

PREVISÃO DE DADOS DE SÉRIES TEMPORAIS DOS PIBS BRASILEIRO E AMERICANO ATRAVÉS DO MODELO ARIMA

Juliana Helena Daroz Gaudencio (UNIFEI)

ju_gaudencio@hotmail.com

Giseli Valentim Rocha (UNIFEI)

giseli.rocha@gmail.com

Alexandre Correa Grassi Bissacot (UNIFEI)

alexandrecgb@hotmail.com

Pedro Paulo Balestrassi (UNIFEI)

ppbalestrassi@gmail.com



A previsão de eventos futuros é um elemento crítico em diversos processos de planejamento e tomada de decisão e, em sua maioria, os problemas de previsão se remetem à utilização de dados de séries temporais. Tais modelos empregam propriedades estatísticas de dados históricos para especificar um modelo formal e estimar os parâmetros desconhecidos. O objetivo deste trabalho é analisar a aplicação do modelo ARIMA para a previsão de valores em uma série temporal que utiliza dados anuais dos PIBs brasileiro e americano, per capita, no período de 1960 até 2012. A manipulação dos dados foi baseada na análise gráfica e em testes estatísticos da própria metodologia, através da qual, observou-se que o modelo ARIMA (1,1,0) foi o mais adequado para a modelagem de ambos os conjuntos de dados. Como resultado, constatou-se que o PIB brasileiro levaria 22 anos para superar o PIB americano de 2012 e, além disso, se considerássemos que a taxa de crescimento de ambos os países se comportasse da mesma maneira que aquela observada nas séries de dados analisadas, o PIB brasileiro superaria o americano em 74 anos.

Palavras-chaves: Modelo ARIMA, Séries Temporais, Previsão do PIB

1. Introdução

Saber o que deverá acontecer no dia de amanhã é uma das perguntas que os seres humanos buscam responder a séculos. De acordo com Barbancho (1970), previsão é uma manifestação relativa a sucessos desconhecidos em um futuro determinado. Além disso, ela não constitui um fim em si, mas um meio de fornecer informações e subsídios para uma consequente tomada de decisão, visando atingir determinados objetivos (MORETTIN e TOLOI, 1981). E, no mundo dos negócios, a curiosidade humana vai além do uso de simples técnicas empíricas, pois cada dia mais, as previsões são mais importantes para o sucesso empresarial.

Assim, previsão pode ser usada como insumo base para a função de planejamento, que conduzirá as atividades e dos meios apropriados de desempenho organizacional à efetividade. Um exemplo é a previsão realizada por governos de países com relação ao crescimento do seu PIB (Produto Interno Bruto) que é um dos principais indicadores que demonstram a realidade econômica de um país ou região. Tal indicador nada mais é do que a mensuração de todos os bens e serviços, ou seja, de toda a riqueza produzida.

Em 2012, o PIB per capita do Brasil foi de US\$ 11.339.521 bilhões, e o dos EUA, de US\$ 49.965.271 bilhões, segundo o Fundo Monetário Internacional (FMI). Esse número foi calculado com base na paridade do poder de compra, uma cotação do dólar que desconta a diferença do poder aquisitivo em cada país.

Para se prever o valor do PIB de ambos os países em anos futuros existem modelos de previsão, dentre os diversos modelos de séries temporais, os modelos ARIMA (*Autoregressive Integrated Moving Average*) são aplicados no caso específico de séries não estacionárias ou estacionárias se compostos por três filtros: auto regressivo, média móvel e diferenciação. Para estruturar-se um modelo ARIMA em uma série temporal, há três estágios a serem considerados: identificação, estimação e verificação ou diagnóstico (MORETTIN, 2002). O modelo foi sistematizado em 1976 pelos estatísticos George Box e Gwilym Jenkins, o que torna o modelo conhecido também por Modelo de Box-Jenkins.

Assim sendo, o presente trabalho tem como objetivo analisar a aplicação do modelo ARIMA para a previsão de valores em uma série temporal utilizando dados dos PIBs brasileiro e americano *per capita*, considerando seus valores anuais desde 1960 até 2012. Como objetivos

específicos: será analisado em quantos anos o PIB brasileiro se igualaria ao dos EUA atualmente (2012) e mantendo-se a taxa atual de crescimento anual de ambos os PIBs, em quantos anos o PIB brasileiro ultrapassaria o PIB americano. As séries de dados com os valores anuais dos PIBs foram obtidas no site *datamarket.com*. Ainda, analisou-se o comportamento estacionário de cada série e, também, definiu-se o modelo ARIMA mais adequado para os resultados das previsões avaliadas.

O artigo está estruturado da seguinte forma: a seção 2 discutirá os modelos utilizados, com destaque para o *ARIMA*, a seção 3 apresentará o procedimento metodológico; a seção 4 mostrará os resultados obtidos, sua interpretação, discussão e, por fim, a seção 5 apresentará as conclusões.

2. Referencial teórico

2.1 Métodos qualitativos e métodos quantitativos

As previsões por métodos qualitativos baseiam-se em opiniões pessoais, ou seja, esses métodos baseiam-se no julgamento e ou opiniões de pessoas sobre tendências futuras, preferências e mudanças tecnológicas (SCHWITZK, 2001). Portanto, consideram o relato ou a posição das pessoas, com experiência em casa ramo de atividade como gerentes, vendedores, clientes e fornecedores.

Para Armstrong e Collopy (1998), os métodos qualitativos são úteis porque os especialistas geralmente têm conhecimento de eventos recentes dos quais os efeitos ainda não podem ser observados em séries temporais, ou de eventos que ocorram no passado, mas são esperados no futuro. Apesar da subjetividade deste método e do uso informal, ele tem sido usado em casos específicos.

Para Armstrong e Collopy (1998), os métodos quantitativos ou estatísticos são menos tendenciosos e usam de forma mais eficiente os dados passados. Além disso, esses métodos são confiáveis, pois, com os mesmos dados, eles produzirão as mesmas previsões. Ainda segundo esses autores, os processos estatísticos possuem certa miopia, pois eles representam apenas os dados a que estão relacionados.

Uma série temporal ou série histórica pode ser entendida como uma sequência de dados obtidos em intervalos regulares de tempo durante um período determinado. Para a análise de

uma série temporal, inicialmente deve-se modelar o fenômeno estudado, sendo que, a partir deste ponto, pode ser feita a descrição do comportamento da série, as suas estimativas e finalmente a avaliação dos fatores que influenciam no comportamento da série, tendo em vista a definição da relação causa e efeito.

Segundo Montgomery (2008), as atividades no processo de previsão são:

- a) Definição do problema;
- b) Coleta de dados;
- c) Análise de dados;
- d) Seleção e ajuste do modelo;
- e) Validação do modelo;
- f) Aplicação da previsão;
- g) Monitoramento da performance.

2.2 ARIMA

De acordo com DeLurgio (1998) nenhum outro modelo de previsão univariável tem sido tão amplamente discutido quanto a construção dos modelos ARIMA. Esse modelo foi popularizado na década de 1970, por George Box e Gwilym Jenkins, os nomes são geralmente utilizados como sinônimos na aplicação geral de modelos ARIMA (Box-Jenkins) para analisar séries temporais e fazer previsões.

O propósito da análise ARIMA é encontrar um modelo que represente precisamente os padrões passados e futuros das séries temporais. Assim, os modelos ARIMA são representados como $ARIMA(p, d, q)$ referindo-se, respectivamente, às ordens de auto regressão, de integração e de média móvel:

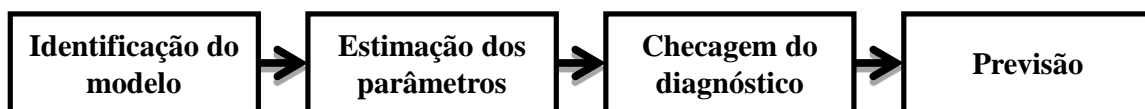
- p é o número de termos auto regressivos;
- d é o número de diferenças; e
- q é o número de termos da média móvel.

No caso de $d = 0$, temos o modelo $ARMA(p, q)$, e no caso de também $q = 0$, temos os modelo $AR(p)$. O modelo $ARIMA(0, 1, 0)$ é o passeio aleatório (*randomwalk*).

O modelo ARIMA, portanto, é um caso geral dos modelos propostos por Box e Jenkins (1976), que é apropriado para descrever séries não estacionárias. Ou seja, séries em que a média não é constante no período de análise, nas quais os parâmetros quase sempre são pequenos, apresentando tendência e ou sazonalidade.

Segundo Zhang (2003), a metodologia de Box-Jenkins inclui três processos iterativos que consistem na identificação do modelo, estimação dos parâmetros e a checagem do diagnóstico. Na fase de identificação, é necessária a transformação dos dados para tornar a série temporal estacionária. Quando o modelo é especificado, a estimação dos parâmetros é direta. Os parâmetros são estimados de maneira que a medida geral dos erros seja minimizada. O último passo é a checagem do diagnóstico para verificar a adequabilidade do modelo. As fases são tipicamente repetidas até que um modelo satisfatório seja selecionado e utilizado para, finalmente, realizar previsões. O esquema de todo o procedimento para previsões é ilustrado na Figura 1.

Figura 1 – Modelo para realização de previsões



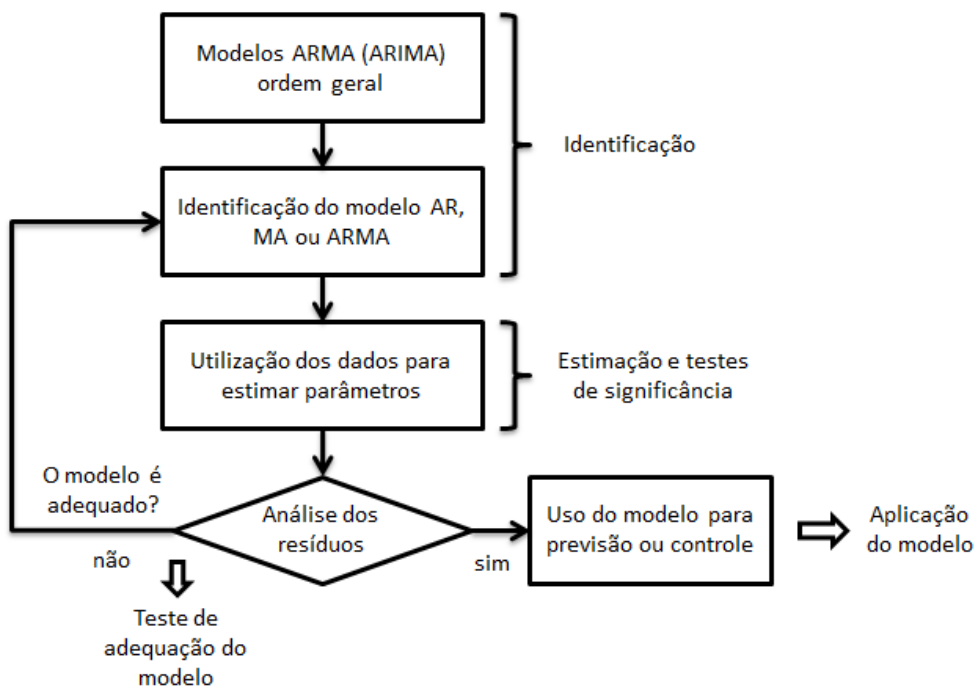
Fonte: Adaptado de Zhang (2003)

3. Procedimentos metodológicos

A metodologia utilizada neste estudo, a ARIMA, parte do princípio de que os modelos econométricos podem ser desenvolvidos a partir da informação contida nos próprios dados. Como mencionado anteriormente, essa metodologia foi desenvolvida inicialmente por Box e Jenkins (1976).

Os dados coletados referem-se ao PIB per capita brasileiro, entre os períodos históricos de 1960 a 2012, totalizando 52 observações, obtidas por meio do banco de dados do site *datamarket.com* (2013). Foram seguidos os seguintes passos conforme demonstrado pela Figura 2. Para a manipulação dos procedimentos estatísticos do presente estudo, foram utilizados neste estudo os seguintes softwares computacionais: Minitab® e Microsoft Excel®.

Figura 2- Esquema para utilização da metodologia Box & Jenkins



Fonte: DeLurgio (1998)

4. Análise dos resultados

A Tabela 1 mostra a série de dados do PIB per capita brasileiro entre os anos de 1960 e 2012.

Tabela 1 – Série de dados do PIB per capita brasileiro

Ano	PIB per capita - Brasil (\$ - dólares)	Ano	PIB per capita - Brasil (\$ - dólares)
1960	208.387	1987	2.074.410
1961	203.190	1988	2.287.237
1962	257.823	1989	2.893.658
1963	289.058	1990	3.086.915
1964	258.628	1991	2.677.150
1965	258.239	1992	2.526.337
1966	312.063	1993	2.791.519
1967	343.526	1994	3.426.102
1968	370.725	1995	4.749.814
1969	399.733	1996	5.107.794
1970	440.635	1997	5.219.098

1971	499.904	1998	4.979.145
1972	580.671	1999	3.411.866
1973	767.995	2000	3.694.463
1974	994.726	2001	3.128.145
1975	1.143.132	2002	2.810.695
1976	1.377.857	2003	3.039.672
1977	1.552.781	2004	3.607.192
1978	1.728.643	2005	4.739.305
1979	1.891.706	2006	5.787.976
1980	1.930.538	2007	7.194.079
1981	2.115.074	2008	8.622.709
1982	2.208.833	2009	8.373.340
1983	1.558.419	2010	10.978.094
1984	1.567.316	2011	12.575.979
1985	1.636.601	2012	11.339.521
1986	1.928.720		

Fonte: Fundo Monetário Internacional (FMI)

A previsão desses dados foi realizada através das três fases do método ARIMA já mencionadas anteriormente. Tais fases são detalhadas de acordo com um conjunto de passos descritos a seguir.

O primeiro passo foi estabilizar a variância dos dados da série e, por isso, realizou-se a transformação logarítmica dos dados através do Logaritmo Neperiano (Ln). Na Figura 3 é possível visualizar a série original plotada e, também, a série original transformada com a aplicação do Ln.

Visualizando-se a Figura 3, nota-se que a série original e, também, a série transformada não são estacionárias, ou seja, em intervalos de tempos diferentes a média dos dados não é constante.

O segundo passo foi a transformação da série não estacionária em estacionária que é realizado através da diferença da série transformada com lag1. Em seguida, ao se plotar a série com diferença de lag1 (Figura 4) notou-se que a série se tornou estacionária. A estacionaridade da

série também pôde ser notada através dos gráficos obtidos pela carta de controle (Figura 5) na qual apenas dois pontos estão localizados fora do intervalo desejado. Isto se deve pela ascensão do PIB brasileiro entre os anos de 1978/1979 e, também, pela abrupta queda entre os anos de 1998/1999.

Figura 3 – Séries original e transformada (Ln)

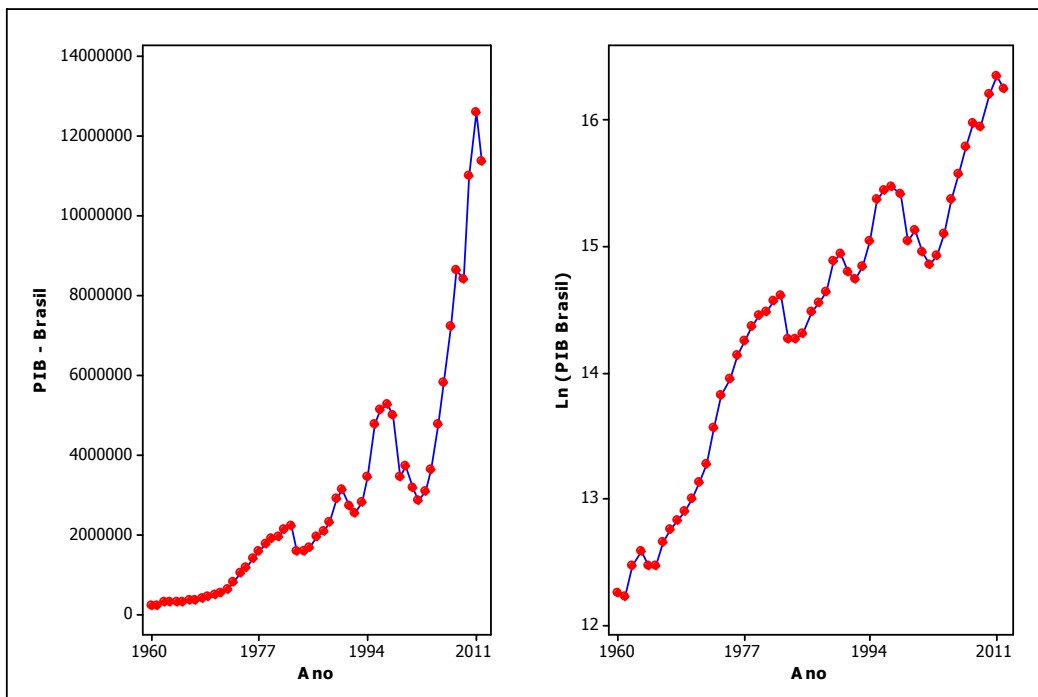


Figura 4 – Série com diferença de lag1

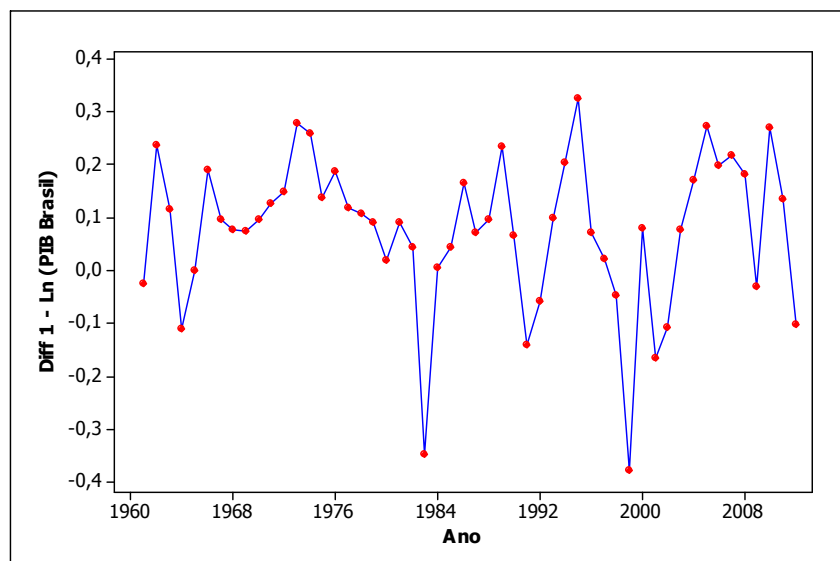
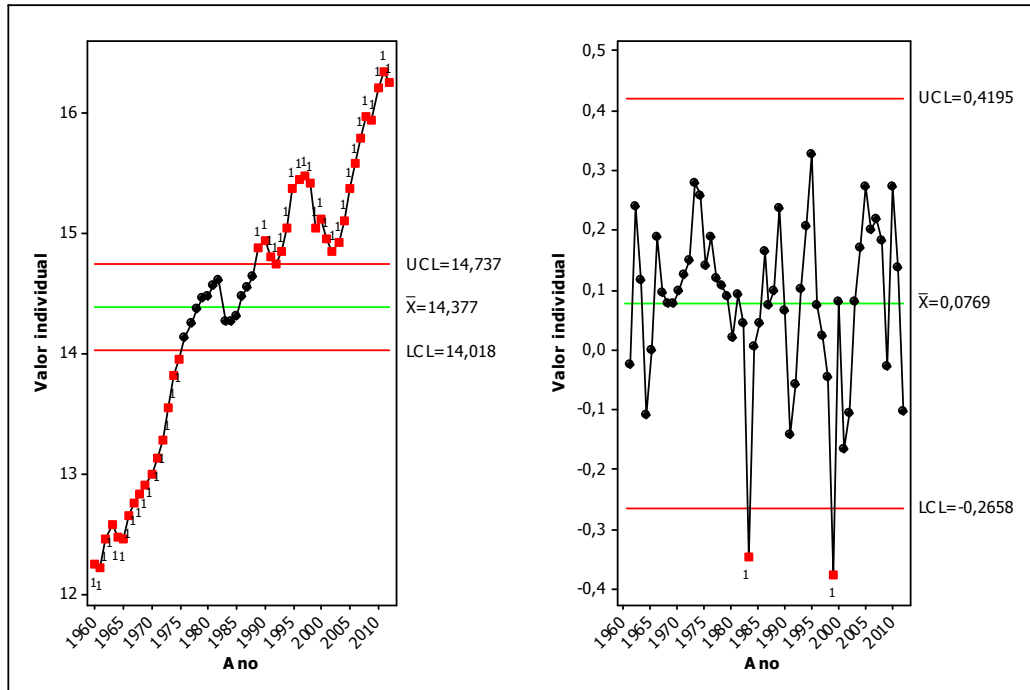
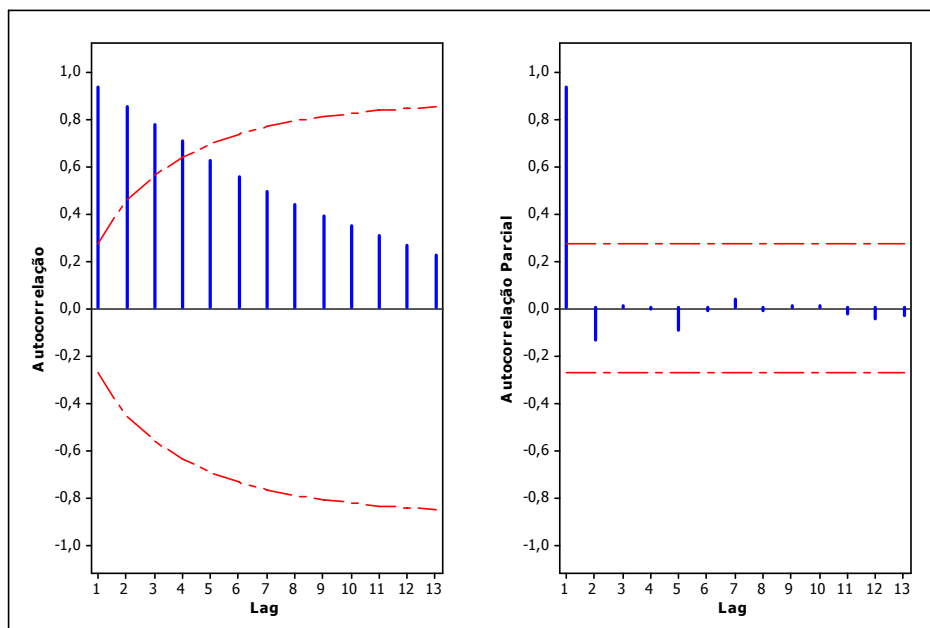


Figura 5 – Carta de controle da série transformada e da série com diferença



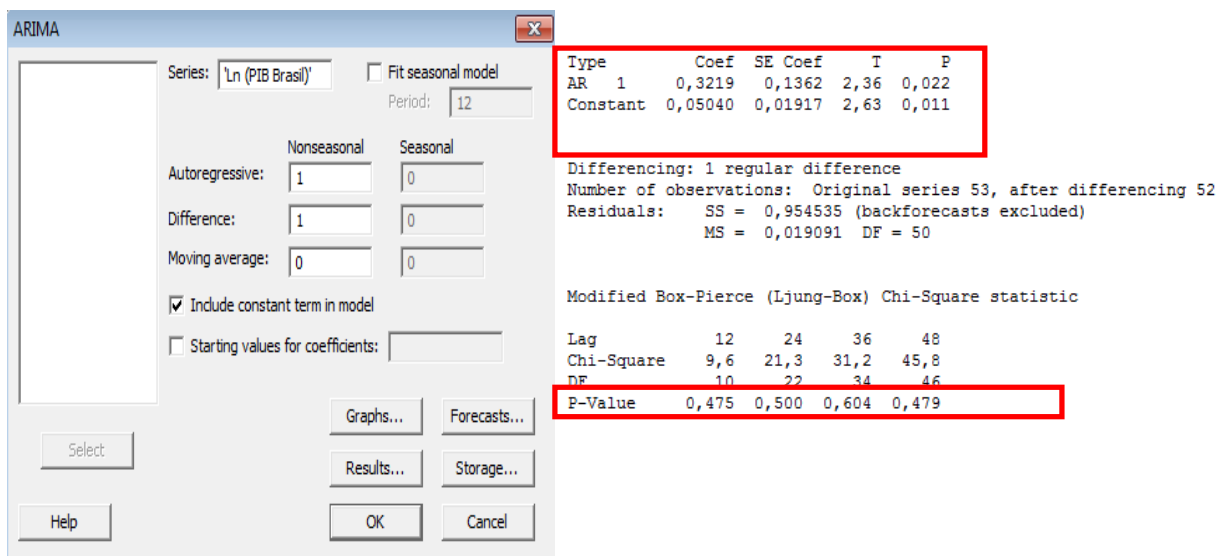
O terceiro passo foi a identificação do potencial modelo para a previsão através das análises dos gráficos ACF (Autocorrelação) e PACF (Autocorrelação Parcial) conforme a Figura 6.

Figura 6 – Gráficos ACF e PACF



Conforme visto na Figura 6, concluiu-se que o potencial modelo ARIMA identificado é o AR(1), D(1) e MA(0), ou seja, com o parâmetro $p=1$, $d=1$ e $q=0$ (1,1,0). Com estas informações, iniciou-se o quarto passo que foi a inserção dos parâmetros. A Figura 7 mostra os campos preenchidos no software e o respectivo resultado do modelo gerado (1,1,0).

Figura 7 – Resultado do modelo ARIMA (1,1,0)



De acordo com a Figura 7, pôde-se concluir que os parâmetros estimados caracterizam um bom modelo, pois o valor de p -value do AR é menor do que 0,05 e isso representa que o termo é significativo. Pôde-se observar, também, que os resultados de p -value do teste de Ljung-Box estão adequados, pois os valores são maiores do que 0,05.

Após a confirmação de que os parâmetros escolhidos para o modelo ARIMA caracterizam-no como um bom modelo, realizou-se o quinto passo que foi a análise dos resíduos como visto nas Figuras 8 e 9.

Figura 8 – Gráficos dos resíduos de ACF e de PACF

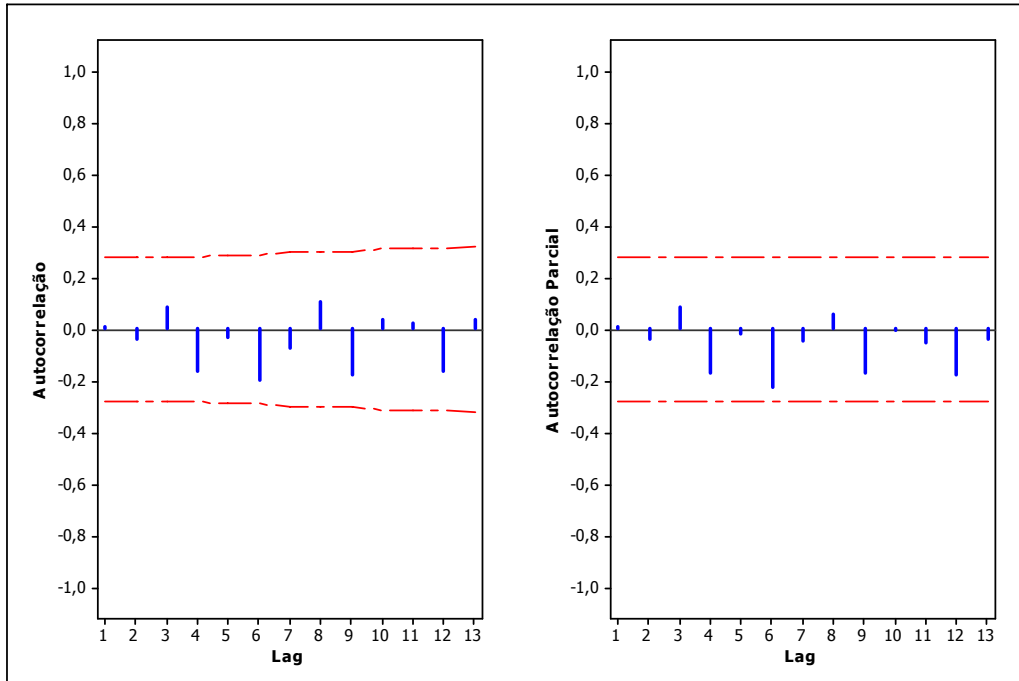
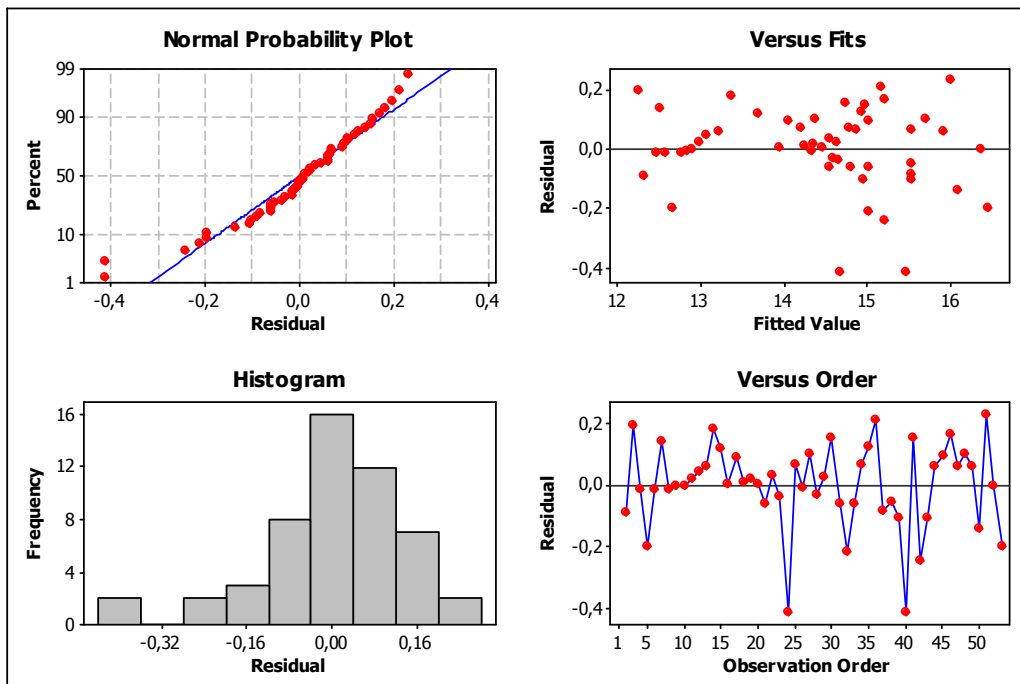


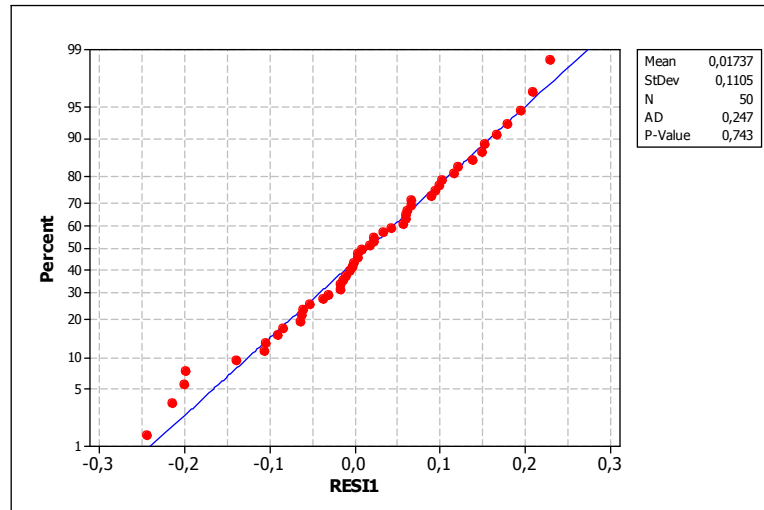
Figura 9 – Gráfico dos resíduos da série transformada (Ln)



A Figura 8 mostra que tanto os resíduos de ACF como o de PACF possuem um comportamento adequado, sem problemas aparente. Já a Figura 9, mostra dois resíduos que podem ser determinados como *outliers* e a presença destes faz com que o histograma não tenha uma aparência de distribuição normal. Ao se retirar os dois pontos caracterizados como

outliers, foi realizado o teste de normalidade de Anderson Darling. A Figura 10 mostra o resultado do teste.

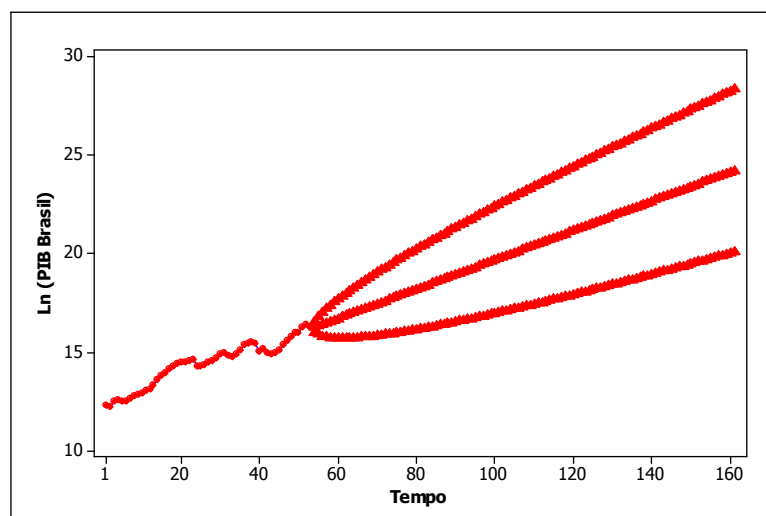
Figura 10 – Teste de normalidade dos resíduos da série transformada



Com o teste de Anderson Darling, obteve-se o valor de *p-value* de 0,743, ou seja, maior do que 0,05 e, assim, comprovando que os resíduos da série transformada são normalmente distribuídos.

Por fim, foi realizado o sexto passo que calcula os valores previstos do PIB per capita brasileiro. A quantidade de pontos que se desejava prever foi inserida. Para este presente trabalho, foram previstos 108 pontos com o intuito de analisarmos em longo prazo o comportamento crescente de ambos os PIBs, como ilustrado na Figura 11.

Figura 11 – Série transformada plotada com os 108 pontos previstos (PIB Brasil)



Com o intuito de se comparar o valor do PIB per capita brasileiro com o dos EUA, realizaram-se os mesmos passos para a série de dados obtida do PIB americano entre os anos de 1960 e 2012. A Tabela 2, a seguir, mostra os valores da série de dados do PIB americano.

Tabela 2 – Série de dados do PIB per capita americano

Ano	PIB per capita - EUA (\$ - dólares)	Ano	PIB per capita - EUA (\$ - dólares)
1960	2.881.100	1987	19.393.782
1961	2.934.553	1988	20.703.152
1962	3.107.937	1989	22.039.227
1963	3.232.208	1990	23.037.941
1964	3.423.396	1991	23.443.263
1965	3.664.802	1992	24.411.143
1966	3.972.123	1993	25.326.736
1967	4.152.020	1994	26.577.761
1968	4.491.424	1995	27.559.167
1969	4.802.642	1996	28.772.356
1970	4.997.757	1997	30.281.636
1971	5.360.178	1998	31.687.052
1972	5.836.224	1999	33.332.139
1973	6.461.736	2000	35.081.923
1974	6.948.198	2001	35.912.333
1975	7.516.680	2002	36.819.445
1976	8.297.292	2003	38.224.739
1977	9.142.795	2004	40.292.304
1978	10.225.307	2005	42.516.393
1979	11.301.682	2006	44.622.642
1980	12.179.558	2007	46.349.115
1981	13.526.187	2008	46.759.560
1982	13.932.678	2009	45.305.052
1983	15.000.086	2010	46.615.511
1984	16.539.383	2011	48.112.600

1985	17.588.810	2012	49.965.271
1986	18.427.288		

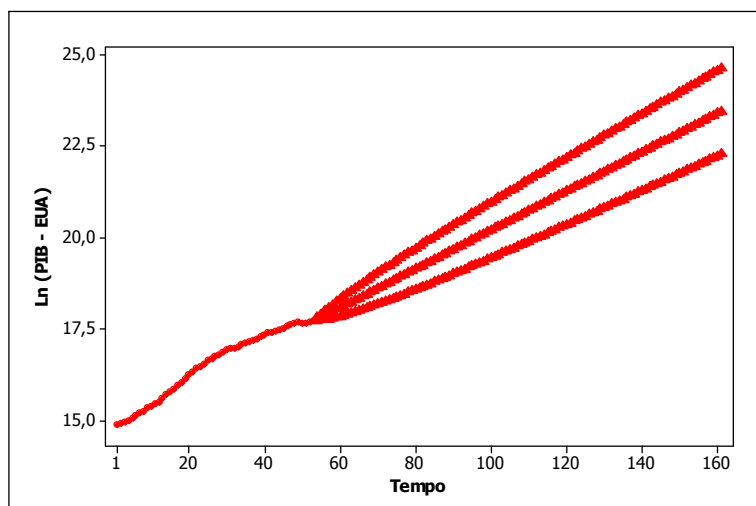
Fonte: Fundo Monetário Internacional (FMI)

Como mencionado anteriormente, a série de dados do PIB americano foi prevista seguindo os mesmos passos da série de dados do PIB brasileiro como a realização da transformação logarítmica para minimizar a variância da série, a diferença da série para se obter uma série estacionária, a análise dos gráficos de ACF e PACF para se identificar o potencial modelo ARIMA que, também, foi identificado como (1,1,0), a análise dos resíduos e, por fim, a previsão da série. A Figura 12 ilustra os 108 pontos previstos para o PIB per capita americano.

Como resultado, obtiveram-se os 108 pontos previstos do PIB per capita brasileiro e, também, os 108 pontos previstos do PIB per capita americano, ou seja, o valor do PIB foi previsto até o ano de 2120 para ambos os países com o objetivo de comparar a evolução do PIB brasileiro com o dos EUA.

O Anexo 1 mostra os valores dos pontos previstos após a transformação exponencial, já que, como mencionado anteriormente, a série original foi transformada pelo Ln com o intuito de minimizar a variância, por isso, realizou-se a transformação inversa que é a exponencial.

Figura 12 - Série transformada plotada com os 108 pontos previstos (PIB – EUA)



5. Conclusão

A economia dos países está exposta a inúmeros eventos que influenciam o seu crescimento. Ainda, quanto maior o PIB per capita, mais difícil é de se manter um ritmo de expansão e

nenhum país mantém o mesmo ritmo de crescimento desse indicador por períodos tão longos, mas a previsão realizada neste presente trabalho nos mostra para onde a economia de cada nação está apontando para os anos futuros.

Assim sendo, o presente trabalho objetivou aplicar o modelo ARIMA para analisar o ritmo de crescimento do Brasil e compará-lo com o dos EUA por se tratar de uma das maiores potências da economia mundial. Como resultado, conclui-se que o PIB brasileiro ultrapassaria o PIB americano em 74 anos caso a taxa de crescimento de ambos os países se comportasse da mesma maneira que as séries de dados obtidas pelo FMI durante os anos de 1960-2012, também, se conclui que o PIB brasileiro demoraria 22 anos (2034) para ultrapassar o atual PIB americano de US\$ 49.965.271 bilhões.

Agradecimentos

Os autores agradecem a Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG), a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes) e ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPQ) pelo apoio concedido a esta pesquisa.

Referências

- ARMSTRONG, J. S.; COLLOPY, F. **Integration of statistical methods and judgment for time series forecasting: Principles from empirical research.** Forecasting with judgment. New York: John Wiley & Sons, Inc. p 269-293, 1998.
- BARBANCHO, A. G. **Fundamentos e possibilidades da econometria.** Rio de Janeiro: Fórum Editora. p. 18-32, 1970.
- BOX, G.E.P.; JENKINS, G.M. **Time Series Analysis: Forecasting and Control.** Holden Day, San Francisco, 1976.
- DeLURGIO, S. A. **Forecasting principles and applications.** 1st Edition. Singapore: McGraw-Hill. 802 p, 1998.
- MONTGOMERY D.C.; JENNINGS, C.L; KULAHCI, M. **Introduction to Time Series Analysis and Forecasting.** John Wiley & Sons, 2008.
- MORETTIN, P. A. *Séries temporais em finanças.* (Time Series in Finance.) Lima, Peru: Instituto de Matemática y Ciencias Afines, 2002. 156p. (em Português).
- MORETTIN, P. A.; TOLOI, C. M. de C. Modelos para previsão de Séries Temporais. In: **13º Colóquio Brasileiro de Matemática.** Rio de Janeiro, 1981.
- SCHWITZK, M. (2001). **Acuracidade dos métodos de previsão e sua relação com o dimensionamento dos estoques de produtos acabados.** Florianópolis. 137f.. Dissertação (Mestrado em Engenharia de Produção) – Universidade Federal de Santa Catarina.

ZHANG, G. P.; **Time series forecasting using a hybrid ARIMA and neural network model.**
Neurocomputing, v.50, p. 159–175, 2003.

<http://datamarket.com> acessado em 25 de novembro de 2013

ANEXO

Anexo 1 – Resultado da previsão dos PIBs brasileiro e americano

Ano	PIB per capita - Brasil (\$ - dólares)	PIB per capita - EUA (\$ - dólares)
2013	11.535.040,95	52.189.923,61
2014	12.198.212,71	54.717.820,48
2015	13.061.767,38	57.494.508,68
2016	14.042.513,99	60.496.737,75
2017	15.116.538,99	63.713.052,16
2018	16.280.847,91	67.140.634,51
2019	17.534.834,44	70.773.840,08
2020	18.887.294,40	74.618.572,85
2021	20.346.103,96	78.680.035,29
2022	21.915.396,84	82.979.155,86
2023	23.605.729,12	87.504.431,87
2024	25.426.436,56	92.285.722,34
2025	27.390.313,86	97.337.998,48
2026	29.502.925,92	102.656.600,46
2027	31.778.483,52	108.276.640,91
2028	34.232.977,78	114.192.936,36
2029	36.873.363,80	120.444.545,30
2030	39.717.402,51	127.025.701,76
2031	42.780.801,64	133.979.853,55
2032	46.085.088,69	141.314.717,49
2033	49.639.626,79	149.036.231,68
2034	53.468.326,04	157.195.372,42
⋮	⋮	⋮



Curitiba, PR, Brasil, 07 a 10 de outubro de 2014.

2086	2.550.869.198,01	2.507.870.942,35
-------------	-------------------------	-------------------------

Fonte: Autores