

**MODELOS DE PREVISÃO PARA O ICMS DO ESTADO DO CEARÁ**

**Ivan Castelar**

*Professor do Curso de Pós-graduação em Economia (CAEN) da  
Universidade Federal do Ceará (UFC)*

**Roberto Tatiwa Ferreira**

*Mestrando em Economia pelo CAEN*

**Fabício Carneiro Linhares**

*Aluno do Curso de Economia da UFC e Bolsista do Centro Nacional de  
Pesquisa (CNPq)*

**Resumo:** O presente estudo teve como objetivo central apresentar previsões mensais da principal rubrica de receita tributária do estado do Ceará, que é o Imposto sobre Circulação de Mercadorias e Serviços (ICMS). Optou-se, neste estudo, pela utilização de três tipos de abordagens de séries temporais - Alisamento Exponencial, modelo ARIMA e Função de Transferência - com o intuito de obterem-se previsões derivadas de diferentes processos metodológicos. Posteriormente, a abordagem mais moderna de combinar previsões foi realizada, corroborando em resultado já bem estabelecido na literatura: por utilizar diferentes informações modelos combinados apresentam maior eficiência que modelos individuais.

**Palavras-chave:** Arrecadação Tributária; ICMS; Previsão de Receitas; Brasil-Região Nordeste-Ceará.

## 1 INTRODUÇÃO

A atmosfera de incerteza que cerca toda atividade econômica, aliada ao fato de que decisões que envolvem períodos futuros são necessárias ao gerenciamento das instituições, leva à necessidade de previsões sobre valores de certas variáveis econômicas. Por esse motivo, os modelos de previsão vêm-se tornando freqüentes em várias áreas, tanto do setor privado como do setor público. A razão da popularidade destes modelos é que a nova visão de planejamento estratégico requer conhecimento antecipado de variáveis que determinam o desempenho das unidades econômicas (HANKE, REITSCH, 1989)<sup>(6)</sup>.

No caso de um órgão arrecadador, torna-se crucial a previsão da sua receita, pois esta será uma variável-chave na determinação do financiamento do custeio da máquina estatal, dos investimentos públicos, do serviço da dívida, etc.

Essas previsões tanto podem ter uma natureza de curto como de longo prazo. As duas prestam-se a diferentes formas de planejamento. O propósito deste trabalho é apresentar esses dois tipos de previsões para a principal rubrica de receita da Secretaria da Fazenda do estado do Ceará, que é o Imposto sobre Circulação de Mercadorias e Serviços (ICMS). É certo que outros itens da receita, como o Fundo de Participação do Estado, possuem um peso significativo na receita total do Estado, mas esse item será objeto de um outro trabalho. Já outros itens, como a receita patrimonial e o IPVA, elevam as suas participações em períodos inflacionários ou em certos períodos do ano. No entanto, essas participações, tomando todo o período de análise, não chegam a ser significativas.

As previsões de curto prazo referem-se à previsão de valores mensais do ICMS; já a previsão de longo prazo refere-se a valores anuais. O primeiro tipo de previsão tem como propósito orientar o administrador público na gerência mais imediata do seu fluxo de receita e como indicador de estratégias com relação a mudanças de conjuntura que atendam aos objetivos de um futuro imediato. Já o segundo tipo fornece subsídios macroeconômicos relacionados à opção por possíveis medidas de política econômica que intervenham estruturalmente com a economia estadual.

A natureza das previsões apresentadas neste trabalho é de cunho essencialmente quantitativo, embora introduza-se no estágio final de previsão elementos de julgamento, conhecidos na literatura como "ajustamento do

termo constante", que dependem fundamentalmente da experiência e intuição do pesquisador.

Este trabalho é apresentado em três partes distintas. Na primeira seção, apresentam-se os fundamentos dos modelos de previsão utilizados e a forma de avaliá-los. Na segunda seção, apresentam-se os resultados empíricos da identificação, estimação e previsão geradas pelos modelos para o ICMS, com uma argumentação sobre a escolha de cada modelo específico. Finalmente, na última seção, é apresentado um sumário dos principais resultados encontrados.

## 2 METODOLOGIA

A metodologia aqui proposta é a de uso mais freqüente na teoria da previsão, que é a busca de modelos ótimos de previsão, quer estes sejam baseados em noções de causalidade da variável a ser predita, quer estes sejam baseados meramente no comportamento passado da variável em questão. Nas previsões de curto prazo, recorre-se à combinação de modelos preditivos, com o objetivo de incorporar informações que de outro modo seriam perdidas, caso se usasse um modelo único de previsão.

Assim, nas previsões de curto prazo para o ICMS, pretende-se utilizar três tipos de modelos. Dois de cunho nitidamente de curto prazo, que são os modelos ARIMA (ou de Box-Jenkins) e de Alisamento Exponencial (ou de Holt-Winters), mais um modelo de Função de Transferência, que se presta tanto a previsões de curto como de longo prazo. Já a previsão que pode ser chamada de longo prazo, referente a valores anuais das variáveis discutidas acima, será tratada com modelos clássicos de regressão, complementados por modelagem dos erros.

### 2.1 O ALISAMENTO EXPONENCIAL

As técnicas de Alisamento Exponencial consideram que a importância dos valores passados na formação do valor presente da série diminui exponencialmente no tempo. O modelo de Holt-Winter é considerado como uma generalização do Alisamento Exponencial simples - o qual baseia-se em uma média móvel ponderada dos valores passados de uma série, atribuindo às observações pesos que decrescem exponencialmente - adicionando a este duas

equações: uma para estimar a tendência e outra para alisar o efeito sazonal, conforme mostrado abaixo:

$$(I.1) S_t = \gamma \frac{y_t}{A_t} + (1-\gamma)S_{t-1}$$

$$(I.2) \hat{y}_{t+p} = (At + \beta Tt) S_{t-l+p}$$

$$(I.3) A_t = \alpha \frac{y_t}{S_{t-L}} + (1-\alpha)(A_{t-1} + T_{t-1}).$$

onde:

$A_t$  = série alisada exponencialmente;

$S_t$  = estimativa da sazonalidade;

$\alpha, \beta$  e  $\gamma$  = constantes de alisamento para a estimativa de sazonalidade ( $0 < \gamma < 1$ );

$L$  = duração da sazonalidade.

## 2.2 O MODELO ARIMA

Ao contrário das técnicas de Alisamento Exponencial, os modelos ARIMA (*autoregressive integrated moving-average*) trabalham com a hipótese de que as séries temporais são geradas por um processo estocástico, com estruturas que podem ser caracterizadas e descritas algebricamente (PINDYCK, RUBINFELD, 1981)<sup>(16)</sup>. Para essa classe de modelos, os elementos que explicam o comportamento de uma série temporal residem nos seus valores passados, bem como em um conjunto de erros aleatórios passados e correntes.

### 2.2.1 Modelos ARIMA

Em alguns casos, as séries estacionárias formadas por um processo estocástico têm características auto-regressivas e de média móvel. O processo misto de ordem (p,q), ARMA (p,q), é representado pela equação (II.18).

$$(I.10) \quad y_t = \Phi_1 y_{t-1} + \Phi_2 y_{t-2} + \dots + \Phi_p y_{t-p} + \delta + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q}$$

Os modelos ARIMA (p,d,q) são aplicados em séries não estacionárias que apresentam tendência. Uma série não-estacionária, em média, pode obter tal qualidade ao se calcular sua diferença. A série " $y_t$ " é chamada de homogênea não-estacionária de ordem "d" quando:

$$(I.11) w_t = \Delta^d y_t,$$

onde:

$w_t$  = uma série estacionária;

$\Delta$  = diferença;

d = número de diferenças aplicadas em " $y_t$ ", para a obtenção de " $w_t$ ".

Então, a série estacionária " $w_t$ " pode ser modelada pelo processo ARMA (p,q). Neste caso, " $y_t$ " é um processo integrado auto-regressivo com média móvel de ordem (p,d,q), ou apenas ARIMA (p,d,q).

A aplicação de um modelo ARMA envolve três etapas para formulação e escolha de um modelo adequado para gerar previsões. Na etapa de identificação do modelo, o primeiro passo é verificar se a série é estacionária. A estacionaridade existe quando: a) a média e a variância da série original são estáveis ao longo do tempo e b) quando a covariância entre dois valores distintos da série (de diferentes períodos) é função apenas da distância temporal entre esses valores.

Formalmente:

$$\begin{array}{ll} \text{a) } E(y_t) = u & \forall t = 1, 2, \dots, T \\ \text{b) } E(y_t - u) = \sigma^2 & \forall t = 1, 2, \dots, T \\ \text{c) } E[(y_i - u)(y_j - u)] / \sigma^2 = \rho_{ij} & \forall i, j = 1, 2, \dots, T \text{ e } i \neq j \end{array}$$

A maioria das séries não-estacionárias podem ser estacionarizadas pela técnica das diferenças; já quando o problema for de instabilidade na variância, a solução geralmente adotada consiste na transformação logarítmica dos dados. O fundamento da transformação logarítmica é a da compressão da escala dos dados, fazendo com que a diferença em variabilidade diminua sensivelmente entre os valores logaritmizados em contraste com a maior variabilidade dos dados originais.

Através da função de autocorrelação, pode-se identificar se uma determinada série temporal apresenta tendência ou não. A função de autocorrelação fornece uma medida de quanta correlação existe entre os dados da série “ $y_t$ ” e seus respectivos valores vizinhos. A equação (I.12) define a autocorrelação com defasagem de “ $k$ ” períodos e a equação (I.13) representa a função de autocorrelação amostral.

$$(I.12) \rho_k = \frac{E[(y_t - u_y)(y_{t+k} - u_y)]}{\sigma_y^2} = \frac{Cov(y_t, y_{t+k})}{\sigma_y^2}$$

$$(I.13) \hat{\rho}_k = \frac{\sum_{t=1}^{T-k} (y_t - \bar{y})(y_{t+k} - \bar{y})}{\sum_{t=1}^T (y_t - \bar{y})^2}$$

onde:

$T$  = número de observações na série ( $y_1, \dots, y_T$ ).

Se  $\rho_k$  decresce exponencialmente quando “ $k$ ” aumenta, a série é estacionária. Outro importante papel da função de autocorrelação, já mencionado anteriormente, refere-se à identificação da ordem do processo ARMA( $p, q$ ), que será utilizado para modelar a série “ $y_t$ ” ou, no caso de um modelo ARI-MA( $p, d, q$ ), a série “ $w_t$ ”. Em síntese, o comportamento teórico de um processo AR( $p$ ) caracteriza-se por uma função de autocorrelação que começa com “ $\rho_0=1$ ” e declina exponencialmente a partir deste. Mais ainda, a função de autocorrelação deve ser acompanhada por uma função de autocorrelação parcial, a qual apresenta os primeiros “ $p$ ” coeficientes de correlação significativamente diferentes de zero e os demais iguais a zero.

De modo contrário, o processo MA( $q$ ) tem um comportamento teórico caracterizado por uma função de autocorrelação que apresenta “ $q$ ” valores diferentes de zero, e uma função de autocorrelação parcial que decresce exponencialmente a partir do primeiro coeficiente de correlação\*. Após a etapa de identificação, os próximos estágios são o da estimação do modelo e do seu teste de adequação.

---

\* Gráficos do comportamento teórico dos processos AR( $p$ ), MA( $q$ ) e ARMA( $p, q$ ), bem como informações adicionais, podem ser encontrados em HANKE, REITSCH, 1989, p. 383-385<sup>(6)</sup> e em PINDYCK, RUBINFELD, 1981, p. 517-28<sup>(13)</sup>.

Depois de estimados os parâmetros, aplicam-se os testes usuais de significância dos coeficientes. Para isso, utiliza-se o teste "t" de Student. O "R<sup>2</sup>", usado como medida de ajustamento na regressão linear, perde seu significado e aplicabilidade em modelos ARMA, em virtude da utilização do processo de estimação não-linear. O seu significado é preservado apenas para a última linearização ocorrida no processo iterativo de estimação. Contudo, mesmo que um modelo seja aprovado pelo teste dos resíduos, um "R<sup>2</sup>" próximo de zero indica limitação em termos de poder de previsão. Uma possível alternativa para a medida de ajustamento do modelo consiste no uso da variância dos resíduos (MADALLA, 1992).

Uma vez especificado e estimado, o modelo ARMA (ou ARIMA) é submetido a um processo de verificação que testa se a sua especificação atual está correta. O procedimento de diagnóstico constitui-se, geralmente, de duas partes. Na primeira, compara-se a função de autocorrelação da série temporal gerada pelo modelo com a função de autocorrelação amostral da série original. Quando as duas funções apresentam diferenças consideráveis, significa que a ordem do modelo deve ser reformulada. Caso contrário, quando inexistente tal diferença, inicia-se a segunda parte.

Na segunda fase, é feita uma análise quantitativa dos resíduos gerados pelo modelo. Recorre-se ao pressuposto, mencionado anteriormente, de que os erros aleatórios são normalmente e independentemente distribuídos. Em outras palavras, o comportamento dos resíduos deve ser similar ao gerado pelo processo de "ruído branco" (*white noise*).

A análise quantitativa utiliza, comumente, a estatística de Box-Pierce (Q), que possui uma distribuição qui-quadrado. A estatística "Q" é dada pela equação (I.14).

$$(I.14) \quad Q = (T - d) \sum_{k=1}^K \hat{r}_K^2$$

onde:

T = número de observações na série temporal;

d = grau de diferenciação necessário para obter-se uma série estacionária;

K = número de defasagens checadas;

$\hat{r}k$  = função de autocorrelação amostral do k-ésimo termo residual.

$\hat{r}k$  é obtido pela equação (I.15).

$$(I.15) \hat{r}_k = \frac{\sum_i \hat{\varepsilon}_i \hat{\varepsilon}_{i-k}}{\sum_i \hat{\varepsilon}_i^2}$$

Caso a estatística calculada "Q" seja maior que o valor teórico correspondente à distribuição qui-quadrado com "k-p-q" graus de liberdade, o modelo é considerado inadequado. Nesse caso, deve-se voltar aos estágios anteriores, com o objetivo de elaborar um modelo alternativo. Esse procedimento continua até que um modelo satisfatório tenha sido encontrado. Após esse procedimento de três estágios, o resultado é a obtenção de um modelo ARIMA estimado e testado, com potencial de fazer previsões.

### 2.3 FUNÇÃO TRANSFERÊNCIA

A análise de séries temporais pode ser combinada com a técnica de regressão linear, com o objetivo de gerar melhores previsões do que o uso isolado de um desses dois métodos.

A inclusão de variáveis explicativas nos modelos ARIMA produz os modelos de função de transferência. A idéia básica é adicionar informações que não fazem parte dos valores passados da própria série em estudo, através da inserção de variáveis explicativas. Assume-se, portanto, uma relação causal entre as variáveis dependente e independente que forneça as informações mencionadas. Formalmente, a função de transferência é definida como:

$$(1.16) Y_t = \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k + \varepsilon_t$$

onde,

$$\varepsilon_t = \phi_1 \varepsilon_{t-1} + \phi_2 \varepsilon_{t-2} + \dots + \phi_p \varepsilon_{t-p} + \delta + \theta_1 v_{t-1} + \theta_2 v_{t-2} + \dots + \theta_q v_{t-q} + v_t.$$

Isto é, " $\varepsilon_t$ " é um processo ARMA (p,q).

A formulação de um modelo de função de transferência consiste em dois passos. Primeiro, a estrutura do modelo de regressão linear deve ser determinada e, então, estima-se o modelo. Depois, aplica-se um modelo



ARIMA para os resíduos da regressão, seguindo os estágios descritos anteriormente: identificação, estimação, diagnóstico e seleção do modelo.

Vale ressaltar que existe na literatura sobre esse assunto uma controvérsia sobre o uso de séries não-estacionárias em modelos de função de transferência. Alguns autores defendem o uso das séries originais, mesmo que não-estacionárias, alegando que a transformação das séries, com o intuito de torná-las estáveis, gera, muitas vezes, perda na relação causal entre as variáveis. Outros defendem o uso de séries estacionárias (ou estacionarizadas), para que os resíduos da regressão apresentem, também, tal qualidade (MALINVAUD (1966), PIERCE (1977))<sup>(12,14)</sup>.

Outro ponto que deve ser lembrado refere-se a uma consideração prática. De uma maneira geral os modelos discutidos nesta seção requerem o conhecimento dos valores futuros da variável independente, para a geração de previsões. Surge, portanto, a dificuldade de gerarem valores futuros da variável (ou variáveis) independente para utilizá-los como dados de entrada na função de previsão da variável dependente. Um maneira de contornar parcialmente este problema consiste no uso de variáveis explicativas defasadas, quando é possível estabelecer essa relação.

Neste estudo, algumas variáveis explicativas da arrecadação do ICMS agregado são disponíveis na frequência mensal, como consumo de energia elétrica, FPE e faturamento do comércio varejista. Todas essas variáveis foram usadas na busca do melhor modelo explicativo do ICMS mensal, produzindo todas resultados muito similares entre si. Isto é, a parte mais significativa da explicação nesses modelos pode ser atribuída ao processo ARMA modelador do erro. No entanto, não deixa de ser informativa a parte causal do modelo que recaiu sobre o FPE defasado um período pela facilidade de informações sobre esta variável e pela sua própria importância no modelo, que conjugada com o processo ARMA mencionado acima, gera um melhor ajustamento do que incluindo o consumo de energia elétrica ou um índice de faturamento do comércio varejista como variável explicativa.

## 2.4 MEDIDAS DE EFICIÊNCIA E COMBINAÇÃO DE PREVISÕES

Ao se considerar a predição como objetivo do modelo, deve-se priorizar a preocupação com o ajustamento da previsão *ex-post* com os valores disponíveis, mais do que o ajustamento estatístico - um modelo e suas equações podem apresentar estatísticas "t" significantes e um alto "R<sup>2</sup>" e, mesmo assim, ser fraco em termos de predição.

Para a avaliação da previsão *ex-post*, existem várias medidas de ajustamento disponíveis na literatura. Nesta pesquisa, utilizou-se o erro percentual absoluto médio (EPAM), que é representado da seguinte forma:

$$(I.17) \text{ EPAM} = \frac{\sum_{i=1}^n |y_i - \hat{y}_i|}{n \cdot y_i}$$

Utilizando essa medida, é possível apresentar a razão entre valores previstos e valores reais, o que tem como principal vantagem indicar para o usuário a magnitude individual do erro de cada previsão.

Já no que se refere à técnica de combinação de previsões, seu objetivo consiste na possibilidade de agregar informações contidas em diferentes métodos de predição. Vários pesquisadores já demonstraram, empiricamente, que as técnicas de combinação de previsões são capazes de produzir resultados superiores às previsões individuais (JOLLY, WONG, 1987)<sup>(8)</sup>.

Estimados os modelos que serão utilizados para realizar as previsões individuais, escolhem-se pesos baseados no desempenho de cada um deles, e usam-se estes pesos para construir um modelo geral englobando cada modelo individual. Existem várias maneiras de determinar esses pesos e uma breve resenha desses métodos é apresentada adiante. O ponto crucial, no entanto, na determinação desses pesos é a manutenção de uma relação lógica entre magnitude do coeficiente de ponderação e desempenho do modelo na previsão.

Os principais métodos de combinação linear são os seguintes (Ferreira):

a) Método com Média Simples:

$$(I.18) C_n = [ F_n^{(1)} + F_n^{(2)} + \dots + F_n^{(p)} ] / p$$

onde:

$C_n$  = previsão composta para o período "n";

$F_n^{(i)}$  = previsão do i-ésimo método para a variável "Yn";

p = número de previsões incluídas na combinação.

b) Métodos com Pesos Diferentes

b.1) Modelo com Restrição:

$$(I.19) C_n = W_1 F_n^{(1)} + W_2 F_n^{(2)} + \dots + W_L F_n^{(p)}$$

Onde:

W's = pesos e " $W_1 + W_2 + \dots + W_L = 1$ "

b.2) Modelo sem Restrição com Termo Constante:

$$(I.20) C_n = W_1 F_n^{(1)} + W_2 F_n^{(2)} + \dots + W_p F_n^{(p)} + W_{p+1} M$$

onde:

M = média da série atual.

A avaliação dos resultados das possíveis combinações entre diferentes modelos preditivos pode ser feita através das medidas de eficiência descritas acima.

Um outro método usual de determinar o valor das ponderações é a regressão. Neste caso, os coeficientes das equações de regressão são estimados pelo método dos mínimos quadrados.

MAKRIDRAKIS, WINKLER (1983)<sup>(11)</sup>, apresentam resultados empíricos de uma pesquisa sobre cinco procedimentos sugeridos por

NEWBOLD, GRANGER (1974), para estimação dos pesos\* de uma determinada combinação linear. Após a comparação do desempenho de cada procedimento, concluíram que dois métodos se destacam entre os demais. Esses procedimentos são representados pelas equações (I.21) e (I.22).

$$(I.21) \quad W_i = \frac{\left( \sum_{s=t-v}^{t-1} e_s^{(i)2} \right)^{-1}}{\sum_{j=1}^p \left( \sum_{s=t-v}^{t-1} e_s^{(j)2} \right)^{-1}}$$

onde:

$i = (1 \dots p)$ ;

$v$  = número de períodos incluídos no processo de determinação dos pesos.

$$(I.22) \quad W_i = \beta W_{i,t-1} + (1 - \beta) \left[ \frac{\left( \sum_{s=t-v}^{t-1} e_s^{(i)2} \right)^{-1}}{\sum_{j=1}^p \left( \sum_{s=t-v}^{t-1} e_s^{(j)2} \right)^{-1}} \right]$$

onde:

$\beta$  = constante no intervalo  $[0,1]$ ;

$W_{i,t-1}$  = peso atribuído ao método de previsão "i", baseado nos dados do período "t-1".

O procedimento utilizado neste trabalho, para elaboração da combinação de previsões, é composto de duas etapas. Na primeira etapa, obtém-se o complemento do erro do erro percentual absoluto médio; posteriormente, obtém-se o peso para cada modelo como a razão entre o complemento e a soma destes complementos.

---

\* Para MAKRIDAKIS, WINKLER (1983)<sup>(1)</sup> os pesos apropriados são os que proporcionam uma combinação de previsões com menor variância dos erros do que todas as previsões individuais.

$$(I.23) \text{CEPM}_j = 1 - \frac{\sum |Y_i - \hat{Y}_{i,j}|}{\sum Y_i}$$

$$(I.24) W_j = \frac{\text{CEPM}_j}{\sum \text{CEPM}_j}$$

onde:

$\text{CEPM}_j$  = complemento do erro percentual absoluto médio referente às pre-

visões do modelo "j";

$\hat{Y}$  = previsões geradas pelo modelo "j";

$W_j$  = peso atribuído ao modelo "j".

Já o ajustamento do termo constante, que consiste em adicionar à previsão combinada, para os últimos quatro meses do ano de 1995, a média dos erros dos primeiros oito meses do mesmo ano. Esse procedimento corresponde, na verdade, a um elemento qualitativo da previsão, já que não existe base teórica para justificá-lo. Constitui-se esse procedimento em elemento de julgamento do pesquisador, levando-o a ajustar a trajetória do valor predito quando identificam-se sinais de que aquela está divergindo sistematicamente do valor real e ainda não é necessário reestimar o modelo.

### 3 RESULTADOS EMPÍRICOS

#### 3.1 MODELOS DE PREVISÃO DO ICMS MENSAL

Nesta seção, apresentam-se os resultados empíricos dos modelos do ICMS mensal, em nível agregado, a partir da metodologia descrita acima. Inicialmente, construiu-se um modelo ARIMA e outro de alisamento exponencial do tipo Holt-Winters, onde todos os parâmetros do modelo foram livremente estimados. Uma investigação preliminar indicou que o período amostral gerador dos melhores modelos de Alisamento Exponencial e ARIMA foi o de Janeiro/1991 a Dezembro/1994. Os seis primeiros meses de 1995 foram inicialmente reservados para geração da previsão *ex post* e comparação com os valores previstos. Numa etapa posterior, esses modelos foram atualizados, utilizando-se todas as observações mensais de 1995.

Para o modelo ARIMA, verificou-se que não existem tendência e sazonalidade nesse período. Já o correlograma da variável em questão indicou

que o modelo mais apropriado é um ARIMA(1,0,4). A análise do correlograma dos resíduos desse modelo comprova que estes se distribuem como um ruído branco, indicando que não é possível melhorar a modelagem. Tem-se assim uma estrutura pronta para gerar previsões que servirão para estabelecer um diagnóstico do modelo. Os resultados da estimação desse modelo, com sua correspondente equação, são apresentados na TABELA 1\*.

TABELA 1  
Estimação do Modelo ARIMA

*LS// Dependent Variable is ICMS*

*SMPL range: 1991.01 - 1994.12*

*Nuber of Observations: 48*

*Convergence achieved after 5 iterations*

<i>VARIABLE</i>	<i>COEFFICIENT</i>	<i>STD. ERROR</i>	<i>T-STAT.</i>
C	54430173.	2917124.2	18.658847
MA(4)	0.5821009	0.1680230	3.4644114
AR(1)	0.7667393	0.0957703	8.0060264
<i>R-squared</i>	0.666559	<i>Mean of depend.</i>	52084098
<i>Adjusted R</i>	0.651739	<i>S.D. of depend.</i>	7360251
<i>S.E. of regression</i>	4343551.	<i>Sum of squared</i>	8.4E+14
<i>Durbin-Watson</i>	1.763295	<i>F-statistic</i>	44.97816
<i>Log likelihood</i>	-800.2019		
$ICMS_t = 54430173 + 0.7667393 \epsilon_{t-1} + 0.5821009 V_{t-4}$			
<div style="display: flex; justify-content: space-around; width: 100%;"> <span>(2917124.2)</span> <span>(0.1680230)</span> <span>(0.0957703)</span> </div>			

O modelo de Alisamento Exponencial, do tipo Holt-Winters, estimado livremente para incluir tendência e sazonalidade, coerentemente indicou que os coeficientes desses dois termos são de fato nulos e o coeficiente de alisamento produziu uma estimativa de 0.88. Na TABELA 2 apresentam-se os resultados da estimação das constantes de alisamento. Previsões baseadas nesse modelo, que vão servir como referencial de comparação com as previsões do modelo ARIMA, são apresentadas na TABELA 4, seguidas da correspondente equação.

\* As tabelas com resultados dos modelos foram diretamente extraídas do *print-out* do TSP 7.0.

TABELA 2

Modelo de Alisamento Exponencial

SMPL range: 1991.01 - 1994.12

Number of Observations: 48

Smoothing Method: Holt-Winters - additive seasonal

Original Series: ICM

Forecast Series

SAICM

Parameters: ALPHA		0.880
BETA (trend)		0.000
GAMA (seasonal)		0.000
Sum of squared residuals		7.9E+14
Root mean squared error		4062589.
End of period levels: MEAN		69069613
TREND		29519.50
SEASONALS	1994.01	1532113.
	1994.02	-4215607
	1994.03	-7443406
	1994.04	-4429527
	1994.05	-858559.6
	1994.06	-1417094.
	1994.07	-20981.40
	1994.08	1849440.
	1994.09	5178883.
	1994.10	2151537.
	1994.11	4144587.
	1994.12	3528615.

$$ICMS'_{t+1} = 0.880 ICMS_t + (1-0.880) ICMS'_t$$

A função de transferência, após estimativas preliminares, indicou que o FPE defasado um período é a variável explicativa mais apropriada para o ICMS mensal no período de jan/1991 a dez/1994. Aparentemente, pode não parecer muito plausível a inclusão do FPE como variável explicativa do ICMS, principalmente por aquele ser uma variável determinada exogenamente à economia estadual. No entanto, um momento de reflexão indicará várias conexões do FPE com o ICMS do estado. Inicialmente, deve-se notar que o FPE varia com o nível de atividade econômica no País, acontecendo o mesmo com o ICMS. Haveria, assim, um movimento co-integrado das duas variáveis, embora não se pudesse dizer rigorosamente, baseado nesse movimento comum, que uma variável determine a outra. No entanto, deve-se notar, também, que o FPE constitui-se em um volume considerável de recursos injetado na economia estadual (ARAUJO, HORTA, Considera, 1973)<sup>(1)</sup>. A razão entre

o valor do FPE e o ICMS, a partir de 1978, excede a 40%, chegando em certos anos, como 1985 e 1986, a atingir 63%. Já a parte de modelagem do erro dessa equação, ainda de acordo com o correlograma dos resíduos, indicou que o modelo indicado para essa parte da equação é um ARIMA(2,0,1). Os resultados da estimativa dessa função podem ser vistos na TABELA 3, seguida da correspondente equação estimada.

TABELA 3

Modelo da Função de Transferência

*LS// Dependent Variable is ICM**SMPL range: 1990.01 - 1994.12**Number of Observations: 60**Convergence achieved after 5 iterations*

<i>VARIABLE</i>	<i>COEFFICIENT</i>	<i>STD. ERROR</i>	<i>T-STAT.</i>
C	46679026.	2962069.9	15.758921
FP(-1)	0.1936955	0.0718002	2.6977013
MA(1)	0.7713187	0.1355776	5.6891315
AR(2)	0.5397806	0.1026192	5.2600378
<i>R-squared</i>	0.555541	<i>Mean of depend. var.</i>	52188642
<i>Adjusted R-squared</i>	0.531730	<i>S.D. of depend. var.</i>	8583954.
<i>S.E. of regression</i>	5874019.	<i>Sum of squared resid</i>	1.93E+15
<i>Durbin-Watson stat</i>	1.965849	<i>F-statistic</i>	23.33194
<i>Log likelihood</i>	-1018.229		

$$ICMS_t = 46679026 + 0.1936955 FPE_{t-1} + 0.7667393 \varepsilon_{t-2} + 0.5821009 V_{t-1}$$

(2962069.9)    (0.0718)            (0.13558)            (0.10262)

Estimados os modelos, previsões do ICMS, no período de janeiro a junho de 1995, foram geradas a partir de cada um deles. Essas previsões foram comparadas aos valores reais do ICMS no período, e o poder de previsão de cada modelo foi aquilatado. Caracterizou-se o período como de teste de previsão dos modelos. Durante esse tempo, o modelo que melhor previu o ICMS individualmente foi o ARIMA. No entanto, dado que os outros modelos poderiam conter informações complementares, no sentido de melhorar a previsão, procedeu-se a várias combinações dos modelos, escolhendo-se a mais eficiente.

Procedido o período de teste, os modelos foram atualizados incorporando-se tanto as informações dos seis primeiros meses, como dos meses subsequentes do ano de 1995. Um fato digno de nota é a grande estabilidade



mostrada por todas as estimações, mostrando que as formulações de cada modelo são robustas com respeito a mudanças do período amostral. Previsões atualizadas foram feitas mês a mês e o poder de previsão de cada modelo foi novamente aquilatado. Mais uma vez, constatou-se que a combinação poderia levar a melhores resultados que as previsões individuais, e o melhor método de ponderar foi novamente pesquisado.

A combinação de previsões pode ser obtida por diferentes métodos, como discutido anteriormente. Um método indicado para a obtenção das ponderações para formar a combinação é o uso de regressão linear das previsões individuais, nos valores reais da variável a ser predita. No entanto, devido ao tamanho do período amostral utilizado neste estudo e do agudo problema de multicolinearidade presente nessa forma de regressão, o método torna-se inapropriado, no contexto, pois costuma produzir, quando aplicado, valores inaceitáveis das ponderações (CASTELAR, 1995)<sup>(2)</sup>. Um outro método popular é encontrar os pesos como proporcionalidade da variância dos resíduos de cada modelo.

Neste trabalho, como citado na parte metodológica, optou-se por uma medida que aproxima a variância da predição.

As previsões individuais e suas combinações são apresentadas na TABELA 4. A previsão do modelo ARIMA é denotada como ARMAICM; a do Alisamento Exponencial como SAICM; e a da Função de Transferência como FTICM. Já "RA", "RS" e "RFT" denotam as respectivas razões entre os valores previstos por cada modelo e o valor correspondente real do ICMS.

Usando o método de combinação, com pesos determinados pela variância da predição, apresenta-se na TABELA 4, sob a designação de "CJ1", as predições do modelo combinado. A razão entre o estimado e o real é mostrada na coluna "RCJ1". Na coluna "CJJ", a média do erro dos oito primeiros meses é acrescida à previsão dos quatro últimos meses de 1995, gerando a previsão final, a qual incorpora um elemento de julgamento ou de revisão do termo constante. Os valores em "RCJJ" são as respectivas razões entre o valor previsto e o valor real do ICMS. Os pesos encontrados foram 0,336 para o modelo ARIMA, 0,329 para o modelo de alisamento exponencial, e 0,333 para a função de transferência. A previsão final apresenta um Erro Percentual Absoluto Médio (EPAM) de 4,8 % para os quatro últimos meses de 1995, com valores mínimo e máximo variando de 2,7 % a 9,6%, tendo o maior erro de previsão ocorrido em setembro de 1995.

Ainda utilizando esses pesos e o ajustamento do termo constante, previsões *ex-ante* do ICMS para os três primeiros meses de 1996 foram geradas. Posteriormente, com a realização destas receitas mensais, tornou-se possível fazer uma comparação entre os valores previstos e os valores reais. Todos os valores são mostrados na TABELA 4 e, como pode ser notado, o erro de previsão é diminuto, principalmente em fevereiro e março, quando correspondem a 1,7% e 1,1%, respectivamente.

### 3.2 MODELO DE PREVISÃO ANUAL

Já existem experiências prévias com a previsão anual do ICMS. (CASTELAR, ARRAES (1991); PESSOA (1992))<sup>(3,13)</sup> apresentam um modelo de previsão do ICMS com duas equações e duas variáveis endógenas, onde a PIB do setor serviços do estado é usado como variável explicativa do ICMS. No entanto, nesse modelo as variáveis são logaritmizadas.

Existem razões específicas para o uso de modelos com variáveis logaritmizadas. Entre elas, a mais justificável é a constatação de que as variáveis explicativas relacionam-se com a variável dependente de forma exponencial. A transformação logarítmica, portanto, linearizaria o modelo nos parâmetros. Uma segunda razão seria a facilidade no cálculo de elasticidades proporcionado pelos modelos logaritmizados. No entanto, neste último caso é negligenciado o fato de que a transformação logarítmica pode violar a especificação mais correta do relacionamento original entre variáveis, bem como impor elasticidades constantes, o que nem sempre é aceitável.

Em modelos de variáveis logaritmizadas, por outro lado, nem sempre a previsão do valor original da variável tem o mesmo nível de qualidade da previsão do logaritmo, embora o ajustamento desses modelos tendam, em geral, a ser melhor do que o de modelos usando variáveis que não foram transformadas. Na verdade, a transformação inversa do valor predito do logaritmo da variável dependente não é a transformação ideal a ser aplicada neste caso. A inversão deve ser procedida sobre o valor predito do logaritmo, acrescido de metade da variância do erro. Por isso, procurou-se construir um modelo para previsão do ICMS que tratasse as variáveis em valores originais, neste trabalho.

**TABELA 4**  
**Valores Reais e Previstos do ICMS Mensal com Diferentes Combinações**  
**de Modelos: jan/95-mar/96**

	ICMS	ARMAICM	RA	FTICM	RFT	SICM	RS	CJI	RCJI	CJJ	RCJJ
95.01	92847616	92337824	0.9945	83635976	0.9008	86373536	0.9303	87470344	0.9421	87470344	0.9421
95.02	83193256	87649128	1.0536	87379232	1.0503	85418104	1.0267	86823704	1.0436	86823704	1.0436
95.03	84081168	82358712	0.9795	79679648	0.9477	79470384	0.9452	80513336	0.9576	80513336	0.9576
95.04	79049384	80595872	1.0196	77883688	0.9853	87343208	1.1049	81915760	1.0363		1.0363
95.05	74136672	76070464	1.0261	76444792	1.0311	84663024	1.1420	79027728	1.0660	79027728	1.0660
95.06	81644512	69557320	0.8520	72554456	0.8887	73876080	0.9049	71980272	0.8816	71980272	0.8816
95.07	77867488	79085160	1.0156	80915480	1.0391	82401520	1.0582	80788640	1.0375	80788640	1.0375
95.08	73331888	74059288	1.0099	73840880	1.0069	80983008	1.1043	76268808	1.0400	76268808	1.0400
95.09	85823960	70357872	0.8198	70584472	0.8224	78489592	0.9145	73113968	0.8519	77595880	0.9041
95.10	85429928	87436656	1.0235	81711760	0.9565	80739464	0.9451	83320168	0.9753	87802080	1.0278
95.11	84378664	80995344	0.9599	81428536	0.9650	87087304	1.0321	83147936	0.9854	87629848	1.0385
95.12	88429968	80673944	0.9123	80686640	0.9124	84180128	0.9519	81833960	0.9254	86315872	0.9761
96.01	1.01E+06	91186000	0.9048	84236000	0.8359	90178000	0.8948	88358000	0.8768	89493600	0.8861
96.02	91782000	94449000	1.0290	94767000	1.0325	91574000	0.9977	93420000	1.0178	93642000	1.0202
96.03	89705000	89646000	0.9993	88147000	0.9826	88797000	0.9899	88868000	0.9887	89091500	0.9931

Existem vários fatores que, a princípio, seriam causadores básicos do ICMS. O PIB total do estado, o PIB do setor serviços, o PIB do subsetor comercial, o PIB do Brasil ou o PIB do Nordeste são todos fatores que, em diferentes níveis, explicam o nível de arrecadação do ICMS. A questão, portanto, é de escolher entre essas variáveis a que melhor explica este comportamento.

Uma modelagem preliminar revelou que o PIB do setor serviços, complementado por um processo ARIMA(0,0,2) para o erro, fornece a melhor equação de previsão do ICMS anual. Ainda é necessário levar em consideração que a mudança constitucional de 1988, por aumentar a base tributária, causa um certo descolamento entre a trajetória da arrecadação e dos indicadores de atividade econômica. Para capturar a mudança, uma variável qualitativa, assumindo o valor zero no período 1970-1989 e valor 1 no restante do período, foi incluída na equação. A razão da variável qualitativa assumir o valor 1, apenas a partir de 1990, deve-se ao fato de que existe uma defasagem entre a promulgação da constituição e o aparecimento dos seus efeitos tributários. Os resultados da estimação do modelo, com sua correspondente equação, podem ser vistos na TABELA 10.

Deve-se salientar que o mesmo modelo foi estimado inicialmente usando-se o PIB total do estado como variável explicativa, em lugar do PIB dos serviços, no entanto sem atingir a mesma qualidade de previsão. A razão do melhor ajustamento produzido pelo PIB dos serviços deve-se, obviamente, ao fato de que este setor é o maior responsável pela geração de ICMS no estado. Como modelos de regressão capturam preferencialmente variáveis que possuem uma correlação mais significativa, segue-se que o PIB total não é a variável mais adequada para explicar o ICMS, sobretudo por incluir o produto agrícola, o que deve causar sérios danos à correlação. Saliente-se ainda que a série de dados disponíveis do PIB dos serviços só vai até 1993; espera-se, portanto, que a qualidade das previsões do modelo melhorem, à medida que a amostra seja atualizada.

Uma extensão natural da equação do ICMS anual seria uma segunda equação especificando o PIB dos serviços como função do nível de atividade econômica no País. No entanto, o erro cumulativo de previsão usando duas equações no modelo excluiu esta formulação. Como pode ser visto na TABELA 11, resolveu-se simular, de uma maneira *ad hoc*, diferentes cenários para a economia estadual e derivar, a partir destas hipóteses, previsões para o ICMS sob expectativas diferentes. Assim, apresentam-se três possíveis cená-

rios (pessimista, otimista e conservador), para a taxa de crescimento do PIB dos

serviços e as respectivas previsões do ICMS anual. Os cenários não são totalmente arbitrários, pois foram gerados a partir da observação das trajetórias de comportamento da taxa de crescimento do PIB dos serviços e da taxa de crescimento do PIB do Brasil, durante o período amostral. Verificou-se uma taxa otimista de 7% de crescimento, durante o período de 1971 a 1979, uma taxa conservadora de 5%, durante o período 1980 a 1986 e, finalmente, uma taxa pessimista de 3%, que vai de 1987 a 1993.

A previsão deste modelo, utilizando um período de estimação até 1993, atinge um erro de previsão de 9,2% para 1994 e um erro de 12% para o ano de 1995, utilizando-se o cenário denominado de conservador. Na TABELA apresentam-se as previsões *ex-ante* deste modelo até 1998, sob as diferentes hipóteses de crescimento da economia.

TABELA 5  
Estimação do Modelo Anual do ICMS

LS// Dependent Variable is I

SMPL range: 1970 - 1993

Number of Observations: 24

Convergence achieved after 6 iterations

VARIABLE	COEFFICIENT	STD. ERROR	T-STAT.
C	76601.745	20816.622	3.679836
S	5.405E-05	3.376E-06	16.013543
DUM	274391.66	22844.731	12.011158
MA(2)	-0.5263687	0.2511140	-2.096134
R-squared	0.967247	Mean of depend.	443425.4
Adjusted R-squared	0.962334	S.D. of depend.	194186.8
S.E. of regression	37687.26	Sum of squared	2.84E+10
Durbin-Watson stat	1.640958	F-statistic	196.8765
Log likelihood	-284.7565		

Obs: I=ICMS, S=PIB DO SETOR SERVIÇOS

$$\text{ICMS}_t = 76601.745 + 5.045\text{E-}05 S_t + 274391.66 \text{DUM}_t - 0.563687 V_{t,2}$$

(20816.6) (3.376E-06) (22844.7) (0.25111)

**TABELA 6**  
**Simulação e Previsão do ICMS Anual Para Taxas de Crescimento**  
**Simuladas do PIB Estadual**

Anos	Pessimista	Conservador	Otimista
1995	855368000	87406800	893129200
1996	842216600	871390800	901697700
1997	856953300	897410600	940247000
1998	872132100	924731500	981494800

#### 4 CONCLUSÕES

Neste trabalho apresentaram-se modelos de previsão para o Imposto sobre Circulação de Mercadorias e Serviços (ICMS). Para isso, foram construídos modelos de previsão de curto prazo (nível mensal) e de longo prazo (nível anual).

Nas previsões de curto prazo (dados mensais) do ICMS arrecadado no estado do Ceará, utilizou-se uma combinação de modelos de séries temporais: ARIMA, Alisamento Exponencial e Função de Transferência. A previsão, utilizando a combinação ótima, tendo os seis primeiros meses de 1995 como período de teste, e valendo-se do recurso de "ajustamento do termo constante", apresentou um Erro Percentual Absoluto Médio de 4,8%, com erros individuais variando de um mínimo de 2,7% a um máximo de 9,6%. Este último valor ocorreu para setembro de 1995. Como pode ser constatado, a previsão deste modelo demonstra um ótimo desempenho no período de teste e um desempenho plenamente satisfatório quando o modelo é atualizado.

Previsões *ex-ante*, só posteriormente comparadas com os valores reais, também mostraram um ótimo desempenho nos três primeiros meses de 1996. Tal desempenho é muito mais conclusivo do que previsões *ex-post* puras. Deve-se, no entanto, notar que por razões institucionais (imposto por substituição) e mudança de regime econômico (Plano Real), a tendência e a sazonalidade da série não estão presentes no período amostral. Pelo menos a tendência deverá ser reintroduzida em algum ponto do tempo, o que requer cautela quanto às previsões envolvendo um período de tempo mais longo.

No tocante à previsão do ICMS anual, ou de longo prazo, utilizando-se uma função de transferência - onde o PIB do setor serviços do Estado é a

variável explicativa, incluindo-se também uma variável qualitativa para capturar as alterações constitucionais e um processo ARIMA para o erro - constatou-se que a magnitude do erro de predição para dois anos à frente foi de 9,2% e 12%. Utilizando-se este modelo, três cenários alternativos foram simulados para a economia estadual e as correspondentes previsões para o ICMS até 1998 foram geradas. Assumiu-se um cenário pessimista (crescimento do PIB de 3% ao ano), um cenário conservador (com taxa de crescimento do PIB de 5% ao ano) e, finalmente, um cenário otimista (com taxa de crescimento do PIB de 7% ao ano). Acredita-se que a qualidade destas previsões recomendam o seu uso num melhor planejamento da receita estadual.

**Abstract:** The main purpose of this paper was to predict VAT State revenue. Three different time series were used: Exponential Smoothing, Box - Jenkins and a Transfer Function model. To avoid loss of information, the more modern approach of combining individual forecasts was used. The results of combining were more robust than individual forecasts, a well established fact in the literature.

**Key Words:** Tax Revenue; Revenue Forecasting; Value Added Tax; Brazil-Northeastern Region-Ceará.

## REFERÊNCIAS BIBLIOGRÁFICAS

1. ARAÚJO, A., HORTA, M. H., CONSIDERA, C. M. *Transferências aos estados*. Brasília: IPEA; 1973. (Relatórios de pesquisa).
2. CASTELAR, I. Some aspects of the theory of econometric prediction with an empirical application. Tese (Professor Titular) - CAEN-UFC. 1995 (mimeo.).
3. CASTELAR, I., ARRAES, R. O Efeito das secas sobre as finanças públicas do ceará. In: MAGALHÃES, A., R., BEZERRA, E., ed. *Impactos sociais e econômicos de variações climáticas*. Fortaleza: IOCE, 1991.
4. COSTA, R. N. Previsão do mercado de energia elétrica de curto prazo: uma abordagem de série de tempo. Dissertação (Mestrado) USP, 1994. (mimeo.).

5. GRANGER, C.W.J. *Forecasting in business and economics*. [s.l.] Academic Press, 1980.
6. HANKE, J., REITSCH, A. *Business forecasting*. [s.l.] Allyn and Bacon, 1989.
7. JOHNSTON, J. *Econometric methods*. [s.l.] McGraw-Hill, 1972.
8. JOLLY, L. O. e WONG, G. Composite forecasting: some empirical results using BAE short-term forecasts. *Review of Marketing and Agricultural Economics*. v. 55, n. 1, 1987.
9. MADALLA, G. S. *Introduction to econometrics*. [s.l.] Macmillan Publishing, 1992.
10. MAKRIDAKIS, S. *Forecasting: methods and applications*. [s.l.] John Wiley & Sons, 1983.
11. MAKRIDAKIS, S. WINKLER, R. L. The Combination of forecasts. *Journal of the Royal Statistical Society A.*, pt. 2, 146, 1983.
12. MALINVAUD, E. *Statistical methods of econometrics*. [s.l.] North-Holland Publishing, 1966.
13. PESSOA, M. N. Análise das Receitas e Despesas Públicas do Estado do Ceará sob a Influência da Seca e Nível de Atividade Econômica. CAEN-UFC. Tese de Mestrado. Mimeo. 1992.
14. PIERCE, D. Relationships - and the lack thereof - between economic time series and, with special reference to money and interest rates. *Journal of the American Statistical Association*. v. 72, 1977.
15. PINDYCK, R., RUBINFELD, D. *Econometric models and economic forecasts*. [s.l.] McGraw-Hill, 1981.

---

Recebido para publicação em 06.08.96.