

# **AValiação de desempenho de modelos de séries temporais para previsões da produção de óleo diesel nacional**

**Ricardo Fagundes Bastos**  
**Universidade Federal do Rio de Janeiro - UFRJ**  
[rfagundesbastos@yahoo.com.br](mailto:rfagundesbastos@yahoo.com.br)

## **Resumo**

O objetivo deste trabalho é estudar os determinantes de diferentes tipos de métodos de previsões, assim como realizar previsões da produção nacional de óleo diesel para o período de 2014 a 2015.

A metodologia utilizada seguiu todos os passos para análise da série temporal, produção mensal de diesel, avaliação de seus componentes, nível de estacionaridade, aplicação dos modelos de previsão e por fim análise dos erros de cada um. O período analisado foi de janeiro de 2000 a agosto de 2014. Os modelos de Decomposição Clássica, Suavização Exponencial corrigido pela tendência (Método de Holt), Suavização Exponencial corrigido pela tendência e sazonalidade (Método de Holt-Winter) e ARIMA foram estimados com o intuito de identificar qual o mais adequado à previsão da produção de óleo diesel nacional. Após a identificação da melhor configuração de cada modelo, os erros na previsão associados a cada um foram estimados e analisados. Nesta situação a metodologia de Suavização Exponencial de Holt-Winter apresentou erros menores na previsão em relação às outras metodologias.

**Palavras Chaves:** Séries Temporais, Óleo diesel, Previsão, Erros de Previsão.

## **Abstract**

The aim of this paper is to study the determinants of different methods of forecasting, as well as to forecast the domestic production of diesel oil for the period 2014-2015.

The methodology followed all the steps for analysis of time series, monthly production of diesel, evaluation of its components, stationary level, application of predictive models and finally analysis of the errors of each. The period analyzed was January 2000 to August 2014. The models of Classical decomposition, exponential smoothing adjusted for trend (Holt Method), exponential smoothing adjusted for trend and seasonality (Holt-Winter Method) and ARIMA were estimated in order to identify the most appropriate for forecasting the national diesel oil production. After identifying the best configuration of each model, errors in forecasting associated with each were estimated and analyzed. In this situation the Exponential Smoothing methodology Holt-Winter had minor errors in the forecast for other methodologies.

**Keywords:** Time series, Diesel, Forecasts, Forecasts errors.

## **1 Introdução**

Em 2013, o óleo diesel foi a fonte com maior consumo energético, com 48,8 Mtep e crescimento de 5,6% com relação ao ano anterior (EPE, 2014).

O setor de transportes foi o principal setor demandante de energia, com 83,2 milhões de tep, o que representou 32% do total de energia consumida em 2013. Deste total, 46,4% correspondeu ao consumo de óleo diesel.

A produção de derivados nas refinarias nacionais atingiu o valor de 107,8 milhões de tep em 2013, crescimento de 6,5% em relação ao ano anterior, com destaque para óleo diesel que participou com 39,2% da produção total (EPE, 2014).

Entretanto, devido a forte expansão da demanda deste combustível, especialmente pelo setor de transportes, este aumento na produção ainda não é suficiente para atender o consumo interno, o que torna o país fortemente dependente de importações do óleo diesel, que possui considerável importância para a economia.

Este trabalho tem como objetivo verificar, dentre os modelos de decomposição clássica, suavização exponencial de Holt, sazonal de Holt-Winter e o modelo autorregressivo integrado de médias móveis (ARIMA) de Box e Jenkins, o comportamento dos mesmos para previsão da produção de óleo diesel. Deste modo, será feita uma comparação entre os métodos propostos, com base no menor erro fornecido para a previsão. Foram realizadas previsões com os modelos requeridos para a produção doze passos à frente, ou seja, para os meses de setembro de 2014 a agosto de 2015.

A análise comparativa entre os quatro métodos de previsão será realizada com base na comparação entre as principais medidas de erros utilizadas para avaliar a precisão de cada um na previsão de valores futuros.

## **2 Base de dados**

A série proposta é a da produção mensal de óleo diesel, em metros cúbicos (m<sup>3</sup>), no período de janeiro de 2000 a agosto de 2014.

Estes dados foram extraídos do Anuário Estatístico Brasileiro do Petróleo, Gás Natural e Biocombustíveis 2014, disponível no site <http://www.anp.gov.br>, da Agência Nacional de Petróleo, Gás Natural e Biocombustíveis (ANP).

A avaliação da amostra de dados foi realizada por meio de rotinas computacionais organizadas pelo *Software Eviews 7* para Windows e pelo Microsoft Excel 2010, importantes ferramentas para a análise matemática e solução de problemas estatísticos, que apresentam uma interface amigável no estudo de séries temporais.

## **3 Metodologia de cálculos**

A metodologia utilizada na realização deste trabalho é baseada na análise e utilização de métodos de estudos de séries temporais.

Os modelos para a execução das previsões foram escolhidos de acordo com as características da série original, com o objetivo de identificar suas principais componentes e verificar qual método de previsão melhor se ajusta às suas características. Com as previsões realizadas pelos diferentes tipos de modelos, foi feita a verificação do nível de erro com intuito de determinar a potencialidade de cada modelo testado.

### **3.1 Decomposição clássica**

A previsão baseada em modelos clássicos de decomposição de séries temporais é construída sobre a filosofia de que, existindo um padrão na série histórica, este pode ser distinguido da aleatoriedade e ser decomposto em sub padrões que permitam identificar cada componente da série em separado.

Com isso, a série original foi decomposta em três componentes básicas: Tendência (T), Sazonalidade (S) e o Componente aleatório ou ruído branco (e).

A função utilizada para relacionar estes três componentes será uma estrutura aditiva, descrita pela equação:

$$Y_t = T_t + S_t + e_t \quad (1);$$

assumindo índices de sazonalidade mensais. A variação cíclica<sup>1</sup> não será considerada, devido a dificuldade de conseguir isolar esta componente.

Para a determinação da tendência foi utilizado o método de filtragem denominado de Média Móvel Simples (MMS).

Neste trabalho foi utilizada uma média móvel centrada de ordem 12 (MMC 2x12), visto que foi considerada uma sazonalidade mensal para o cálculo das previsões, segundo esta metodologia.

Foi assumido que os índices sazonais são constantes, repetindo-se regularmente a cada ano, o que resultou em um conjunto de doze valores para o componente sazonal.

A seguir segue os principais passos adotados para a realização de previsões a partir da análise de decomposição clássica da série original.

- 1 Decompor a série em seus componentes:  $Y_t = T_t + S_t + e_t$
- 2 Obter a série livre de sazonalidade:  $\tilde{T} = Y_t - S_t = T_t + e_t$
- 3 Determinar um modelo de regressão linear ou não linear para estimar a tendência através da série livre de sazonalidade.
- 4 Previsão da série temporal para os n períodos futuros, correspondentes ao mês j:  
 $Y_{t+n,j} = \tilde{T}_{t+n} + S_j + \varepsilon; j = 1,2,3,...,12$

## 3.2 Métodos de Suavização Exponencial

### 3.2.1 Suavização Exponencial de Holt (SEH)

Este método é adequado quando as séries temporais apresentam tendência, além de suavizar o nível, utiliza uma nova constante de suavização para “modelar” a tendência da série. Os valores do nível e da tendência são estimados pelas equações (2) e (3) a seguir:

$$\bar{Y}_t = \alpha Y_t + (1 - \alpha)(\bar{Y}_{t-1} + \hat{T}_{t-1}) \quad (2)$$

$$\hat{T}_t = \beta(\bar{Y}_t - \bar{Y}_{t-1}) + (1 - \beta)(\hat{T}_{t-1}) \quad (3)$$

Onde:  $0 < \alpha < 1$  é a constante de suavização para o nível e;  $0 < \beta < 1$  é a constante de suavização para a tendência.

Desta forma, a previsão para os n períodos futuros é expressa pela equação (4).

$$\hat{Y}_{t+n} = \bar{Y}_t + n\hat{T}_t \quad (4)$$

Ou seja, a previsão é feita adicionando-se ao valor básico  $\bar{Y}_t$  a tendência multiplicada pelo número de períodos a frente que se deseja prever (n). Assim como na SES, as previsões podem ser atualizadas a cada nova observação  $Y_{t+1}$ , a partir das equações (2) e (3).

---

<sup>1</sup>Variações cíclicas correspondem às oscilações de longo prazo relativamente ao nível da série. A diferença principal para as variações sazonais está na sua duração que é irregular e normalmente superior a um ano. Portanto, estas flutuações são irregulares e tem diferentes durações e intensidades.

### 3.2.2 Suavização Exponencial Sazonal de Holt-Winter (HW)

Este método é adequado para séries que apresentam, além da tendência, uma componente sazonal. É baseado em três equações com constantes de suavização diferentes, que são associadas a cada uma das componentes do padrão da série: nível, tendência e sazonalidade.

Como descrito no item 3.1, este estudo utilizará uma estrutura aditiva para a série histórica. Neste caso, as três equações de suavização são dadas por:

$$\hat{S}_t = \gamma(Y_t - \bar{Y}_t) + (1 - \gamma)\hat{S}_{t-s} \quad (5)$$

$$\bar{Y}_t = \alpha(Y_t - \hat{S}_{t-s}) + (1 - \alpha)(\bar{Y}_{t-1} + \hat{T}_{t-1}) \quad (6)$$

$$\hat{T}_t = \beta(\bar{Y}_t - \bar{Y}_{t-1}) + (1 - \beta)(\hat{T}_{t-1}) \quad (7)$$

Onde:  $0 < \alpha < 1$  é a constante de suavização para o nível;  $0 < \beta < 1$  é a constante de suavização para a tendência e;  $0 < \gamma < 1$  é a constante de suavização para a sazonalidade.

A previsão para os  $n$  períodos futuros considerados na análise é expressa pela equação (8).

$$\hat{Y}_{t+n} = \bar{Y}_t + n\hat{T}_t + \hat{S}_{t+n} \quad (8)$$

Assim como nos casos anteriores, as previsões no modelo HW também podem ser atualizadas a cada nova observação  $Y_{t+1}$ , a partir das equações (5), (6) e (7).

Estas variantes do método de suavização exponencial apresentam como desvantagem a dificuldade em determinar os valores mais apropriados para as constantes de suavização,  $\alpha$ ,  $\beta$  e  $\gamma$ . Entretanto, uma forma adequada para determinar os melhores valores para estas constantes é através do cálculo de uma medida de qualidade de ajustamento do modelo, dada pela minimização da soma dos quadrados dos erros de previsão (Morettin e Toloi, 2004) descrito pela equação (9):

$$EQM = \frac{1}{n} \sum_{t=1}^n (Y_t - \bar{Y}_t)^2 \quad (9)$$

### 3.3 Métodos de Box e Jenkins

Box e Jenkins (1970) sugerem uma abordagem consagrada para a construção de um modelo, que consiste num ciclo iterativo com os seguintes passos, que devem ser repetidos até que o modelo se mostre adequado:

1. Uma classe geral de modelos é considerada para a análise (especificação);
2. Há identificação de um modelo, com base na análise de autocorrelações, autocorrelações parciais e outros critérios;
3. A seguir vem a fase de estimação, na qual os parâmetros do modelo identificado são estimados;
4. Finalmente, há a verificação ou diagnóstico do modelo ajustado, através de uma análise de resíduos, para validar a adequação do modelo.

Dentre os modelos lineares de Box e Jenkins, será escolhido o modelo que apresentar os melhores critérios para a realização das previsões.

O modelo linear contemplado neste trabalho, assim como os métodos utilizados para a realização das etapas de modelagem serão explicitados a seguir.

#### 3.3.1 Modelos autorregressivos integrados de médias móveis (ARIMA)

A maioria dos modelos de previsão de séries temporais supõe que estas sejam estacionárias. Se a série for não estacionária, é necessário torná-la estacionária por meio de sua diferenciação. Desta forma, o modelo ARIMA ( $p, d, q$ ) é um caso especial de

processo integrado no sentido de que supõe que a  $d$ -ésima diferença de uma série não estacionária pode ser representada por um modelo ARMA ( $p, q$ ), estacionário (Santos, 2012).

Para verificar se a série é estacionária, seguiram-se três passos. O primeiro consiste na análise gráfica. O segundo passo baseia-se na análise dos correlogramas de autocorrelação (AC). A função de autocorrelação mede a razão entre a covariância da observação com defasagem e a variância da amostra.

A terceira forma para confirmar se a série é estacionária ou não, consiste na aplicação do teste de Dickey-Fuller aumentado (ADF), que verifica a existência de pelo menos uma raiz unitária na série temporal, caso em que ela é não estacionária.

Depois de testada a estacionaridade da série, o próximo passo é analisar os correlogramas de autocorrelação e de correlações parciais da série a ser modelada, isto é, a série estacionária.

A identificação dos modelos a serem avaliados para o processo de previsão aplicada na série temporal estudada foi feito com base na tabela 1 a seguir:

**Tabela 1:** Escolha de modelos Box e Jenkins através das funções de autocorrelação serial e parcial.

Modelo Box & Jenkins	Autocorrelação	Autocorrelação Parcial
<b>MA(<math>q</math>)</b>	Picos nas defasagens até $q$ e após apresenta truncamento abrupto.	Picos nas defasagens até $q$ e decaimento exponencial ou senoidal amortecido (rápido) após.
<b>AR(<math>p</math>)</b>	Picos nas defasagens até $p$ e após decaimento exponencial ou senoidal amortecido (rápido).	Picos nas defasagens até $p$ e após apresenta truncamento abrupto.
<b>ARMA(<math>p, q</math>)</b>	Picos nas defasagens até $p$ e após decaimento exponencial ou senoidal amortecido (rápido).	Picos nas defasagens até $q$ e após decaimento exponencial ou senoidal amortecido (rápido).

**Fonte:** Carpio (2014).

A partir dessa análise surgem modelos candidatos a serem testados para realizar a previsão. A escolha dos modelos neste trabalho foi baseada nos critérios de Akaike e Schwarz. Os modelos candidatos devem ser rodados e aquele que apresentar os menores valores de Akaike e Schwarz será escolhido.

A validação do modelo é feita através da análise do correlograma dos resíduos. Se todos estiverem dentro do intervalo de confiança significa que podemos considerar todos os resíduos estatisticamente iguais a zero, ou seja, não haverá correlação entre os resíduos, indicando que o modelo escolhido é adequado. Caso se verifique a existência de correlação, deve-se voltar a etapa de escolha do modelo e testar outro.

Após encontrar o modelo mais adequado, foi feita a previsão pelo processo dinâmico, ou seja, uma previsão de longo prazo, quando se deseja prever mais valores além do próximo período.

### 3.4 Medições de erro de previsão

A estimação e avaliação dos erros de previsão irão determinar se o modelo utilizado está prevendo com precisão o componente sistemático<sup>2</sup> da série, ou seja, indica o quão próximo os dados reais seguem as previsões.

A tabela 2 a seguir mostra a definição dos principais tipos de erros utilizados na determinação de modelos de previsão.

**Tabela 2:** Principais medidas de erros de previsão.

Erro de Previsão	Erro Quadrático Médio	Desvio Médio Absoluto	Erro % Absoluto Médio
$E_t = F_t - Y_t$	$EQM = \frac{1}{n} \sum_{t=1}^n E_t^2$	$DMA_n = \frac{1}{n} \sum_{n=t}^n  E_t $	$EPAM_n = \frac{\sum_{t=1}^n \left  \frac{E_t}{Y_t} \right  \cdot 100}{n}$

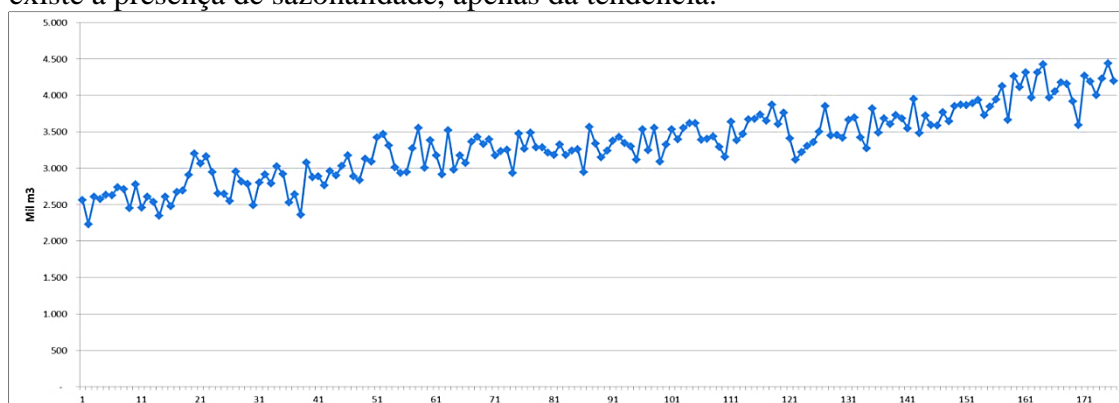
**Fonte:** Elaboração própria com base em Biral (2011).

Após submeter a amostra de dados às metodologias de previsão propostas, foi realizada uma análise comparativa entre cada um dos modelos de previsão utilizando como medida de desempenho os erros de previsão explicados anteriormente. A análise realizada buscou identificar o nível de erro fornecido por cada modelo na previsão, na tentativa de determinar qual modelo melhor se ajusta à série analisada.

## 4 Resultados e discussão

### 4.1 Método de decomposição clássica

Na figura 1 é apresentada a série original da produção mensal de óleo diesel (m<sup>3</sup>) no Brasil. A análise preliminar do comportamento desta série não permite afirmar que existe a presença de sazonalidade, apenas da tendência.



**Figura 1:** Produção Mensal de óleo diesel no Brasil: 2000-2014 (10<sup>3</sup> m<sup>3</sup>).

**Fonte:** Elaboração própria com base em ANP (2014).

Primeiramente a série foi submetida ao processo de decomposição clássica. Com isso foi feita a previsão para os doze meses seguintes (setembro de 2014 a agosto de 2015). A componente  $T_t$  foi estimada pelo processo de filtragem de Média móvel simples MMS 2x12.

Para o cálculo das componentes sazonais, primeiramente calculou-se a série livre de tendência, a partir da equação (10):

$$Y_t^{LT} = Y_t - T_t = S_t + e_t \quad (10)$$

Onde:  $Y_t^{LT}$  é o termo da série livre de tendência.

Assim, as componentes sazonais foram calculadas obtendo-se a média referente a cada mês da série livre de tendência.

<sup>2</sup> Chopra e Menindl (2011) definem componente sistemático como o valor esperado de uma série, composto pelos seguintes elementos: nível (valor da série dessazonalizada), tendência e sazonalidade.

Os resultados calculados das previsões são mostrados na tabela 3 abaixo. Para a estimativa da tendência através da série livre de sazonalidade, foi adotado um modelo de regressão linear entre a produção dessazonalizada e o tempo, com  $R^2 = 0,8366$ .

**Tabela 3:** Previsão - método de decomposição clássica.

	Sazonalidade	Tendência livre de sazonalidade	Previsão
Meses (2014-2015)	Índice Sazonal	$\hat{T}$ (Regressão Linear) - $R^2 = 0,8366$	$\hat{T} + S_j$
set/14	24.193	4.052.295	4.076.487
out/14	112.887	4.060.583	4.173.470
nov/14	-45.391	4.068.872	4.023.481
dez/14	3.981	4.077.161	4.081.142
jan/15	-82.542	4.085.450	4.002.908
fev/15	-275.975	4.093.739	3.817.764
mar/15	104.459	4.102.027	4.206.486
abr/15	-43.436	4.110.316	4.066.880
mai/15	33.798	4.118.605	4.152.403
jun/15	-55.803	4.126.894	4.071.091
jul/15	87.623	4.135.183	4.222.806
ago/15	92.810	4.143.471	4.236.282

Fonte: Elaboração própria.

#### 4.2 Método de Suavização Exponencial de Holt (SEH)

A estimativa do nível no período 0 e a tendência inicial foi obtida por meio de uma regressão linear entre a produção observada e o período  $t$ , onde obteve-se os seguintes resultados:  $Y_{t0} = 2.565.956$  e  $T_{t0} = 8.532$ .

O valor do vetor  $(\alpha, \beta)$  das constantes de suavização do nível e da tendência que fornece o menor erro quadrático médio (EQM), igual a 39.228.699.871 é (0,141; 0,001). Estes valores foram obtidos variando as constantes  $\alpha$  e  $\beta$  no intervalo 0,001 a 1, a fim de minimizar o EQM.

#### 4.3 Método de Suavização Exponencial Sazonal de Holt-Winter (SEH)

Primeiramente, a estimativa do nível e da tendência para o período 0 e dos fatores sazonais para os períodos de 1 a  $s = 12$  foram realizadas exatamente como no caso da decomposição clássica. O nível inicial e a tendência foram estimados executando uma regressão entre a produção dessazonalizada e o tempo, obtendo-se os valores  $Y_{t0} = 2.585.177$  e  $T_{t0} = 8.289$ . Os fatores sazonais (figura 3) foram estimados obtendo-se a média referente a cada mês da série livre de tendência. Em seguida estes fatores foram atualizados a partir da equação (5).

O valor do vetor  $(\alpha, \beta, \gamma)$  das constantes de suavização do nível, tendência e da componente sazonal que fornece o menor EQM, igual a 39.228.699.87 é (0,237; 0,001; 0,001).

As previsões para valores subsequentes a agosto de 2014 segundo os métodos de suavização de Holt e Holt-Winter são apresentadas na tabela 4.

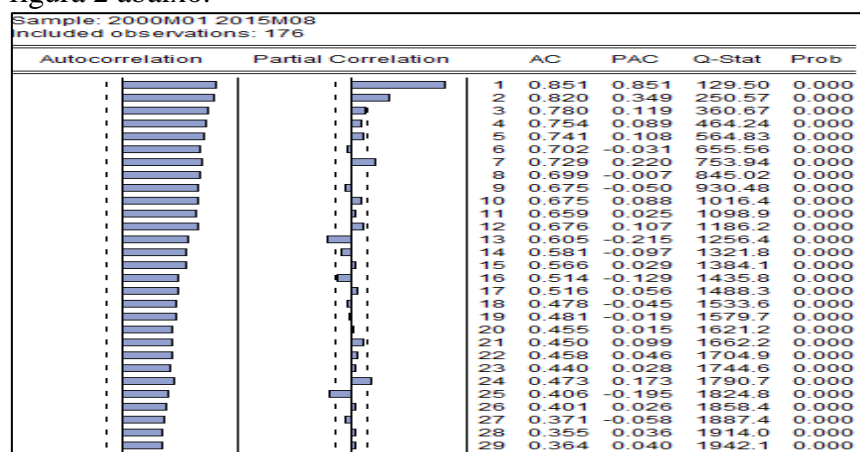
**Tabela 4:** Previsão - métodos de suavização de Holt e Holt-Winter.

Período	Meses (2000-2014)	Previsão	
		Método Holt ( $Y_{t+n} = Y_t + nT_t$ )	Método Holt-Winter ( $Y_{t+n} = Y_t + nT_t$ )
177	set/14	4.197.324	4.218.900
178	out/14	4.205.978	4.322.371
179	nov/14	4.214.633	4.161.222
180	dez/14	4.223.287	4.222.553
181	jan/15	4.231.941	4.144.218
182	fev/15	4.240.595	3.959.102
183	mar/15	4.249.249	4.356.043
184	abr/15	4.257.903	4.205.355
185	mai/15	4.266.558	4.296.699
186	jun/15	4.275.212	4.208.987
187	jul/15	4.283.866	4.371.290
188	ago/15	4.292.520	4.385.045

**Fonte:** Elaboração própria.

#### 4.4 Método autorregressivo integrado de médias móveis (ARIMA)

Primeiramente foi testada a estacionariedade da série original. A partir de inspeção visual da figura 1, pode-se deduzir que a série não é estacionária. Entretanto, para reforçar esta ideia, analisou-se o correlograma da série original, apresentado na figura 2 abaixo.

**Figura 2:** Correlograma da série original.

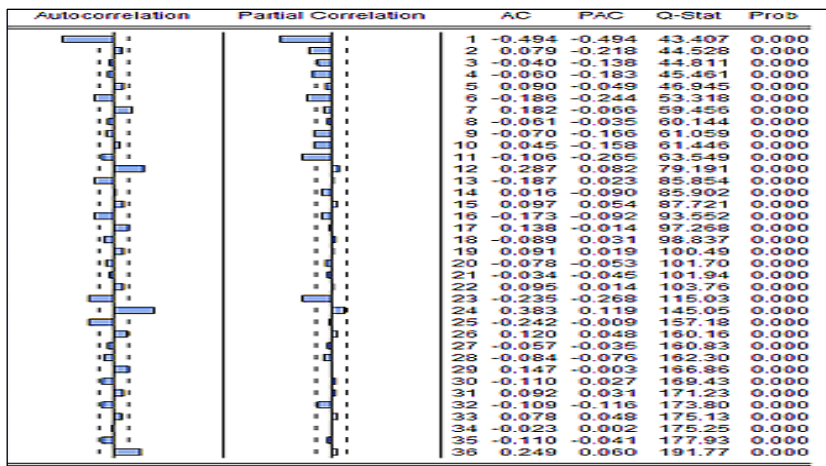
**Fonte:** Elaboração própria.

O padrão de decaimento lento observado no correlograma de autocorrelações indica a não estacionariedade da série, que foi confirmada com a aplicação do teste de Dickey - Fuller aumentado.

A estatística de teste (-0,21354) calculada é maior do que as estatísticas correspondentes aos níveis de significância usuais: -3,470427 (1%); -2,879045 (5%) e -2,576182 (10%). Assim, a hipótese nula de que a série tem raiz unitária não será rejeitada, concluindo que a mesma é não estacionária.

Com isso, é necessário diferenciar a série. O correlograma de autocorrelações e correlações parciais da série em primeira diferença mostrou o seguinte comportamento:





**Figura 3:** Correlograma da série em primeira diferença.

**Fonte:** Elaboração própria.

A análise da figura 3 indica uma série com comportamento similar ao de uma série estacionária.

Nesse sentido, recorreu-se mais uma vez ao Teste Dickey - Fuller aumentado, agora aplicado à série diferenciada, onde foi confirmada sua estacionaridade, já que a estatística de teste apresentou valor inferior às estatísticas correspondentes aos níveis de significância usuais.

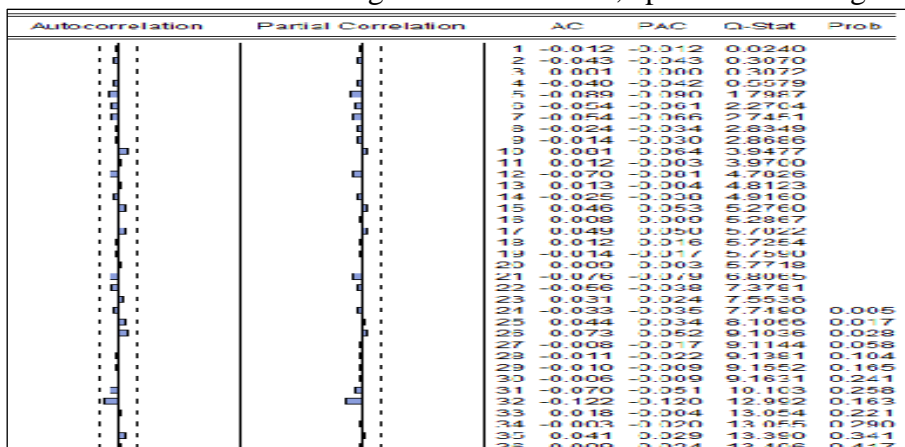
A próxima etapa consiste na determinação dos melhores modelos a serem usados na previsão. Com base na análise do correlograma da figura 3 e tabela 1, serão testados três modelos, ARIMA(12,1,0), ARIMA(16,1,0) e ARIMA(24,1,0). A tabela 5 apresenta os valores correspondentes a cada um dos modelos testados.

**Tabela 5:** Critério Akaike e Schwarz para os modelos escolhidos.

Modelo/Critério	Akaike	Schwarz
ARIMA(12,1,0)	27,26423	27,51097
ARIMA(16,1,0)	27,30805	27,63617
<b>ARIMA(24,1,0)</b>	<b>27,26117</b>	<b>27,74073</b>

**Fonte:** Elaboração própria.

Com base nestes resultados, escolheu-se o modelo ARIMA (24,1,0) para realizar as previsões, pois apresenta o menor valor de Akaike. A validação do modelo é feita através da análise do correlograma dos resíduos, apresentado na figura 4.



**Figura 4:** Correlograma dos resíduos: modelo ARIMA(24,1,0).

**Fonte:** Elaboração própria.

Avaliando a figura 4, vemos que todos os resíduos estão dentro do intervalo de confiança de 95% igual a  $\frac{1,96}{\sqrt{176}} = \pm 0,148$ , o que significa que não haverá correlação entre os resíduos, portanto o modelo escolhido é adequado.

A equação que descreve este modelo é dada por:

$$\begin{aligned} \text{PDiesel}_t = & 8898,925 - 0,7 \text{PDiesel}_{t-1} - 0,4868 \text{PDiesel}_{t-2} - 0,44 \text{PDiesel}_{t-3} - 0,331 \text{PDiesel}_{t-4} \\ & - 0,285 \text{PDiesel}_{t-5} - 0,304 \text{PDiesel}_{t-6} - 0,134 \text{PDiesel}_{t-7} - 0,165 \text{PDiesel}_{t-8} - 0,160 \text{PDiesel}_{t-9} \\ & - 0,023 \text{PDiesel}_{t-10} + 0,151 \text{PDiesel}_{t-11} + 0,005 \text{PDiesel}_{t-12} - 0,062 \text{PDiesel}_{t-13} + 0,023 \text{PDiesel}_{t-14} \\ & - 0,0102 \text{PDiesel}_{t-15} - 0,0728 \text{PDiesel}_{t-16} - 0,0318 \text{PDiesel}_{t-17} - 0,0937 \text{PDiesel}_{t-18} - 0,134 \text{PDiesel}_{t-19} \\ & - 0,098 \text{PDiesel}_{t-20} - 0,1325 \text{PDiesel}_{t-21} - 0,224 \text{PDiesel}_{t-22} + 0,158 \text{PDiesel}_{t-23} + \varepsilon \end{aligned} \quad (11)$$

Por fim, são feitas as previsões para os períodos requeridos. A tabela 6 a seguir apresenta as previsões para o período futuro.

**Tabela 6:** Previsão - modelo ARIMA (24,1,0).

Ano/Mês	Produção Diesel (m³)
2014M9	4.174.982
2014M10	4.184.133
2014M11	4.198.848
2014M12	4.197.807
2015M01	4.218.141
2015M02	4.215.473
2015M03	4.232.506
2015M04	4.240.642
2015M05	4.244.003
2015M06	4.262.396
2015M07	4.259.187
2015M08	4.278.910

**Fonte:** Elaboração própria.

#### 4.5 Comparação dos modelos: Estimativa dos erros de previsão

Os quatro métodos testados para executar as previsões para os períodos requeridos foram submetidos a uma análise comparativa baseada na avaliação da estimativa de erro de previsão. Com isso, pode-se determinar qual modelo melhor se ajusta à série original.

A tabela 7 abaixo mostra os resultados das medidas de erro para cada modelo testado.

**Tabela 7:** Estimativas de erro de previsão da produção de óleo diesel\*.

Métodos de previsão	DMA	EQM	EPAM (%)
Decomposição	137	30.495	4
Modelo de Holt	160	39.229	4,95
Modelo de Holt-Winter	123	25.809	3,75
Modelo de Box e Jenkins	167	44.206	5,00

**Fonte:** Elaboração própria.

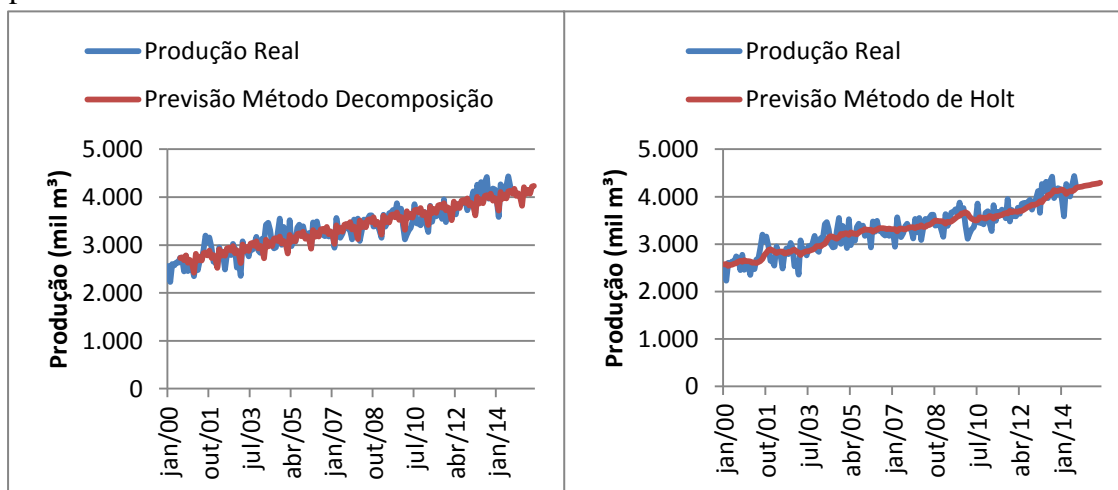
(\*) Estes valores das estimativas de erros são calculados com base nos valores de produção em 10<sup>3</sup> metros cúbicos.

Analisando os resultados oferecidos pela tabela 7, pode-se indicar uma melhor previsão por parte do modelo de Suavização Exponencial de Holt-Winter, quando observado os indicadores de EQM, DMA e EPAM.

Entretanto, o modelo de decomposição apresentou desempenho bem semelhante, não havendo vantagens significativas de um modelo de previsão em relação ao outro. O

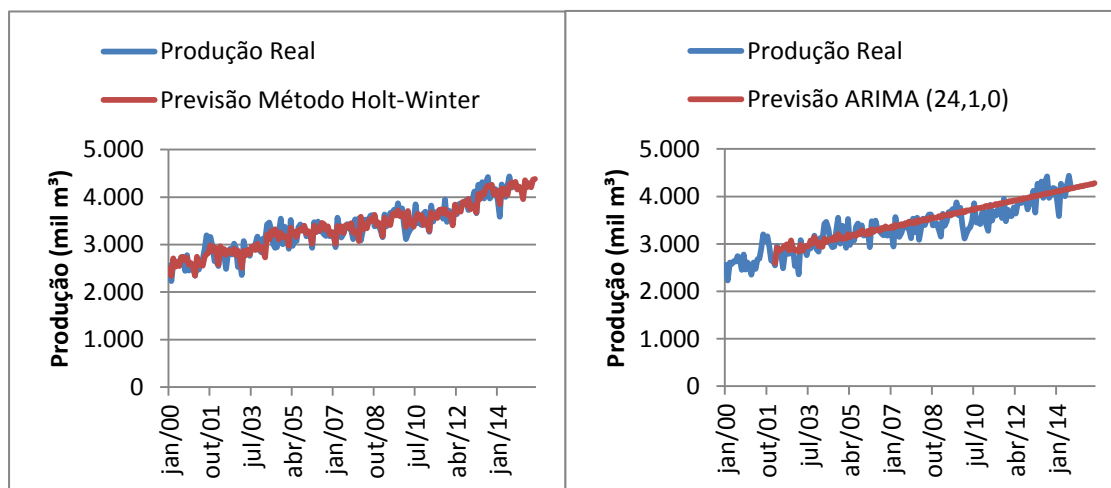
modelo ARIMA (24,1,0) de Box e Jenkins foi o que apresentou os maiores indicadores de erro. Isto pode ser um indicativo da presença de uma componente sazonal na série temporal, que não é considerada no modelo ARIMA de Box e Jenkins e no modelo de Holt.

O desempenho dos métodos testados, assim como os valores previstos para os doze meses futuros podem ser visualizados nas figuras 5 e 6 a seguir, onde foram colocados em comparação os valores reais da série temporal com os valores previstos pelos modelos.



**Figura 4:** Comparativo entre os dados reais e previstos - modelo de decomposição e modelo de Holt.

**Fonte:** Elaboração própria.



**Figura 5:** Comparativo entre os dados reais e previstos - modelo de Holt-Winter e modelo de Box e Jenkins.

**Fonte:** Elaboração própria.

Observando as figuras acima pode-se verificar que os valores estimados pelos modelo de Suavização de Hol-Winter e de Decomposição Clássica ajustam-se melhor às variações observadas na série original. A similaridade entre os desempenhos destes dois modelos pode de certa forma, ser compreendida pelo fato de que os valores previstos

pelo modelo de Holt-Winter, da forma que foi aplicado neste estudo, consiste basicamente nos valores previstos no modelo de decomposição atualizados a cada novo valor observado. Esta atualização foi feita a partir das constantes de suavização para o nível, tendência e sazonalidade, determinadas como  $\alpha$ ,  $\beta$  e  $\gamma$  respectivamente.

As previsões para os valores futuros a partir dos modelos de suavização de Holt e Holt-Winter também podem ser atualizadas na medida em que cada novo valor for observado, aprimorando assim a sua capacidade preditiva.

## 5 Conclusões

O objetivo proposto nessa pesquisa foi analisar o comportamento de modelos de previsão de séries temporais para aplicação em previsões de produção de óleo diesel nacional, derivado de petróleo mais consumido no país.

Os modelos de decomposição clássica, suavização exponencial de Holt, sazonal de Holt-Winter e o modelo ARIMA de Box e Jenkins foram escolhidos de acordo com as características da série original, com o objetivo de identificar suas principais componentes e verificar qual método de previsão melhor se ajusta as suas componentes.

Partindo do pressuposto que o melhor modelo de previsão para a série temporal seria o que tivesse o menor erro, foram analisadas três medidas de erro de previsão: erro quadrático médio (EQM), desvio absoluto médio (DMA) e erro percentual absoluto médio (EPAM). Frente a estes indicadores, os modelos de Suavização Exponencial de Holt-Winter e Decomposição Clássica apresentaram uma vantagem sobre o método de Suavização de Holt e o modelo ARIMA, onde o método de Holt-Winter foi o que apresentou o melhor desempenho, com base na informação de erro da tabela 7.

## 6 Referências bibliográficas

ANP. **Anuário Estatístico Brasileiro do Petróleo e do Gás Natural 2014**. Agência Nacional do Petróleo, Gás Natural e Biocombustíveis – ANP, 2014 Disponível em: <<http://www.anp.gov.br>>.

ANP. **Estudo Temático – Evolução do mercado de combustíveis e derivados: 2000-2012**. Agência Nacional do Petróleo, Gás Natural e Biocombustíveis – ANP, 2013 Disponível em: <<http://www.anp.gov.br>>.

BIRAL, G. Z. **Forecasting: avaliação de desempenho de modelos de séries temporais para previsões de preços do barril de petróleo da corrente OPEC Basket**. Fundação Instituto Capixaba de Pesquisas em Contabilidade, Economia e Finanças (FUCAPE). Vitória. 2011. Dissertação de Mestrado.

BOX, G. E. P., PIERCE, D. A. Distribution of Residual Autocorrelations in Autoregressive Integrated Moving Average Time Series Models, **Journal of the American Statistical Association**, 64, 1509-1526, 1970.

CARPIO, L. G. T. Tópicos em modelos quantitativos para Planejamento Energético e Ambiental. Apresentação da disciplina. 2014.

CHOPRA, S., Menindl, PETER. **Gestão da Cadeia de Suprimentos – Estratégia, Planejamento e Operações**. 4 ed. São Paulo: Pearson Prentice Hall, 2011. 520 p.

EPE. **Balanço Energético Nacional 2014 – Ano base: 2013**. EPE - Empresa de Pesquisa Energética e Ministério de Minas e Energia. Rio de Janeiro, Brasil. 2014. Disponível em: <<http://www.epe.gov.br>>.

GUJARATI, D. N. **Econometria Básica**. 4 ed. Rio de Janeiro: Elsevier, 2006. 812 p.

MORETTIN, P. A., TOLOI, C. M. C. **Análise de séries temporais**. São Paulo: Edgard Blucher, 2004. 535 p.

SANTOS, L. H. P. **Uso de modelos autoregressivos e gráficos de controle para monitorar volatilidade de ativos financeiros**. Escola Politécnica da Universidade de São Paulo. São Paulo. 2012. Trabalho de Formatura.