

Utilização de Modelagem Estatística Temporal na Previsão do Preço do Barril de Petróleo

Use of Temporal Statistical Modeling in Oil Barrel Price Forecasting

Uso del modelado estadístico temporal en el pronóstico del precio del barril de petróleo

Raiza Gabriele Lima dos Santos¹

Bacharel em Ciências Contábeis.
raiza-gabriele@hotmail.com
<http://lattes.cnpq.br/1418536287657672>

Eric Gardel de Oliveira Santos¹

Mestrando em Ciências Contábeis.
ericsantos_3@hotmail.com
<http://lattes.cnpq.br/2946495032106153>

Gilvan Farias da Silva¹

Mestrando em Engenharia Elétrica.
gilvan1991@gmail.com
<http://lattes.cnpq.br/4767842842233385>

Universidade Federal da Bahia, UFBA, Brasil¹

Resumo

Este artigo apresenta uma análise da previsão dos preços do barril de petróleo, utilizando modelagem estatística temporal. O petróleo, juntamente com o gás natural, representa a principal fonte de energia primária consumida no mundo e diariamente movimenta bilhões de dólares. A obtenção de uma estimativa mais precisa do seu valor, auxilia a mitigar eventuais impactos que possam surgir. Para tanto, utilizou-se dados extraídos do *Energy Information Administration (EIA)*, que contém o preço diário do barril de petróleo. Os dados utilizados nessa pesquisa correspondem ao período entre janeiro de 1987 a maio de 2019 e foram divididos em duas partes: teste e treino. A parte de teste foi utilizada para criação do modelo estatístico e a parte de treino, que corresponde ao mês de maio de 2019, foi utilizada para comparar as previsões feitas pelo modelo com os valores reais. Achados dessa pesquisa podem contribuir de várias maneiras para literatura, tais como, desenvolvimento e um maior conhecimento de técnicas que possam atuar junto à Contabilidade e formas de minimizar possíveis variações. Os resultados obtidos podem ser considerados relevantes, em virtude de demonstrar que o modelo utilizado (ARIMA), apresenta uma alta capacidade de prever preços, podendo, portanto, afirmar que a prática do uso de modelagem estatística temporal pode ser agregada ao processo decisório referente ao comportamento dos preços de *commodities*.

Palavras-chave: Modelagem Estatística; Preço do Petróleo; ARIMA; *Commodities*.

Abstract

This article presents an analysis of the prediction of oil barrel prices, using temporal statistical modeling. Petroleum, together with natural gas, represents the main source of primary energy consumed in the world and generates billions of dollars a day. Obtaining a more accurate estimate of its value helps to mitigate any impacts that may arise. For this purpose, data extracted from the Energy Information Administration (EIA), which contains the daily price of a barrel of oil, was used. The data used in this research correspond to the period between January 1987 and May 2019 and were divided into two parts: test and training. The test part was used to create the statistical model and the training part, which corresponds to the month of May 2019, was used to compare the predictions made by the model with the actual values. Findings of this research can contribute in several ways to literature, such as, development and a greater knowledge of techniques that can work with Accounting and ways to minimize possible variations. The results obtained can be considered relevant, in view of demonstrating that the model used (ARIMA), has a high capacity to predict prices, being able, therefore, to affirm that the practice of using temporal statistical modeling can be added to the decision process regarding the behavior of commodity prices.

Keywords: Statistical Modeling; Oil Price; ARIMA; Commodities

Resumen

Este artículo presenta un análisis del pronóstico de los precios del barril de petróleo, utilizando modelos estadísticos temporales. El petróleo, junto con el gas natural, representa la principal fuente de energía primaria consumida en el mundo y genera miles de millones de dólares por día. Obtener una estimación más precisa de su valor ayuda a mitigar cualquier impacto que pueda surgir. Para este propósito, se utilizaron datos extraídos de la Energy Information Administration (EIA), que contiene el precio diario de un barril de petróleo. Los datos utilizados

en esta investigación corresponden al período comprendido entre enero de 1987 y mayo de 2019 y se dividieron en dos partes: prueba y capacitación. La parte de prueba se utilizó para crear el modelo estadístico y la parte de capacitación, que corresponde al mes de mayo de 2019, se utilizó para comparar las predicciones realizadas por el modelo con los valores reales. Los resultados de esta investigación pueden contribuir de varias maneras a la literatura, como el desarrollo y un mayor conocimiento de las técnicas que pueden funcionar con la contabilidad y las formas de minimizar las posibles variaciones. Los resultados obtenidos pueden considerarse relevantes, en vista de demostrar que el modelo utilizado (ARIMA), tiene una alta capacidad para predecir precios, pudiendo, por lo tanto, afirmar que la práctica de usar modelos estadísticos temporales se puede agregar al proceso de decisión con respecto al comportamiento de los precios de los productos básicos.

Palabras clave: modelado estadístico; Precio del petróleo; ARIMA Productos básicos

1 INTRODUÇÃO

Um dos objetivos da Contabilidade é reconhecer, mensurar e evidenciar os fatos que provocam alterações no patrimônio das entidades. Além disso, a Ciência Contábil analisa eventuais fenômenos que venham a impactar os resultados da entidade e, por conseguinte, afetar a tomada de decisão, muitas vezes, utilizando informações no presente para influenciar o futuro. Dentro desse contexto, o Comitê de Pronunciamento Contábeis (CPC) 38, que trata de Instrumentos Financeiros: Reconhecimento e Mensuração, estabelece os princípios para reconhecimento inicial de ativos financeiros, passivos financeiros e alguns tipos de contratos relacionados à compra e venda de itens não financeiros. Ademais, tal norma aborda a avaliação posterior de todos os instrumentos financeiros e da baixa de ativos e passivos financeiros.

O petróleo, juntamente com o gás natural, representa a principal fonte de energia primária consumida no mundo e diariamente movimenta bilhões de dólares em atividades industriais, passando a ser imprescindível às facilidades e comodidades da vida moderna (SANTOS; SILVA; SANCOVSKI, 2006). Além disso, mesmo com o iminente surgimento de energias consideradas renováveis, o petróleo se manteve como a principal matriz energética mundial (RIBEIRO; ALBA NETO; SENE, 2018).

Acerca dos seus preços, Adrangiet *al.* (2001) e Bernabe *et al.* (2004) ressaltam a importância de fatores econômicos, políticos e ambientais na sua formulação. Nessa perspectiva, Godoy (2004) aponta que a atividade de exploração e produção de petróleo e, conseqüentemente, de sua contabilização e evidênciação, é caracterizada por vários problemas teóricos e técnicos decorrentes. Ainda segundo o autor, as seguintes particularidades são observadas: (a) alto risco envolvido na atividade de exploração; (b) alto custo dos investimentos; (c) falta de correlação entre o valor dos gastos efetuados e o valor resultante das reservas encontradas; (d) longo espaço de tempo entre os gastos incorridos e os benefícios gerados.

Em virtude dessa importância, torna-se válido para as organizações tentarem prever o comportamento dos preços dessa *commodity*. Holden, Peel e Thompson (1990) defendem a ideia de que as previsões são necessárias por causa da incerteza do futuro e do impacto de muitas dessas decisões tomadas no presente que não são sentidas no futuro, ou seja, apesar da tentativa de controlar o futuro, ele ainda permanece incerto. Vale ressaltar que os recursos são finitos e, para tanto, as organizações devem levar em consideração vários fatores para a tomada de decisão. Corrar e Garcia (2001) corroboram esse pensamento, pois acreditam que os gestores precisam se valer de instrumentos mais apropriados de forma a administrar esses fatores limitativos para tomar decisões, controlar as operações e simular desempenhos.

Segundo Matias (2006), na medida em que teorias ou novas abordagens matemáticas foram desenvolvidas, estas foram utilizadas como referência para o desenvolvimento de novos estudos sobre comportamento de preços de *commodities*. Entende-se que a análise do comportamento dos preços de um ativo é importante, sob o ponto de vista de sua utilidade, para a previsão de receitas futuras, análise do comportamento passado de uma série de preços e estudo do preço histórico de um produto.

Sendo assim, a importância deste artigo é justamente obter uma estimativa do preço do petróleo mais próxima possível da realidade a fim de mitigar eventuais impactos que possam surgir futuramente, sobretudo, na tomada de decisão dos usuários (internos e externos). Certamente, as organizações que utilizam essa *commodity* como base de seus produtos serão impactadas

negativamente (aumento no seu passivo) com a alta do valor do barril do petróleo, bem como as que produzem e vendem na sua forma bruta e/ou refinada, serão impactadas positivamente (aumento no valor do seu ativo) caso o barril fique em alta.

Achados dessa pesquisa podem contribuir de várias maneiras para as Ciências Contábeis e à literatura, dentre as quais podemos evidenciar: (i) desenvolvimento e validação de um código de programação do modelo ARIMA (Autorregressivos, Integrados e de Médias Móveis), o qual poderá ser usado por outros pesquisadores, que desejam realizar pesquisas com temática semelhante relacionada à modelagem estatística temporal; (ii) aumento da literatura contábil a respeito de técnicas que possam atuar junto à Contabilidade, como forma de auxiliar na tomada de decisão, das quais, Guth (2013), retrata em sua pesquisa que a utilização de metodologias e instrumentos na tomadas de decisões, agregam mais confiabilidade e relevância nas informações e (iii) formas de minimizar possíveis variações não previstas que venham impactar nas demonstrações contábeis, que Dolabela (2008), evidencia como fatores que afetam a receita (demanda do mercado e previsão de vendas; capacidade instalada; retorno financeiro planejado e esperado do negócio; preço de venda).

Diante do exposto, este artigo visa demonstrar o impacto do valor do petróleo na tomada de decisão, utilizando à aplicação de modelagem estatística temporal nos preços diários de comercialização do barril do petróleo, aplicados entre janeiro de 1987 a maio de 2019, com a finalidade de auxiliar os diversos grupos de interesses (*stakeholders*) que são direta e indiretamente impactados com eventuais variações.

O artigo está estruturado em cinco seções, iniciando com essa introdução. Na sequência, faz-se o desenvolvimento de uma revisão de literatura sobre alguns trabalhos relacionados à previsão de preços de *commodities* e uma incursão teórica que aborda o modelo estatístico utilizado. Em seguida, descreve-se o método, o procedimento da pesquisa, a descrição e a análise dos dados. Por fim, apresentam-se as conclusões da pesquisa realizada.

2 REVISÃO DE LITERATURA

2.1 Trabalhos Relacionados ao Tema

Segundo Castro (2015), muitas pesquisas têm sido realizadas para prever preços de *commodities*, utilizando técnicas de séries temporais, em resposta à complexa formação de preços desse tipo de produto. Já Matias (2006), mostra que no Brasil são poucas as pesquisas realizadas sobre o comportamento de preços de *commodities* e, a maioria delas, se concentram em *commodities* agrícolas.

Uma das técnicas utilizadas no contexto das séries temporais é o Modelo ARIMA, que foi desenvolvido por Box e Jenkins (1976). Diversos trabalhos usaram tal modelo para tentar prever os preços de diversas *commodities*. A Tabela 1 demonstra alguns desses trabalhos.

Tabela 1 – Estudos que utilizaram o modelo ARIMA para prever preços de *commodities*.

AUTORES	COMMODITY
Cas (2018)	Milho
Oliveira (2016); Tularam e Saeed (2016)	Petróleo
Felipe (2012)	Soja

Souza <i>et al.</i> (2007),	Madeira
Marques <i>et al.</i> (2010), Carrasco-Gutierrez e Almeida (2013)	Café
Castro (2015)	Alumínio
Coelho Júnior <i>et al.</i> (2019)	Carvão
Aredes e Pereira (2018)	Trigo

Fonte: Autores (2019).

Acerca do petróleo, torna-se relevante compreender seu comportamento, em virtude de oferecer elevado risco macroeconômico à conjuntura de países, impactando direta e indiretamente às atividades econômicas. No Brasil, sobretudo após a descoberta do pré-sal, criou-se a expectativa de que o país tornar-se-ia um dos grandes polos mundiais de produção e de tecnologia da indústria do petróleo (Furtado, 2019).

Dada a importância dessa *commodity*, a literatura tem tratado essa temática a partir de modelos de séries temporais para tentar explicar as flutuações da série de preços do petróleo. Tularam e Saeed (2016), utilizaram séries temporais para determinar os preços futuros do petróleo, utilizando os modelos de Suavização Exponencial (ES), *Holt-Winters* (HW) e o ARIMA. Os resultados encontrados utilizando o último modelo superou os demais.

Um estudo feito por Fonseca, Marques e Santos (2019) no qual investigou a contabilidade aplicada ao setor petrolífero, com a utilização de modelo próprio de reconhecimento e mensuração, teve por objetivo geral avaliar a relevância do conjunto de informações reportado nas demonstrações contábeis das empresas petrolíferas para o mercado de capitais, por meio da técnica de estudo de eventos. Os resultados demonstraram que as informações contábeis das empresas petrolíferas apresentam relevância (FONSECA; MARQUES; SANTOS, 2019).

3 METODOLOGIA

Nesta seção, são descritos os procedimentos metodológicos utilizados para a construção do modelo ARIMA para previsão dos preços de barril do petróleo em dólar. As etapas estão de acordo com as recomendações de Hair *et al.*(2005), em que são apresentados cinco aspectos fundamentais envolvidos na utilização de uma modelagem estatística temporal: (i) preparação dos dados, (ii) definição da estrutura do modelo, (iii) estimativa do modelo, (iv) avaliação dos resultados do modelo e (v) validação do modelo.

De acordo com Gujarati (2000), existem cinco modelos para realizar previsões utilizando séries temporais: (1) métodos de suavização exponencial, (2) modelos de regressão uniequacionais, (3) modelos de regressão de equação simultânea, (4) processo autorregressivo integrado de média móvel (ARIMA) e (5) modelos de vetores autorregressivos (VAR).

Neste artigo, os modelos de séries temporais Autoregressivos Integrados de Média Móvel são usados para previsão. A metodologia utilizada se divide em quatro etapas: **Etapa 1** – Análise preliminar: a série histórica do preço do petróleo é analisada em relação à estacionariedade; **Etapa 2** – preparação dos dados: a série histórica do preço do petróleo é dividida em duas partes: a primeira, corresponde aos dados de 1987 a abril de 2019 e foi utilizada para gerar o modelo ARIMA e, a segunda parte, corresponde ao mês de maio de 2019 e foi utilizada para testar a capacidade do modelo em prever o preço corretamente; **Etapa 3** – Criação do modelo ARIMA: foram testadas várias combinações de **p**, **d** e **q**, em que **p** é o termo autoregressivo (AR), **q** o termo média móvel (MA) e **d** o número de diferenças para tornar a série estacionária;

Etapa 4 – Teste: foi feita a previsão do preço do barril de petróleo para o mês de maio de 2019 para verificar o potencial de previsão do modelo.

Antes de apresentar a metodologia propriamente dita, é necessário definir alguns conceitos sobre modelos autorregressivos e médias móveis que servem como base para definição dos modelos.

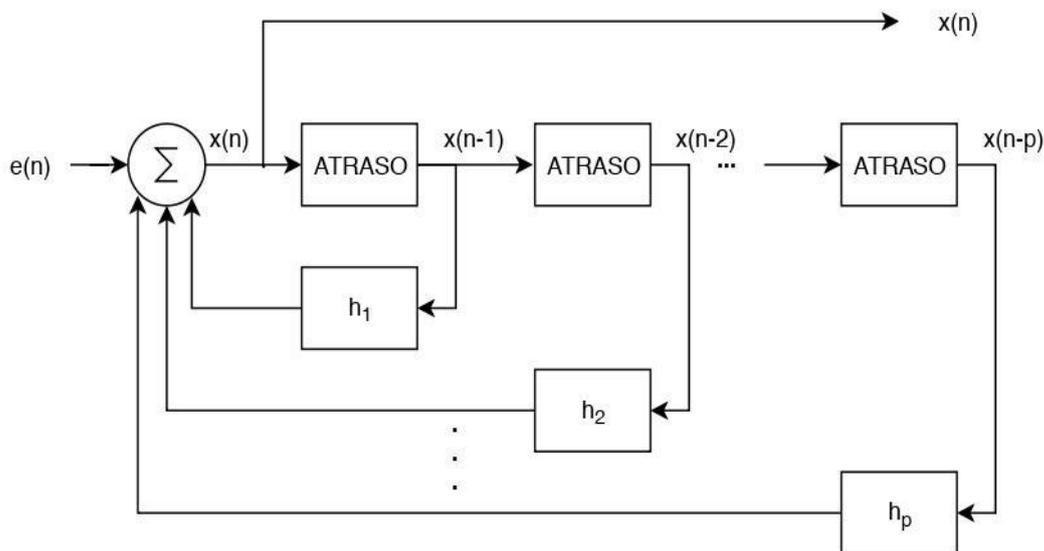
3.1 Modelo Autoregressive Process

Um processo autorregressivo (AR – *Autoregressive Process*) é aquele em que o valor atual é uma combinação linear dos seus elementos anteriores e de uma sequência de variáveis aleatórias com distribuição gaussiana de média zero (SHANMUGAN, 1988). Matematicamente, o modelo AR é descrito por:

$$X(n) = \sum_{i=1}^p h_i X(n-i) + e(n)$$

Em que $X(n)$ é o valor atual da sequência aleatória X ; $X(n-i)$ é o i -ésimo valor passado da sequência X ; h_i é o parâmetro de ponderação dos valores; $X(n-i)$ e $e(n)$ é uma sequência de variáveis aleatórias com distribuição gaussiana de média zero. A sequência $e(n)$ é denominada ruído branco gaussiano. Para uma sequência em que o somatório acima seja de ordem p é chamada sequência de modelo autorregressivo de ordem p . O diagrama de blocos da Figura 1 esquematiza o funcionamento de uma sequência de modelo AR:

Figura 1 - Diagrama de blocos esquematizando o funcionamento do modelo AR.



Fonte: Autores (2019), baseados em Shanmugan (1988).

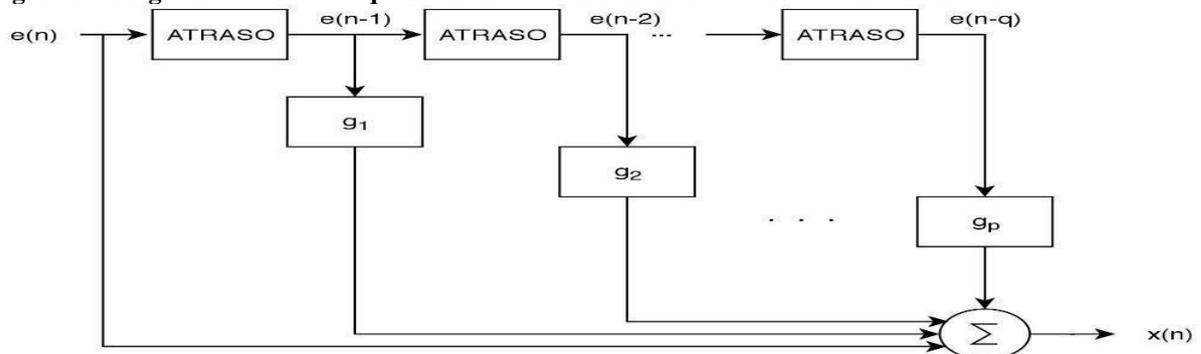
3.1.2 Modelo Moving Average Model

Um processo Média Móvel (MA – *Moving Average Model*) é aquele que pode ser representado pela seguinte equação:

$$X(n) = \sum_{i=1}^q g_i e(n-i) + e(n)$$

Essa equação mostra que o valor atual $X(n)$ é uma média ponderada de valores passados de ruído branco gaussiano. Nesse caso, temos que $g_0 = 1$ e o somatório dos demais g_i é igual a 1 . Um processo modelado como a média de q elementos é chamado de processo Média Móvel de ordem q (SHANMUGAN, 1988). A Figura 2 ilustra o funcionamento da modelagem média móvel.

Figura 2 - Diagrama de blocos esquematizando o funcionamento do modelo média móvel.



Fonte: Autores (2019), baseados em Shanmugan (1988).

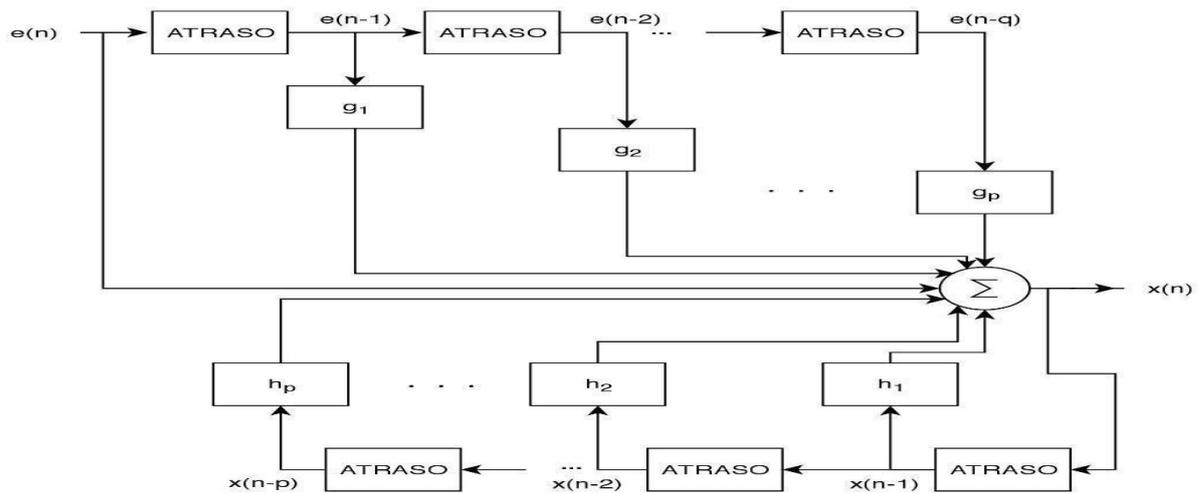
3.1.3 Modelo *Autorregressive Moving Average*

Um modelo autorregressivo com média móvel (ARMA – *Autorregressive Moving Average*) é um modelo que une os modelos AR e MA, portanto a sequência aleatória é descrita da seguinte forma:

$$X(n) = \sum_{i=1}^p h_i X(n-i) + \sum_{k=1}^q g_k e(n-k) + e(n)$$

Desse modo, a sequência é descrita como sendo formada por uma parte influenciada por seus valores passados e parte influenciada pela média do ruído branco gaussiano. Um modelo formado com p elementos autorregressivos e q elementos com média móvel é chamado ARMA (p, q) (SHANMUGAN, 1988). A Figura 3 ilustra o diagrama de blocos de um processo ARMA.

Figura 3 - Diagrama de Blocos Esquematizando o Funcionamento do Modelo ARMA.



Fonte: Autores (2019), baseados em Shanmugan (1988).

3.1.4 Modelo *Autorregressive Integrated Moving Average*

O modelo autorregressivo integrado com média móvel (ARIMA – *Autorregressive Integrated Moving Average*) é a generalização do modelo ARMA com a inclusão do conceito de integração. Neste modelo, o sinal que se deseja modelar é diferenciado para torná-lo estacionário. A notação padrão de um modelo ARIMA pode ser feita da seguinte forma: ARIMA ($\mathbf{p}, \mathbf{d}, \mathbf{q}$), em que \mathbf{p} indica o número de elementos da parte autorregressiva, \mathbf{d} indica o número de vezes que o sinal é derivado e \mathbf{q} indica o número de elementos de média móvel (CADENAS, 2016).

Neste trabalho foi utilizado o modelo ARIMA para modelar o comportamento da série histórica do preço do petróleo em dólares por barril. Os dados da série histórica foram obtidos do EIA (2019) e contém o preço diário do barril do petróleo desde 1987, com duas classes de preços: RWTC (*Cushing, OK WTI Spot Price FOB, Dollars per Barrel*) e RBRTE (*Europe Brent Spot Price FOB, Dollars per Barrel*). A classe RBRTE foi escolhida por apresentar menor falta de dados.

Para realizar a modelagem da série foi escrito um programa com a linguagem *Python* (APÊNDICE A) em que várias combinações de \mathbf{p}, \mathbf{d} e \mathbf{q} foram testadas.

Os dados foram divididos em duas partes: teste e treino. A parte de teste foi utilizada para criação do modelo ARIMA e a parte de treino, que corresponde ao mês de maio de 2019, foi utilizada para comparar as previsões feitas pelo modelo em relação aos valores reais. Para comparar os valores previstos com os reais foi calculado o erro quadrático médio, que é definido pela equação abaixo (HYNDMAN, 2006):

$$EQM = \frac{1}{N} \sum_{i=1}^N (R_i - P_i)^2$$

Em que \mathbf{R}_i são os valores reais para o mês de maio de 2019, \mathbf{P}_i são os valores previstos e \mathbf{N} é o número de dias válidos para o mês de maio de 2019. O melhor modelo foi escolhido como sendo aquele com menor valor de erro quadrático médio.

4 ANÁLISE DE RESULTADOS

A Figura 4 mostra o gráfico da série histórica do preço do barril do petróleo em dólares desde o ano 1987 até 2019. O preço manteve-se aproximadamente constante até 1997, quando houve uma queda no preço, que somente voltou a crescer em 1999. Entre os anos 2000 e 2019, o preço apresentou maior oscilação, variando entre US\$ 20 e US\$ 140 dólares.

Figura 4 - Série histórica do preço do barril do petróleo em dólares entre os anos de 1987 e 2019.



Fonte: Autores (2019), baseados em dados da pesquisa.

A Tabela 2 mostra algumas medidas estatísticas referentes à série histórica. Houve grande variação de preços, com valor mínimo de 9,1; valor máximo de 143,95; valor médio de 46,2 e desvio padrão 32,8.

Tabela 2 - Dados da série histórica do preço do barril de petróleo em dólares.

MEDIDA	VALOR
Número de amostras	8112
Média	46,222871055226825
Valor mínimo	9,1
Valor máximo	143,95
Variância	1078,7139853819656
Desvio padrão	32,843781532916786

Fonte: Autores (2019), baseados em dados da pesquisa.

Nota-se que a série histórica não é estacionária, ou seja, não apresenta valores oscilando em torno de uma média. Esta condição é mais facilmente percebida ao aplicar-se um filtro chamado média móvel de janela de 365 valores à série, conforme pode ser visto na Figura 5.

Figura 5 - Série histórica do preço do barril de petróleo em dólares, entre 1988 e maio de 2019, filtrada por filtro média móvel.



Fonte: Autores (2019), baseados em dados da pesquisa.

Um filtro média móvel diminui o ruído de uma série $\mathbf{x}[\mathbf{n}]$, definindo uma janela temporal de largura $(M_2 + 1)$ cuja média define um valor da série filtrada $\mathbf{y}[\mathbf{n}]$ (MARTINS, 2003). Matematicamente, um filtro média móvel pode ser descrito pela equação abaixo:

$$y[n] = \frac{1}{M_2 + 1} \sum_{k=0}^{M_2} x[n-k]$$

Como foi utilizada uma janela temporal de 365 dias o gráfico da Figura 5 não contém dados referentes ao ano de 1987. Essa é uma característica dos filtros média móvel, existe uma perda de dados no início da série igual à largura da janela temporal.

Esta janela temporal é movida, variando valores de n e k , de forma a excursionar toda a série, gerando uma série de valores filtrados. A variável \mathbf{n} é a abscissa dos gráficos de $\mathbf{x}[\mathbf{n}]$ e $\mathbf{y}[\mathbf{n}]$ e a variável \mathbf{k} é utilizada para referenciar valores deslocados de \mathbf{n} . Como a série não é estacionária, tem-se um forte indício de que os modelos AR ou ARMA não são capazes de modelar esta série. Apesar deste indício, tais modelos foram testados para confirmar se essa hipótese é verdadeira, chegando a confirmação da não capacidade de modelar a série.

Para tornar a série estacionária, é preciso realizar uma transformação. Neste trabalho, a transformação realizada foi a derivação da série, pois ela é uma etapa da modelagem ARIMA em que o parâmetro $d > 0$. A derivada da série é mostrada na Figura 6. Nota-se que os valores oscilam em torno de zero. Isso é um indicativo que o melhor modelo ARIMA deve ter o parâmetro $d = 1$. Essa hipótese é reforçada pelo gráfico da Figura 7, que mostra o gráfico da derivada da série histórica filtrada por um filtro média móvel de janela de 365 valores. Conforme ocorreu na Figura 5, o gráfico da Figura 7 não apresenta dados para o ano 1987.

Figura 6 - Derivada da série histórica do preço do barril do petróleo entre 1988 a maio de 2019.



Fonte: Autores (2019), baseados em dados da pesquisa.

A Tabela 3 mostra algumas medidas estatísticas referentes a derivada da série histórica. Houve pouca variação nos valores, com valor mínimo de -0,209, valor máximo de 0,251, valor médio de 0,007 e desvio padrão 0,059.

Tabela 3 - Dados da derivada da série histórica do preço do barril do petróleo.

MEDIDA	VALOR
Número de amostras	8112
Média	0,006900548334220405
Valor mínimo	-0,20893150684931505
Valor máximo	0,2511232876712329
Variância	0,003446650439898403
Desvio padrão	0,05870818034906552

Fonte: Autores (2019), baseados em dados da pesquisa.

Figura 7 - Derivada da série histórica do preço do barril do petróleo entre 1988 a maio de 2019 filtrada por filtro média móvel.



Fonte: Autores (2019), baseados em dados da pesquisa.

O modelo ARIMA(p,d,q) é a generalização dos modelos AR, MA e ARMA. Para os casos em que $p > 0$ e $q = 0$ temos um modelo AR. Para casos em que $p = 0$ e $q > 0$ temos um modelo MA. Para o caso em que $p > 0$, $d = 0$ e $q > 0$ temos um modelo ARMA. O modelo ARIMA está completo se $p > 0$, $d > 0$ e $q > 0$.

Neste artigo, todos os modelos testados são referidos como modelo ARIMA(p,d,q). Caso trate-se de um caso particular, AR, MA ou ARMA, ele será referido com os devidos índices iguais a zero.

Para definir o melhor modelo ARIMA foram testadas 27 combinações de valores de p, d e q. Com o modelo resultante, foram feitas previsões para o mês de maio de 2019. Então, o erro quadrático médio entre as previsões e os valores reais foi calculado.

A Tabela 4 apresenta todas as combinações utilizadas e o respectivo erro quadrático médio resultante. Em sete casos, o algoritmo não conseguiu convergir para uma solução factível, portanto, no local referente ao erro quadrático médio foi escrito NC para tais casos.

Tabela 4 - Combinações utilizadas e erro quadrático médio.

p	d	q	EQM
0	0	1	174.05723310
0	0	5	NC
0	1	5	1.74303813
0	2	5	NC
0	0	10	NC
0	1	10	1.75341940
0	2	10	NC
1	0	0	1.75527459
1	1	0	1.74919510
1	2	0	2.82123452
1	0	1	NC
1	1	1	1.74537827
1	2	1	1.74905409
5	0	0	1.72422941
5	1	0	1.74286555
5	2	0	2.01511582
5	0	5	1.65303021
5	1	5	NC
5	2	5	1.67054171
10	0	0	1.72916014
10	1	0	1.75906155
10	2	0	1.85628839
10	0	5	1.65596682
10	1	5	NC
10	2	5	1.85628839
20	0	0	1.78306659
20	1	0	1.79429165

Fonte: Autores (2019), baseados em dados da pesquisa.

O modelo ARIMA (5,0,5) resultou no menor erro quadrático médio, sendo igual a 1.65303021, ou seja, trata-se de um modelo ARMA (5,5). Contrariando a hipótese inicial de que o melhor modelo deveria ter o parâmetro $d = 1$, devido a não estacionariedade da série histórica. Uma hipótese que explica este resultado é que no mês de maio de 2019 o preço do barril do petróleo apresentou comportamento aproximadamente estacionário, com baixo desvio padrão, facilitando as previsões feitas pelo modelo ARIMA (5,0,5).

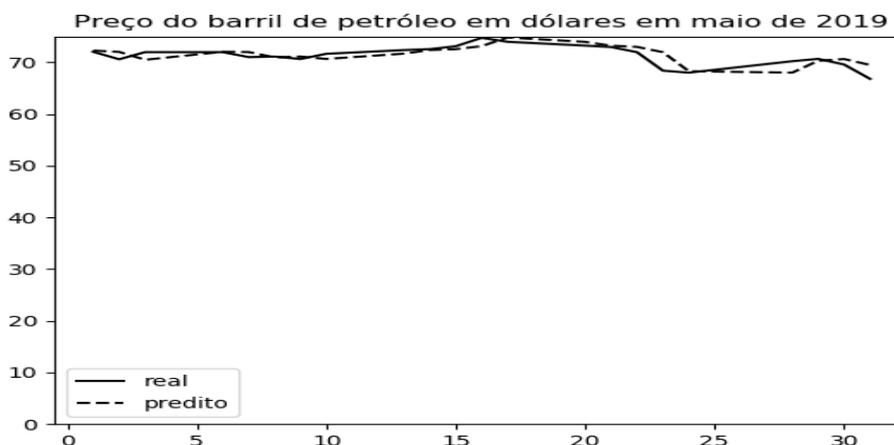
Tabela 5 - Dados da série histórica do preço do barril de petróleo para o mês de maio de 2019.

MEDIDA	VALOR
Número de amostras	22
Média	71.31772727272727
Valor mínimo	66.78
Valor máximo	74.7
Variância	3.7374088744588714
Desvio padrão	1.9332379249484195

Fonte: Autores (2019), baseados em dados da pesquisa.

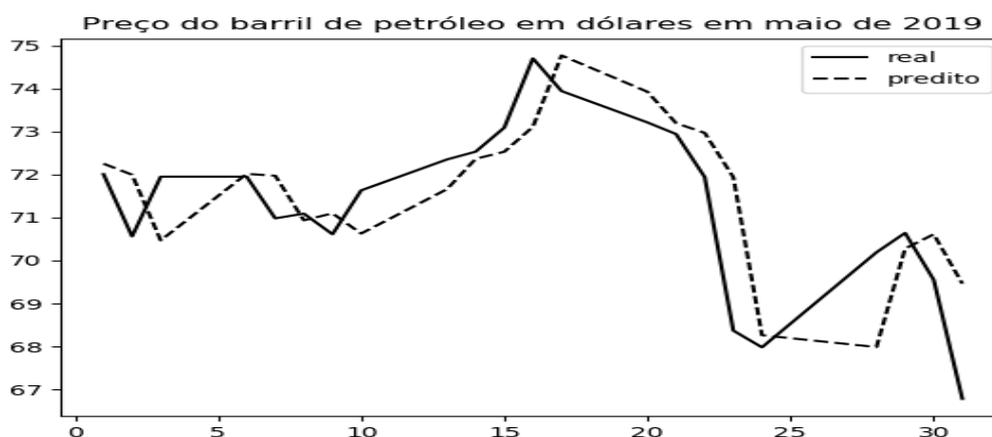
As figuras abaixo mostram os valores reais e previstos pelo modelo ARIMA. Na Figura 8, é apresentado pela escala vertical completa, enquanto na Figura 9, a escala é ampliada para melhor visualização.

Figura 8 - Validação do modelo para previsão do barril de petróleo em escala vertical.



Fonte: Autores (2019), baseados em dados da pesquisa.

Figura 9 - Validação do modelo para previsão do barril de petróleo em escala ampliada.



Fonte: Autores (2019), baseados em dados da pesquisa.

5 CONSIDERAÇÕES FINAIS

Esse trabalho utilizou a modelagem estatística temporal ARIMA na previsão dos preços do barril de petróleo para o período de maio de 2019. Os resultados obtidos podem ser considerados relevantes, em virtude de demonstrar que modelo apresenta uma alta capacidade de prever preços, podendo, portanto, afirmar-se que a prática do uso de modelagem estatística temporal, pode ser agregada ao processo decisório referente ao comportamento dos preços de *commodities*.

Além disso, pode-se verificar que a estratégia apresentada neste trabalho é uma boa alternativa aos métodos estatísticos tradicionais para previsão de séries temporais, tais como, média móvel e modelo autorregressivo. Para pesquisas futuras sugere-se a utilização de outros tipos de *commodities* de igual importância econômica – por exemplo, ouro – bem como, de modelos estatísticos diferentes ao ARIMA, a fim de verificar a semelhança/diferença dos resultados encontrados. Ademais, este estudo disponibiliza o código de programação que pode ser utilizado por pesquisadores para replicação da respectiva previsão.

Por fim, vale destacar ainda que em uma economia globalizada parece importante que as empresas brasileiras busquem entender o comportamento do mercado internacional de *commodities*, principalmente com o objetivo de buscar maior competitividade. Logo, buscar estudos e pesquisas sobre esse fenômeno parece ser fonte enriquecedora para avaliação estratégica de tomada de decisões, assim como, para a Ciências Contábeis.

REFERÊNCIAS

- Adrangi, B., Chatrath, A., Dhanda, K. K., & Raffiee, K. (2001). Chaos in oilprices? Evidencefrom futures markets. *Energy Economics*, 23(4), 405-425.
- Arêdes, A. F., & Pereira, M. W. G. (2008). Potencialidade da utilização de modelos de séries temporais na previsão do preço do trigo no Estado do Paraná. *Revista de Economia Agrícola*, v. 55, n. 1, p. 63-76.
- Bernabe, A., Martina, E., Alvarez-Ramirez, J., & Ibarra-Valdez, C. (2004). A multimodel approach for describingcrudeoilprice dynamics. *Physica A: StatisticalMechanicsand its Applications*, 338(3), 567-584.

- Box, G. P., & Jenkins, G. M. (1976). Time series analysis, forecasting and control. Holden-Day, San Francisco.
- Cadenas, E. (2016). Wind speed prediction using a univariate ARIMA model and a multivariate NARX model, *Energies*, Multidisciplinary Digital Publishing Institute, v. 9, n. 2, p. 15.
- Gutierrez, C. E. C., & de Moura Almeida, F. M. (2013). Modelagem e previsão do preço do café brasileiro. *Revista de Economia*, 39(2).
- Cas, Carlos Gonçalves (2008). Application of The ARIMA Model to Forecast the Price of the Commodity Corn. *Revista Gestão da Produção Operações e Sistema*, [s.l.], v. 11, n. 1, p.263-279. A Fundação para o Desenvolvimento de Bauru (FunDeB).
<<https://revista.feb.unesp.br/index.php/gepros/article/view/2040>>.
- Castro, João Bosco Barroso de (2015). Projeção de preços de alumínio: modelo ótimo por meio de combinação de previsões. 2015. Tese (Doutorado em Administração) - Faculdade de Economia, Administração e Contabilidade, University of São Paulo, São Paulo.
- Coelho Junior, Luiz Moreira *et al.* (2019). Análise longitudinal dos preços do carvão vegetal, no Estado de Minas Gerais. *Rev. Árvore*, Viçosa, v. 30, n. 3, p. 429-438, June 2006. <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S01007622006000300013&lng=en&nrm=iso>. accesson 06 Aug. <http://dx.doi.org/10.1590/S0100-67622006000300013>.
- Comitê de Pronunciamentos Contábeis (Brasil). Pronunciamento Técnico 38: Instrumentos Financeiros: Reconhecimento e Mensuração. Correlação às Normas internacionais de Contabilidade IAS 39. Disponível: <http://www.cvm.gov.br/export/sites/cvm/menu/regulados/normascontabeis/cpc/CPC_38_rev06_Consolidado.pdf>.
- Corrar, Luiz, & Garcia, Editinete André da Rocha (2001). Programação linear: uma aplicação à contabilidade de custos no processo de tomada de decisão. Anais. Léon: Universidade de Léon.
- Dolabela, F. (2008). O segredo de Luiza. Rio de Janeiro: Sextante, 2008.
- EIA, Petroleum & Other Liquids (2019). Disponível em: <<https://www.eia.gov/>>.
- Felipe, I. J. S (2012). Aplicação de modelos Arima em séries de preços de soja no Norte do Paraná. *Tekhne e Logos*, v. 3, n. 3.
- Fonseca, Raphael da; Marques, José Augusto Veiga da Costa, & Santos, Odilanei Moraes dos (2019). RELEVÂNCIA DA INFORMAÇÃO CONTÁBIL: ESTUDO DE EVENTOS NO SETOR DE PETRÓLEO E GÁS. *Revista Universo Contábil*, Blumenau, p.47-65, mar. 2019. ISSN 1809-3337. Disponível em: <<https://proxy.furb.br/ojs/index.php/universocontabil>>.
- Furtado, André Tosi (2019). Indústria do Petróleo e Gás Natural: uma vocação desperdiçada. *CADERNOS do DESENVOLVIMENTO*, Rio de Janeiro, vol. 14, n. 24, p. 155-175, jan.-jun.
- Hair, Jr., J. F., Anderson, R. E., Tatham, R. L., & Black, W. C (2005). Análise multivariada de dados. 5ª ed. Porto Alegre: Bookman.
- Godoy, C.R. (2004). Evidenciação contábil e as avaliações pelo fluxo de caixa descontado e pela teoria de opções: um estudo aplicado à indústria petrolífera mundial. São Paulo,

- SP. Tese (Doutorado em Contabilidade e Controladoria). FEA/USP, Departamento de Contabilidade.
- Guth, Sergio Cavagnoli (2003). As decisões financeiras e gestão do controle. *Revista Brasileira de Gestão e Inovações*, v. 1, n. 1, set./nov.
- Holden, K, Peel, D.A., & Thompson, J.L (1990). *Economic forecasting: an introduction*. Cambridge University Press.
- Hyndman, R. *et al.* (2006). Another Look at Forecast-Accuracy Metrics for Intermittent Demand. *Foresight: The International Journal of Applied Forecasting*, International Institute of Forecasters, v. 4, n. 4, p. 43-46.
- Lamounier, W. M (2007). Tendência, ciclos e sazonalidade nos preços do café. *Revista Gestão e Produção*, v. 14, n. 1, p. 13-23.
- Marques, D. V., Souza, G.S., Souza, M. O., & Marra, R (2010). Análise do mercado de café via modelos de previsão. In: SIMPÓSIO DE PESQUISA OPERACIONAL. 42. Anais... Bento Gonçalves, RS.
- Martins, Carlos F. *et al.* (2003). APLICAÇÃO DE UM FILTRO DIGITAL IMPLEMENTADO EM MICROCONTROLADOR PARA SENSORIAMENTO ÓTICO. SENAI-SC/CTAI. Disponível em: <https://wp.ufpel.edu.br/nimeq/files/2011/04/artigo_SBAI-2003.pdf>. Acesso em: 01 ago. 2019.
- Matias, Márcia Athayde (2006). Análise do comportamento de preços da Commodity Cobre: uma abordagem sob a ótica da teoria dos fractais. 94 f. Dissertação (Mestrado em Ciências Contábeis) - Universidade de Brasília, Programa Institucional e Inter-Regional de Pós-Graduação em Ciências Contábeis (UnB/UFPB/UFPE/UFRN), Brasília.
- Oliveira, Alex Sandro (2016). Uma Análise do Mercado do Petróleo Utilizando Aprendizado de Máquina. 36 f. Trabalho de Conclusão de Curso (Tecnólogo em Sistemas de Computação) - Universidade Federal Fluminense, Rio de Janeiro.
- Ribeiro, Cássio. Alba Neto, Henrique Belluzzo, & Sene, Túlio Silva. (2018). A oscilação do preço do petróleo: uma análise sobre o período entre 2010-2015. *Estudos internacionais: revista de relações internacionais da PUC Minas*. 6. 87-106. 10.5752/P.2317-773X.2018v6n1p87.
- Santos, Odilanei Moraes dos, Silva, Paula Danyelle Almeida da, & Sancovschi, Moacir (2006). CONTABILIDADE DAS EMPRESAS PETROLÍFERAS: O CUSTO DE ABANDONO. *Revista Base: Administração e Contabilidade da Unisinos, São Leopoldo*, p.5-14, abr. 2006. Quadrimestral. Disponível em:<<http://revistas.unisinos.br/index.php/base>>.
- Shanmugan, N.S., Breipohl, A. M. (1988). *Random Signals: Detection, Estimation and Data Analysis*. v. 1. John Wiley and Sons, 1988.
- Souza, V. S., Nunes, B. H. C., Almeida, A. N., & Einfeld, C. L (2007). Utilização da metodologia de Box & Jenkins na previsão do preço futuro pago as exportações paranaenses de madeira serrada. In: ENCONTRO DE ECONOMIA PARANAENSE. 5, 2007. Anais... ECOPAR, UFPR, Paraná, PR.
- Tularam, Anand, & Saeed, Tareq. (2016). Oil-Price Forecasting Based on Various Univariate Time-Series Models. *American Journal of Operations Research*. 06. 226-235. 10.4236/ajor.2016.63023.

APÊNDICE A

CÓDIGO EM PYTHON PARA CRIAR MODELO ARIMA, E PREVER VALORES DO BARRIL DE PETRÓLEO.

```

# -*- coding: utf-8 -*-
from pandas import read_csv
from pandas import datetime
from matplotlib import pyplot
from statsmodels.tsa.arima_model import ARIMA
from sklearn.metrics import mean_squared_error

dias = [1,2,3,6,7,8,9,10,13,14,15,16,17,20,21,22,23,24,28,29,30,31]

def parser(x):
    # Configura string de datas
    return datetime.strptime(x, '%d-%m-%Y')

# Lê arquivo com preço do barril do petróleo desde 1987
serie = read_csv('PET_PRI_SPT_S1_D.csv',
                header=0, parse_dates=[0],
                index_col=0, squeeze=True,
                date_parser=parser)

# Seleciona o RBRTE
X = serie.RBRTE

# Seleciona dados para treino e teste
indice = 8090 # Até Abril
treino, teste = X[0:indice], X[indice:len(X)]
historico = [x for x in treino]
predicoes = list()

for t in range(len(teste)):
    model = ARIMA(historico, order=(10,2,5))
    model_fit = model.fit(dispatch=1)
    output = model_fit.forecast()
    yhat = output[0]
    predicoes.append(yhat)
    obs = teste[t]
    historico.append(obs)
    print('predicted=%f, expected=%f' % (yhat, obs))

error = mean_squared_error(teste, predicoes)
print('teste MSE: %.3f' % error)
real = [x for x in teste]
pyplot.plot(dias, real)
pyplot.plot(dias, predicoes, color='red')
pyplot.show()

```