

Eixo Temático: Estratégia e Internacionalização de Empresas

**PREVISÃO DA ARRECADAÇÃO DE ICMS DO ESTADO DO RIO GRANDE DO
SUL: UMA ABORDAGEM COM MODELOS ARIMA**

**PREDICTING THE COLLECTION OF ICMS OF RIO GRANDE DO SUL:
APPROACH WITH MODELS ARIMA**

Deise Scheffer, Adriano Mendonça Souza e Roselaine Ruviano Zanini

RESUMO

O objetivo deste trabalho foi realizar previsão para a série de arrecadação do Imposto sobre Circulação de Mercadorias e Serviços (ICMS) do Estado do Rio Grande do Sul, no período de janeiro de 2000 a março de 2014. Para tanto, aplicou-se a metodologia de Box-Jenkins que estima os modelos de séries temporais por meio de modelos conhecidos como autorregressivos integrados e de médias móveis (ARIMA). A obtenção do modelo mais adequado foi realizada com base na análise de resíduo e dos testes de significância dos parâmetros. Após a análise de vários modelos concorrentes, o modelo ajustado foi um ARIMA (2, 1, 0). Destaca-se a importância de se analisar variáveis por meio de modelos de séries temporais de Box e Jenkins, visando conhecer o comportamento e projetar valores, embasando futuras tomadas de decisão.

Palavras-chave: Metodologia Box-Jenkins, Previsão, ICMS, Séries Temporais.

ABSTRACT

The objective of this study was to forecast the number of collection of Tax on Goods and Services (ICMS) of the State of Rio Grande do Sul, in the period from January 2000 to March 2014. Therefore, we applied the methodology Box-Jenkins estimating models for time series using models known as autoregressive and integrated moving average (ARIMA). To obtain the most appropriate model was based on the analysis of residues and tests of significance of parameters. After examining several competing models, the fitted model was a ARIMA (2, 1, 0). Highlights the importance of analyzing variables through time series models of Box and Jenkins, aiming to understand the behavior and design values, basing future decision making.

Keywords: Box-Jenkins Methodology Forecasting ICMS, Time Series.

1. INTRODUÇÃO

O Imposto sobre Circulação de Mercadorias e Serviços (ICMS) é o principal tributo estadual e o de maior arrecadação no Brasil (REZENDE, 2009). O ICMS tem um objetivo fiscal, e um dos principais fatos geradores é a circulação de mercadoria, até mesmo as que iniciam no exterior. Além disso, incide sobre vários tipos de serviços, como transporte intermunicipal e interestadual, importação, telecomunicação, prestação de serviços entre outros, sendo responsável por praticamente 90% da arrecadação estadual.

Devido a sua importância no montante de arrecadação no país, autores como Marques e Uchôa (2006), Camargo (2008) e Santos e Costa (2008), utilizaram em seus trabalhos métodos de prever econometricamente o valor arrecadado.

Para Castanho *et al.* (2011), a tarefa de estimação, previsão e controle de receitas fiscais têm tido um crescimento notável tanto no ambiente acadêmico como a atenção de políticos, institutos de pesquisas e do setor público contribuinte. Em pesquisa realizada pelo autor, são identificadas características estatísticas da série do ICMS do Estado do Espírito Santo e elaboradas previsões com a metodologia de Box-Jenkins, análise e intervenção e de um modelo econométrico causal com estrutura dinâmica.

Autores como Marques e Uchôa (2006), utilizaram modelos ARIMA para analisar a série temporal da Bahia. No Espírito Santo, Ribeiro (2009) utiliza modelos univariados e multivariados para prever a arrecadação do ICMS no estado. No estado de Pará, Passos *et al.* (2011), realizaram um pesquisa utilizando a metodologia de Box-Jenkins que permite estimar modelos de séries temporais denominados autorregressivos integrados e de médias móveis.

Esta pesquisa tem como objetivo ajustar um modelo de séries temporais que permita identificar o processo gerador da série assim como realizar previsões no curto prazo para a arrecadação de ICMS do Estado do Rio Grande do Sul no período de janeiro de 2000 a março de 2014, por meio da classe de modelos propostos por Box e Jenkins. O ajuste de um modelo apropriado permite realizar melhorias na elaboração orçamentária e na arrecadação do Estado, sendo fundamental para prever as arrecadações futuras de seus tributos, sendo importante para o auxílio de tomada de decisões futuras.

2. METODOLOGIA

Inicialmente pretende-se ajustar um modelo da classe geral ARIMA de Box e Jenkins à série temporal do ICMS, considerando se o processo em análise é estacionário ou não, a identificação do modelo, entre outros fatores, de acordo com os itens a seguir.

2.1 Caracterização de uma série temporal em relação a sua estacionariedade

Um processo Z é estacionário se ele se desenvolve no tempo de modo que a escolha de uma origem dos tempos não é importante, ou seja, as características de $Z(t + f)$, para todo f , são as mesmas de $Z(t)$. Portanto, se a série temporal desenvolve-se durante um determinado tempo em torno de um valor constante, não necessitando transformações matemáticas para estacionarizá-la, pode-se dizer que esta é uma série estacionária (Morettin, 2006).

Um processo estacionário poderá ser fortemente ou fracamente estacionário. Para Morettin (2006), tecnicamente, há duas formas de estacionariedade: fraca (ou ampla ou de segunda ordem) e estrita (ou forte). Um processo estocástico $Z = \{Z(t), t \in T\}$ será fracamente estacionário (ou estacionário de segunda ordem) se, e somente se:

- i) $E\{Z(t)\} = \mu(t) = \mu$; constante, para todo $t \in T$;
- ii) $E\{Z^2(t)\} < \infty$, para todo $t \in T$;
- iii) $\gamma(t_1, t_2) = \text{cov}\{Z(t_1), Z(t_2)\}$, é uma função $t_1 - t_2$. Um processo Z tal que (ii) esteja satisfeito diz-se um processo de segunda ordem.

Ao mesmo tempo, sua função de autocovariância apresenta comportamento como:

$$\gamma(t) = \text{Cov}\{y(t), y(t+t)\} = \text{cov}\{y(0), y(t)\}$$

Porém, a definição de processo estocástico choca-se com o fato de ser de caráter contínuo, ou seja, apresentar um número elevado de momentos. Limitando o espaço a momentos de primeira e segunda ordem, obtém-se o conceito de estacionariedade fraca, que apresenta as seguintes propriedades:

$$\begin{aligned} E\{y(t)\} &= \mu(t) = \mu, \text{ constante, para todo } t \in T \\ E\{y^2(t)\} &< \infty, \text{ para todo } t \in T \\ \gamma(t_1, t_2) &= \text{Cov}\{y(t_1), y(t_2)\} \text{ é uma função apenas de } |t_1 - t_2| \end{aligned}$$

Pode-se dizer que um processo de estacionariedade fraca como uma situação especial que ameniza a robustez imposta pela condição de estacionariedade forte (Pinto *et al.*, 2008).

2.2 Modelo da classe geral ARIMA (p, d, q)

Segundo Werner e Ribeiro (2003), em um modelo autorregressivo AR(p), a série de dados anteriores Z_t é descrita por seus valores passados regredidos e pelo ruído aleatório a_t .

$$\begin{aligned} Z_t &= \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \dots + \phi_p Z_{t-p} + a_t, \text{ onde} \\ Z_t &= Z_t - \mu; \end{aligned}$$

O modelo autorregressivo de ordem 1 ou AR(1) é a versão mais simples desta classe e seu modelo algébrico é dado por:

$$Z_t = \phi_1 Z_{t-1} + a_t,$$

Para o modelo ser estacionário é preciso que $|\phi| < 1$ além de que as autocovariâncias (γ_k) sejam independentes. Dado um modelo AR(1), suas autocovariâncias são dadas por:

$$\gamma_k = \phi_1^k \gamma_0$$

Já as autocorrelações de ρ_k são dadas por:

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \phi_1^k \text{ onde, } k = 0, 1, 2, \dots$$

Quando ϕ_1 é positivo, a função de autocorrelação decai exponencialmente; quando ϕ_1 é negativo, a função de autorrelação também decai exponencialmente, porém apresenta alternância de sinais positivos e negativos (Werner e Ribeiro, 2003).

Em relação a um processo de médias-móveis, conhecido também como MA(q) e a representação algébrica é dada por:

$$\tilde{Z}_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q}$$

Onde, $\tilde{Z}_t = Z_t - \mu$; θ_i sendo o parâmetro que descreve como Z_t se relaciona com o valor de ε_{t-i} , para $i = 1, 2, \dots, q$.

As autocorrelações ρ_k , que são autocovariâncias divididas pela variância são dadas:

$$\rho_1 = \frac{\gamma_1}{\gamma_0} = \frac{-\theta_1 \sigma_\varepsilon^2}{(1+\theta_1^2)\sigma_\varepsilon^2} = \frac{-\theta_1}{(1+\theta_1^2)} e;$$

$$\rho_k = 0 \quad k > 1$$

Segundo Werner e Ribeiro (2003), uma série temporal é não-estacionária quando apresenta média e variância dependentes do tempo. A não-estacionariedade da série implica que:

a) existe uma inclinação nos dados e eles não ficam ao redor de uma linha horizontal no decorrer de um período;

b) a variação dos dados não se mantém essencialmente constante sobre o tempo, ou seja, as flutuações aumentam ou diminuem no decorrer do tempo, indicando que a variância está se alterando.

Como a maioria dos procedimentos de análise estatística de séries temporais supõe que estas sejam estacionárias, será preciso transformá-las caso ainda não sejam. Segundo Morettin e Tolo (2006), estas séries precisam ser defasadas para atingir a condição de estacionariedade. O processo mais comum consiste em tomar diferenças sucessivas da série original até conseguir uma série estacionária.

Há casos em que será suficiente tomar uma ou duas diferenças para que a série se torne estacionária. O número necessário de diferença para tornar uma série estacionária é denominado ordem de integração (d). A inclusão do termo de ordem de integração permite que sejam utilizados os modelos ARIMA (p,d,q) dados pela equação:

$$W_t = \Delta^d Z_t,$$

podemos representar W_t por um modelo ARMA (p,q), sendo,

$$\phi(B) W_t = \theta(B) \varepsilon_t$$

Diz-se que Z_t segue um modelo autorregressivo, integrado, de médias móveis ou um ARIMA,

$$\phi(B) \Delta^d Z = \theta(B) \varepsilon_t$$

de ordem (p, q, d), se p e q são as ordens de $\phi(B)$ e $\theta(B)$, respectivamente.

2.3 Etapas da metodologia de Box & Jenkins

A metodologia de Box-Jenkins é bastante utilizada na análise de modelos paramétricos, onde se refere ao método sistemático de identificação e ajuste de modelos autorregressivos integrados e média móvel ou, simplesmente, modelos ARIMA (p,d,q), a um conjunto de dados.

Para se obter um resultado mais apurado por meio da utilização da metodologia de Box-Jenkins (ARIMA), há três hipóteses que necessariamente precisam ser tratadas com cuidado. Primeiramente, o tamanho inicial da amostra deve ter no mínimo 50 observações (Box-Jenkins, 1976). Além disso deve ser considerada a suposição de que a série seja estacionária, isto é, que a série varie em torno de uma média e variância constantes. E por fim, que a série seja homocedástica, ou seja, que tenha uma variância constante ao longo do tempo.

Segundo Morettin e Tolo (2006), a estratégia para a construção do modelo será baseada em um ciclo iterativo, onde a escolha da estrutura do modelo é baseada nos próprios dados. Os estágios do ciclo iterativo são:

- (a) uma classe geral de modelos é considerada para análise (*especificação*);
- (b) há *identificação* de um modelo, com base na análise de autocorrelações, autocorrelações parciais e outros critérios;
- (c) *estimação*, na qual os parâmetros do modelo identificado são estimados;
- (d) e finalmente, há a verificação ou diagnóstico do modelo ajustado, por meio de uma análise de resíduos, para se saber se este é adequado ou não.

Caso o modelo não seja adequado, o ciclo é repetido, voltando-se a fase de identificação. Um procedimento que muitas vezes é utilizado é identificar não só um único modelo, mas alguns modelos concorrentes que serão então estimados e verificados. Logo, quando se obtém um modelo satisfatório, passa-se para a última etapa da metodologia de Box-Jenkins, ou seja, a realização de previsões.

Por fim faz-se a análise dos resíduos para verificar se os mesmos são ruído branco, isto é, se possuem média zero, variância constante e não são autocorrelacionados, além de se utilizar as estatísticas de ajustes AIC e BIC, para determinar o melhor modelo.

$$\text{AIC} = \ln \hat{\sigma}_\varepsilon^2 + \frac{2(p+q)}{n}$$

$$\text{BIC} = \ln \hat{\sigma}_\varepsilon^2 + \frac{2(p+q) \ln(n)}{n}$$

2.4 Realizar as previsões utilizando os modelos adequados

Uma das razões para a popularidade da modelagem ARIMA é o seu sucesso na previsão (GUJARATI, 2006). A previsão é a última etapa a ser executada, sua função é testar o potencial do modelo e prever os potenciais valores futuros para a série estudada (PINTO *et al.*, 2008).

A previsão pode assumir dois aspectos, primeiramente a ela pode ser usada para prever valores futuros, que ainda não existem, essa previsão é conhecida como *ex-ante*. Logo, pode haver também a previsão acerca dos valores já existentes dentro da série estudada, este tipo conhecido como *ex-post*. Porém, a melhor previsão será aquela que apresentar o menor Erro Quadrado Médio (PINTO *et al.*, 2008).

As métricas de comparação usadas correspondem ao erro percentual médio absoluto (MAPE), que penaliza mais os erros maiores; e o coeficiente U de Theil, que avalia o desempenho da previsão em relação à previsão ingênua ou trivial, ou seja, previsão ingênua ou trivial significa que a estimativa do valor futuro é igual ao valor atual.

A métrica MAPE, *Mean Absolute Percentage Error*, indica o valor médio do erro percentual das previsões e é definido pela seguinte fórmula:

$$\text{MAPE} = \frac{100}{N} \times \sum_{i=1}^N \left| \frac{x_i - \hat{x}_i}{x_i} \right|$$

Onde, N é o número de previsões realizadas e x_i e \hat{x}_i os valores observado e previsto para o índice i , respectivamente.

O coeficiente U de Theil é definido como:

$$U = \frac{\sqrt{\sum_{i=1}^N (a_k - y_k)^2}}{\sqrt{\sum_{i=1}^N (a_k - a_{k-1})^2}}$$

O coeficiente acima analisa a qualidade de uma previsão através dos seguintes valores:
 $U > 1$, significa que o erro do modelo é maior do que da previsão ingênua;

$U < 1$, significa que o erro do modelo é menor que da previsão ingênua.

O coeficiente U de Theil menor do que 1 já indica uma previsão melhor que a previsão ingênua; quanto mais próximo o mesmo for de zero, melhor será o resultado da previsão (BARROSO, 2010).

3 ANÁLISE E DISCUSSÃO DOS RESULTADOS

Com o intuito de gerar um modelo de previsões para a série estudada, primeiramente reuniu-se os dados históricos de janeiro de 2000 a maio de 2014 da arrecadação do ICMS, obtidos pela Secretaria da Fazenda do Estado do Rio Grande do Sul e corrigidos pelo Índice Geral dos Preços – IGP-DI, trazendo assim, os valores da época aos valores da moeda corrente.

Apresenta-se, na Figura 1, a série em nível, a qual apresenta as características tendência, variância não constante e de sazonalidade, desta forma a série revela se é ou não estacionária.

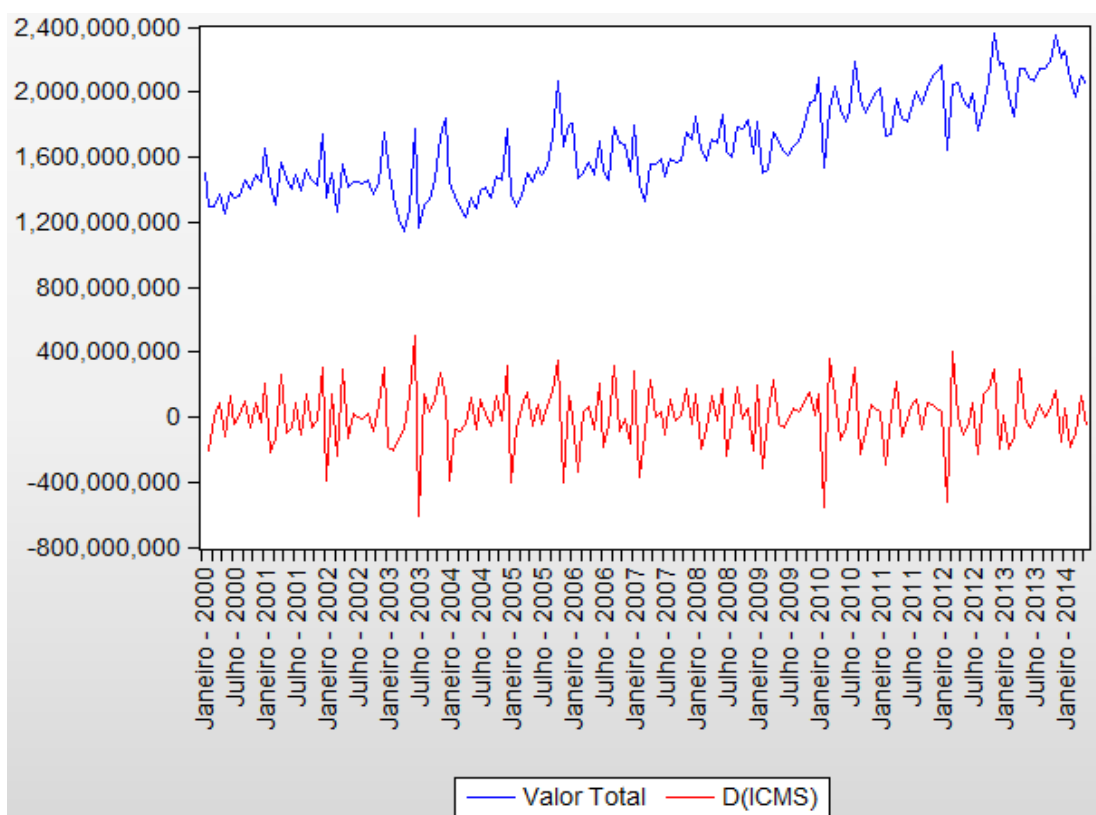


Figura 1: Gráfico da Série ICMS do estado do Rio Grande do Sul (2000-2014) em nível e em primeiras diferenças

O próximo passo foi analisar as funções de autocorrelação (FAC) e de autocorrelação parciais (FACP) da série ICMS em nível, sendo que o comportamento dessas funções auxilia na verificação da estacionariedade e na proposição do modelo, conforme Figura 2.

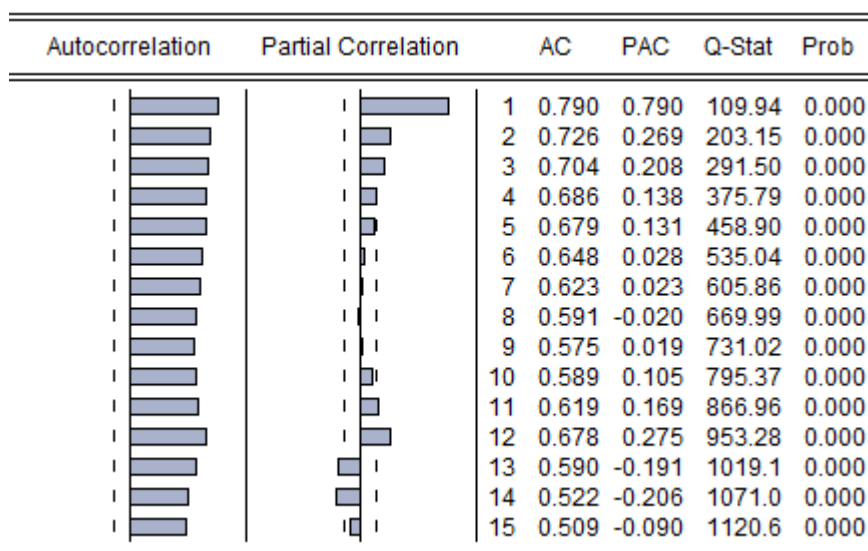


Figura 2: Correlogramas das FAC e FACP da série ICMS

Na Figura 2 pode-se observar a FAC e a FACP da série de arrecadação do ICMS no Estado do Rio Grande do Sul, em primeiras diferenças. Na série em nível, após serem obtidas as ACF e PACF observou-se um decaimento lento, o que caracteriza uma não-estacionariedade, indicando, portanto, que um modelo ARMA não poderá ser ajustado. Nas ACF e PACF em diferenças, Figura 3, observa um decaimento rápido da PACF e um lag significativo de ordem 1 e 11 na ACF. Desta forma o modelo geral ajustado será um ARIMA.

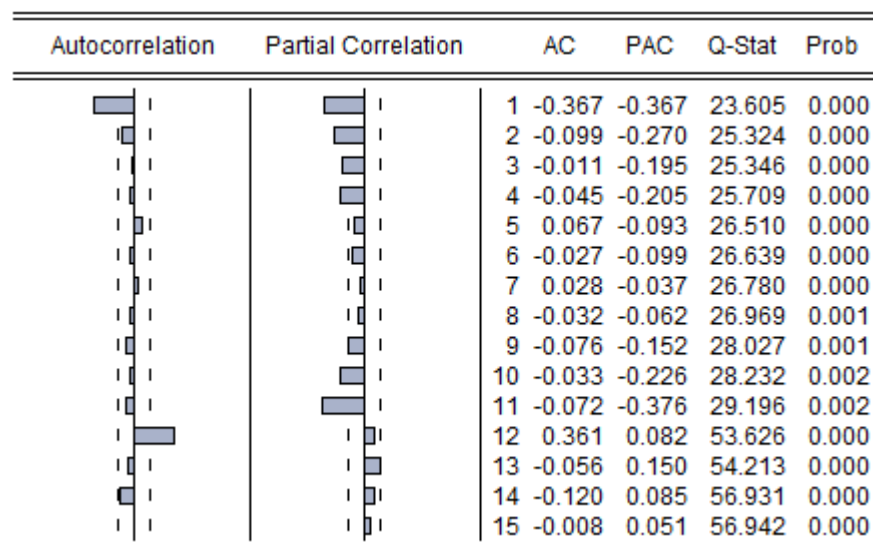


Figura 3: FAC e a FACP da série de arrecadação do ICMS no estado do Rio Grande do Sul, em primeiras diferenças.

Foram analisados vários modelos concorrentes, porém foram escolhidos os mais representativos, cujos principais resultados podem ser observados na Tabela 1.

Tabela 1 – Principais resultados dos modelos ajustados

ICMS	Coeficiente	STD. ERROR	T-EST	p-valor	AIC	BIC	Presença de Ruído Branco
ARIMA(1, 1, 0)	0,99	0,00	123,36	0,00	40,86	40,87	SIM
AR(2, 1, 0)	0,63	0,07	8,87	0,00	40,72	40,76	SIM
	0,36	0,07	5,13	0,00			
MA(0, 1, 1)	0,89	0,03	26,44	0,00	44,16	44,17	NÃO
MA(0, 1, 2)	1,27	0,04	30,36	0,00	43,33	43,36	NÃO
	0,83	0,04	19,99	0,00			
ARIMA(0, 1, 1)	0,99	0	138,56	0,00	40,86	40,89	NÃO
	-0,11	0,07	-1,44	0,15			
SARIMA(1, 1, 0)(1, 1, 0) ₁₂ SAR(12)	0,24	0,07	3,16	0,00	40,7	40,74	SIM
	1,02	0,01	98,07	0,00			
SARIMA(0, 1, 1)(0, 1, 1) ₁₂ SMA(12)	0,88	0,03	24,08	0,00	43,15	43,19	NÃO
	0,88	0,01	45,51	0,00			

Tabela 1. Os demais modelos não relacionados juntamente com os abaixo foram excluídos por apresentar coeficientes não significativos ou porque tinham variância residual maior que os relacionados.

O modelo selecionado entre os modelos concorrentes para o ICMS – RS foi um modelo AR (2) com uma diferença simples, ou seja, um ARIMA (2, 1, 0). Este modelo proporcionou que os resíduos gerados apresentassem características de ruído branco, isto é, média zero, variância constante e não eram autocorrelacionados, conforme se observa nas Figuras 4.

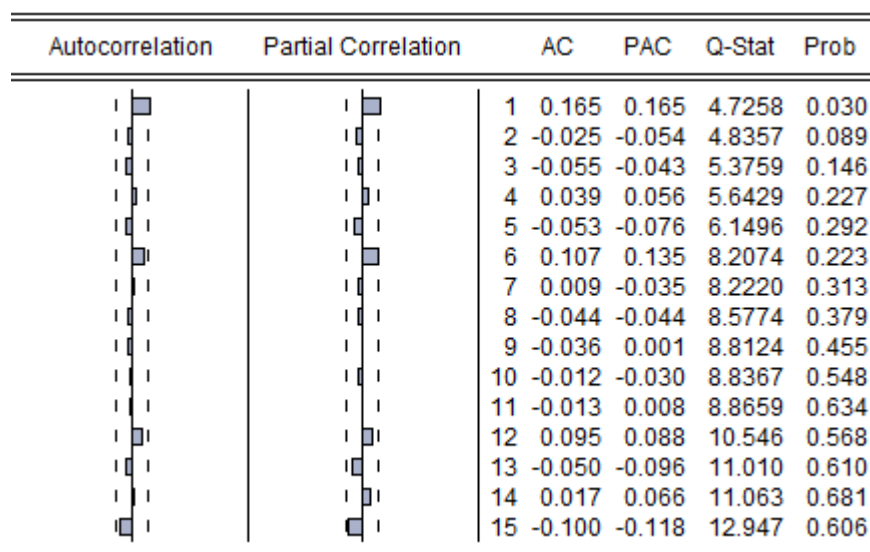


Figura 4 - FAC e FACP da série de resíduos do modelo AR (2)

A FAC residual do modelo mostra que os valores de autocorrelação dos resíduos estão dentro do limite de confiabilidade de 95%, assim como a FACP, indicando que estes resíduos não são correlacionados.

Como forma de observar o comportamento da série ajustada e dos resíduos, observa-se, na Figura 6, que a série ajustada converge para a série real, o que gera resíduos próximos de zero, indicando que o modelo ajustado é um bom representante do ICMS-RS e validando o mesmo para que se façam previsões futuras.

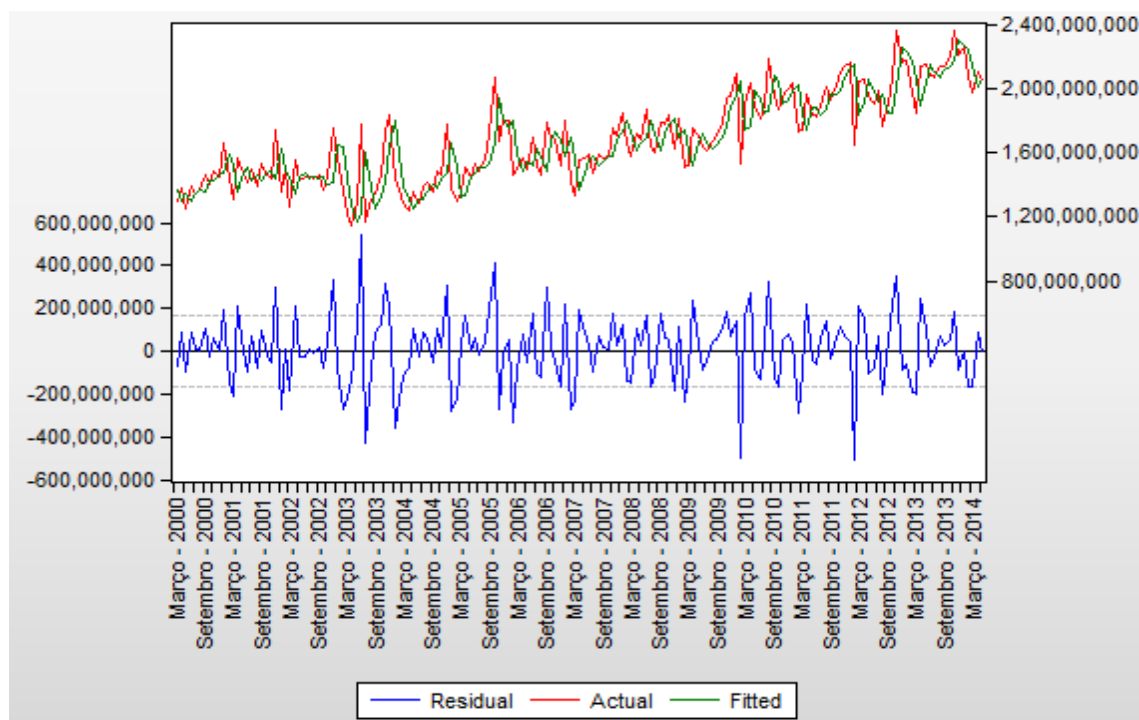


Figura 6 – Gráfico representativo da série real e da série ajustada

Utilizando o modelo selecionado, realizou-se a previsão dentro da amostra para o período de análise, conforme Figura 7.

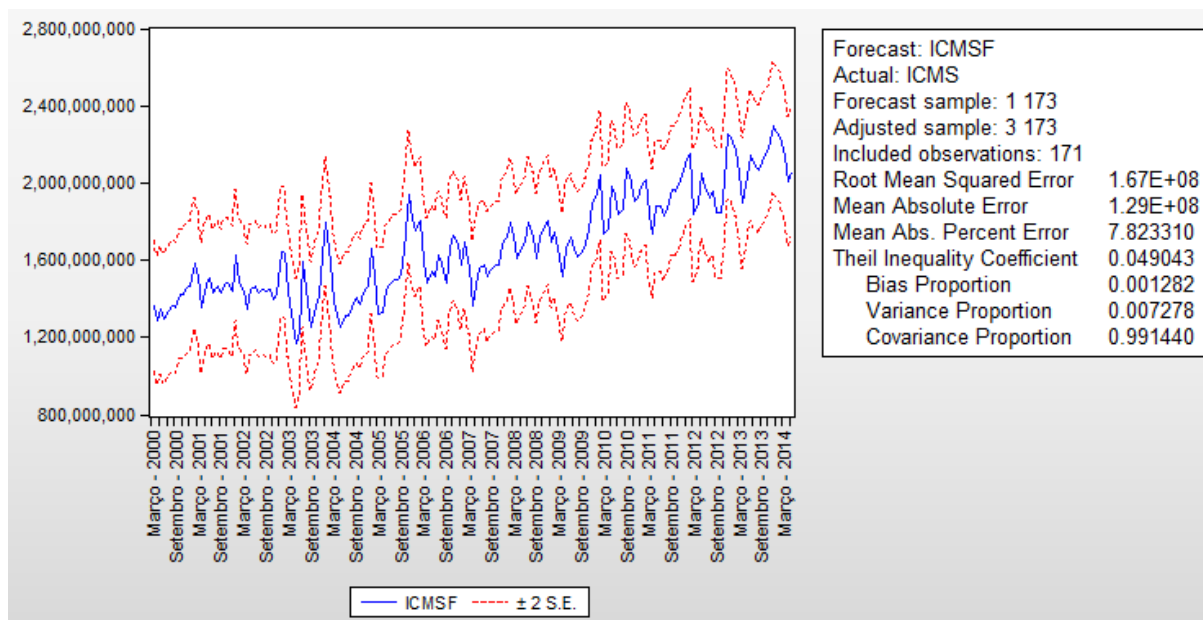


Figura 7 – Gráfico de Previsão

Observa-se que os valores previstos apresentam o comportamento semelhante à série original com estatísticas de ajustes de Erro Médio Absoluto de $1,29 \times 10^8$, erro percentual médio absoluto de 7,8233% e coeficiente U-Theil = 0,049.

2 CONCLUSÃO

Nesta pesquisa, buscou-se ajustar um modelo de previsão da série de ICMS do Estado do Rio Grande do Sul com base na metodologia de Box-Jenkins, considerando que um bom modelo e previsões adequadas podem ser uma ferramenta muito útil para auxiliar nas tomadas de decisões tanto de gestores públicos quanto da sociedade.

Aplicando os modelos ARIMA propostos por Box e Jenkins verificou-se que o melhor modelo foi um ARIMA (2, 1, 0).

Por meio do ajuste deste modelo foi possível verificar e prever uma tendência crescente na arrecadação, fator crucial para o desenvolvimento estadual, visto que o ICMS é um dos principais tributos cobrados no estado do Rio Grande do Sul.

REFERÊNCIAS:

BARROSO, S. N. O. **Abordagem HPA para a modelagem de dados de alta frequência: aplicação à previsão horária de carga elétrica.** 2010. 24f. Dissertação (Mestrado em Engenharia Elétrica) – PUC-Rio, Rio de Janeiro, 2010.

CASTANHO, B. ; BRASIL, G. H. ; SAMOHYL, R. W. . **Previsão de receitas tributárias: o ICMS do Estado do Espírito Santo.** In: SBPO 2011, 2011. SBPO 2011. Ubatuba, 2011.

FAVA, V. L. . **Metodologia de Box-Jenkins para modelos univariados.** In: Marco A. S.Vasconcellos; Denisard C. O. Alves. (Org.). Manual de Econometria: Nível Intermediário. SÃO PAULO: ATLAS, 1999, v. 1, p. 205-232.

GUJARATI, D. N. **Econometria Básica.** Rio de Janeiro: 4ª Ed. AMGH Editora Ltda, 2006.

MORETTIN, P.A . & TOLOI, C.M.C. **Análise de Séries Temporais.** São Paulo, Edgard Blücher, 2006.

PASSOS, J. J.; RAMOS, E. M. L. S.; ALMEIDA, S. S. **Utilização de modelos ARIMA para previsão da arrecadação de ICMS do estado do Pará.** Disponível em: <<http://www.ufpa.br/abe/programacao/resumos/com-jairopasso2.pdf>>.

PINTO, P. L. A. ; MAIA, S. F. . **Previsão do preço das commodities agrícolas brasileiras: uma abordagem por modelos univariados.** In: XLVI Congresso da Sociedade Brasileira de Economia e Sociologia Rural - SOBER Amazônia mudanças globais e agronegócio: O desenvolvimento em questão. 2008, Rio Branco - Acre. Anais do XLVI Congresso da Sociedade Brasileira de Economia e Sociologia - SOBER Rural Amazônia mudanças globais e agronegócio: O desenvolvimento em questão., 2008.

RIBEIRO, Livio. **Modelos mensal e trimestral para projeção de arrecadação do ICMS para o estado do Espírito Santo.** Instituto Jones dos Santos Neves, Texto para discussão, ISBN 978-85-62509-14-8. Vitória, ES, 25p. II, 2009. Disponível em: http://www.ijsn.es.gov.br/attachments/416_td10.pdf.

SALOMAO, M.F. **A arrecadação de ICMS no Estado do Espírito Santo: análise da evolução recente e modelos econométricos para previsão de receita.** In: I Encontro de Economia do Espírito Santo, 2010, Vitória, ES. Anais do I Encontro de Economia do Espírito Santo, 2010.

WERNER, Liane; Ribeiro, José Luis Duarte. **Previsão de demanda: uma aplicação dos modelos Box-Jenkins na área de assistência técnica de computadores pessoais.** Gestão & Produção (UFSCAR. Impresso), São Carlos - SP, v. 10, n.1, p. 47-68, 2003.