} - K_t Z_t \quad t = 1, \dots, n$$

O Conjunto de equações (9) a (11) deixa claro o processo de aprendizagem que é realizado pelo filtro de Kalman. A cada momento no tempo, a atualização do vetor de estados é feita com base no erro de previsão um passo à frente descontado pelo ganho do filtro (K_t). O ganho do filtro embute os pesos relativos que o modelo atribui ao passado da série e a mais recente observação. Estes pesos, por sua vez, são função das variâncias estimadas da perturbação aleatória presentes nas equações de medida e transição.

A partir da implementação do filtro, utilizamos o método de *disturbance smoother*, o qual é passado em toda a amostra iniciando-se com $r_n = N_n = 0$ e realizando-se as recurssões dadas por (12) e (13).

$$e_t = F_t^{-1} v_t - K_t' r_t \quad (12)$$

$$r_{t-1} = Z_t' F_t^{-1} v_t + L_t r_t \quad (13)$$

$$D_t = F_t^{-1} + K_t' N_t K_t$$

$$N_{t-1} = Z_t' F_t^{-1} Z_t + L_t' N_t L_t \quad t = n, \dots, 1$$

4 Uma explicação mais detalhada do Filtro de Kalman e da forma de estimação dos hiperparâmetros, assim como a derivação das equações de previsão e atualização, pode ser encontrada em Portugal (1993).

Em particular, podemos observar que $H_t e_t$ é a estimativa de ε_t e $Q_t r_{t-1}$ a de η_t com os erros quadráticos médios dados respectivamente por $H_t - H_t D_t H_t'$ e $Q_t - Q_t N_{t-1} Q_t'$. Essas estimativas são também denominadas *smoothed estimates* ou resíduos auxiliares (*auxiliary residuals*). De acordo com Harvey e Koopman (1992) podemos utilizar a simples inspeção gráfica destes resíduos, em conjunto com a análise dos testes de normalidade, como indicativo da existência de quebra estrutural, bem como da presença de *outliers* na série a ser analisada. Na existência de quebras estruturais e *outliers* são realizadas intervenções no modelo com a especificação de variáveis *dummy* para captar o comportamento irregular da série. No presente trabalho utilizamos esta metodologia para a especificação dos períodos de ocorrência da quebra e modelagem das irregularidades apresentadas.⁵ A inclusão das variáveis *dummy* como variáveis explicativas é relativamente simples e requer a utilização das recursões apresentadas por De Jong (1989).⁶

Deve-se destacar, ainda, que mesmo quando o modelo é corretamente especificado os resíduos são serialmente correlacionados e, de acordo com Harvey e Koopman (1992), seguem um processo ARIMA(2,0,0), influenciando, assim, os testes de normalidade apresentados. Contudo, no presente trabalho tais testes são ajustados para a presença autocorrelação serial.

2.2 Redes neurais

As Redes Neurais Artificiais (RNA) surgiram como uma área da informática cujo objetivo básico era o de criar modelos artificiais do cérebro humano, de forma a permitir que computadores "pensem". Em outras palavras, uma RNA é, basicamente, um modelo não paramétrico do cérebro humano.⁷ Uma RNA funciona de forma similar ao cérebro humano, tentando reconhecer padrões e regularidades nos dados que lhe são apresentados. Elas podem aprender por experiência e fazer generalizações com base em seu conhecimento passado.

5 Uma metodologia diferenciada de determinação do período de ocorrência de quebra estruturais pode ser encontrada em Atkinson, Koopman e Shephard (1997).

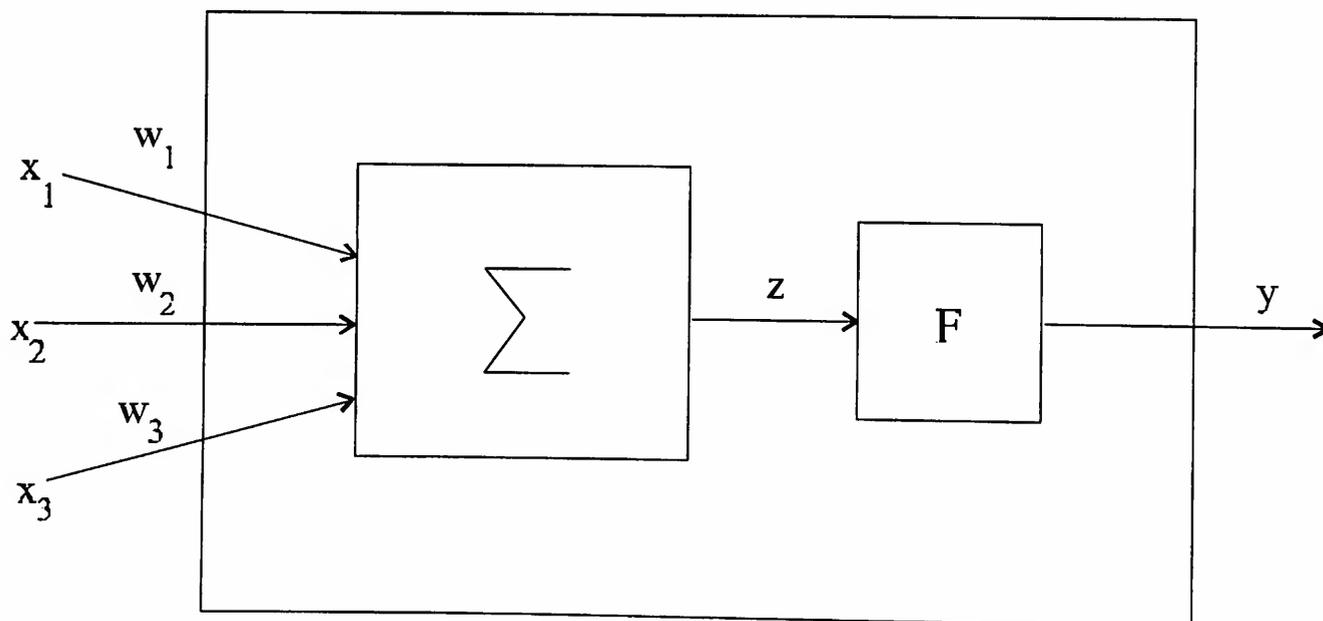
6 Por motivo de espaço omitiremos as mesmas, mas uma análise detalhada pode ser encontrada em Koopman, Harvey, Doornick e Shephard (1995) e Atkinson, Koopman e Shephard (1997).

7 Para uma introdução geral às redes neurais, ver Haykin (1994).

O cérebro humano, com seus bilhões de neurônios interligados em rede, é bem distinto de um computador. As experiências e sensações, que são observadas por meio dos sentidos humanos, geram sinais elétricos que passam pelos neurônios criando elos de ligação entre eles. Estas ligações entre neurônios são novamente sensibilizadas sempre que uma outra experiência ou sensação similar ocorre, permitindo, então, o seu reconhecimento. O aprendizado de algo novo é sempre feito em relação àquilo que já era conhecido, isto é, ele confirma ou modifica as ligações preexistentes entre os neurônios. Na verdade, diferentes experiências ou sensações são armazenadas no cérebro com diferentes ligações, ou diferentes intensidades de ligações, entre neurônios.

Cada neurônio, individualmente, funciona de maneira bastante simples. O neurônio recebe sinais elétricos das outras células do corpo humano por meio de pontos de contato (sinapses), processa estes sinais e, se o resultado deste processamento em um dado período de tempo excede um certo valor, produz outros sinais que são, então, transmitidos a outras células. Aparentemente, não é o funcionamento do neurônio, mas sim o alto grau de ligação entre eles, que é responsável pelo grande poder "computacional" do cérebro humano.⁸

Figura 1



8 Ver, por exemplo, Wasserman (1989, p. 194). O processamento em paralelo efetuado pelo cérebro humano permite que sejam realizadas diversas tarefas de forma simultânea. Neste sentido, é possível, por exemplo, falar e caminhar ao mesmo tempo.

Um neurônio artificial atua exatamente como um neurônio comum, como pode ser visto na Figura 1. O neurônio artificial recebe vários sinais entrada (x), calcula uma média ponderada destes sinais (z), e quando esta média é aplicada a uma função de ativação (F) produz então um sinal de saída (y). Neste sentido, usando o exemplo da Figura 1, temos que

$$z = \sum_{i=1}^3 x_i w_i.$$

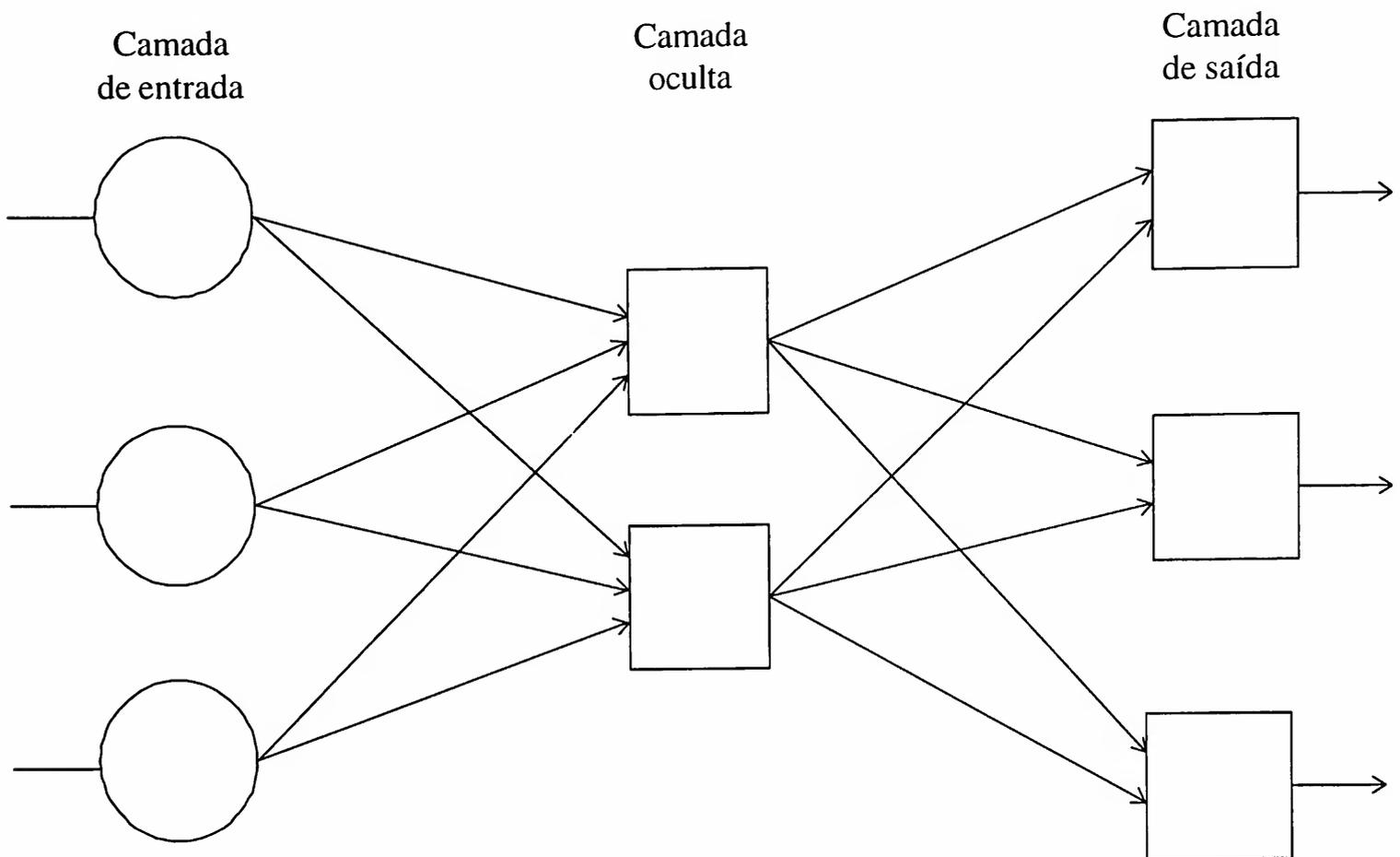
A função de ativação pode, em princípio, tomar diversas formas. Ela pode ser uma função logística, degrau, ou mesmo uma simples função linear. A função mais usada em aplicações práticas é uma função logística do tipo

$$F(z) = \frac{1}{1 + e^{-z}}$$

Uma RNA é composta por um certo número de neurônios artificiais organizados em diferentes camadas, como pode ser visualizado pela Figura 2. A primeira camada, chamada usualmente de camada de entrada, serve apenas para propiciar a entrada dos dados, não executando qualquer computação. É nas camadas intermediárias ou ocultas e na camada de saída que ocorre o efetivo processamento dos dados. A rede mostrada na Figura 2 é usualmente chamada de *feed forward network*, pois apresenta todas as ligações em um só sentido. É possível, contudo, conceber redes em que as informações movimentam-se em várias direções.

A questão principal para operacionalizar uma RNA é como obter os vetores de pesos (w), que correspondem às sinapses que se estabelecem entre os neurônios biológicos. No cérebro humano as sinapses são estabelecidas, e fortalecidas ou não, por meio de um processo de aprendizado. Para uma RNA precisamos também constituir um processo de aprendizado para a determinação dos pesos que ligam os neurônios artificiais. Para a realização de previsões, nós vamos usar apenas algoritmos de aprendizado supervisionado. A idéia básica é treinar a RNA até que esta seja capaz de reconhecer padrões e regularidades nos dados, para então extrapolar quanto ao comportamento futuro.

Figura 2



Para treinar a RNA nós podemos usar um algoritmo de aprendizado supervisionado onde, dado um conjunto de dados de entrada, os pesos são ajustados continuamente de forma a gerar uma saída desejada. O treinamento de uma RNA consiste, portanto, no ajuste dos pesos de forma a produzir um resultado desejado. O algoritmo de treinamento mais usado na literatura é conhecido como *backpropagation*.⁹ Por meio dele os pesos são ajustados de forma a minimizar os erros de previsão dentro do período amostral. Em cada interação os pesos são revisados de acordo com a inclinação da superfície de erros, tentando atingir um mínimo global. As equações (14) a (16) apresentam o algoritmo de *backpropagation*.

$$\Delta w_{ij,k}^n = \eta \delta_{j,k} y_{i,k-1} \quad (14)$$

9 Na verdade, o algoritmo de *backpropagation* é apenas uma das várias alternativas disponíveis para a realização do aprendizado da rede. O maior problema na sua utilização diz respeito à possibilidade de obtenção de mínimos locais. Na prática, contudo, existem diferentes formas de tentar evitar este tipo de ocorrência. Ver, a este respeito, Portugal e Fernandes (1996, p. 68).

$$w_{ij,k}^{n+1} = w_{ij,k}^n + \Delta w_{ij,k}^n \quad (15)$$

$$\delta_{j,k} = \left(\frac{dy}{dz} \right)_{i,k} (y_{i,k} - y_{i,k}^T); \quad (16)$$

para $i = 1, 2, \dots, I$; $j = 1, 2, \dots, J$ e $k = K$

onde; η é o coeficiente de aprendizado (usualmente entre 0.01 e 1.0);

$y_{i,k}$ é a saída do neurônio i na camada k ;

$y_{i,k}^T$ é o valor desejado para $y_{i,k}$;

$w_{ij,k}^n$ é o peso ligando o neurônio i na camada $k-1$ ao neurônio j na camada k no passo n .

Note que o algoritmo acima só pode ser aplicado para a última camada da rede, pois ele pressupõe o conhecimento do valor desejado de saída, informação esta que só está disponível para a última camada ($k = K$). Para as camadas ocultas o valor de $\delta_{i,k}$ na equação (16) tem de ser obtido de outra forma. O algoritmo de *backpropagation* deriva seu nome do fato de propagar para trás o valor de $\delta_{i,k}$ usando os pesos em cada camada, como é indicado pela equação (16a).

$$\delta_{j,k} = \left(\frac{dy}{dz} \right)_{i,k} \left(\sum_j \delta_{j,k} w_{ij,k} \right); \quad (16a)$$

para $k = 1, 2, \dots, K-1$.

Para atingir a convergência para o mínimo de forma mais rápida, pode ser utilizada uma transformação auto-regressiva das equações (14) e (15).¹⁰

$$\Delta w_{ij,k}^{n+1} = \eta (\delta_{j,k} y_{i,k-1}) + \alpha (\Delta w_{ij,k}^n) \quad (17)$$

10 Ver Rumelhart, Hinton e Williams (1986).

$$w_{ij,k}^{n+1} = w_{ij,k}^n + \Delta w_{ij,k}^{n+1} \quad (18)$$

onde α é o *momentum coefficient* (normalmente em torno de 0.9).

3 Dados utilizados e metodologia

A partir destas considerações iniciais sobre a aplicação e uso do filtro de Kalman e das RNAs, vamos iniciar a análise das séries com uma breve apresentação das suas características básicas. Num segundo momento faremos uma discussão a respeito da metodologia de análise empregada para a realização de previsões considerando a ocorrência das quebras estruturais e sua influência na realização destas últimas.

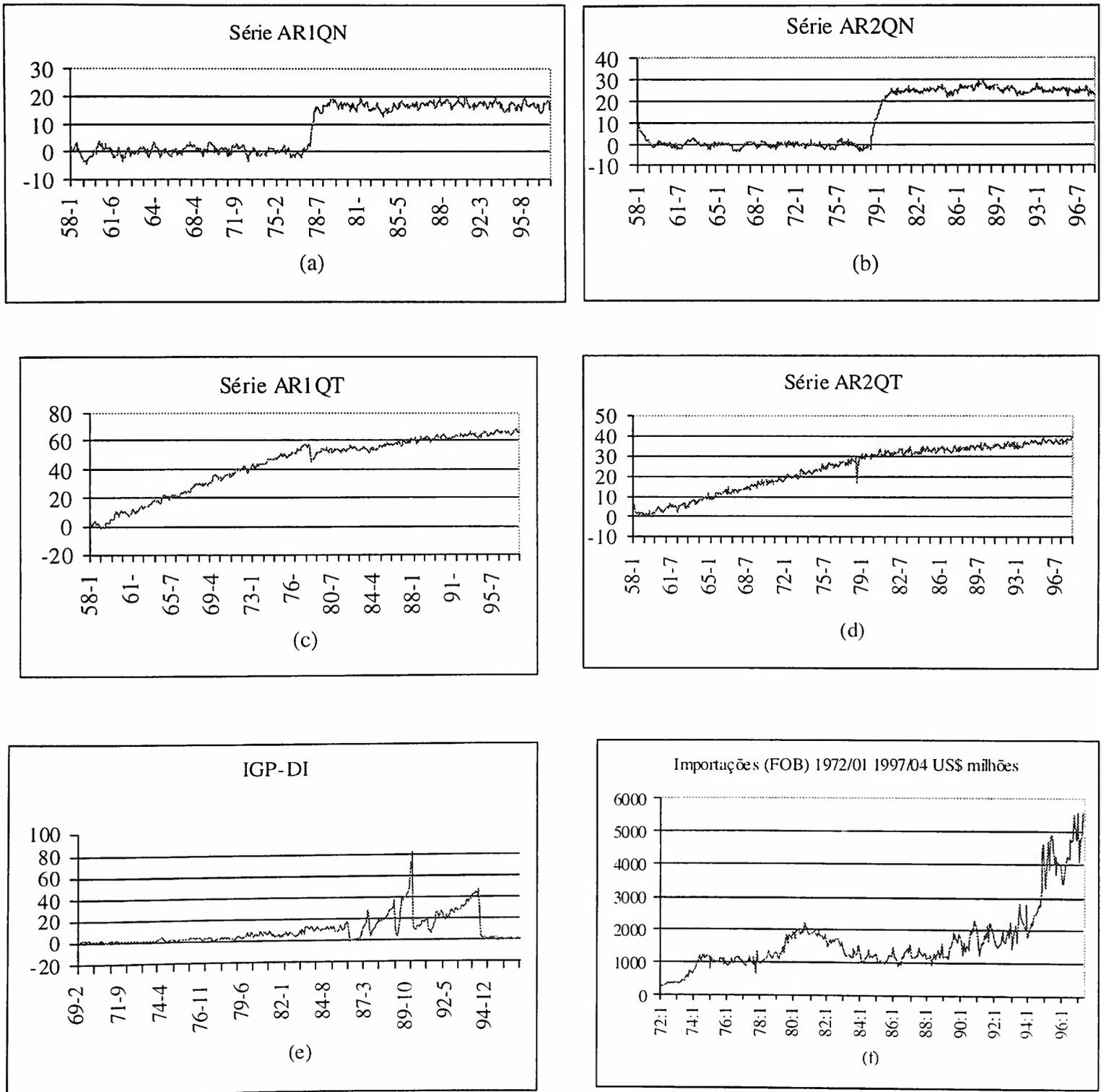
As séries passíveis de análise neste trabalho podem ser subdivididas em dois grupos distintos. O primeiro constitui-se de séries construídas artificialmente a partir de um processo estocástico ARIMA, e o segundo pelas séries denominadas “reais”, as quais correspondem ao total das importações brasileiras, no período compreendido entre janeiro de 1972 e abril de 1997, e a taxa de inflação medida pelo IGP-DI para o período de fevereiro de 1969 até junho de 1987

As séries geradas pelo processo ARIMA procuram emular dois comportamentos distintos, o de quebra estrutural no nível da série e o de quebra na inclinação da tendência. Desta forma temos 4 séries para dois processos ARIMA distintos AR(1) e AR(2). As seis séries estudadas são apresentadas nos Gráficos 1(a) a 1(f).

No que concerne às séries reais, cabe aqui apenas um comentário a respeito da restrição do período de análise do IGP-DI até junho de 1987. Este inclui a primeira tentativa de estabilização econômica com o Plano Cruzado, quando as taxas de inflação apresentaram uma queda substancial nos seus valores de março a dezembro de 1986, com a ocorrência inclusive de uma deflação em abril de 1986. Posteriormente, durante o primeiro semestre de 1987, a inflação retoma a sua trajetória de crescimento anterior, ultrapassando inclusive os valores observados pré-implementação do Plano. Devemos ressaltar aqui, que apesar do Plano Real também representar uma quebra estrutural significativa, como podemos visualizar no Gráfico 1(e), esta se assemelha bastante à ocorrência da quebra estrutural no nível das séries geradas pelos processos ARIMA, onde existe uma abrupta alteração no nível da série e a permanência destes valores em torno da nova média. Neste sentido, a análise comparativa das previsões resultantes da utilização dos modelos estruturais

(MEST) e das RNAs para a inflação pós-Plano Real implicaria apenas uma duplicação dos resultados já obtidos.¹¹

Gráfico 1



¹¹ Além disso, as frequentes quebras estruturais resultantes dos diversos planos de estabilização de preços dificultam a obtenção das estimativas de máxima verossimilhança confiáveis para os hiperparâmetros.

Considerando que o objetivo central deste trabalho consiste em realizar uma avaliação empírica da performance dos MEST e das RNAs na presença de quebras estruturais no DGP (*Data Generation Process*), concentraram-se esforços no sentido determinar o período de ocorrência da quebra na série. A partir desta determinação, realizaram-se previsões procurando investigar a capacidade preditiva e, fundamentalmente, de correção desta previsão com a introdução de novas informações. Nosso estudo baseia-se então na avaliação da capacidade relativa de ambos os modelos em reconhecer alterações no DGP mediante o aprendizado com os erros de previsão passados. No caso dos MEST, este aprendizado ocorre via decomposição dos erros de previsão, e para as RNAs, pelo processo de treinamento da rede com a minimização dos erros de previsão dentro do período amostral em questão.

Em particular, para a determinação do período de quebra foram utilizados os resíduos auxiliares gerados pelo Filtro de Kalman. Valores padronizados destes resíduos maiores do que 2 em módulo são considerados como um indicativo da presença de quebras estruturais, como discutido anteriormente.¹² A validade desta metodologia de determinação do período de ocorrência da quebra estrutural no DGP pode ser avaliada considerando-se as quatro séries geradas pelos processos ARIMA, onde evidentemente os períodos de quebra são conhecidos. Na Tabela 1 são apresentados os resultados obtidos para a determinação dos pontos de quebra.¹³

Tabela 1

Série	Modelo Especificado	Ponto de Quebra Estimado ¹	Ponto de Ocorrência da Quebra
AR2QN	Tendência Fixa	1978/07	1978/09
AR1QT	Tendência Fixa	1978/06	1978/06
AR1QN	Tendência Fixa	1977/12	1977/12
AR2QT	Tendência Fixa	1978/06	1978/06

Notas: (1) Para a determinação do ponto de quebra utilizamos o maior valor absoluto dos resíduos auxiliares estimados.

12 Nas estimações deste artigo foram utilizados os pacotes econométricos STAMP 5.0, para a estimação dos MEST, e RATS 4.0, para estimação das RNAs.

13 Uma apresentação mais detalhada dos modelos utilizados e dos testes de especificação é apresentada na seção 4.

Fica claro, portanto, a eficácia empírica da metodologia aplicada na determinação dos períodos de ocorrência de quebra. Contudo, existe a necessidade de considerarmos a simplicidade dos modelos estimados e o padrão relativamente regular no comportamento dos dados, à exceção da ocorrência da quebra. Dada a simplicidade dos modelos, fica relativamente facilitada a determinação do período de quebra pela decomposição dos resíduos auxiliares em irregulares e de nível, uma vez que a tendência é fixa. Não obstante a maior complexidade do modelo estrutural básico (BSM) estimado para a série de importações, os resultados se mostraram compatíveis com as informações externas disponíveis sobre a data de quebra, que neste caso não são conhecidas com certeza. Na verdade, quanto mais sofisticado for o modelo maior a dificuldade em se diferenciar entre os vários tipos de quebra estrutural. Os valores dos resíduos auxiliares para as importações indicaram uma quebra no nível em novembro de 1994 e outra em julho de 1995, além de uma quebra na inclinação da tendência em abril de 1992, resultados coerentes com o comportamento das importações brasileiras diante das alterações ocorridas em seu volume e valores a partir da abertura comercial iniciada no começo da década de 90.

Por fim, antes de nos determos na discussão dos resultados obtidos com as previsões, algumas observações se tornam importantes com relação à forma de realização das mesmas. Em particular, para os modelos estruturais a ocorrência da quebra é considerada em uma análise de previsão efetuada dentro da amostra, isto é, a estimação do vetor de estados é feita para todo o período amostral. Neste caso podemos efetuar as previsões de duas maneiras distintas. A primeira consiste em produzir estimativas para os valores da série a partir do ponto de ocorrência da quebra sem utilizarmos o processo de intervenção no modelo, que consiste basicamente na introdução de variáveis *dummy* na tentativa de captar a alteração no DGP, a segunda considerando a referida intervenção. Podemos definir três tipos de variáveis *dummy* procurando captar, respectivamente, alterações causadas pela presença de *outliers*, quebra no nível e quebra na inclinação da tendência.

A título de validar a necessidade de intervenção nos modelos estimados, com vistas à realização das previsões, investigou-se o comportamento do Erro Absoluto Médio (EAM) para ambas as situações. Embora, como veremos adiante, a inclusão das variáveis *dummy* melhore substancialmente os resultados dos testes de especificação dos modelos estimados, principalmente o teste para normalidade dos resíduos, elas têm pouco efeito em termos de redução do EAM. O EAM foi calculado a partir das previsões um passo à frente geradas por cada modelo estimado.

$$\text{impulse dummy } D = \begin{cases} 1 & t = t^* \\ 0 & t \neq t^* \end{cases}$$

$$\text{step dummy } D = \begin{cases} 1 & t \geq t^* \\ 0 & t < t^* \end{cases}$$

$$\text{staircase dummy } D = \begin{cases} (T - t^*) & t \geq t^* \\ 0 & t < t^* \end{cases}$$

De maneira semelhante, consideramos o EAM obtido com as previsões realizadas pelas RNAs como fator de comparação entre a capacidade preditiva de ambos os modelos (RNA e MEST). Em particular, para as RNAs realizamos as previsões um passo à frente treinando inicialmente com toda a amostra, de tal forma que a rede pudesse reconhecer o padrão de mudança no DGP. Deste modo podemos estabelecer uma comparação entre o desempenho de ambos os modelos na capacidade de reconhecer as alterações ocorridas.

4 Resultados empíricos

Dando prosseguimento à análise das séries “reais” e “construídas” vamos considerar inicialmente os resultados obtidos com as previsões para as séries geradas pelos processos ARIMA. Para facilitar a compreensão denominamos as séries em questão com o DGP utilizado na construção de cada uma. Neste caso temos AR1QN para a série gerada por um processo auto-regressivo de ordem 1 com uma quebra no nível, AR1QT para um processo auto-regressivo de ordem 1 com quebra na inclinação da tendência, e assim por diante. Apresentamos na Tabela 2 os resultados obtidos para a estimação dos diferentes modelos estruturais para as respectivas séries.

Como mencionado anteriormente, podemos observar que a utilização de intervenções nos MEST não melhora significativamente a sua capacidade de previsão, alterando, porém, a especificação do mesmo. A título de exemplificação apresentamos no Gráfico 2 as previsões um passo à frente com intervenção (2 a) e sem intervenção (2b), intervenção para a variável AR1QN, onde claramente não se observam diferenças, demonstrando que a ordem de grandeza desta é mínima.

Tabela 2

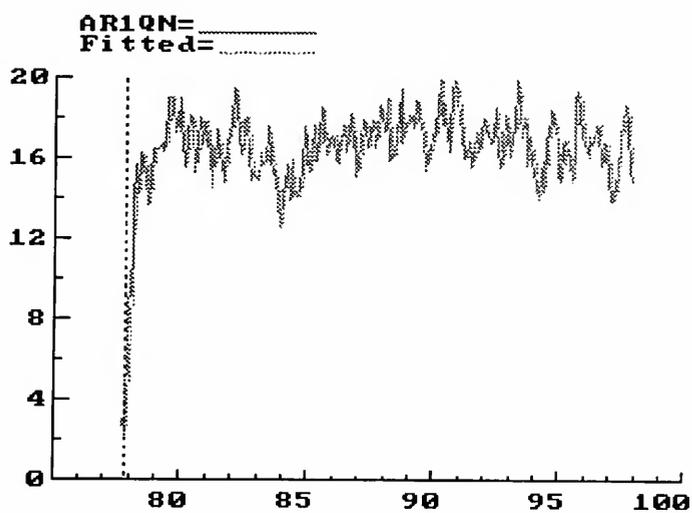
Variável	Testes de Especificação ¹	Modelo Escolhido	Critério de Ajuste	Erros Absoluto Médio	Coefficiente da Variável de intervenção ²
AR1QN Sem intervenção	Normalidade=13.48 H(159) = 0.96338 Q(20,19) = 32.55	Tendência Fixa $\sigma_\zeta = 0$	$R^2 = 0.9817$ $R_d^2 = 0.0141$ $R_s^2 = -0.013$	EAM=0.962414 t = t* em 77/11	
Com intervenção	Normalidade=2.735 H(159) = 0.96604 Q(20,19) = 39.56	Tendência Fixa $\sigma_\zeta = 0$	$R^2 = 0.9825$ $R_d^2 = 0.0737$ $R_s^2 = 0.0473$	EAM=0.933858 t = t* em 77/11	Nível 1977/12 6.1171 (5.5916)
AR2QN Sem intervenção	Normalidade=32.49 H(159) = 1.353 Q(20,19) = 107.4	Tendência Fixa $\sigma_\zeta = 0$	$R^2 = 0.9907$ $R_d^2 = 0.2179$ $R_s^2 = 0.2070$	EAM=0.99791 t = t* em 78/10	
Com intervenção	Normalidade=9.710 H(155) = 1.373 Q(20,18) = 61.70	Tendência Fixa $\sigma_\zeta = 0$	$R^2 = 0.9916$ $R_d^2 = 0.2702$ $R_s^2 = 0.2618$	EAM=0.991097 t = t* em 78/10	Nível 1978/09 4.950 (5.0934)
AR1QT Sem intervenção	Normalidade=3873 H(159) = 0.9668 Q(20,19) = 34.84	Tendência Fixa $\sigma_\zeta = 0$	$R^2 = 0.9963$ $R_d^2 = 0.0133$ $R_s^2 = -0.0108$	EAM=0.9167584 t = t* em 78/7	
Com intervenção	Normalidade=3.885 H(159) = 0.9663 Q(20,19) = 45.31	Tendência Fixa $\sigma_\zeta = 0$	$R^2 = 0.9969$ $R_d^2 = 0.1934$ $R_s^2 = 0.1737$	EAM=0.9267144 t = t* em 78/7	Nível 1978/6 -11.033 (-10.23) Inclinação 1979/9 -0.1939 (-2.3945)
AR2QT Sem intervenção	Normalidade=4890 H(159) = 0.8733 Q(20,19) = 53.00	Tendência Fixa $\sigma_\zeta = 0$	$R^2 = 0.9882$ $R_d^2 = 0.4183$ $R_s^2 = 0.3980$	EAM=0.950283 t = t* em 78/7	
Com intervenção	Normalidade=1.366 H(159) = 0.8780 Q(20,19) = 54.96	Tendência Fixa $\sigma_\zeta = 0$	$R^2 = 0.9906$ $R_d^2 = 0.5363$ $R_s^2 = 0.5201$	EAM=0.928755 t = t* em 78/7	Irregular 1978/6 -11.789 (-11.057)

Notas: (1) Os testes aqui apresentados correspondem aos testes de Bowman-Shenton para normalidade, que tem distribuição χ^2_2 ; H(h) é o teste para heterocedasticidade, com distribuição aproximadamente F(h,h); Q(p,d) é o teste Q de Box-Ljung baseado nas primeiras p autocorrelações dos resíduos com distribuição aproximadamente χ^2_{d-1} .
(2) Os valores entre parênteses indicam a estatística t dos coeficientes.

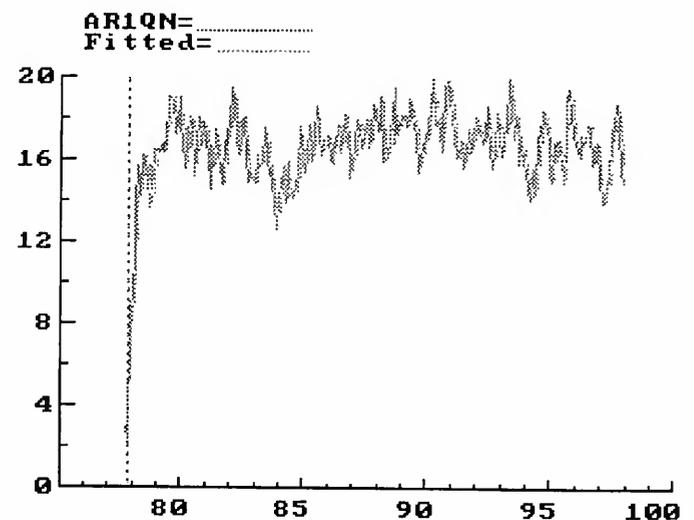
Para as séries geradas a partir de processos ARIMA observamos que o EAM chega inclusive a ser ligeiramente inferior quando não utilizamos intervenção, apesar de uma deterioração sensível dos testes de especificação do modelo, principalmente no que diz

respeito aos testes de normalidade dos resíduos. Os critérios de ajuste R_d^2 e R_s^2 ¹⁴ apresentam ganhos substanciais nos modelos com intervenção, em relação aos modelos sem intervenção. Contudo, devemos levar em conta que estes critérios representam resultados calculados dentro da amostra, enquanto que o EAM constitui um critério de análise baseado em extrapolações realizadas pelas previsões um passo à frente. De acordo com Harvey (1989), não obstante os resultados obtidos com o EAM, os modelos sem intervenção devem ser preteridos em relação aos modelos com intervenção, com base nos resultados obtidos nos testes de especificação do modelo. Em particular, os critérios de ajuste R_d^2 e R_s^2 também devem ser considerados. Neste caso modelos que apresentem ambos os critérios de ajuste com valores negativos devem ser rejeitados em favor da especificação mais simples apresentada na Tabela 2 seguindo critérios de parcimônia.

Gráfico 2



(a)



(b)

A boa performance dos modelos estruturais sem a utilização de intervenção demonstra sua poderosa capacidade preditiva, a qual está intimamente relacionada às equações de atualização e previsão bem como com o ganho do filtro, uma vez que estas envolvem parâmetros que são variáveis no tempo. Neste sentido, os modelos estruturais se mostraram capazes de prever e detectar corretamente os pontos de mudanças estruturais sem a necessidade de uma intervenção subjetiva no modelo.

14 Os critérios de ajuste R_d^2 e R_s^2 são definidos, respectivamente, por $(1 - SSE) / \sum_{t=2}^T (\Delta y_t - \hat{\Delta y}_t)^2$ e $1 - SSE/SSDSM$, onde SSE é definido como a soma dos quadrados dos resíduos e SSDSM é a soma dos quadrados dos resíduos da primeira diferença em torno das médias sazonais ($\Delta y_t = \sum_{j=1}^4 \gamma_j^* Z_{tj} + \eta_t, t = 2, \dots, T$).

Com relação às séries de importações e de inflação, o resultado se mostrou bastante diverso do anterior, como podemos observar na Tabela 3. Considerando-se inicialmente a série de importações, podemos observar uma evidente melhora nos testes de especificação do modelo. Novamente esta melhora está relacionada com o teste para normalidade dos resíduos e com critérios de ajuste para a especificação do modelo BSM com tendência fixa e a realização de uma intervenção em 94/09 no nível e em 92/04 na inclinação da tendência. Os resultados apontaram ainda para um ganho marginal em termos de EAM com a especificação das intervenções. Esta conclusão fica evidenciada quando consideramos as previsões um passo à frente para ambos os modelos que são apresentadas no Gráfico 3, onde temos o Gráfico 3(b) representando o modelo com intervenção e novamente reforçando a hipótese de independência destas com a capacidade preditiva do modelo.

Tabela 3

Variável	Testes de Especificação ¹	Modelo Escolhido	Critério de Ajuste	Erros Absoluto Médio	Coefficiente da Variável de intervenção ²
Importações Sem intervenção	Normalidade=283.2 H(97) = 7.621 Q (16,13) = 40.36 Chi ² (11) = 81.5152	Estrutural Básico (BSM)	$R^2 = 0.9500$ $R_{dl}^2 = 0.3442$ $R_s^2 = 0.1917$	EAM=459.062 t = t* em 94/11	
Com intervenção	Normalidade=175.7 H(97) = 7.305 Q (16,14) = 34.18 Chi ² (11) = 82.6465	Estrutural Básico com tendência fixa (BSM)	$R^2 = 0.9528$ $R_{dl}^2 = 0.3815$ $R_s^2 = 0.2377$	EAM=457.8173 t = t* em 94/11	Inclinação 47.055 em 92/4 (3.0982) Nível 572.00 em 94/9 (2.9565)
Com intervenção	Normalidade=202.3 H(97) = 7.160 Q (16,14) = 33.852 Chi ² (11) = 86.3604	Estrutural Básico com tendência fixa (BSM)	$R^2 = 0.9532$ $R_{dl}^2 = 0.3866$ $R_s^2 = 0.2440$	EAM=353.807 t = t* em 95/08	Inclinação 53.955 em 92/4 (3.3444) Nível 533.97 em 94/9 (2.7338) Nível -346.50 em 95/4 (-1.770)
IGP-DI Sem intervenção	Normalidade=294.5 H(73) = 18.15 Q (13,12) = 25.62	Nível Local <i>Random Walk plus Noise</i>	$R^2 = 0.8277$ $R_{dl}^2 = 0.0036$ $R_s^2 = -0.09178$	EAM=2.598021 t = t* em 86/04	
Com intervenção	Normalidade=345.9 H(73) = 14.22 Q (13,12) = 13.76	Nível Local <i>Random Walk plus Noise</i>	$R^2 = 0.8593$ $R_{dl}^2 = 0.1864$ $R_s^2 = 0.1085$	EAM=2.705026 t = t* em 86/04	<i>Dummy</i> para o período 86/03 - 86/12 -9.3261 (-8.3614)

Notas: (1) Chi²(s-1) é o teste de sazonalidade com distribuição aproximadamente χ_{s-1}^2 .

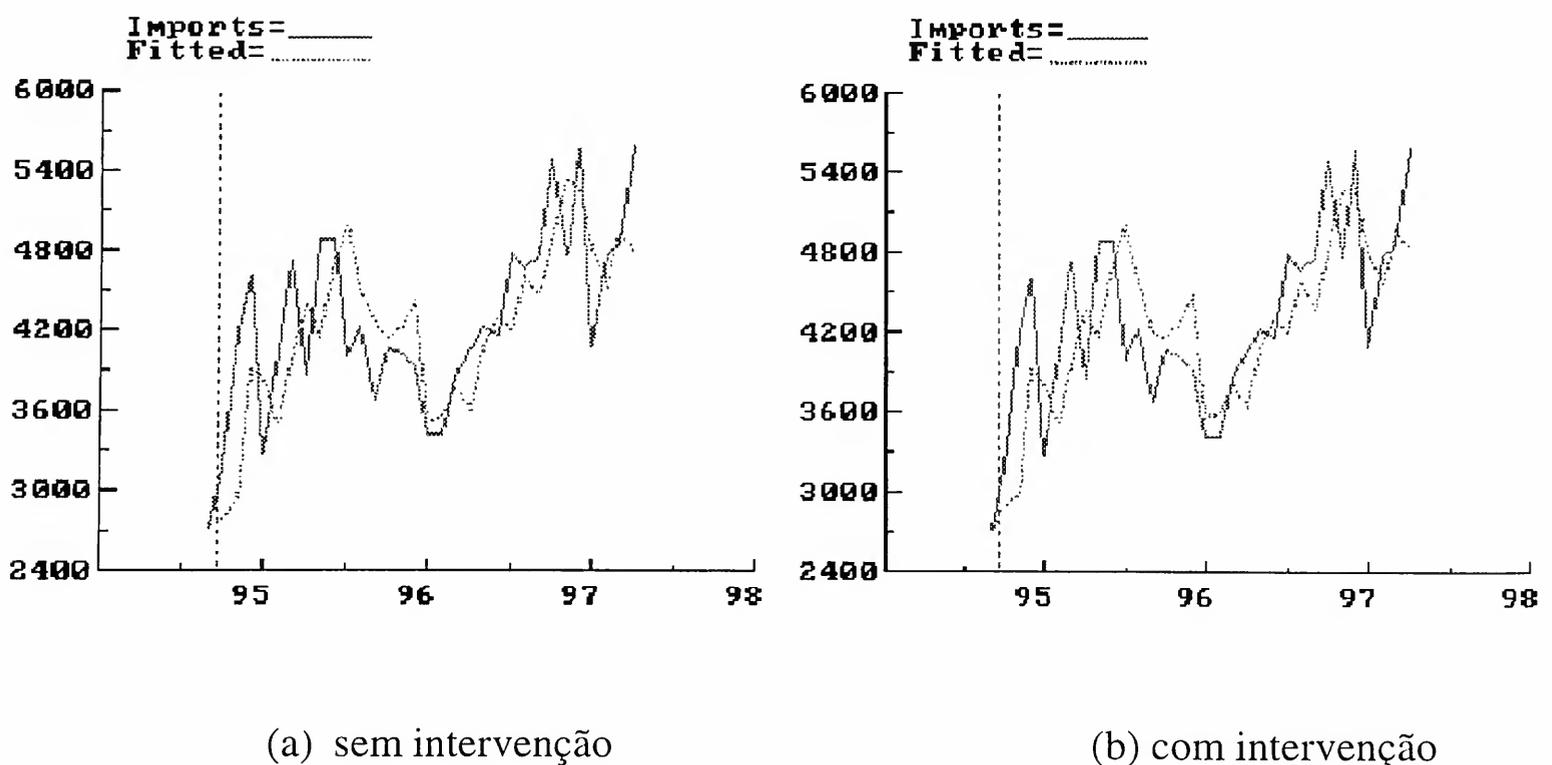
(2) Os valores entre parênteses indicam a estatística t dos coeficientes.

Não obstante a melhora no teste de normalidade dos resíduos, como ocorrera com as séries geradas por processos ARIMA, esta não se deu na mesma dimensão daquelas, no sentido em que o modelo especificado não é capaz de satisfazer este pressuposto básico.

Em particular, isto se deve ao comportamento bastante irregular da série de importações, observado a partir do início da década de 90. Nos Gráficos 4 apresentamos os componentes sazonal 4(b), irregular 4(a) e de tendência 4(c) para a série de importações.

Claramente a decomposição apresentada pelos modelos estruturais nos permite visualizar as mudanças ocorridas na série de importações, com destaque para o forte aumento da sazonalidade da série a partir da década de noventa, atingindo um patamar de 500 milhões de dólares, e a abrupta mudança no nível da série ocorrida em meados da década de 90 para algo em torno de 4 bilhões de dólares. Estas mudanças na série de importações estão intimamente ligadas à liberalização comercial que foi deflagrada a partir dos anos 90. O controle quantitativo das importação implicava não apenas em uma redução artificial do nível, mas também das características sazonais da série.

Gráfico 3



O reflexo destas alterações encontra-se na distribuição de frequência dos resíduos do modelo especificado, onde observamos a ocorrência de valores em módulo muito distantes da média esperada destes com destacada frequência, violando o pressuposto de normalidade dos mesmos. A título de exemplo apresentamos no Gráfico 4 as distribuições de frequência dos resíduos para o modelo em questão 4(e), e para a série AR1QN 4(d), a qual apresentou resíduos seguindo, aproximadamente, uma distribuição normal.

Tabela 4

Variável	EAM / EPAM ¹ Redes Neurais	EAM/EPAM Modelos Estruturais	Período Observado	Configuração Básica da Rede ²
AR2QN	1,0111 / 4.32%	0.9911 / 4.28%	78/10 - 97/12	y_{t-1} y_{t-2} , <i>dummy</i> , 4 neurônios na camada oculta
AR1QT	0.9910 / 1.70%	0.9313 / 1.57%	77/12 - 97/12	y_{t-1} , <i>dummy</i> , 4 neurônios na camada oculta
AR1QN	0.8685 / 5.48%	0.9354 / 5.79%	77/12 - 97/12	y_{t-1} , <i>step dummy</i> , 5 neurônios na camada oculta
AR2QT	1.1003 / 3.19%	0.9287 / 2.87%	78/07 - 97/12	y_{t-1} , tendência, <i>staircase dummy</i> , 3 neurônios na camada oculta
Importações	544.236 / 11.87%	457.8173 / 10.29%	74/10 - 97/04	y_{t-1} , tendência, <i>step dummy</i> , <i>dummy</i> Sazonais (11), 3 neurônios na camada oculta
Inflação	2.0289 / 60.85%	2.5246 / 110.97%	86/04 - 87/04	y_{t-1} y_{t-2} , <i>dummy</i> para o período 86/3-86/12
Média dos EPAM/EQMP	14.57%	22.63%		
Média dos EPAM/EQMP sem o IGP-DI	5.31%	4,96%		

Notas: (1) As abreviações correspondem, respectivamente, ao Erro Absoluto Médio e ao Erro Porcentual Absoluto Médio.

(2) As variáveis *dummy* inseridas no processo de treinamento e geração de resultados na rede seguem os resultados obtidos com os modelos estruturais. A função de ativação utilizada foi sempre uma logística.

Gráfico 4

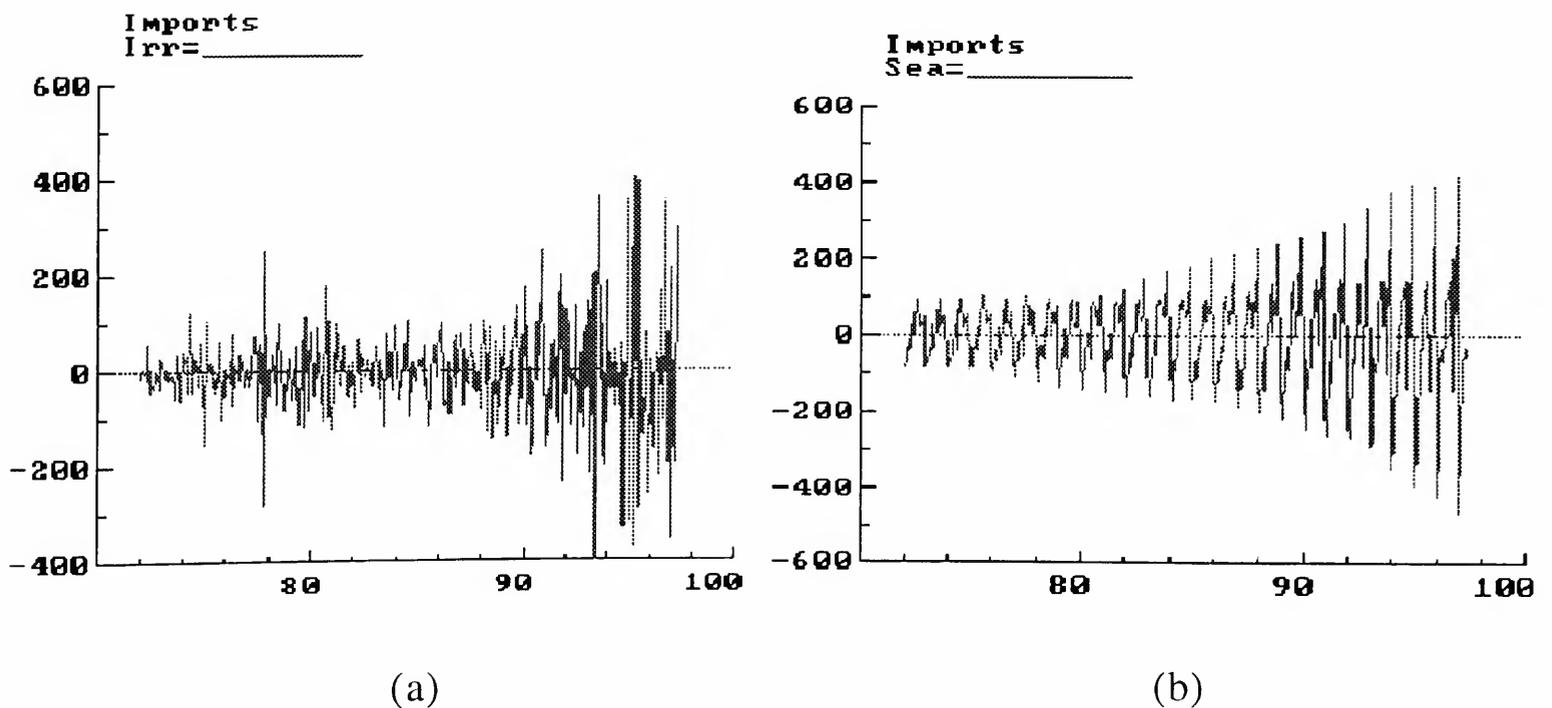
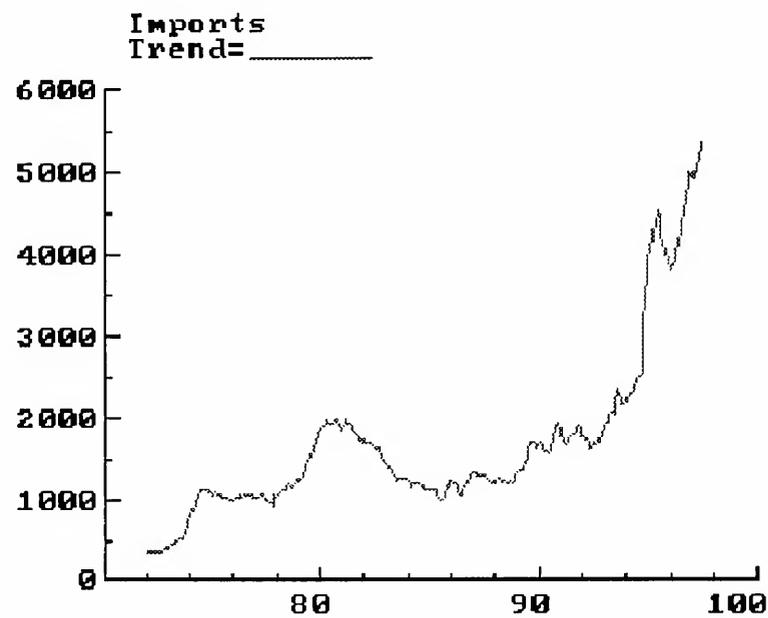
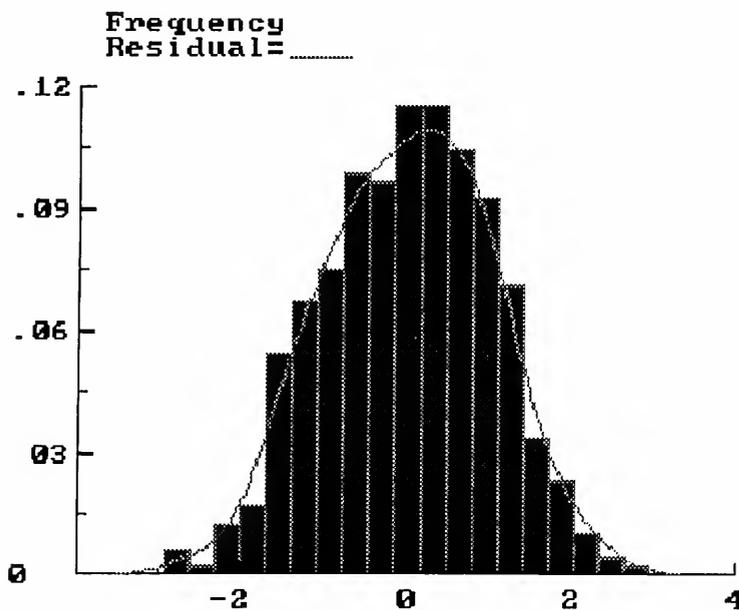


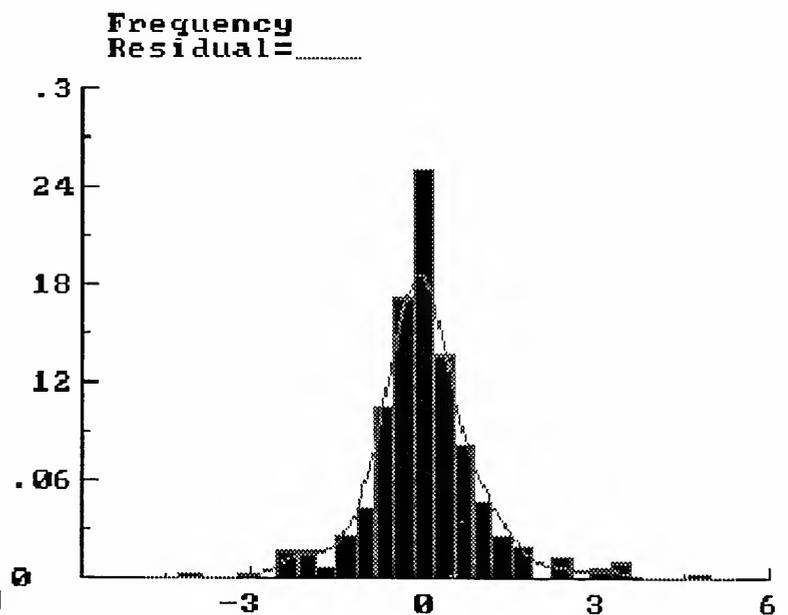
Gráfico 4 (Continuação)



(c)



(d)



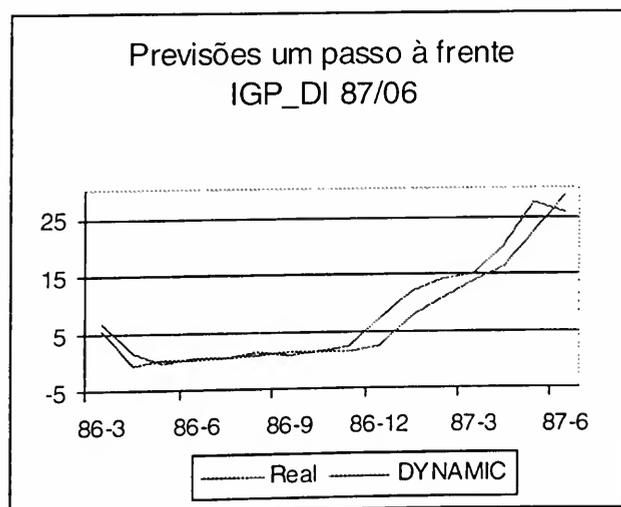
(e)

Apesar das irregularidades presentes na série, principalmente no que diz respeito ao nível da mesma, que apresenta um significativo aumento a partir de meados de 1995, a especificação de uma nova intervenção no nível em 1995/04 não resultou em melhoras nos testes de especificação, sendo que as melhoras obtidas nos critérios de ajuste se mostraram marginais. Desta forma podemos dizer que a esta nova especificação do modelo estrutural básico com intervenção mostrou-se inócua.

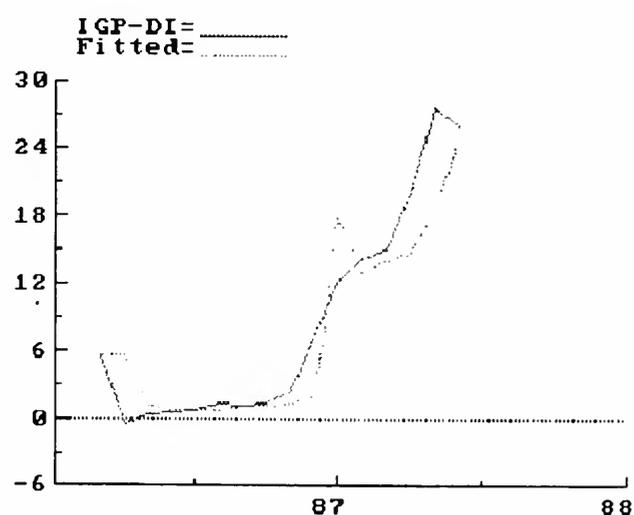
Para a série de inflação observamos um movimento distinto com a especificação do modelo com intervenção, representado pela variável *dummy* no período de março a dezembro de 1986. Neste caso observou-se uma piora na especificação dos modelos em termos dos testes de normalidade, contrapondo-se à sensível melhora observada no critério de ajuste, o qual, por sua vez, indica a rejeição do modelo inicial sem intervenção em favor do modelo com intervenção, seguindo os critérios de parcimônia, como discutido anteriormente. Novamente observamos uma perda em termos de EAM para o modelo especificado com intervenção; contudo, esta se mostrou justificável devido aos resultados obtidos com o critério de ajuste.

No que concerne à avaliação dos resultados obtidos pelas RNAs, que são modelos não-lineares com estimadores não paramétricos, e aqueles gerados pelo MEST efetuamos uma comparação com base no EPAM. Os resultados desta comparação são apresentados na Tabela 4, juntamente com a arquitetura ótima obtida para a rede neural. Levou-se em consideração as previsões um passo à frente realizadas pelas diversas configurações das redes para as mesmas variáveis e para as mesmas datas de início da geração de resultados. Devemos ressaltar que as comparações entre as RNAs e os MEST devem se limitar à capacidade preditiva de cada modelo, uma vez que para as RNAs não existe um conjunto de testes de especificação a ser utilizado para se obter a parametrização ótima do modelo. Neste sentido, para obtermos a arquitetura ótima da rede podemos utilizar características observadas a partir da estimação dos modelos estruturais, bem como do DGP utilizado para a geração das séries seguindo processos ARIMA. Outra característica que deve ser ressaltada é a de que as RNAs são modelos não-lineares, em contraposição à linearidade imposta pelo filtro de Kalman, o que não permite, portanto, maiores comparações entre as especificações dos modelos presentes em um ou outro.

Gráfico 5



(a) RNA



(b) MEST

É importante considerar ainda que para as RNAs a previsão um passo à frente é caracterizada pelo treinamento da rede na totalidade do período amostral, o que permite que esta reconheça o padrão de mudanças no DGP. Também para os modelos estruturais, as previsões são baseadas na totalidade do período amostral para o algoritmo, com a estimação do vetor de estados realizada por meio do processo de alisamento (*smoothing*). Para os MEST as previsões um passo à frente são obtidas a partir do ponto determinado de quebra com a incorporação de novas observações a cada ponto no tempo. Este procedimento possui semelhanças com o processo de treinamento da rede, o qual é realizado para todo o período amostral, e a partir da fixação dos pesos atribuídos para cada sinapse procede-se a geração do vetor de resultados (camada de saída), com a incorporação da nova informação ocorrendo com a inclusão de uma observação a mais no vetor de entradas da rede (camada de entrada).

Tabela 5
Erros Absoluto e Porcentual Médio Posteriores à Quebra (3 períodos)

Variável	EAM / EPAM	EAM / EPAM	Períodos	
	Redes Neurais	Modelos Estruturais		
AR2QN	1.8265 / 16.47%	2.8846 / 26.68%	78/10	78/12
AR1QT	0.7781 / 1.64%	0.8520 / 1.80%	78/07	78/09
AR1QN	2.1907 / 25.36%	0.8186 / 14.28%	77/12	78/02
AR2QT	1.3980 / 5.17%	1.6555 / 5.84%	78/07	78/09
Importações	596.2881 / 14.66%	735.9333 / 17.82%	94/10	94/12
Inflação	0.8853 / 173.51%	2.3763 / 448.98%	86/04	86/06
Média dos EPAM	39.47%	85.90%		
Média dos EPAM sem o IGP-DI	12.66%	13.28%		

Os resultados apresentados na Tabela 4 indicam que a diferença medida em termos de erros percentuais entre as RNAs e os MEST não se mostrou significativa, com exceção da série de inflação, o que nos encorajou a detalhar mais a capacidade de previsão de ambos os modelos a partir da ocorrência da quebra. Para isto apresentamos nas Tabelas 5 e 6 o EAM e o Erro Porcentual Absoluto Médio (EPAM) para, respectivamente, as três e seis primeiras observações posteriores à data de ocorrência da quebra. Levando-se em consideração as três observações posteriores à quebra, observamos um desempenho superior das RNAs em cinco das seis séries estudadas. A única cujo resultado dos MEST se mostrou superior foi a série AR1QN, com uma diferença pequena, porém não desprezível, em termos percentuais. Considerando-se os resultados obtidos para as seis observações posteriores à quebra, apresentados na Tabela 6, não é mantido o comportamento anterior, com as RNAs mostrando-se superiores em apenas 3 das seis séries apresentadas. Ainda assim, para as três séries onde os MEST são superiores, AR2QT,

importações e AR1QN, os ganhos em termos percentuais se mostram marginais. Esses resultados demonstram uma melhor capacidade das RNAs em perceber alterações no DGP nos períodos imediatamente posteriores à quebra, utilizando-se da sua reconhecida capacidade de aprender por experiência e fazer generalizações com base no seu conhecimento passado.

Tabela 6
Erros Absoluto e Porcentual Médio Posteriores à Quebra (6 períodos)

Variável	EAM / EPAM Redes Neurais	EAM / EPAM Modelos Estruturais	Período
AR2QN	2.2931 / 17.72%	2.5411 / 20.95%	78/10 - 79/03
AR1QT	0.9513 / 1.67%	1.069 / 2.18%	78/07 78/12
AR1QN	1.6344 / 16.33%	1.5100 / 15.14%	77/12 - 78/05
AR2QT	1.3905 / 4.86%	1.3504 / 4.76%	78/07 78/12
Importações	702.3402 / 17.80%	673.0333 / 16.15%	94/10 - 95/03
Inflação	0.6207 / 102.01%	1.3169 / 234.72%	86/04 - 86/09
Média dos EPAM	26.73%	48.98%	
Média dos EPAM sem o IGP-DI	11.67%	11.83%	

Em particular, para a série de inflação observamos um EPAM extremamente alto, tanto para as previsões dos três primeiros períodos quanto para aquelas geradas para os seis períodos após a quebra. Estes valores altos se deveram à ocorrência de uma deflação em abril de 1986, a qual representou uma queda muito abrupta no valor da série, visto que em fevereiro de 1986 a inflação alcançava os 14%. Os resultados obtidos demonstram uma evidente incapacidade de ambos os modelos em conseguir prever este tipo de alteração no DGP. Tal fato é representado pela existência de um erro de previsão extremamente alto para abril de 1986, como pode ser visualizado no Gráfico 5, onde são apresentadas as previsões um passo à frente para a série de inflação para as RNAs e para os MEST.

Quando observamos apenas a média dos resultados obtidos pelos modelos em consideração nas Tabelas 5 e 6, notamos uma sensível superioridade das RNAs em relação aos MEST, independente do período analisado. Contudo, este resultado está enviesado pela série de inflação, onde a ocorrência da deflação mencionada acima reduz severamente a capacidade de previsão de ambos os modelos. Ao excluirmos a série de inflação os resultados mostram uma equivalência entre as RNAs e os MEST para a previsão na totalidade do período, bem como para os períodos imediatamente posteriores à ocorrência da quebra, como exposto nas Tabelas 4, 5 e 6. Para o período de previsão como um todo os MEST mostram-se marginalmente superiores, enquanto que para as previsões próximas do momento da quebra são as RNAs que se mostram marginalmente superiores.

5 Conclusão

Neste trabalho procurou-se investigar o comportamento dos modelos estruturais e das RNAs com relação à sua capacidade em reconhecer mudanças estruturais no DGP

Os MEST conseguiram realizar previsões de maneira eficiente, ainda que desconsiderando o processo de intervenção, resultando em uma independência entre capacidade de previsão e intervenções nos modelos especificados. Em outras palavras, o processo de atualização do vetor de estados a cada período de tempo, que é realizado pelo filtro de Kalman, foi capaz de ajustar muito rapidamente o modelo à quebra estrutural sem a necessidade de uma intervenção subjetiva. Contudo, esta independência presente nos MEST tem como resultado, para todas as variáveis analisadas, uma sensível perda em termos de má especificação do modelo, o que desencoraja a utilização dos MEST sem intervenção em um estudo comparativo com outros modelos, ainda que não-paramétricos.

Considerando-se as séries de maneira isolada, nossos resultados indicam ainda uma leve superioridade na capacidade de previsão e, portanto, reconhecimento da alteração no DGP, por parte das RNAs nos momentos imediatamente posteriores à ocorrência da quebra estrutural. No momento da quebra as RNAs geram melhores previsões para 5 das 6 séries utilizadas. Contudo, este resultado não se mantém quando estendemos as previsões para a totalidade do período, onde os ganhos obtidos pelos MEST em relação às redes é marginal. Este resultado é válido para todas as séries analisadas neste trabalho, à exceção da série de inflação. Nesta última a ocorrência de uma deflação em abril de 1986 produz uma quebra estrutural no nível da série de tal ordem que ambos os modelos se mostraram incapazes de percebê-la, resultando em uma perda não desprezível na capacidade de previsão. Apesar deste fenômeno, as RNAs se mostraram capazes de se ajustar à ocorrência da quebra estrutural na série de inflação, de maneira muito superior aos MEST. Neste sentido, podemos concluir que as RNAs conseguem identificar mais rapidamente as quebras estruturais.

Além dos MEST e das redes neurais, a questão da mudança estrutural poderia ser também modelada por meio de mudanças de regime por cadeias de Markov (Markov Switching). Uma apresentação desta metodologia com exemplos ilustrativos pode ser encontrada em Hamilton (1994, cap. 22).

Por fim, é preciso considerar uma limitação da análise presente neste trabalho com relação ao número de séries modeladas. A extensão das conclusões obtidas exigiria a realização de um experimento de Monte Carlo. Contudo, a realização deste tipo de

experimento foge ao escopo inicial deste trabalho, que procura incluir em sua análise, não somente séries construídas, mas também séries observadas. A realização de um experimento de Monte Carlo fica, portanto, como uma sugestão para um trabalho posterior.

Referências bibliográficas

- Alexander, I. and Morton, H. *An introduction to neural computing*. London: Chapman and Hall, 1990.
- Atkinson, A. C., Koopman, S. J. and Shephard, N. Detecting shocks: outliers and breaks in time series. *Journal of Econometrics*, 80, p. 387-422, 1997
- Box, G. E. P. & Jenkins, G. M. *Time series analysis, forecasting and control*. San Francisco: Holden-Day, 1976.
- De Jong, P. Smoothing and interpolation with the state space model. *Journal of the American Statistical Association*, 84, 1085-1088, 1989.
- Diderrich, G. T. The Kalman filter from the perspective of Goldberg-Theil estimators. *The American Statistician*, 39, 193-198, 1985.
- Duncan, D. B. and Horn, S. D. Linear dynamic estimation from the viewpoint of regression analysis. *Journal of the American Statistical Association*, 67, 815-821, 1972.
- Fernandes, L. G. L., Navaux, P. O. A. & Portugal, M. S. Previsão de séries tempo: redes neurais e modelos estruturais. *Pesquisa e Planejamento Econômico*, v. 26, n. 2, p. 253-276, 1996.
- Hamilton, J. D. *Time series analysis*. Princeton: Princeton University Press, 1994.
- Harvey, A. C. Applications of the Kalman filter in econometrics. In: Ewley, T. F. (ed.), *Advances in econometrics Fifth World Congress*, v. I, Cambridge: Cambridge University Press, 1987
- _____. *Forecasting structural time series models and the Kalman filter*. Cambridge: Cambridge University Press, 1989.
- Harvey, A. C. and Koopman, S. J. Diagnostic checking of unobserved components time series models. *Journal of Business and Economic Statistics*, 10, 377-389, 1992.
- Haykin, S. *Neural networks: a comprehensive foundation*. Nova York: Macmillan, 1994.

- Kalman, R. E. A new approach to linear filtering and prediction problems. Transactions of the ASME, series D. *Journal of Basic Engineering*, 35-45, 1960.
- Koopman, S. J, Harvey, A. C., Doornick, J. A. and Shephard, N. *STAMP 5.0: structural time series analyser, modeller and predictor*. Londres: Chapman & Hall, 1995.
- Meinhold, R. J. and Singpurwalla, N. D. Understanding the Kalman filter. *The American Statistician*, 37, 123-127, 1983.
- Portugal, M. S. Modelos de parâmetros variáveis: uma resenha crítica. *Pesquisa e Planejamento Econômico*, v. 23, n. 1, p. 99-134, 1993.
- _____ Neural networks versus time series methods: a forecasting exercise. *Revista Brasileira de Economia*, v. 49, n. 4, p. 611-629, 1995.
- Portugal M. S. & Fernandes L. G. L. Redes neurais artificiais e previsão de séries econômicas: uma introdução. *Nova Economia*, v. 6, n. 1, p. 51-74, 1996.
- Refenes, A. N. Constructive learning and its application to currency exchange rate Forecasting. In: Turban, E. and Trippi, R. (eds.), *Neural networks applications in investment and finance services*. New York: Probus Publishing, 1991.
- Rumelhart, D. E., Hinton, G. E. and Williams, R. J. Learning internal representations by error propagation. *Parallel Distributed Processing*, Cambridge: MIT Press, v. I, p. 318-362, 1986.
- Souza, R. C. and Zandonade, E. Forecasting via neural networks: a comparative study. PUC-RJ: Departamento de Engenharia Elétrica, 1993, *mimeo*.
- Wasserman, P. D. *Neural computing: theory and practice*. New York: Van Nostrand Reinhold, 1989.
- West, M. and Harrison, J. *Bayesian forecasting and dynamic models*. 2^a ed. New York: Springer Verlag, 1997.