



# XI Congresso UFPE de Ciências Contábeis

## Previsão da arrecadação de ICMS para o Estado de Sergipe: Modelagem a partir da metodologia de Box e Jenkins

**Tiago de Moura Soeiro**

Universidade Federal de Pernambuco

**Rodrigo Vicente dos Prazeres**

Universidade Federal de Pernambuco

**João Gabriel Nascimento de Araújo**

Universidade Federal de Pernambuco

**Wilton Bernardino da Silva**

Universidade Federal de Pernambuco

### Resumo

Este trabalho teve por objetivo conhecer o comportamento da arrecadação e contribuir, através da modelagem de Box e Jenkins (1976), com o melhor modelo de previsão das receitas do Imposto sobre Circulação de Mercadorias e Serviços (ICMS) do Estado de Sergipe, tomando por base o período de janeiro de 2005 a dezembro de 2014. Como metodologia, foram utilizados os modelos SARIMA, estimados por dois pacotes do R: o "forecast", que contém o algoritmo "auto.arima", e o "seasonal" que contém o algoritmo "seas", ambos para modelagem e previsão de series temporais através de um modelo autoregressivo integrado de média móvel. As análises evidenciaram que cada algoritmo elegeu um modelo diferente. O "auto.arima" elegeu (1,0,1) (2,0,0) e o "seas" (0,1,1) (0,1,1). Com a finalidade de avaliar a superioridade de um dos modelos em relação ao outro, foram avaliados os critérios de Raiz Quadrada do Erro Quadrado Médio de Previsão (RQEMP), Erro Absoluto Médio de Previsão (EAMP) e Coeficiente de Desigualdade de Theiler (CDT). Em todos os critérios, o modelo que se mostrou superior foi o estimado pelo algoritmo "seas".

**Palavras-chave:** Previsão; ICMS; SARIMA;

### 1 Introdução

Dentro das atribuições do governo, como gestor dos recursos públicos, projetar a arrecadação de impostos constitui uma tarefa básica no processo de planejamento de orçamentos, possibilitando ao Estado fixar de forma mais adequada o montante a ser gasto com despesas e investimentos, com a finalidade de evitar ou minimizar possíveis déficits nas contas públicas, alocando os recursos de forma mais eficiente.

Na esfera estadual, o Imposto sobre Operações Relativas à Circulação de Mercadorias e de Serviços de Transporte Interestadual e Intermunicipal e de Comunicações (ICMS) responde pela maior parcela da arrecadação no que diz respeito às receitas tributárias (Rezende, 2009).

Diante da representatividade do ICMS nas receitas tributárias estaduais, diversas pesquisas foram realizadas com a finalidade de tentar prever de forma mais acurada a arrecadação do ICMS através do emprego de modelos temporais, a exemplo dos estudos de Arraes e Chumvichitra (1996), Passos e Ramos (2005), Marques e Uchôa (2006), Cruz (2007) e Camargos (2008), Passos, Ramos e Almeida (2011) concluíram sobre a superioridade dos modelos de autorregressivos em relação a outros modelos de previsão de arrecadação do ICMS.



# XI Congresso UFPE de Ciências Contábeis

Neste estudo, se propõe a utilização dos modelos autoregressivos integrados de médias móveis com sazonalidade (SARIMA), seguindo a metodologia de Box e Jenkins (1976) para uma série temporal com observações mensais da arrecadação do ICMS do estado de Sergipe para o período de janeiro de 2005 a dezembro de 2014, sendo geradas previsões para a arrecadação do ICMS para os anos de 2015 e 2016.

Assim, seguindo a linha dos estudos supramencionados, o presente trabalho objetiva conhecer o comportamento da arrecadação de ICMS para o estado de Sergipe com a finalidade de contribuir com o melhor modelo de previsão das receitas de ICMS através da modelagem de Box e Jenkins (1976).

Após esta introdução, o restante do trabalho apresenta o referencial teórico, seguido da metodologia de pesquisa e da apresentação dos resultados. Por fim, as conclusões.

## 2 Referencial Teórico

Este tópico está subdividido em duas partes. Na primeira, é feita uma breve explicação sobre métodos de previsão, enquanto que na segunda se apresenta, de forma simplificada, o método Box-Jenkins. Esses tópicos servem para elucidar a escolha metodológica do estudo.

### 2.1 Métodos de previsão

Conforme Teti (2009), prever arrecadação é pressupor, permitir antecipar resultados futuros, tomando como base dados históricos e modelando-os com ferramentas de análises estatísticas. Existe uma grande variedade de modelos de previsão, podendo ser divididas em dois grandes grupos: modelos qualitativos e modelos quantitativos.

Os métodos qualitativos baseiam-se na presunção, isso é, em hipóteses ou conjecturas, a chamada aproximação não-extrapolativas. Esses métodos não apresentam uma formalidade matemática mínima para os resultados encontrados (Teti, 2009).

Dentre os métodos qualitativos mais utilizados estão o de previsão conjectural e o de previsão de consenso. Ambos os métodos consistem em análises baseadas na experiência prática do analista ou grupo responsável pela previsão. Esses modelos são mais indicados em casos de poucos dados históricos e em ambientes de muita instabilidade política, econômica e/ou administrativa, uma vez que os modelos quantitativos não são, por vezes, capazes de captar as principais informações relevantes que afetam a projeção. Entretanto, Siqueira (2002) afirma que há uma série de desvantagens para esses modelos por serem baseados em métodos subjetivos.

Os métodos quantitativos se utilizam de dados históricos para prever valores, com a premissa de que as séries seguirão o mesmo padrão de comportamento, sendo essa sua principal desvantagem ou limitação. Tais métodos podem ser divididos em modelos casuais e não casuais.

Enquanto os modelos casuais se utilizam de variável dependente e variáveis explicativas (independentes), os não casuais consistem em metodologias que, segundo Gujarati (2006), contêm todas as informações necessárias e que, portanto, podem falar por si. De acordo com Siqueira (2002), os métodos quantitativos, em geral, apresentam resultados mais precisos que os qualitativos e que, dentre os métodos quantitativos, os não casuais, em geral, possuem maior acurácia que os casuais.

As previsões também podem ser classificadas em *ex-post* e *ex-ante*. As previsões *ex-post* se utilizam de variáveis que já são conhecidas com certeza para o período de previsão, enquanto a *ex-ante* usa variáveis que ainda não são conhecidas com precisão (TETI, 2009).

### 2.2. Metodologia Box e Jenkins



# XI Congresso UFPE de Ciências Contábeis

O método desenvolvido por Box e Jenkins (1976) permite, num estudo de séries temporais, prever em qual modelo a série melhor se ajusta, dentre o processo autoregressivo puro (determinar o valor de  $p$ ), processo de média móvel genuíno (encontrar o valor de  $q$ ), processo ARMA (determinando  $p$  e  $q$ ) e o modelo ARIMA (encontrando os valores de  $p$ ,  $d$ ,  $q$ ).

Nessa metodologia, conforme Gujarati (2006) é necessário seguir quatro passos, quais sejam:

- Identificação: esta etapa consiste em determinar os valores de  $p$ ,  $d$ ,  $q$  que são, respectivamente, a defasagem da série, a ordem de integração e a defasagem dos erros aleatórios. Esses valores são encontrados através da função de autocorrelação (ACF - *Autocorrelation Function*), função de autocorrelação parcial (PACF - *Partial Autocorrelation Function*) e do correlograma;
- Estimação: Estudados os ACF, PACF e o correlograma, são definidos os valores de  $p$ ,  $d$ ,  $q$  e se estima o modelo mais adequado;
- Verificação do diagnóstico: estimado o modelo, deve-se verificar se ele de fato é adequado para a série temporal ou não. Para o diagnóstico, podem ser usados os critérios de Akaike (AIC) e de Schwarz (BIC);
- Realização das previsões: cumpridas as etapas anteriores, utiliza-se o modelo para prever os valores futuros da série temporal.

Na seleção do modelo, portanto, deve-se fazer uso do princípio da parcimônia, recomendado por Box e Jenkins (1976), que tem como preceito minimizar a quantidade de parâmetros e defasagens. Outro ponto relevante é verificar a significância dos regressores e também se os resíduos encontrados são ruídos brancos.

### 3. Metodologia

Para a realização deste estudo foram coletados no sítio eletrônico do Instituto de Pesquisa Economica Aplicada (IPEA): (i) a série mensal de ICMS do estado de Sergipe entre os anos de 2005 e 2014 resultando em 160 observações; (ii) e para realizar a correção dos valores do ICMS para a mesma base temporal, usou-se o índice IGP-DI mensal. Por fim, a série histórica da arrecadação realizada do estado de Sergipe para os anos de 2015 e 2016 foram coletados no sítio eletrônico do Convênio Nacional de Política Fazendária (Confaz).

Para realizar a modelagem SARIMA, plotagem dos gráficos, e os demais testes foi utilizado o *software* R versão 3.3.2, com a interface *RStudio* 0.99.489.0.; e as tabelas com resultados pelo *Microsoft Excel* 2016.

Os pacotes necessários para a análise no R foram: (i) “*tseries*”, para a reconhecimento/plotagem de séries temporais; (ii) “*urca*”, para a realização do teste de Dirkey-Fuller Aumentado com a finalidade de verificar a existência de raízes unitárias, (iii) “*portes*”, para a realização dos testes Box-Pierce e Ljung-Box para identificação de autocorrelação, (iv) “*seasonal*” para avaliação do período sazonal, e (v) “*forecast*” para a modelagem da série SARIMA e realização das previsões.

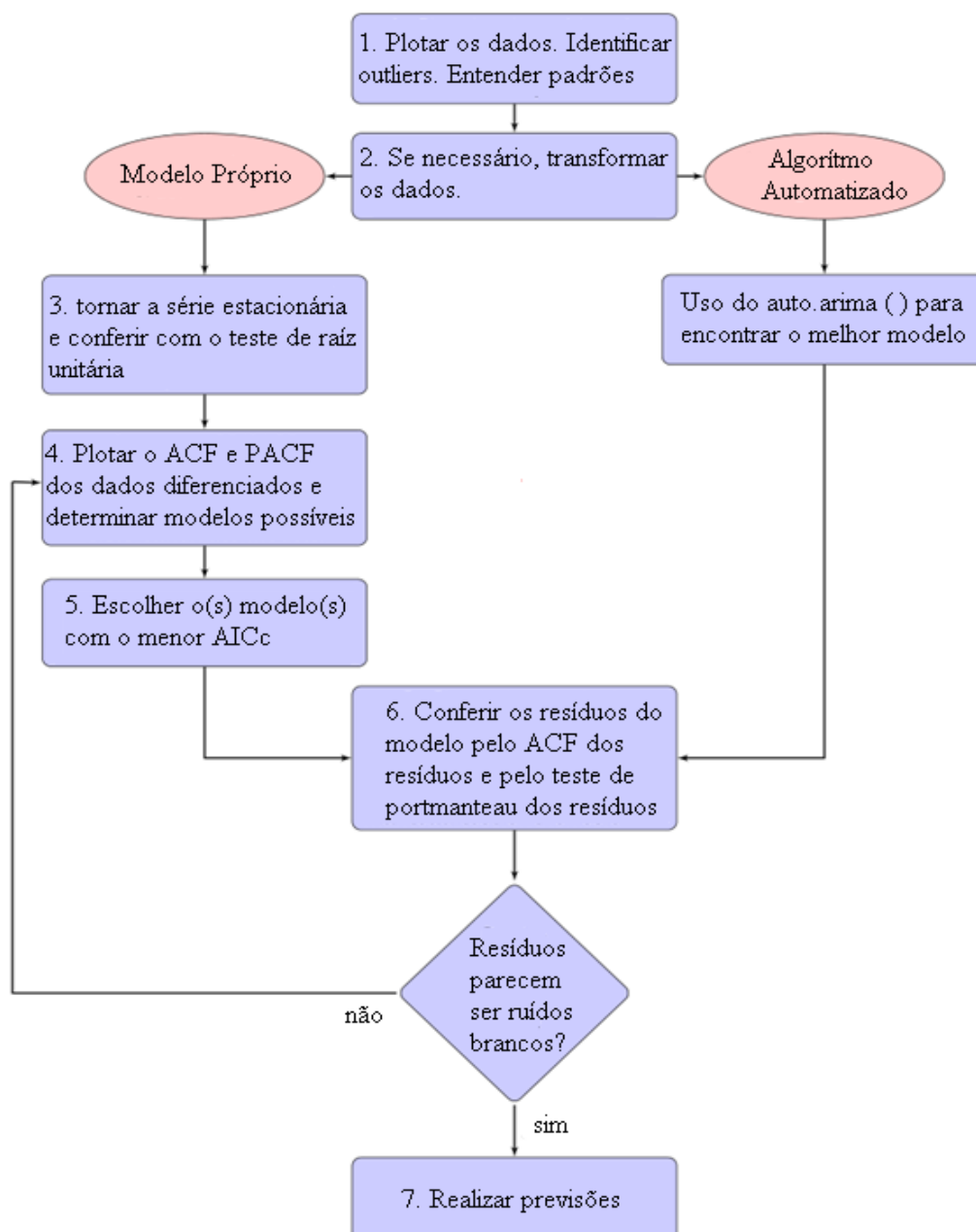
Para a realização da modelagem, atentou-se para os seguintes passos elencados, apresentados de forma gráfica na Figura 1:

1. Plotagem dos dados, correção dos valores pelo IGP-DI e identificação de observações discrepantes;
2. Transformação dos dados através de técnicas como a transformação Box-Cox para estabilizar a variância;



# XI Congresso UFPE de Ciências Contábeis

3. Julgamento sobre a estacionariedade: por não ser estacionário, realizou-se a primeira diferença para tornar os dados estacionários;
4. Examinar o ACF / PACF, para avaliar o modelo apropriado.
5. Realizar a modelagem para encontrar o Menor AIC que resultará no melhor modelo;
6. Conferir se os resíduos do modelo pelo ACF e teste de Portmanteau para resíduos, para identificar se são ruídos brancos, caso contrário o modelo precisa ser alterado.
7. Identificados os ruídos brancos, realizam-se as previsões.



**Figura 1-** esquema da modelagem da série temporal

Para avaliar o desempenho do modelo estimado, são calculadas quatro medidas de avaliação utilizadas na pesquisa de Pessoa, Coronel e Lima (2013). Desta forma, dado  $Y_t$  e  $\hat{Y}_t$  como valores observados e previstos, respectivamente, para um horizonte de previsão de  $T+1$ ,  $T+2$ , ...,  $T+h$ , os indicadores podem ser definidos como:



1. Raiz Quadrada do Erro Quadrado Médio de Previsão (RQEMP 1);

$$RQEMP = \sqrt{\sum_{t=T+1}^{T+h} \frac{(Y_t - \hat{Y}_t)^2}{h}} \quad (1)$$

2. Erro Absoluto Médio de Previsão (EAMP);

$$EAMP = \sum_{t=T+1}^{T+h} \frac{|Y_t - \hat{Y}_t|}{h} \quad (2)$$

3. Erro Absoluto Médio Percentual de Previsão (EAMPP); e,

$$EAMPP = \sum_{t=T+1}^{T+h} \frac{\frac{|Y_t - \hat{Y}_t|}{Y_t}}{h} \times 100 \quad (3)$$

4. Coeficiente de Desigualdade de Theil (CDT).

$$CDT = \frac{\sqrt{\sum_{t=T+1}^{T+h} \frac{(Y_t - \hat{Y}_t)^2}{h}}}{\sqrt{\sum_{t=T+1}^{T+h} \frac{Y_t^2}{h}} + \sqrt{\sum_{t=T+1}^{T+h} \frac{\hat{Y}_t^2}{h}}} \quad (4)$$

Assim, o critério de avaliação do desempenho do modelo estimado é o menor valor dos indicadores. Quanto menor, melhor.

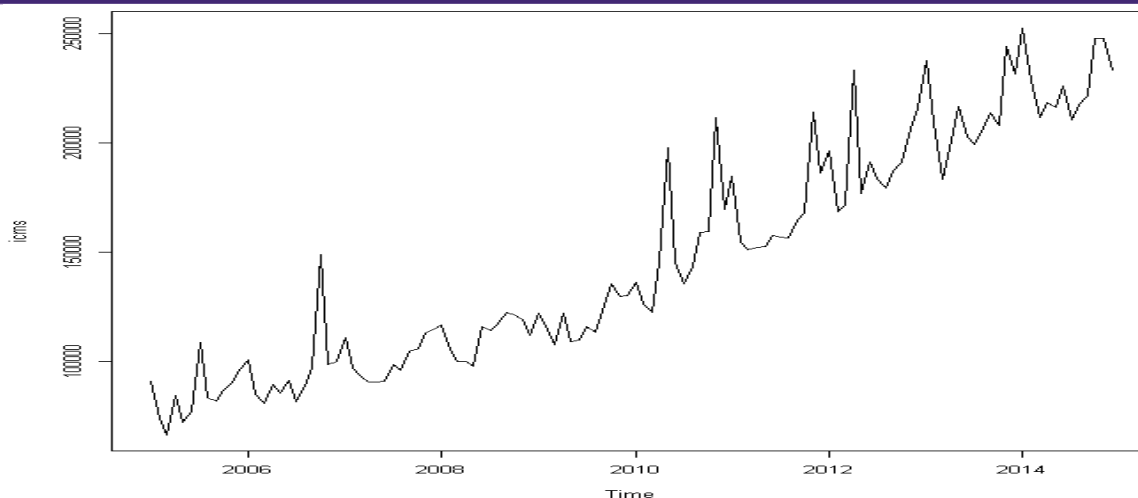
## 4 Apresentação dos Resultados

### 4.1 Análise da série de ICMS

A figura 1 apresenta a série de ICMS do Estado de Sergipe do período de janeiro de 2005 a dezembro de 2014. A partir da série é possível perceber uma tendência ascendente que inicialmente parece uma tendência determinística, o que leva a necessidade de remover o efeito na modelagem da série. A série também apresenta períodos de sazonalidade, que mostram uma volatilidade elevada.

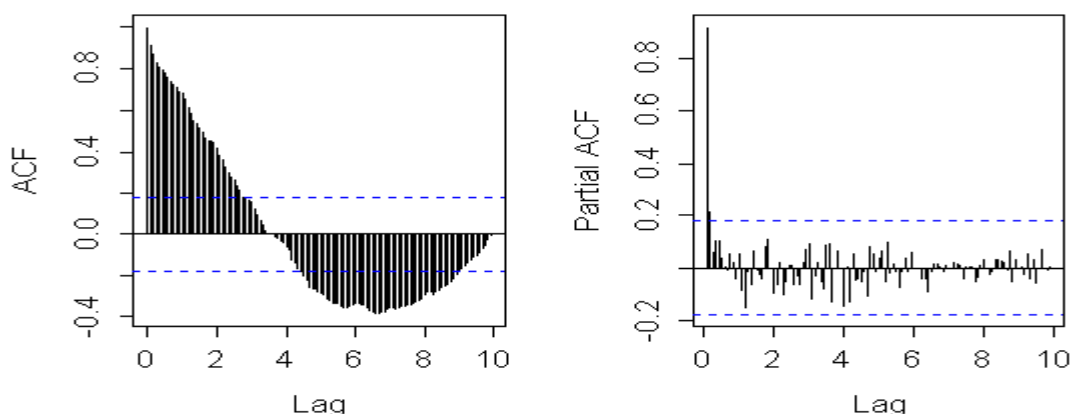


# XI Congresso UFPE de Ciências Contábeis



**Figura 1-** Série do ICMS do Estado de Sergipe no período de janeiro de 2005 a dezembro de 2014.

Conforme comentado anteriormente, na metodologia, procedeu-se a estimação do ACF e PACF para verificar a defasagem da série, a ordem de integração e a defasagem dos erros aleatórios.



**Figurass 2 e 3 – ACF e PACF da Série**

Como se observa nas figuras 2 e 3, com base no ACF da série percebe-se um decaimento exponencial, e o PACF apresenta comportamento de decaimento oscilatório que sugere a presença de padrões autoregressivos e média-móvel, apesar de picos além das bandas que sugerem autocorrelação dos dados.

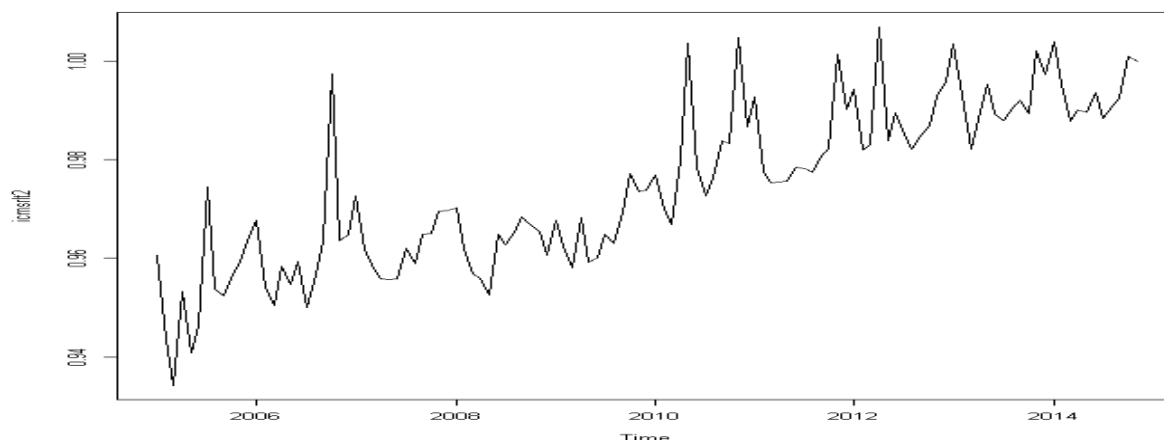
O teste Dickey-Fuller Aumentado (ADF) e o teste de Box-Pierce, expostos na Tabela 1, mostram, respectivamente, que a série não é estacionária, sendo necessária a diferenciação para torná-la estacionária e que a série é autocorrelacionada.

**Tabela 1 – Teste de raiz unitária e autocorrelação da série ICMS**

	<b>Estatística Calculada</b>	<b>Valor Crítico a 1%</b>
<b>Dickey-Fuller Aumentado (ADF)</b>	0.1134	-4,015341
<b>Box-Pierce</b>	Valor p	
	0,0000	

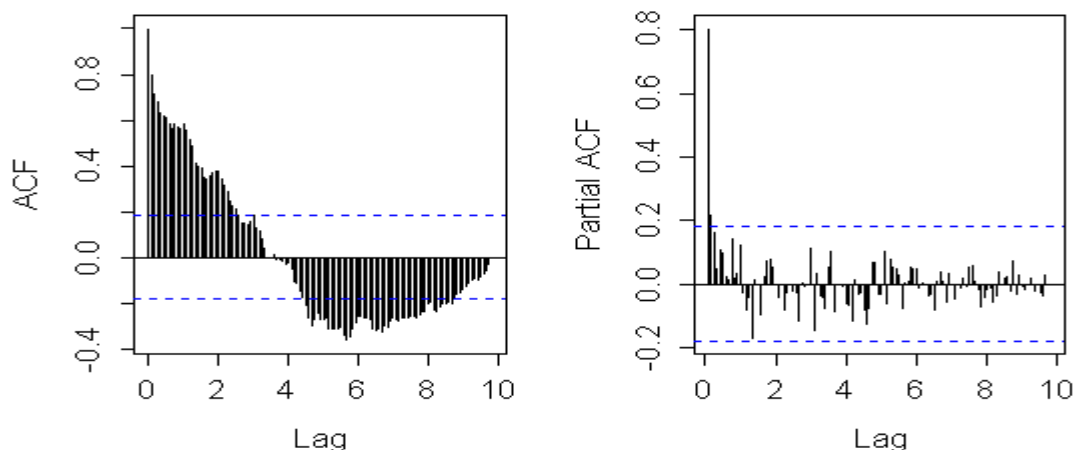
**Fonte:** dados da pesquisa

Ao se transformar a série (figura 4), buscou-se reduzir a variância da série. Além disso, a série se comporta com o mesmo padrão da série original, tendo seus picos mais evidenciados.



**Figura 4** – Série ICMS transformada

Da mesma forma que a série original, observadas nas figuras 1, 2 e 3, as figuras 5 e 6, remetidos aos ACF e PACF da série transformada apresentam, decaimento exponencial, e comportamento de decaimento oscilatório que sugere a presença de padrões autoregressivos e média-móvel respectivamente. Da mesma forma os picos além das bandas que sugerem autocorrelação dos dados.



**Figuras 5 e 6** – ACF e PACF da série de ICMS transformada

Os testes ADF e Box-Pierce, expostos na Tabela 2, mostram que a série é autorrelacionada e estacionária, não sendo necessária a diferenciação.

**Tabela 2** – Teste de raiz unitária e autocorrelação da série ICMS Transformada

	Estatística Calculada	Valor Crítico a 1%
Dirkey-Fuller Aumentado (ADF)	-11,8328	-4,015341
Box-Pierce	Valor p	
	0,0014	

Fonte: dados da pesquisa

Para confirmação, ao se decompor a série de ICMS (figura 7), percebe-se um padrão sazonal e de tendência que variam entre períodos. Com base nessas características apresentadas, julga-se que o modelo a ser aplicado é o SARIMA.



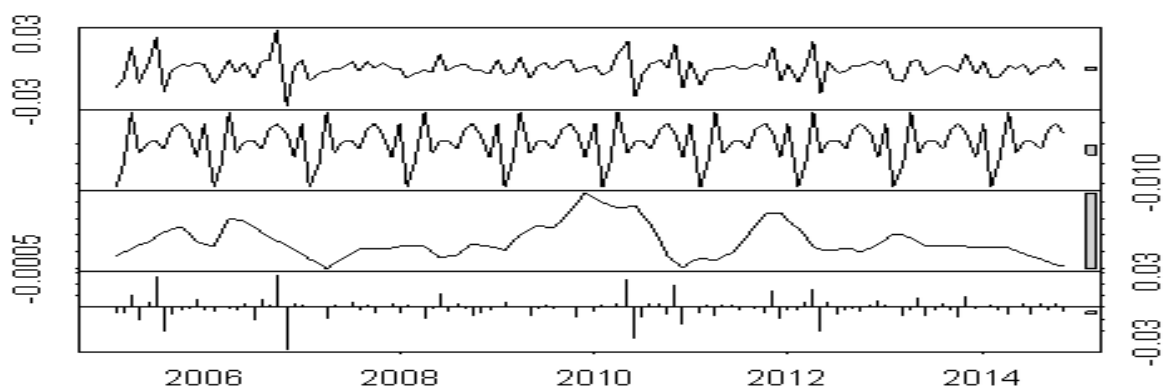


Figura 7 - Decomposição da Série de ICMS

## 4.2 Elegendo o melhor modelo

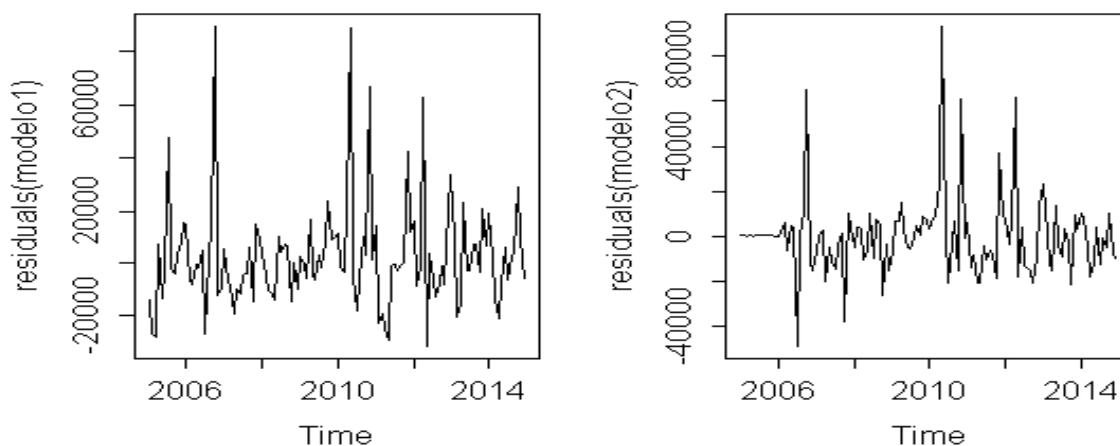
Para a seleção do melhor modelo estimou-se inicialmente pelo "auto.arima" selecionando um modelo SARIMA (1,0,1) (2,0,0). Contudo com base no que foi evidenciado na série, julgou-se não ser o mais adequado, devido aos períodos sazonais. Sendo assim, utilizou-se do "seas" para avaliar o melhor modelo sazonal para a série, encontrando o (0,1,1) (0,1,1) conhecido também como modelo Airlines. Superando o AIC, AICc e BIC do modelo do "auto.arima". Como pode ser observado na Tabela 3.

Tabela 3 – Comparativo entre os Modelos

	"auto.arima"	"seas"
<b>Melhor Modelo</b>	(1,0,1) (2,0,0)	(0 1 1) (0 1 1)
<b>AIC</b>	-790,65	-824,80

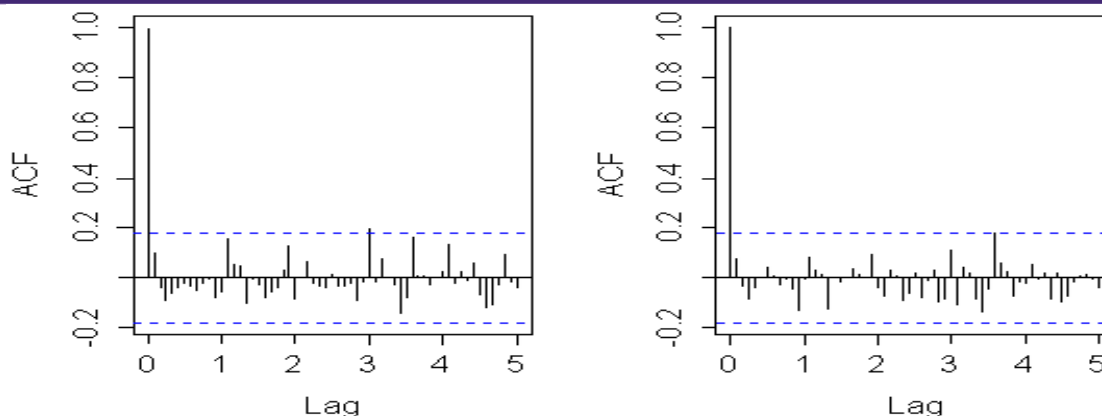
Fonte: Dados da Pesquisa

Para avaliar se os modelos são apropriados, procedeu-se a análise dos resíduos, plotando o gráfico dos resíduos (figuras 8 e 9) e o seu ACF (figuras 10 e 11). Desta forma, observa-se que o Modelo 2 apresenta melhores resultados que o modelo 1. Os gráficos dos resíduos (figuras 8 e 9) aparentaram ser ruídos brancos, mas ao se observar o ACF dos resíduos (figuras 10 e 11) apenas o "seas" manteve-se dentro das bandas.



Figuras 8 e 9 – Resíduos dos modelos

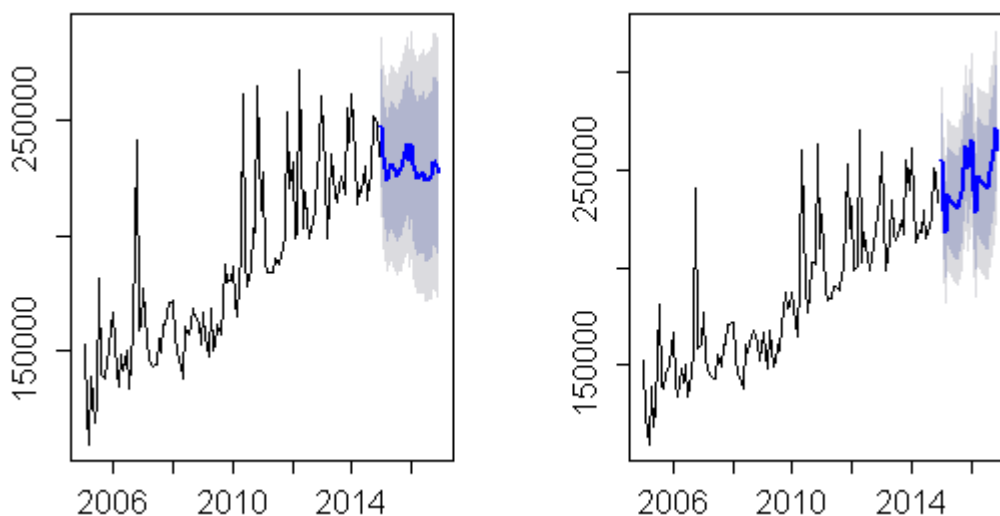




Figuras 10 e 11 - ACF dos resíduos dos modelos 1 e 2

## 4.4 Previsão

Realizada a confirmação dos resíduos, procederam-se as previsões conforme podem ser vistas nas figuras 12 e 13. Conforme se observa na previsão fornecida pelo "auto.arima", haveria uma mudança de tendência no comportamento da série. Por outro lado o "seas" manteve na sua previsão um comportamento sazonal similar ao observado na série de ICMS.



Figuras 12 e 13 – Previsão do ICMS pelos "auto.arima" e "seas"

Visando confirmar a superioridade do modelo 2 sobre o modelo 1, foram calculados os indicadores RQEPM, EAMP, EAMPP e CDT, conforme pode ser observado na Tabela 4. De acordo com o critério de superioridade adotado, quanto menor, melhor o modelo, observa-se que o "seas" foi superior ao "auto.arima" em todos os indicadores utilizados.

Tabela 4 – Indicadores de diferença				
	RQEPM	EAMP	EAMPP	CDT
Modelo1 "auto.arima"	32030032790	32030032790	5,21%	10413756
Modelo2 "seas"	434296444,8	434296444,8	0,61%	139548,6

Fonte: Dados da Pesquisa

## 5 Conclusões



## XI Congresso UFPE de Ciências Contábeis

Este trabalho teve por objetivo realizar uma previsão para a série de arrecadação do Imposto sobre Circulação de Mercadorias e Serviços (ICMS) do Estado de Sergipe, com base na arrecadação do período de janeiro de 2005 a dezembro de 2014.

A metodologia de pesquisa consistiu em estimar modelos SARIMA através de dois pacotes do R: o “*forecast*”, que contém o algoritmo “*auto.arima*”, e o “*seasonal*” que contém o algoritmo “*seas*”, ambos para modelagem e previsão de series temporais através de um modelo autoregressivo integrado de média móvel.

A partir dos procedimentos metodológicos, foi possível observar que cada algoritmo elegeu um modelo diferente. O “*auto.arima*” elegeu (1,0,1) (2,0,0) e o “*seas*” (0,1,1) (0,1,1). Com a finalidade de avaliar a superioridade de um dos modelos em relação ao outro, foram avaliados os critérios de Raiz Quadrada do Erro Quadrado Médio de Previsão (RQEMP), Erro Absoluto Médio de Previsão (EAMP) e Coeficiente de Desigualdade de Theiler (CDT).

Como resultados dos procedimentos metodológicos, o modelo estimado a partir do algoritmo “*seas*” se mostrou superior, modelo este que indicou o crescimento com comportamento sazonal para a arrecadação do ICMS do estado de Sergipe, ao passo que o modelo estimado a partir do algoritmo “*auto.arima*” forneceu uma mudança de tendência na arrecadação do ICMS de Sergipe, indicando um decaimento da série.

### Referências

- Arraes, Ronaldo de Albuquerque & Chumvichitra, Pichai. (1996). Modelos autoregressivos e poder de previsão: uma aplicação com o ICMS. Texto para discussão, n. 152.
- Box, G. E. P. & Jenkins, G. M. (1976). Time series analysis: forecasting and control. Revised Edition. San Francisco: Holden-Day.
- Camargos, A. A. B. (2008). Modelos de previsão de arrecadação tributária do estado de São Paulo: ICMS, IPVA, ICMD e Taxas. 170 f. Dissertação (Mestrado Profissional em Economia) - IBMEC, São Paulo.
- Confaz. (2017). Disponível em: < [www.confaz.fazenda.gov.br](http://www.confaz.fazenda.gov.br)>. Acesso em: 24 de Abr.
- Cruz, C. C. S. (2007). Análise de séries temporais para uma previsão mensal do ICMS: O caso do Piauí. Dissertação (Mestrado em economia). Universidade Federal do Ceará. Ceará.
- Gujarati, D. N. & Porter, D. C. (2011). Econometria Básica. 5. ed. Porto Alegre: AMGH.
- Ipeadata. IGPDI-Mensal. (2016). Disponível em: <[www.ipeadata.gov.br](http://www.ipeadata.gov.br)>. Acesso em: 05 de Dez.
- Marques, C. A. G. & Uchôa, C. F. A. (2006). Estimação e previsão do ICMS na Bahia. Revista Desenhavia, Salvador, v. 3, n. 5, p. 195-211.
- Passos, J. J., Ramos, E.M.L.S. & Almeida, S. S. (2011). Utilização de modelos ARIMA para previsão da arrecadação de ICMS do estado do Pará. Reunião Regional Da Abe, v. 37.
- Passos, J. J. & Ramos, E. M. L. S. (2005). Modelagem estatística para previsão de arrecadação de ICMS do Estado do Pará. In: XI Reunion de Trabajo em Procesamiento de la Informacion y Control, 2005, Rio Cuarto. XI Reunion de Trabajo em Procesamiento de la Informacion y Control. Rio Cuarto: Universidad Nacional de Rio Cuarto, v. 01. p. 248-253.



## XI Congresso UFPE de Ciências Contábeis

Pessoa, F. M. G., Coronel, D. A. & Lima, J. E. (2012). Previsão de arrecadação de ICMS para o estado de Minas Gerais: uma comparação entre modelos Arima e Arfima. *Revista Brasileira de Gestão e Desenvolvimento Regional*. v. 9, n. 2, p. 47-64, mai-ago/2013, Taubaté, SP, Brasil.

Rezende, Fernando. (2009). *A reforma tributária e a Federação*. Rio de Janeiro: Fundação Getúlio Vargas.

Siqueira M. L. (2002). *Modelos de séries temporais para a previsão da arrecadação tributária federal*. Dissertação de Mestrado. Universidade de Pernambuco, Recife.

Teti, A. C. C. (2009). *Modelo de Previsão da Receita Tributária: O Caso do ICMS no Estado de Pernambuco*. Dissertação (Mestrado em Economia). Universidade Federal de Pernambuco. Recife.