

**UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO  
CENTRO DE CIÊNCIAS JURÍDICAS E ECONÔMICAS - CCJE  
PROGRAMA DE PÓS-GRADUAÇÃO EM ECONOMIA**

**TAIZI HONORATO**

**PRODUÇÃO INDUSTRIAL CAPIXABA: UMA ANÁLISE COMPARATIVA DOS  
PRINCIPAIS MÉTODOS ESTATÍSTICOS DE COMBINAÇÃO DE PREVISÃO**

**VITÓRIA - ES**

**2018**

**TAIZI HONORATO**

**PRODUÇÃO INDUSTRIAL CAPIXABA: UMA ANÁLISE COMPARATIVA DOS  
PRINCIPAIS MÉTODOS ESTATÍSTICOS DE COMBINAÇÃO DE PREVISÃO**

Dissertação apresentada ao Programa de Pós-Graduação em Economia da Universidade Federal do Espírito Santo, como requisito para obtenção de título de Mestre em Economia.

Orientador: Prof. Dr. Gutemberg Hespanha  
Brasil

**VITÓRIA - ES**

**2018**

Dados Internacionais de Catalogação-na-publicação (CIP)  
(Biblioteca Central da Universidade Federal do Espírito Santo, ES, Brasil)

---

H774p Honorato, Taizi, 1990-  
Produção industrial capixaba : uma análise comparativa dos principais métodos estatísticos de combinação de previsão / Taizi Honorato. – 2018.  
120 f. : il.

Orientador: Gutemberg Hespanha Brasil.  
Dissertação (Mestrado em Economia) – Universidade Federal do Espírito Santo, Centro de Ciências Jurídicas e Econômicas.

1. Indústrias - Espírito Santo (Estado). 2. Análise de séries temporais. 3. Previsão econômica. I. Brasil, Gutemberg Hespanha. II. Universidade Federal do Espírito Santo. Centro de Ciências Jurídicas e Econômicas. III. Título.

CDU: 330

---

Elaborado por Perla Rodrigues Lôbo – CRB-6 ES-527/O

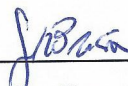
TAIZI HONORATO

**PRODUÇÃO INDUSTRIAL CAPIXABA: UMA ANÁLISE COMPARATIVA DOS  
PRINCIPAIS MÉTODOS ESTATÍSTICOS DE COMBINAÇÃO DE PREVISÃO**

Dissertação apresentada ao Programa de Pós-graduação em Economia da Universidade Federal do Espírito Santo como requisito para a obtenção do título de Mestre em Economia.

Vitória/ES, 26 de junho de 2018.

**COMISSÃO EXAMINADORA**



---

Prof. Dr. Gutemberg Hespanha Brasil  
Universidade Federal do Espírito  
Orientador



---

Prof. Dr. Ricardo Ramallete Moreira  
Universidade Federal do Espírito Santo



---

Prof. Dr. Antônio Fernando Pêgo e Silva  
DEST/Universidade Federal do Espírito Santo

Aos meus avós, Elvira Alves Pianissoli e João Pianissoli.

## **AGRADECIMENTOS**

Agradeço a Deus que me conduziu a alcançar esse sonho, me sustentou e direcionou nos momentos mais difíceis dessa jornada. E a intercessão de Nossa Senhora da Penha, mãe e padroeira no meu querido Estado.

Agradeço aos meus queridos e amados avós, Elvira Alves Pianissoli e João Pianissoli, que são para mim como que presente de Deus. Me conduziram a uma boa educação, ensinando os ricos preceitos da vida, e trazendo valores que nunca poderão ser substituídos. Dedico esse parágrafo também a minha irmã, Laisi Honorato, que tanto amo e quero bem.

Agradeço ao meu esposo, Jardel Coelho dos Reis, por todo apoio e compreensão que foi um imenso diferencial nessa conquista. Obrigada por acreditar nos meus sonhos, e mais que isso, por me impulsionar a conquistá-los.

Agradeço imensamente ao Professor Gutemberg Hespanha Brasil, que antes de orientador dessa dissertação de mestrado, é um amigo, e alguém que admiro como ser humano. Obrigada pela compreensão com a minha jornada de trabalho corporativo que muito concorreu com a elaboração desse estudo. Obrigada por todo aprendizado que me proporcionou.

Agradeço ao Programa de Pós-Graduação em Economia da UFES como um todo, professores e secretaria. Obrigada pela orientação e disponibilidade.

## RESUMO

Este trabalho tem por objetivo especificar modelos de previsão e combinação de previsão aplicados as séries temporais de relativas aos índices na Indústria Extrativa, Indústria de Transformação, e Indústria Geral para o Estado do Espírito Santo. A priori, concentrou-se em avaliar a performance dos principais métodos de previsão dispostos na literatura, tais como Holt-Winters, Box-Jenkins, Modelos de Redes Neurais Artificiais, e Modelos Econométricos, incorporando, neste último método, outras variáveis econométricas, tais como inflação, taxa de juros, taxa de desocupação, índice de confiança do empresário industrial, utilização da capacidade instalada, dentre outras. Em segundo momento, considera-se selecionar o melhor modelo estimado para cada metodologia, e então aplicar métodos de combinação de previsão, com intuito de avaliar se há diferença entre a acurácia de previsões individuais e a de suas combinações. Como técnica de combinação das previsões, foram considerados os métodos da média aritmética, variância mínima simplificado, e regressão por mínimos quadrados ordinários. A avaliação de desempenho das previsões e combinações de previsões é obtida por meio das medidas de acurácia MAE, MSE, RMSE, MAPE, SMAPE, e U de Theil. Como principal resultado obtido, destaca-se as previsões obtidas a partir do método de combinação de previsão por mínimos quadrados ordinários que unanimemente apresentaram desempenho superior as demais previsões para as três series de produção industrial consideradas neste estudo.

**Palavras-Chave:** Produção Industrial, Séries Temporais, Combinação de Previsão.

## ABSTRACT

The objective of this work is to specify forecast models and forecast combinations applied to the time series of indices in the Extraction Industry, Transformation Industry and General Industry for the State of Espírito Santo. A priori, it focused on evaluating the performance of the main forecasting methods available in the literature, such as Holt-Winters, Box-Jenkins, Models of Artificial Neural Networks, and Econometric Models, incorporating other econometric variables in the latter method. such as inflation, interest rate, unemployment rate, industrial entrepreneur confidence index, utilization of installed capacity, among others. Secondly, it is considered to select the best model estimated for each methodology, and then to apply predictive combination methods, in order to evaluate if there is a difference between the accuracy of individual forecasts and their combinations. As technique combination forecast, were considered the methods of arithmetic mean, simplified minimum variance, and ordinary least squares regression. The performance evaluation of predictions and combinations of forecasts is obtained by accuracy measurements know by MAE, MSE, RMSE, MAPE, SMAPE, and U - Theil's. As the main result obtained, we highlight the predictions obtained from the forecast combination method ordinary least squares, which unanimously performed better than the other predictions for the three industrial production series considered in this study.

**Keywords:** industrial production, time series, combination of forecast.



## LISTA DE FIGURAS

Figura 1 - Principais Setores da Indústria Capixaba – Portal da Industria. ....	25
Figura 2 - Produção Industrial Capixaba – Índice para Indústria Geral, Extrativa e da Transformação.....	25
Figura 3 - Esboço das Etapas Previstas do Estudo.....	26
Figura 4 - Taxa de Juros Selic (período 2011 a 2018).....	27
Figura 5 - Taxa de Inflação IGP (período 2011 a 2018).....	27
Figura 6 - Taxa de Cambio (período 2011 a 2018).....	28
Figura 7 - Índice de Confiança do Empresário Indústria – Setor Extrativo (período 2011 a 2018).....	28
Figura 8 - Índice de Confiança do Empresário Industrial – Setor de Transformação (período 2011 a 2018) .....	28
Figura 9 - Utilização da Capacidade Instalada (período 2011 a 2018) .....	29
Figura 10 - Percentual de Desocupação (período 2011 a 2018) .....	29
Figura 11 – Consumo de Energia (período 2011 a 2018) .....	30
Figura 12 – Modelo Matemático de um Neurônio Artificial.....	43
Figura 13 - Redes Multicamadas.....	44
Figura 14 - Séries de Produção Industrial com uma diferença .....	50
Figura 15 - Produção Industrial Geral: Realizado (2002-2017) x Previsto via Box-Jenkins (2017).....	52
Figura 16 - Produção Industrial Geral: Realizado (2002-2017) x Previsto via Holt-Winters (2017).....	53
Figura 17 - Produção Industrial Geral: Realizado (2002-2017) x Previsto via Modelo Econométrico (2017).....	60
Figura 18 - Modelo Redes Neurais Artificiais - Indústria Geral (jan-02 a dez-16) .....	61
Figura 19 - Produção Industrial Geral: Realizado (2002-2017) x Previsto via Redes Neurais (2017).....	61
Figura 20 - Produção Industrial Geral: Realizado (2002-2017) x Previsto via Combinação Previsão (2017) .....	63
Figura 21 – Previsão de Produção Industrial Geral (2017) - combinação de previsão via regressão por MQO.....	64
Figura 22 – Indústria Extrativa: Realizado (2002-2017) x Previsto via Box-Jenkins (2017)..	66

Figura 23 – Indústria Extrativa: Realizado (2002-2017) x Previsto via Holt-Winters (2017).	66
Figura 24 – Indústria Extrativa: Realizado (2002-2017) x Previsto via Modelagem Econométrica (2017) .....	71
Figura 25 - Modelo Redes Neurais Artificiais - Indústria Extrativa.....	72
Figura 26 – Indústria Extrativa: Realizado (2002-2017) x Previsto via Redes Neurais Artificiais (2017) .....	72
Figura 27- Indústria Extrativa: Realizado (2002-2017) x Previsto via Combinação Previsão (2017).....	74
Figura 28 – Indústria Extrativa - combinação de previsão via regressão por MQO .....	74
Figura 29 – Indústria de Transformação: Realizado (2002-2017) x Previsto via Box-Jenkins (2017).....	76
Figura 30 – Indústria de Transformação: Realizado (2002-2017) x Previsto via Holt-Hinters (2017).....	77
Figura 31 - Indústria de Transformação: Realizado (2002-2017) x Previsto via Modelagem Econométrica (2017) .....	79
Figura 32 - Modelo Redes Neurais Artificiais - Indústria de Transformação.....	80
Figura 33 - Indústria de Transformação: Realizado (2002-2017) x Previsto via Redes Neurais (2017).....	81
Figura 34 - Indústria de Transformação: Realizado (2002-2017) x Combinação de Previsão (2017).....	82
Figura 35 - Previsão de Produção Industrial de Transformação - combinação de previsão via regressão por MQO.....	83
Figura 36 - Produção Industrial Geral: Realizado x Previsto via Box-Jenkins com intervalo de confiança de 95% (jan-2017 x dez x 2017) .....	119
Figura 37- Produção Industrial Geral: Realizado x Previsto via Combinação de Previsão com intervalo de confiança de 95% (jan-2017 x dez x 2017).....	120

## LISTA DE TABELAS

Tabela 1 - Números da Indústria Geral Capixaba – Variação % em relação ao mês anterior	23
Tabela 2 - Números da Indústria Geral Capixaba – Variação % interanual	24
Tabela 3 - Utilização de Capacidade Instalada	24
Tabela 4 - Testes de estacionariedade	49
Tabela 5 - Testes de Estacionariedade na Série I(1)	49
Tabela 6 - Modelos Box-Jenkins para Produção Industrial Geral (jan-02 a dez-16)	51
Tabela 7 - Medidas de Acurácia para modelos Box-Jenkins – Indústria Geral (jan-17 a dez-17)	51
Tabela 8 - Medidas de Acurácia para modelos Holt-Winters – Indústria Geral (jan-17 a dez-17)	52
Tabela 9 – Coeficientes Modelo Holt-Winters Selecionado – Indústria Geral (jan-02 a dez-16)	53
Tabela 10 - Modelo Econométrico - Indústria Geral (jan-11 a dez-16)	54
Tabela 11 – Indústria Geral - Realizado x Previsto (2017)	55
Tabela 12 – Taxa de Juros (Selic) - Realizado x Previsto (2017)	56
Tabela 13 - Testes de Estacionariedade - Selic (jan-11 a dez-16)	57
Tabela 14 – Taxa de Cambio - Realizado x Previsto (2017)	57
Tabela 15 - Testes de Estacionariedade - Taxa de Cambio (jan-11 a dez-16)	58
Tabela 16 - Testes de Estacionariedade (Percentual de Desocupação) (jan-11 a dez-16)	58
Tabela 17 – Testes de Estacionariedade – ICE Indústria de Transformação (jan-11 a dez-16)	58
Tabela 18 - Testes de Estacionariedade (UCI Indústria Geral) (jan-11 a dez-16)	59
Tabela 19 – Consumo de Energia - Realizado x Previsto (2017)	59
Tabela 20 - Testes de Estacionariedade (Consumo de Energia) (jan-11 a dez-16)	60
Tabela 21 - Medidas de Acurácia para Modelo Econométrico - Indústria Geral (jan-17 a dez-17)	60
Tabela 22 - Medidas de Acurácia via Redes Neurais - Indústria Geral (jan-17 a dez-17)	61
Tabela 23 - Previsões Indústria Geral (2017)	62
Tabela 24 - Medidas de Acurácia - Todos Modelos- Indústria Geral (jan-17 a dez-17)	62
Tabela 25 - Modelos Box-Jenkins para Indústria Extrativa (jan-02 a dez-16)	65
Tabela 26 - Medidas de Acurácia para modelos Box-Jenkins – Indústria Extrativa (jan-17 a dez-17)	65

Tabela 27 - Acurácia para modelos Holt-Winters – Indústria Extrativa .....	66
Tabela 28 - Coeficientes Modelo Holt-Winters – Indústria Extrativa .....	67
Tabela 29 - Modelo Econométrico - Indústria Extrativa .....	67
Tabela 30 - Indústria Extrativa - Realizado x Previsto (2017) .....	69
Tabela 31 – Taxa de Juros (Selic) - Realizado x Previsto (2017) .....	69
Tabela 32 – Taxa de Cambio - Realizado x Previsto (2017).....	70
Tabela 33 - Testes de Estacionariedade - ICE Indústria Extrativa (jan-11 a dez-16).....	71
Tabela 34 - Medidas de Acurácia - Modelo Econométrico - Indústria Extrativa .....	71
Tabela 35 - Medidas de Acurácia - Redes Neurais - Indústria Extrativa .....	72
Tabela 36 - Previsões Indústria Extrativa.....	73
Tabela 37 - Medidas de Acurácia - Todos Modelos- Indústria Extrativa.....	73
Tabela 38 - Modelos Box-Jenkins para Produção Indústria De Transformação .....	75
Tabela 39 - Medidas de Acurácia para modelos Box-Jenkins – Indústria De Transformação	75
Tabela 40 - Acurácia para modelos Holt-Winters – Indústria De Transformação.....	76
Tabela 41 - Coeficientes Modelo Holt-Winters – Indústria De Transformação.....	77
Tabela 42 - Modelo Econométrico - Indústria de Transformação.....	78
Tabela 43 – Indústria de Transformação - Realizado x Previsto (2017).....	79
Tabela 44 - Medidas de Acurácia - Modelo Econométrico – Indústria De Transformação ....	79
Tabela 45 - Medidas de Acurácia - Redes Neurais - Indústria de Transformação.....	80
Tabela 46 - Previsões Indústria de Transformação .....	81
Tabela 47 - Medidas de Acurácia - Todos Modelos- Indústria de Transformação .....	82

## **LISTA DE SIGLAS**

**ADF:** Teste de Raiz Unitária Dickey-Fuller Aumentado

**AIC:** Critério de Informação de Akaike

**BJ:** Box-Jenkins

**CNAE:** Classificação Nacional de Atividades Econômicas

**CNI:** Confederação Nacional da Indústria

**DF:** Teste de Raiz Unitária Dickey-Fuller

**IDEIES:** Instituto de Desenvolvimento do Espírito Santo

**HW:** Holt-Winters

**IJSN:** Instituto Jones dos Santos Neves

**IBRE:** Instituto Brasileiro de Economia

**IBGE:** Instituto Brasileiro de Geografia e Estatística

**IPEA** - Instituto de Pesquisa Econômica Aplicada.

**FINDES:** Federação das Indústrias do Estado do Espírito Santo

**MAE:** Erro Absoluto Médio

**MSE:** Erro Quadrático Médio

**MAPE:** Erro Percentual Absoluto Médio

**PIA:** Pesquisa Industrial Anual

**PAIC:** Pesquisa Anual da Indústria da Construção

**PP:** Teste de Raiz Unitária Philips Perron

**RMSE:** Raiz do Erro Quadrático Médio

**RNA:** Redes Neurais Artificiais

**SELIC:** Sistema Especial de Liquidação e de Custódia

**SMAPE:** Erro Percentual Absoluto Médio Simétrico

**SIDRA:** Sistema IBGE de Recuperação Automática

**VTI:** Valor da Transformação Industrial

## SUMÁRIO

<b>2 A INDÚSTRIA CAPIXABA.....</b>	<b>20</b>
<b>3 MATERIAIS E MÉTODOS .....</b>	<b>26</b>
<b>3.1 MODELOS BOX-JENKINS.....</b>	<b>31</b>
<b>3.2 MODELO HOLT-WINTERS (HW).....</b>	<b>36</b>
<b>3.3 METODOLOGIA ECONOMÉTRICA TRADIONAL .....</b>	<b>39</b>
<b>3.4 REDES NEURAS ARTIFICIAIS .....</b>	<b>42</b>
<b>3.5 TÉCNICAS DE COMBINAÇÃO DE PREVISÃO.....</b>	<b>45</b>
<b>3.6 MEDIDAS DE ACURÁCIA.....</b>	<b>47</b>
<b>4 RESULTADOS E DISCUSSÃO .....</b>	<b>49</b>
<b>4.1 INDÚSTRIA GERAL .....</b>	<b>50</b>
<b>4.2 INDÚSTRIA EXTRATIVA.....</b>	<b>65</b>
<b>4.3 INDÚSTRIA DE TRANSFORMAÇÃO.....</b>	<b>75</b>
<b>5 CONCLUSÃO .....</b>	<b>84</b>
<b>REFERÊNCIAS BIBLIOGRÁFICAS .....</b>	<b>86</b>
<b>APENDICE A - SEÇÃO B DA CNAE 2.0 .....</b>	<b>90</b>
<b>APENDICE B - SEÇÃO C DA CNAE 2.0 .....</b>	<b>92</b>
<b>APENDICE C – LISTA DE ATIVIDADES COM PESO NORMALIZADO .....</b>	<b>93</b>
<b>APENDICE D – PREVISÃO VARIÁVEIS REGRESSORAS.....</b>	<b>95</b>
<b>APENDICE E – ROTINA R STUDIO (INDÚSTRIA GERAL) .....</b>	<b>100</b>
<b>APENDICE F – INTERVALO DE CONFIANÇA DAS PREVISÕES .....</b>	<b>119</b>

## 1 INTRODUÇÃO

No Brasil, o processo de industrialização tornou-se fator determinante na dinâmica econômica a partir da década de 1930 com o fim da era mercantil, ano este em que o país passou pela conhecida Revolução de 1930 e culminantemente ao Golpe de Estado, onde Getúlio Vargas assume a chefia do Governo Provisório, ocasionando a perda da hegemonia política pela burguesia cafeeira em favor da classe industrial ascendente (DE LACERDA, 2005). Suzigan e Furtado (2006), comentam a industrialização do Brasil de forma definitiva na agenda política após a Segunda Guerra Mundial -1945 -, com a implantação de diversas políticas industriais. A partir de então, a indústria se intensificou, e no ano de 1956 a produção industrial se destaca por superar a produção agrícola.

Versiani (1984) apresenta três fatores que contribuíram para o forte desempenho da indústria a partir de 1930:

- (i) o comportamento do setor cafeeiro,
- (ii) a evolução da taxa de câmbio; e
- (iii) as políticas monetárias e fiscais seguidas/adotadas pelo Governo.

No início do governo de Juscelino Kubitschek, em 1956, a indústria já era formada, e a partir de então, ocorre a consolidação nacional via introdução de planificação setorial, e o aceleramento do desenvolvimento econômico (PEREIRA, 1963).

Bonelli e Pessoa (2010), em seu estudo “Desindustrialização no Brasil: um resumo da evidência” apresenta uma análise da evolução da posição brasileira em relação a um padrão internacional e mostra que em 1970 o Brasil tinha uma indústria muito maior do que seria justificado por um conjunto de variáveis indicativas do grau de desenvolvimento econômico, tecnológico e tamanho, em outras palavras, acima do padrão.

A economia do estado de Espírito Santo, até o início da segunda metade do século XX, esteve apoiada na economia agroexportadora da monocultura cafeeira e se apresentava como uma região periférica e subdesenvolvida, com pouca integração ao mercado nacional (NADER, 2007).



“Localizado entre pólos industriais consolidados, tais como São Paulo, Rio de Janeiro e Belo Horizonte, o Espírito Santo contava com uma população fragmentada e com vazios demográficos. Diferentemente da estrutura urbana formada nos estados de São Paulo e Rio de Janeiro, onde a base necessária para o desenvolvimento manufatureiro se deu de forma expressiva, uma vez que contava com população densa, mão-de-obra especializada e dinheiro disponível, o mercado urbano capixaba não dispunha de tais condições para se desenvolver tanto, sendo obrigado aos empreendimentos da comercialização do café” (NADER, 2007).

Nessa época, grandes foram os esforços assumidos pelo Governo do Estado para apoiar a indústria, tais como:

- (i) realizar doação de terrenos,
- (ii) promover a isenção de impostos e
- (iii) fornecer energia elétrica gratuita como forma de incentivos fiscais para a implantação de indústrias em solo capixaba.

Todavia, observou-se nesse período expansão da cafeicultura nas terras capixabas, e com isso, a economia capixaba precisava modernizar-se e exigia medidas decisivas dos poderes públicos, no sentido de investimento ao processo de industrialização do Estado. Para isso, várias iniciativas foram tomadas e dentre as quais se destacam os projetos de indústrias de siderurgia, mineração e metalurgia, de papel e celulose, além de ter sido criada uma política de transportes naval e rodoviário.

A partir de então, são apresentados os chamados Grandes Projetos de Impacto, ou Grandes Projetos Industriais, os quais foram a conjugação de esforços dos Governos Federal e Estadual que viabilizaram uma mudança radical na economia do Espírito Santo, obrigando-o a deixar sua dependência da monocultura cafeeira secular para ligar-se ao grande capital internacional (NADER, 2007).

Destaca-se como primeiro grande projeto industrial (1942), a Companhia de Ferro e Aço de Vitória (COFAVI). Esse empreendimento, inicialmente financiado pelo capital privado, configura uma das primeiras iniciativas da siderurgia nacional e revela a vocação do Espírito Santo para projetos de grande porte vinculados à indústria de base (FORTUNATO, 2011).

“A industrialização do Espírito Santo, conjuntamente com a crise da economia cafeeira, marcada pela política de erradicação dos cafezais improdutivos, determinou um novo padrão de crescimento do Estado. De fato, no período de 1970-1977, verificou-se um arrefecimento no crescimento do setor agrícola no Estado, que foi de apenas 1,6% contra 5,8% no Brasil; a indústria, no pólo oposto, apresentou crescimento vertiginoso, na ordem de 22%, contra 11,7% do crescimento industrial nacional” (FORTUNATO, 2011).

A partir de 1990, empresas dos Grandes Projetos continuaram a influenciar a dinâmica industrial capixaba, bem como fatores como o renascimento da extração de petróleo e gás, crescimento no setor de alimentos e bebidas, metal-mecânico, rochas e moveleiro, e na construção civil (CAÇADOR, 2009).

De forma geral, a economia capixaba apresentou um desempenho superior ao Brasil neste período. Segundo Caçador (2009), o Espírito Santo cresceu a uma taxa média anual de 3,9% e o país a 2,4%. Despenho esse superior ao verificado na década de 80, apresentada por Mota (2002) como década perdida.

A indústria do Espírito Santo nos anos 2000 até 2007 seguia uma tendência de crescimento similar a média nacional, mas com a Crise Mundial, em setembro de 2008, não somente o Estado Capixaba começa a apresentar grandes déficits mas o país como um todo (JUNIOR, 2014), retomando aos patamares observados no período antes da Crise, somente em meados de 2010.

Em novembro de 2015, com o rompimento da Barragem de Fundão, da mineradora Samarco, em Mariana (Minas Gerais), a produção industrial capixaba encolheu novamente, onde na comparação com igual mês do ano anterior, apresentou queda de 11,1% em novembro, e - 19,1% dezembro de 2015, com três das cinco atividades pesquisadas mostrando queda na produção, onde destaca-se o setor de minérios de ferro pelotizados ou sinterizados (IBGE, 2015).

Considerando os fatos acima destacados, o presente trabalho tem como centro de sua pesquisa o aprofundamento teórico acerca da Produção Industrial no Estado do Espírito Santo, utilizando como fonte primária o Sistema de Indicadores da Produção Industrial tendo como órgão responsável pelo fornecimento e atualização do mesmo, o Instituto Brasileiro de

Geografia e Estatística (IBGE), a partir do Sistema IBGE de Recuperação Automática – SIDRA.

Paralelamente, visa complementar diversos estudos hoje já disponibilizados pelos mais importantes órgãos: IBGE, Instituto Jones dos Santos Neves (IJSN), Federação das Indústrias do Estado do Espírito Santo (Findes), Confederação Nacional da Indústria (CNI), Instituto de Desenvolvimento do Espírito Santo (IDEIES), entre outros.

Nesse sentido, será base desse estudo o índice mensal observado nos últimos 16 anos, no período de janeiro/12 a dezembro/17, totalizando uma série observada de 192 pontos, como principal fonte de dados na aplicação de modelagem estatística dos dados via métodos de previsão, tais como Holt-Winters, Box-Jenkins, Redes Neurais e Modelos Econométricos.

Posteriormente, propõe-se aplicar métodos de Combinação de Previsão alcançando modelos de previsão mais eficientes, e assim o objetivo principal foi prever o Índice de Produção Industrial do Estado do Espírito Santo, em seus níveis: Indústria Geral, Indústria Extrativa e Indústria de Transformação, via modelagem estatística. A seguir é apresentado os objetivos secundários:

- Apresentar Referencial Teórico acerca da economia do Estado do Espírito Santo no que tange ao processo de industrialização;
- Apresentar Análise Descritiva das Séries de Índice de Produção Industrial Geral, Indústria Extrativista e Indústria da Transformação;
- Apresentar Métodos Estatísticos de Previsão: Holt-Winters, Box-Jenkins, Redes Neurais e Modelos Econométricos, com base na série histórica observada de janeiro de 2002 a dezembro de 2016;
- Selecionar modelos de Previsão que apresentaram melhores resultados e Aplicar Métodos de Combinação de Previsão;
- Selecionar melhor método de Combinação de Previsão via critérios de performance, tais como: MAPE (Média Percentual Absoluta do Erro), MASE (Erro Absoluto Médio Escalonado), MSE (Erro Quadrático Médio), MAD (Desvio Médio Absoluto), U de Theil;

- Apresentar previsão estatística de 12 meses (ano de 2017) para os índices de Produção Industrial Geral, Industrial Extrativa e Indústria de Transformação.

A estrutura do trabalho apresentada a seguinte estrutura: No Capítulo 1 estão presentes o problema de pesquisa, seus objetivos, bem como a relevância e limitação do estudo. No Capítulo 2, serão apresentados os indicadores Estaduais para Produção Industrial Geral, Indústria Extrativa e Indústria de Transformação, bem como as séries de dados que subsidiarão esse estudo.

O Capítulo 3 será dedicado a apresentar os métodos de previsão a serem aplicados série de dados, tais como: Modelo Holt-Winters, Box-Jenkins, Redes Neurais, e Modelos Econométricos. Também será detalhada a metodologia dos principais métodos de combinação de previsão e medidas de acurácia para seleção de melhores resultados.

A análise dos dados estará presente no Capítulo 4, onde serão apresentados os dois melhores resultados nas previsões dos índices de produção industrial, e na sequência os três melhores métodos de combinação de previsão, apontando como critério de seleção as medidas de acuraria apresentadas no Capítulo 3.

O Capítulo 5 conterà as conclusões obtidas a partir dos principais resultados da pesquisa, apresentado estrapolações para a economia industrial capixaba, e as recomendações para trabalhos futuros.

## 2 A INDÚSTRIA CAPIXABA

A Indústria Geral compreende as informações econômico-financeiras, de produção, bens e serviços consumidos, e emprego, entre outros aspectos, específicas das atividades das Indústrias extrativas – que pertencem à seção B da CNAE 2.0 (APENDICE A) –, e as Indústrias de Transformação –que pertencem à seção C da CNAE 2.0 (APENDICE B).

Desde 1970, são apurados mensalmente o Índice de Produção Industrial, refletindo em uma medida de variação de volume de produção industrial (temporal). Nota-se que a metodologia de apuração do índice passou por revisões/melhorias ao longo dos anos, sendo estas associadas com a liberação de informações estruturais sobre a composição das atividades e sobre a oferta de produtos, podendo-se resumir por atualização de informações relevantes ao índice (IBGE, 2015).

Tais reformulações ocorridas ao longo do tempo objetivaram atualizar dois principais níveis de informação:

- i) o grupo de produtos;
- ii) e a lista de informantes que compunham o índice, atualizado a partir da Pesquisa Industrial Mensal – Produção Física (PIM-PF), com a finalidade de apresentar a produção física de uma série de produtos previamente selecionados, de modo que uma variação *quantum* dessa cesta de produtos indicaria a taxa de variação mensal da produção das atividades produtivas industriais (IBGE, 2015).

Desse modo, o resultado da metodologia não se atém a gerar informações para cada produto individualmente, mas sim, a partir de um sistema de ponderação que permita agregar as informações individuais de quantidades produzidas, relacionado ao desempenho da atividade industrial (IBGE, 2015).

Atualmente, a nível Brasil, são considerados 944 produtos fabricados em cerca de 7.800 unidades locais industriais, escolhidos segundo os agrupamentos de atividades industriais que o sistema de indicadores se propõe representar. No Espírito Santo, são considerados 30 produtos, com índice de cobertura de 79%.

De acordo com o Sistema de contas nacionais 2008 (apud IBGE 2015) a produção é “uma atividade realizada sob a responsabilidade, o controle e gestão de uma unidade institucional, nas quais se utilizam insumos de mão-de-obra, capital e bens e serviços para obter outros bens ou serviços. [...] A análise econômica da produção se ocupa principalmente de atividades que produzem saídas de um tipo que pode ser entregue ou fornecido a outras unidades institucionais”.

Assim sendo, a produção abrangeria todos os setores da economia, porém, a metodologia de apuração dos índices de produção industrial baseia-se em um conceito de produção num âmbito mais restrito, considerando-se apenas as atividades industriais que contidas na Pesquisa Industrial Anual - PIA e na Pesquisa Anual da Indústria da Construção - PAIC, ambas do IBGE, acrescentando-se ainda os serviços de produção e distribuição de eletricidade, gás e água (IBGE, 2015).

A conjuntura de indicadores apurados pelo IBGE acerca da produção industrial nacional e para algumas Unidades da Federação tem como objetivo teórico central o de refletir a evolução do valor adicionado ao longo do tempo, de modo que o valor adicionado mensurado pela quantidade de bens e serviços produzidos (seja por estabelecimento, unidade local, indústria, etc...) excede os insumos intermediários consumidos (IBGE, 2015).

Considera-se ainda as atividades industriais organizadas de acordo com a CNAE 2.0, particularmente, as seções Indústrias extrativas e Indústrias de transformação. Este trabalho, se atêm inicialmente a investigar apenas as séries em nível geral, atividades extrativas e de transformação, sem se ater individualmente a cada umas das sub-atividades industriais.

A divulgação dos índices de produção industrial se dá a nível do Brasil, e ainda: Região Nordeste e os Estados selecionados de Amazonas, Pará, Ceará, Pernambuco, Bahia, Minas Gerais, Espírito Santo, Rio de Janeiro, São Paulo, Paraná, Santa Catarina, Rio Grande do Sul, Mato Grosso e Goiás. Cujo critério de divulgação é a participação de no mínimo, 1,0% no total do valor da Valor da Transformação Industrial (VTI) no ano de 2010 (IBGE, 2015).

No Estado do Espírito Santo, a PIM-PF considerada as seguintes atividades (APENDICE C – Lista Atividades e Peso)

- Indústria Geral;

- Indústrias Extrativas (CNAE – Seção B);
- Indústrias de Transformação (CNAE – Seção C);
  - Fabricação de produtos alimentícios (CNAE – Seção C – Atividade 10);
  - Fabricação de celulose, papel e produtos de papel (CNAE – Seção C – Atividade 17);
  - Fabricação de produtos de minerais não metálicos (CNAE – Seção C – Atividade 23);
  - Metalurgia (CNAE – Seção C – Atividade 24).

O resultado deste levantamento é a divulgação mensal de cinco tipos de índices:

- (i) Índice Base Fixa Mensal (Número-Índice) [ $I_{BF\_Sem\_Ajuste\_Sazonal}$ ]: responsável por apresentar comparação da produção do mês de referência do índice com a média mensal produzida no ano base da pesquisa (2012);
- (ii) Índice Mês/Mês Anterior [ $I_{BF\_Com\_Ajuste\_Sazonal}$ ]: responsável por apresentar comparação da produção do mês de referência do índice com a do mês imediatamente anterior. As séries são obtidas a partir do índice de base fixa mensal ajustado sazonalmente e são divulgadas somente para a indústria geral;
- (iii) Índice Mensal [ $I_{Mensual}$ ]: responsável por apresentar comparação da produção do mês de referência do índice em relação a igual mês do ano anterior;
- (iv) Índice Acumulado No Ano [ $I_{acum\_no\_ano}$ ]: responsável por apresentar comparação a produção acumulada no ano, de janeiro até o mês de referência do índice, em relação a igual período do ano anterior; e
- (v) Índice Acumulado Nos Últimos 12 Meses [ $I_{acum\_ult\_12\_meses}$ ]: responsável por apresentar comparação a produção acumulada nos últimos 12 meses de referência do índice em relação a igual período imediatamente anterior.

A seguir são apresentados macro números de desempenho industrial da Indústria no Espírito Santo em 2017 e 2018:

- Redução da Produção Física da Indústria Capixaba de 1,1% na comparação fev/18 x jan/18, enquanto o Brasil teve variação positiva de 0,2% nessa mesma base de comparação – série livre de efeitos sazonais;
- Na comparação interanual (fev/18 x fev/17) a indústria capixaba recuou 6,3% com o setor de alimentos sendo o único a apresentar variação positiva (3,7%). E os principais influenciadores negativos foram: -30,0% setor de fabricação de celulose, papel, e produtos de papel; -11,9% setor de minerais não metálicos; -6,3% setor da metalurgia – todos estes vinculados a indústria de transformação, sendo registrado queda de -1,8 na indústria extrativa.
- No acumulado do primeiro bimestre de 2018, a indústria capixaba apresentou forte recuo de 7,8% e, novamente, o setor alimentício foi o único a registrar crescimento, com 1,5% para o período.

**Tabela 1** - Números da Indústria Geral Capixaba – Variação % em relação ao mês anterior

Indústria Geral	mai/17	jun/17	jul/17	ago/17	set/17	out/17	nov/17	dez/17	jan/18	fev/18
Faturamento Real	8,8	15,2	-22,5	10,0	23,8	-8,2	-	2,9	-19,8	-3,0
Horas Trabalhadas na Produção	9,6	-2,9	1,7	2,1	-5,4	1,4	-	-5,5	-1,4	-1,6
Pessoal Empregado Total	0,9	-0,8	0,1	0,0	-0,1	-0,5	-	-3,0	-2,5	-0,5
Massa Salarial Real	6,5	-1,9	-1,1	3,5	-0,1	0,5	-	2,9	-23,4	1,0
Rendimento Médio Real	5,5	-1,2	-1,1	3,5	0,0	1,0	-	6,0	-21,5	1,5

Fonte: IDEIES, 2018.

- O faturamento real da indústria capixaba recuou 3,0% em fev/18 frente a jan/18, mas segue com tendência de melhoria (IDEIES, 2018). Os principais setores responsáveis pela queda observada em fevereiro são: -27,7% setor de impressão e reprodução de gravações; -26,4% setor de borracha e material plástico; -25,0% setor de metalurgia. Em contrapartida, registra-se que o setor de vestuário ampliou 129,8% em consequência ao maior número de pedidos no mercado interno provocado pelo lançamento de novas coleções. Se comparado com fev/17, o faturamento real cresceu 11,6% (fev/18 x fev/17).
- As horas trabalhadas na produção e o pessoal empregado tem apresentado queda nos últimos meses: -5,5% (dez/17 x nov/17) ; -1,4% (jan/18 x dez/17); e -1,6% (fev/18 x jan/18), quando analisado a variação mês de horas trabalhadas, e -3,0% (dez/17 x nov/17) ; -2,5% (jan/18 x dez/17); e -0,5% (fev/18 x jan/18), quando analisado a



variação do pessoal empregado total. O mesmo cenário de queda é verificado ao analisar a comparar interanual disposta na Tabela 2.

**Tabela 2 - Números da Indústria Geral Capixaba – Variação % interanual**

Indústria Geral	mai/17	jun/17	jul/17	ago/17	set/17	out/17	nov/17	dez/17	jan/18	fev/18
	x	x	x	x	x	x	x	x	x	x
	mai/16	jun/16	jul/16	ago/16	set/16	out/16	nov/16	dez/16	jan/17	fev/17
Faturamento Real	8,5	20,3	-7,6	4,3	8,4	26,6	-	8,7	25,4	11,6
Horas Trabalhadas na Produção	-4,0	-5,9	-2,7	-3,5	-2,2	-1,2	-	-6,1	-6,0	-3,6
Pessoal Empregado Total	-7,3	-7,3	-4,8	4,3	-2,5	-1,5	-	-2,4	-5,9	-6,0
Massa Salarial Real	-1,3	-2,1	-11,4	0,9	5,9	-1,2	-	-3,6	-18,0	-4,6
Rendimento Médio Real	6,4	5,6	-6,9	5,4	8,6	0,3	-	-1,3	-12,9	1,6

Fonte: IDEIES, 2018.

- A utilização de capacidade instalada também reflete em queda se comparado com dez/17 – Tabela 3.
- A massa salarial e o rendimento médio real iniciaram 2018 com forte queda, registrando -23,4% e -21,5 respectivamente, e em fevereiro, observou-se uma aceleração positiva frente janeiro, com 1,0% e 1,5% respectivamente, ocasionado principalmente pelo pagamento de rescisões (IDEIES, 2018).

**Tabela 3 - Utilização de Capacidade Instalada**

Mês	Utilização da Capacidade Instalada
jan/17	74,7%
fev/17	76,0%
mar/17	76,0%
abr/17	75,0%
mai/17	76,2%
jun/17	76,3%
jul/17	74,9%
ago/17	75,7%
set/17	76,7%
out/17	75,8%
nov/17	77,4%
dez/17	78,3%
jan/18	75,0%
fev/18	74,0%

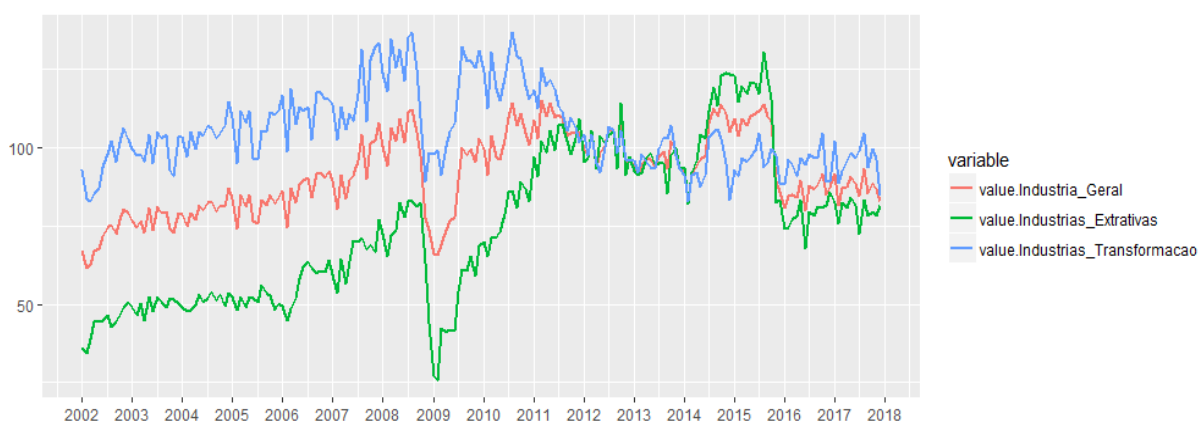
Fonte: IDEIES, 2018.

O Espírito Santo possui PIB industrial de R\$ 31,2 bilhões, representando uma força de 31,1% na economia capixaba, e equivalente a 2,7% da indústria nacional em 2015. A Figura 1 apresenta a participação percentual dos principais setores no PIB industrial (Portal da Indústria, 2018).



**Figura 1** - Principais Setores da Indústria Capixaba – Portal da Industria.

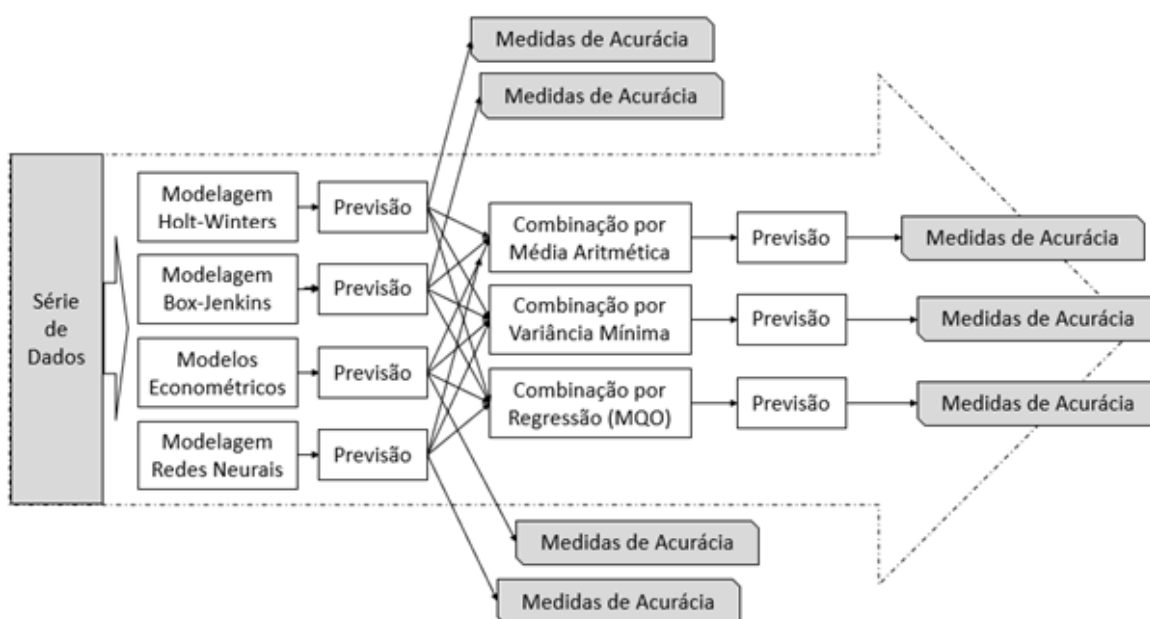
A partir da Figura 2, que traz o gráfico do índice mensal de produção industrial Geral, indústria extrativa e indústria de transformação, é facilmente identificado o efeito da Crise Mundial de 2008 com a queda brusca do índice em 2009. Bem como também é visível o efeito do Rompimento da Barragem da Samarco em Mariana / MG na economia capixaba, devido a redução no setor de minérios de ferro pelotizados ou sinterizados, a partir da parada da Usina em Ubu, em Anchieta / ES.



**Figura 2** - Produção Industrial Capixaba – Índice para Indústria Geral, Extrativa e da Transformação.

### 3 MATERIAIS E MÉTODOS

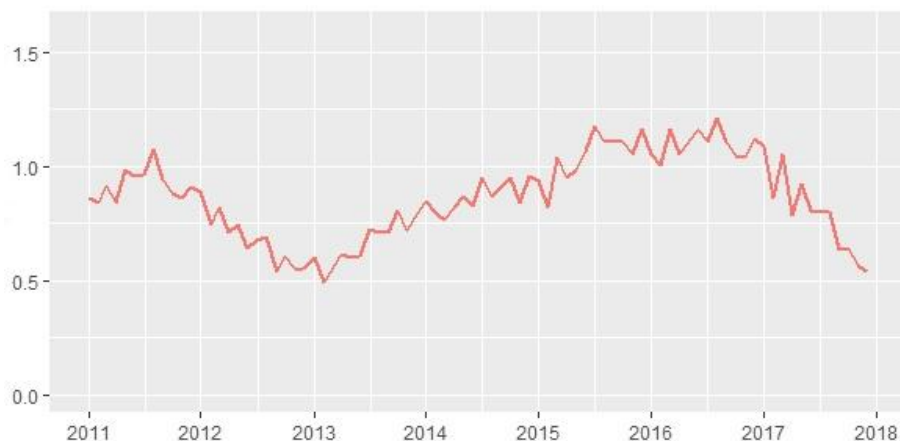
A Figura 3 ilustra o procedimento adotado para cada série de dados da Produção Industrial, onde serão computados quatro métodos de previsão – considerando o período até 2016, e três métodos de combinação de previsão, e a seleção de melhor modelo será realizada a partir do cálculo de medidas de acurácia 12 passos a frente (janeiro e dezembro de 2017).



**Figura 3** - Esboço das Etapas Previstas do Estudo

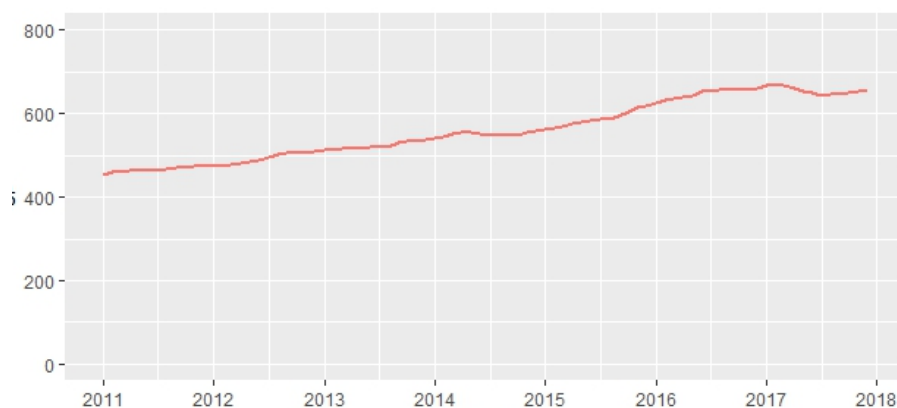
Ressalta-se a priori que na Modelagem Econométrica, se faz necessário a inserção de outras variáveis, conhecidas na literatura como regressoras, ou ainda explicativas. São elas:

- **Taxa de Juros:** Para aferição dessa variável, considerou-se a taxa básica de juros da economia brasileira, taxa SELIC - Sistema Especial de Liquidação e de Custódia, Figura 4. A série histórica foi coletada no site do Banco Central.



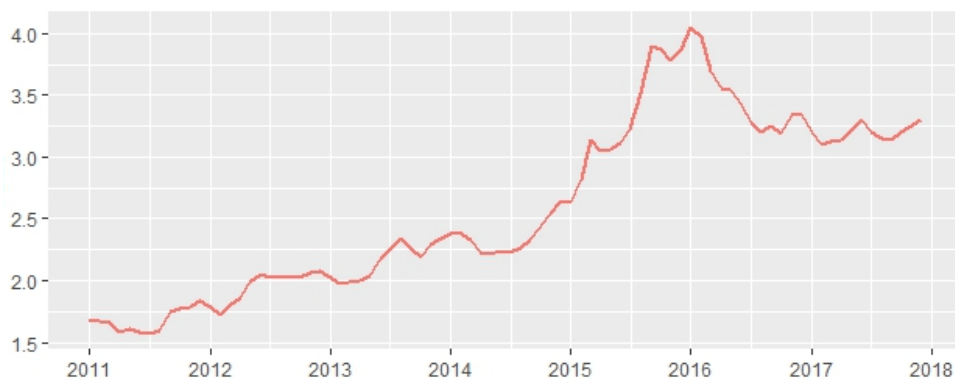
**Figura 4** - Taxa de Juros Selic (período 2011 a 2018)

- **Taxa de Inflação:** a partir dessa variável é possível mensurar as variações de preços de matérias-primas agropecuárias e industriais, de produtos intermediários e de bens e serviços finais, cuja nomenclatura adotada é IGP – índices gerais de preços. Essa medida é atualizada mensalmente pelo IBRE, Instituto Brasileiro de Economia, vinculado a Fundação Getúlio Vargas. Os dados coletados estão dispostos na Figura 5.



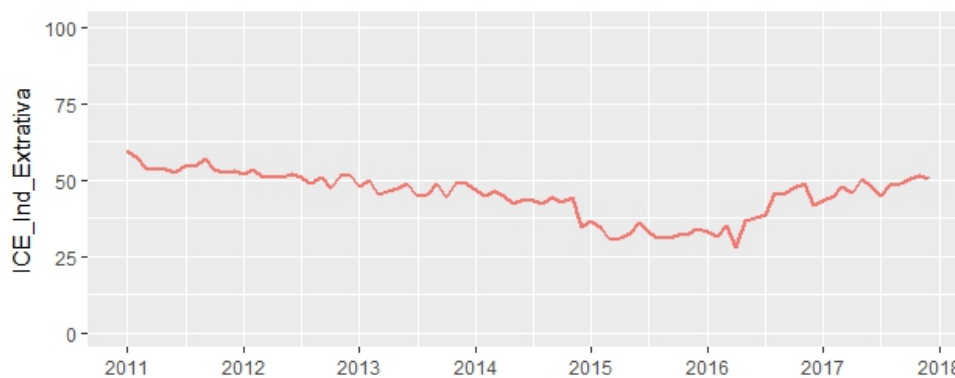
**Figura 5** - Taxa de Inflação IGP (período 2011 a 2018)

- **Taxa de Câmbio:** Reconhecida pelo BACEN como sendo o preço de uma moeda estrangeira medido em unidades ou frações (centavos) da moeda nacional. Os dados disponíveis no site do Banco Central considerando a cotação do Dólar dos Estados Unidos (Figura 6).

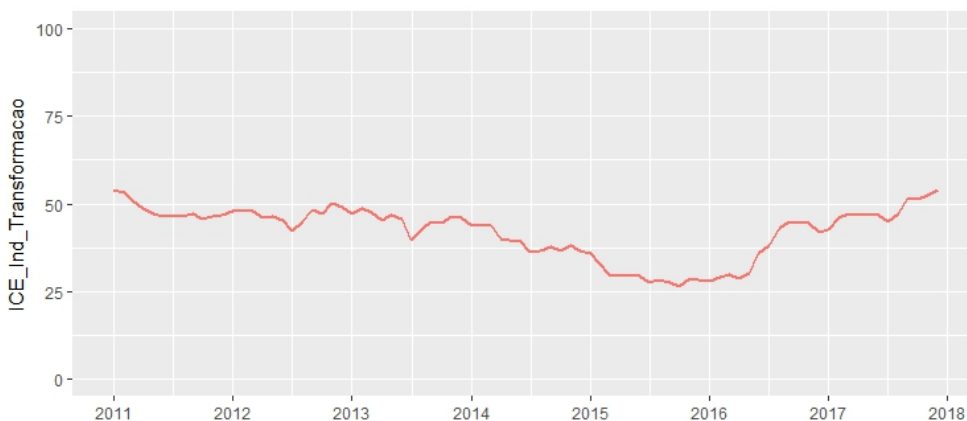


**Figura 6** - Taxa de Cambio (período 2011 a 2018)

- **Índice de Confiança do Empresário Industrial:** Também disponibilizado pelo IPEA, esse indicador foi coletado separando a expectativa do empresário na indústria extrativa e indústria de transformação – Figuras 7 e 8 respectivamente.

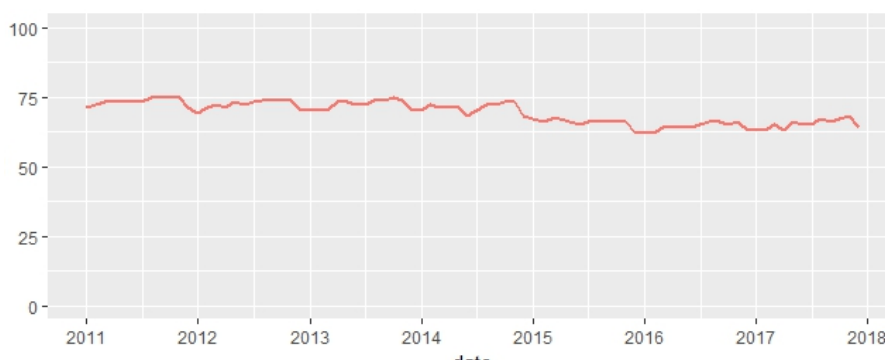


**Figura 7** - Índice de Confiança do Empresário Indústria – Setor Extrativo (período 2011 a 2018)



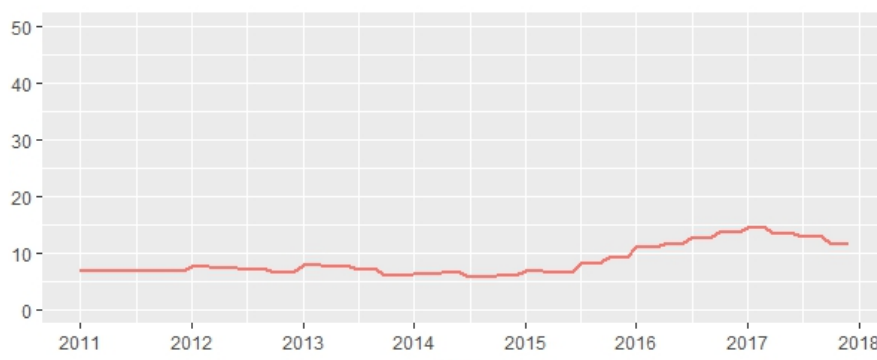
**Figura 8** - Índice de Confiança do Empresário Industrial – Setor de Transformação (período 2011 a 2018)

- **Utilização da Capacidade Instalada:** Essa variável é importante na indústria pois a partir dela pode-se aferir o percentual de ociosidade da capacidade. O IPEA disponibiliza esses dados segmentados por indústria extrativa e indústria de transformação, todavia apenas o índice da indústria geral foi relevante na modelagem – Figura 9.



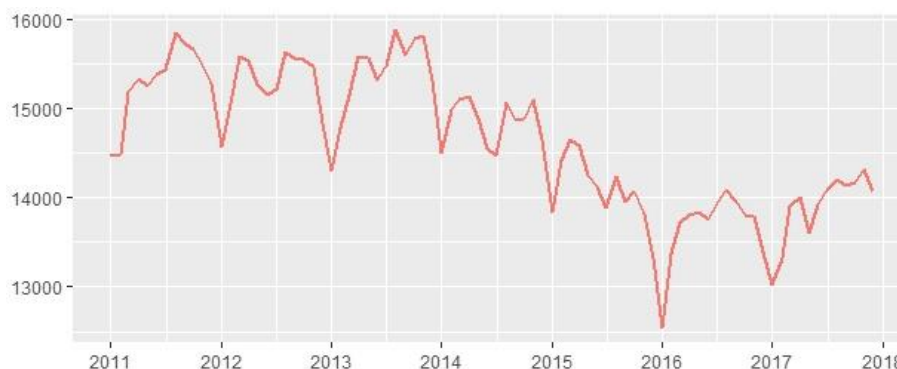
**Figura 9** - Utilização da Capacidade Instalada (período 2011 a 2018)

- **Percentual de Desocupação:** Variável calculada a partir da razão entre a quantidade de pessoas desocupadas sobre a quantidade de pessoas em força de trabalho na indústria. Ambas variáveis são fornecidas pelo IBGE a partir da pesquisa mensal de emprego e salário – Figura 10.



**Figura 10** - Percentual de Desocupação (período 2011 a 2018)

- **Consumo de Energia no Setor Industrial:** Variável medida em GWh, apurado em nível nacional, e divulgada pela Eletrobrás – Figura 11.



**Figura 11** – Consumo de Energia (período 2011 a 2018)

- **Paralização Samarco:** Variável binária (dummy), para indicar o período de paralização da Empresa Samarco, sendo atribuídos os seguintes valores:

$$Paraliza\c{a}o\ Samarco = \begin{cases} 0, & \text{se no respectivo m\^{e}s a Samarco estava em opera\c{c}o} \\ 1, & \text{se no respectivo m\^{e}s a Samarco estava fora opera\c{c}o} \end{cases}$$

Ou seja, no período pós novembro de 2015, a variável recebeu o valor 1, inclusive no período de previsão, no qual não se apresentou nenhuma perspectiva de retorno das atividades desta empresa.

Autores como Silva, Carmo e França (2002), Matos (2008), Hollauer, Issler e Notini (2008), e Prates (2010) desenvolveram estudos na área de produção industrial nos quais obtiveram bons resultados ao considerar variáveis como acima especificadas. Adicionalmente, também apresentou relevância a inserção de uma variável dummy para inserir a informação de paralização da empresa Samarco a partir de novembro de 2015, pois conforme já discutido, afetou a indústria capixaba (IBGE, 2015/2016).

Na revisão bibliográfica, identificaram-se outras variáveis que apresentaram bons resultados na modelagem econométrica de produção industrial, tais como volume de exportações e importações, e horas pagas pela indústria.

### 3.1 MODELOS BOX-JENKINS

Segundo Box & Jenkins (1976), uma série temporal é qualquer conjunto de observações ordenadas no tempo, podendo estas serem contínuas ou discretas. E uma série temporal é uma observação de um processo estocástico (MORETTIN E TOLOI, 2006), e reserva uma área de grande importância na literatura estatística cujo principal interesse é obter previsões baseadas em informações passadas coletadas ao longo do tempo.

Se o processo estocástico que gerou a série de observações é invariante com respeito ao tempo, ou seja, as características de  $Y_{t+\tau}$  são as mesmas de  $Y_t$  para todo  $\tau$ , diz-se que o mesmo é estacionário. Um processo estacionário pode ser classificado em:

- Estritamente Estacionário, quando as séries  $Y_t$  e  $Y_{t+\tau}$  são distribuídas identicamente para qualquer que seja  $\tau$ . Ou seja, todas as distribuições permanecem as mesmas sob translações do tempo (BOX & JENKINS, 1976).
- Fracamente Estacionário (ou simplesmente estacionário), se possui média e variância constante ao longo do tempo, e a covariância só depende do intervalo de tempo. Ou seja, que atenda aos 3 critérios abaixo (BOX & JENKINS, 1976):

1.  $E(Y_t) = \mu_t = \mu$ , constante para todo  $t \in T$ .

2.  $E(Y_t^2) < \infty$ , para todo  $t \in T$

3.  $\gamma(t_1, t_2) = \text{Cov}(Y_{t_1}, Y_{t_2})$  é uma função de  $|t_1 - t_2|$

Na prática, aceita-se que as séries observadas sejam séries fracamente estacionárias, situações nas quais garante-se apenas médias e variâncias invariantes no tempo.

Nesse sentido, a literatura dispõe de alguns testes para verificação de estacionariedade, conhecidos popularmente como testes de raízes unitárias.: Dickey-Fuller Aumentado e Phillips-Perron.

Esses testes baseiam-se na verificação da existência de alguma raiz dos operadores de retardos dentro do círculo unitário, denominada simplesmente por raiz unitária, e define-se as hipóteses:

$H_0$ : a série não é estacionária, ou seja, possui raiz unitária.



$H_1$ : a série é estacionária.

O teste de Dickey-Fuller (DF), inicialmente estudado por Fuller (1976) e Dickey & Fuller (1979), consiste em considerar o seguinte processo

$$\Delta Y_t = \theta Y_{t-1} + \varepsilon_t \quad \text{Equação 3.1.1}$$

Com  $\varepsilon_t \sim N(0, \sigma^2)$  independente e identicamente distribuído.

Se  $|\theta| < 1$  então tem-se um processo estacionário.

Se  $|\theta| = 1$  então o processo apresentará raiz unitária e, portanto, será não estacionário.

Assim sendo, o teste de raiz unitária testa se  $\theta = 1$  ou não. Em geral, segue-se considerando

$$Y_t - Y_{t-1} = \theta Y_{t-1} - Y_{t-1} + \varepsilon_t$$

$$\Delta Y_t = (\theta - 1)Y_{t-1} + \varepsilon_t$$

$$\Delta Y_t = \pi Y_{t-1} + \varepsilon_t \quad \text{Equação 3.1.2}$$

E portanto, deseja-se testar  $H_0: \pi = 0$ , onde se  $\pi = 0$  então  $\theta = 1$  e, conseqüentemente,  $Y_t$  possui raiz unitária. Note ainda que  $H_1: \pi < 0$ , logo, trata-se de um teste unilateral.

O teste DF assume que os erros são não correlacionados, ou seja, que a covariância entre  $\varepsilon_t$  e  $\varepsilon_{t-k}$  para todo  $k > 0$  seja nula. O que implica em garantir que os resíduos não apresentam autocorrelação.

O teste de Dickey-Fuller Aumentado, conhecido na literatura como teste ADF (Augmented Dickey-Fuller), consiste em incluir um número maior de defasagens das diferenças, o que implica em estimar a seguinte regressão:

$$\Delta Y_t = \alpha + \beta_t + \pi Y_{t-1} + \sum_{i=1}^m \gamma \Delta Y_{t-i} + \varepsilon_t \quad \text{Equação 3.1.3}$$

Onde  $\varepsilon_t$  é um ruído branco puro,  $\alpha$  é o intercepto,  $\beta_2$  é o coeficiente de tendência,  $\pi$  é o coeficiente de presença de raiz unitária e  $m$  é o número de defasagens tomadas na série.

Sob as seguintes hipóteses:  $H_0: \pi = 0$ , e  $H_1: \pi \neq 0$ .

A limitação neste teste é a determinação do número de defasagens utilizado no teste, pois considerando que o poder do teste é baixo, queremos utilizar o menor número possível de defasagens que garanta a não autocorrelação dos resíduos.

O Teste Phillips-Perron, conhecido na literatura como teste PP é uma generalização do teste de Dickley - Fuller para os casos em que os erros  $\varepsilon_t$  são correlacionados, ou ainda, heterocedásticos (PHILLIPS & PERRON, 1988).

O teste baseia-se na mesma equação do teste ADF, porém sem a adição de defasagens. Sendo considerada uma abordagem não paramétrica.

E é baseado na seguinte equação:

$$Y_t = \beta_0 + \pi Y_{t-1} + \varepsilon_t \quad \text{Equação 3.1.4}$$

Sob a hipótese  $H_0 = \theta = 1$ , contra  $H_1: \theta < 1$ . Com  $\pi = \theta - 1$ .

Ambos testes são facilmente computados via Software de Modelagem Estatística. No RStudio, os testes ADF e PP são apurados a partir das funções `adf.test` e `pp.test`, respectivamente.

Box & Jenkins (1976) definiram uma metodologia de modelos paramétricos que consiste em ajustar modelos auto-regressivos integrados de médias móveis ARIMA( $p, d, q$ ) a um conjunto de dados.

Morettin e Toloí (2006), apresentam uma estratégia de construção do modelo baseada em um ciclo iterativo:

- (a) Identificação: etapa pela qual verifica-se qual modelo descreve o comportamento da série temporal.
- i. Tratar sazonalidade;
  - ii. Determinar ordem de integração; e
  - iii. Encontrar os valores apropriados de  $p$  e  $q$ .
- (b) Estimação: etapa pela qual estima-se os parâmetros do componente autoregressivo, do componente de médias móveis e da variância.
- (c) Diagnóstico: consiste em verificar se o modelo descreve adequadamente o comportamento da série, por meio de análise de resíduos.
- (d) Previsão: etapa final que representa o objetivo principal, e realizada somente quando as etapas anteriores são satisfatórias.

Deste modo, caso o modelo não seja adequado, o ciclo é repetido, voltando-se à fase de identificação. Essas etapas devem ser repetidas, tantas vezes quanto forem necessárias até que seja obtido um modelo satisfatório que descreva o processo gerador da série. Comumente, estima-se mais de um modelo a partir da análise de identificação, e então, após a etapa de estimação e diagnóstico, seleciona-se entre os modelos ajustados o que apresentar menor erro (Morettin e Toloi, 2006).

Os modelos autoregressivos e de médias móveis consideram em sua especificação o grau de autocorrelação entre os valores da série temporal. A modelagem proposta segue conforme especificação a seguir:

$$Y_t = \mu + \sum_{k=0}^{\infty} \Psi_k u_{t-k} = \mu + \Psi(B)u_t \quad \text{Equação 3.1.5}$$

Sendo o filtro linear  $\Psi(B)$  definido por

$$\Psi(B) = \frac{\theta(B)}{\phi(B)} \quad \text{Equação 3.1.6}$$

Onde  $\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q$  e  $\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$  são polinômios de graus  $q$  e  $p$ , respectivamente.

Ao considerar  $\tilde{Y} = Y_t - \mu$ , tem-se  $\tilde{Y} = \Psi(B)u_t$ , e conseqüentemente,

$$\phi(B)\tilde{Y} = \theta(B)u_t \quad \text{Equação 3.1.7}$$

Onde  $u_t$  é considerado um ruído branco, com  $E(u_t) = \mathbf{0}$  e  $Var(u_t) = \sigma_u^2$ . Segundo Box & Jenkins, o modelo descrito a partir da equação 2.1.3 é denominado ARMA( $p, q$ ), onde assume-se a premissa de  $\tilde{Y} = Y_t$ , com  $\mu = \mathbf{0}$ . Assim sendo, algebricamente é possível verificar a formulação abaixo (Box e Jenkins, 2008):

$$\begin{aligned} (1 - \phi_1 B - \dots - \phi_p B^p)Y_t &= (1 - \theta_1 B - \dots - \theta_q B^q)u_t \\ Y_t - \phi_1 Y_{t-1} - \dots - \phi_p Y_{t-p} &= u_t - \theta_1 u_{t-1} - \dots - \theta_q u_{t-q} \\ Y_t &= \phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p} - \theta_1 u_{t-1} - \dots - \theta_q u_{t-q} + u_t \end{aligned} \quad \text{Equação 3.1.8}$$

Assim, o valor de  $Y$  no período  $t$ , depende do seu valor  $p$  períodos anteriores – processo autoregressivo AR( $p$ ) – e além do termo de erro - ruído branco -, considera uma média móvel dos termos de erro corrente e passado (Gujarati, 2006).

Em geral, deseja utilizar modelos mais parcimoniosos, alcançada numa representação com menor número de parâmetros necessários. Nesse sentido, considera-se o modelo com  $\min(p, q)$ . Etapa esta do ciclo apresentado por Moretim (2006) identificada por fase de identificação, e em que consiste ainda na verificação menores valores obtidos para os critérios AIC e BIC (do inglês, Akaike Information Criterion e Bayesian Information Criterion, respectivamente).

Modelos ARMA ajustam-se melhor a séries de dados cujo nível médio e variância são constantes (séries estacionárias). Todavia, várias séries econômicas/financeiras não possuem tal característica (Morettin e Toloí, 2006), e quando estes pressupostos não são atendidos, é preciso empregar o operador de diferenciação  $\nabla^d$  para remover as mudanças de nível e tornar a série estacionária (Box e Jenkins, 2008).

Por exemplo,  $\nabla^1 Y_t = Y_t - Y_{t-1}$ , ou, analogamente,  $\nabla^1 Y_t = (1 - B)Y_t$ . De modo geral, tem-se que um operador  $\nabla^d$ , de ordem  $d$ , calcula a diferença  $\nabla^d Y_t = \nabla \nabla \dots \nabla Y_t = (1 - B)^d Y_t$ .

Se  $Z_t = \Delta^d Y_t$  for estacionária, então  $Z_t$  pode ser representada por um modelo ARMA( $p, q$ ):

$$Z_t = Y_t - Y_{t-1} = (1 - B)Y_t = \Delta Y_t$$

$$Z_t = \phi_1 Z_{t-1} + \dots + \phi_p Z_{t-p} - \theta_1 u_{t-1} - \dots - \theta_q u_{t-q} + u_t \quad \text{Equação 3.1.9}$$

Neste caso, diz-se que  $Y_t$  segue uma modelo auto-regressivo, integrado, de médias móveis de ordem  $p, d, q$  – ARIMA( $p, d, q$ ):

$$\phi(B)\Delta^d Y_t = \theta(B)u_t \quad \text{Equação 3.1.10}$$

Modelos ARIMA podem ser adaptados para descrever a autocorrelação em séries sazonais, modelos estes caracterizados por tratar a sazonalidade estocástica dos dados, e conhecidos na literatura como SARIMA ou ainda ARIMA sazonal (Brockwell e Davis, 2002).

A seguir, apresentados, sua especificação para o período  $s = 12$  (Morettin e Toloi, 2004).

O modelo denominado SARIMA de ordem  $(p, d, q) \times (P, D, Q)_{12}$  é dado por:

$$\phi(B)\Phi(B^{12})(1 - B^{12})^D(1 - B)^d Y_t = \theta(B)\Theta(B^{12})u_t \quad \text{Equação 3.1.11}$$

Onde  $P$  é a ordem do polinômio autoregressivo sazonal SAR( $P$ ), dado por  $\Phi(B^{12})$  - sendo,  $\Phi(B^{12}) = 1 - \Phi_1 B^{12} - \dots - \Phi_P B^{12P}$  -, e  $D$  é o número de diferenciações,  $Q$  é a ordem do polinômio de médias móveis sazonais SMA( $Q$ ) dados por  $\Theta(B^{12})$  - sendo  $\Theta(B^{12}) = 1 - \Theta_1 B^{12} - \dots - \Theta_Q B^{12Q}$  -, e  $s=12$  é o tamanho do ciclo sazonal.

Na literatura, complementarmente aos autores acima citados, Wei (1990) e Chatfield (1989), dentre outros, apresentam uma discussão detalhada do método, e aplicações.

### 3.2 MODELO HOLT-WINTERS (HW)

O método de *Holt-Winters* foi sugerido por Holt (1957) e Winters (1960), e sua principal característica do é permitir a adaptação da sazonalidade ao longo do tempo, a partir de médias

móveis exponenciais, o qual propicia melhores previsões do que médias móveis simples (Lawton, 1998), podendo ser classificados como aditivo ou multiplicativo,

Na literatura, Chatfield (1978), Chatfield e Yar (1988), Gelper et all (2010), dentre outros, apresentam uma discussão detalhada do método, e aplicações.

No modelo aditivo, a amplitude da variação sazonal é constante ao longo do tempo. Em outras palavras, diz-se que a diferença entre o maior e menor valor de demanda dentro dos ciclos permanece relativamente constante no tempo. Podemos escrever o modelo da seguinte maneira:

$$Y_t = L_t + T_t + S_t + \epsilon_t \quad \text{Equação 3.2.1}$$

Onde,  $Y_t$  representa o valor observado no tempo  $t$ ,  $L_t$  a componente de nível,  $T_t$  a componente de tendência,  $S_t$  a componente de sazonalidade,  $\epsilon_t$  é o erro aleatório ocorrido no tempo  $t$  onde  $E(\epsilon_t) = 0$  e  $VAR(\epsilon_t) = \sigma^2$ .

A função de previsão do modelo aditivo de *Holt-Winters* para  $i$  passos a frente é dada por:

$$\hat{Y}_{t+n} = L_t + nT_t + S_{t-s+i} \quad \text{Equação 3.2.2}$$

Para  $i = 1, 2, \dots, n$ .

As fórmulas de atualização para as três componentes consideram três constantes de alisamento:  $\alpha$ ,  $\beta$ , e  $\gamma$ , onde a constante  $\alpha$  é utilizada como parâmetros para nível,  $\beta$  para inclinação, e  $\gamma$  para o fator sazonal. E as equações de atualização são como dadas a seguir:

- Nível ( $L_t$ )

$$\alpha(Y_t - S_{t-s}) + (1 - \alpha)(L_{t-1} + T_{t-1}) \quad \text{Equação 3.2.3}$$

- Tendência ( $T_t$ )

$$\alpha(Y_t - S_{t-s}) + (1 - \alpha)(L_{t-1} + T_{t-1}) \quad \text{Equação 3.2.4}$$

- Sazonalidade ( $S_t$ )

$$\gamma(Y_t - T_t) + (1 - \gamma)S_{t-l} \quad \text{Equação 3.2.5}$$

A determinação de tais constante se dá a partir da condição de tornar mínima a soma dos quadrados dos erros de ajustamento.

No modelo *Holt-Winters* multiplicativo é obtido multiplicando a amplitude da variação sazonal aumenta ou diminui como função do tempo, ou seja, a diferença entre o maior e o menor ponto nos ciclos cresce com o passar do tempo. Com efeito, a componente de tendência continua tendo uma formulação aditiva. Este modelo pode ser escrito da seguinte maneira:

$$Y_t = L_t S_t + T_t + \epsilon_t \quad \text{Equação 3.2.6}$$

E a função de previsão do modelo multiplicativo de *Holt-Winters* para  $i$  passos a frente é dada por:

$$\hat{Y}_{t+n} = (L_t + nT_t)S_{t-s+i} \quad \text{Equação 3.2.7}$$

Para  $i = 1, 2, \dots, n$ . A seguir, apresenta-se as equações de atualização:

- Nível ( $L_t$ )

$$\alpha \frac{Y_t}{S_{t-s}} + (1 - \alpha)(L_{t-1} + T_{t-1}) \quad \text{Equação 3.2.8}$$

- Tendência ( $T_t$ )

$$\beta(L_t - L_{t-1}) + (1 - \beta)T_{t-1} \quad \text{Equação 3.2.9}$$

- Sazonalidade ( $S_t$ )

$$\gamma \frac{Y_t}{L_t} + (1 - \gamma)S_{t-s} \quad \text{Equação 3.2.10}$$

Morettin e Tolo (2004) descrevem as vantagens desse modelo como sendo:

- (i) facilidade de entendimento;
- (ii) aplicação não dispendiosa;

- (iii) é considerada adequada para estudo de série com padrão de comportamento geral.

Em contrapartida, sua principal desvantagem é a dificuldade em determinar os valores mais apropriados das constantes de suavização e/ou impossibilidade de estudar as propriedades estatísticas, tais como média e variância da previsão e, conseqüentemente, a construção de um intervalo de confiança.

### 3.3 METODOLOGIA ECONOMÉTRICA TRADIONAL

Segundo Gerhard (1968), econometria é a aplicação da estatística matemática aplicada aos dados econômicos para dar apoio empírico aos modelos formulados pela economia matemática e obter resultados numéricos. Genericamente, diz-se que a econometria mede a relação entre duas ou mais variáveis de forma quantitativa, tal como os modelos de regressão linear múltipla - Equação 3.3.1 - comumente utilizados nas mais diversas áreas de conhecimento (KMENTA, 1997). Em outras palavras, diz-se que a partir desse método é possível verificar a relação entre as variáveis explicativas de um determinado evento.

Seja  $X_1, X_2, \dots, X_k$  variáveis independentes, e  $Y$  a variável dependente, ou seja, aquela que se deseja explicar. Considere a amostra de  $n$  observações  $(X_{1i}, X_{2i}, \dots, X_{ki}, Y_i)$ ,  $i = 1, 2, \dots, n$ . O modelo de regressão linear múltiplo será dado por:

$$E(Y_i | X_{1i}, X_{2i}, \dots, X_{ki}) = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_k X_{ki} \quad \text{Equação 3.3.1}$$

Ou ainda

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_k X_{ki} + \varepsilon_i \quad \text{Equação 3.3.2}$$

Sendo  $n > (k + 1)$ .  $\beta_0$  é o intercepto da equação,  $\beta_j$  são os coeficientes (pesos) associado ao  $j$ -ésima variável,  $j = 1, \dots, k$ . O termo  $\varepsilon_i$  é chamado de perturbação aleatória, onde  $\varepsilon_i \sim N(0, \sigma^2)$ , ou seja, possui distribuição normal, com média zero e variância  $\sigma^2$ , são não correlacionados.



A estimação dos coeficientes  $\beta_j$ , pode ser obtida a partir do método de mínimos quadrados, minimizando o erro quadrático total ( $\sum \varepsilon_i^2$ ). Antes disso, é importante apresentar a notação matricial:

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, X = \begin{bmatrix} 1 & x_{11} & x_{12} & \vdots & x_{1k} \\ 1 & x_{21} & x_{22} & \vdots & x_{2k} \\ \dots & \dots & \dots & \vdots & \dots \\ 1 & x_{n1} & x_{n2} & \vdots & x_{nk} \end{bmatrix} \text{ e } \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} \quad \text{Equação 3.3.3}$$

E assim, pode-se escrever:

$$Y = X\beta + \varepsilon \quad \text{Equação 3.3.4}$$

Sendo  $Y$  um vetor ( $n \times 1$ ),  $X$  uma matriz ( $n \times (k + 1)$ ),  $\beta$  um vetor ( $k + 1$ ), e  $\varepsilon$  um vetor ( $n \times 1$ ).

Então, retomando a estimação dos  $k$  coeficientes  $\beta$ , deve-se minimizar a expressão  $(\sum \varepsilon_i^2) = \sum_{i=1}^n [y_i - (\beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_k X_{ik})]^2 = (y - X\beta)'(y - X\beta)$ , derivando-se simultaneamente em termos de  $\beta$ , tem-se:

$$\frac{\partial y'y}{\partial \beta} - 2 \frac{\partial y'X\beta}{\partial \beta} + \frac{\partial \beta'X'X\beta}{\partial \beta} = 0 - 2(y'X)' + 2X'X\beta \quad \text{Equação 3.3.5}$$

No qual o vetor de dimensão  $p + 1$  resulta na expressão  $(X'X)^{-1}X'y = \hat{\beta}$ , em que  $\hat{\beta}$  é estimador não viciado do modelo, desde que  $(X'X)^{-1}$  exista.

Desse modo, o modelo de regressão ajustado e o vetor de resíduos são dados, respectivamente por:

$$\hat{y} = X\hat{\beta} \text{ e } \varepsilon = y - \hat{y} = y - X\hat{\beta} \quad \text{Equação 3.3.6}$$

E o estimador de  $\sigma^2$  é dado por:

$$\frac{y'y - \hat{\beta}'X'y}{n - k - 1} \quad \text{Equação 3.3.7}$$

E a matriz de covariância de  $\hat{\beta}$  é dada por

$$\text{cov}(\hat{\beta}) = \sigma^2(X'X)^{-1} \quad \text{Equação 3.3.8}$$

Para análise de resíduos, tem-se

$$\text{e } \text{cov}(\varepsilon) = \sigma^2[I - X(X'X)^{-1}X'] \quad \text{Equação 3.3.9}$$

A análise de resíduos é importante para validação de um ajuste de regressão, pois mostra a significância do modelo e avalia as contribuições das variáveis regressoras. Complementarmente deve-se analisar a variância do modelo (ANOVA), testando as hipóteses:

$$H_0: \beta_0 = \beta_1 = \beta_2 = \dots = \beta_k = 0$$

$$H_1: \text{ao menos um } \beta \neq 0$$

Logo, se  $H_0$  for verdadeira, tem-se que o modelo não está bem ajustado, pois os coeficientes são estatisticamente iguais a zero.

Outra etapa importante, é a verificação do coeficiente de determinação (Equação 3.3.9) e o coeficiente de determinação ajustado (Equação 3.3.10).

$$R^2 = \frac{SQReg}{SQT} = \frac{y'Hy - n\bar{y}^2}{y'y - n\bar{y}^2} \quad \text{Equação 3.3.9}$$

$$R^2_{ajustado} = 1 - \frac{\frac{SQRes}{(n-k-1)}}{\frac{SQT}{(n-1)}} = 1 - \frac{\frac{y'y - \hat{\beta}'X'y}{(n-k-1)}}{\frac{y'y - n\bar{y}^2}{(n-1)}} \quad \text{Equação 3.3.10}$$

Ambos coeficientes são facilmente computados em softwares estatísticos, tal como o R-Studio, e representam a porcentagem da variação da variável dependente explicada pelas variáveis independentes.

Adicionalmente, nessa etapa do estudo, será aplicado o teste de Dubin-Watson para verificação de existência de autocorrelação de primeira ordem nos resíduos (KMENTA, 1997), em outras palavras, é significa testar a independência dos resíduos.

É comum na metodologia de Modelagem Econométrica avaliar os coeficientes da regressão com as variáveis padronizadas, conhecidos na literatura como coeficientes beta (GUJARATI, 2006). Para calcular estes coeficientes, basta resgatar a Equação 3.3.2, e redefinir as variáveis em termos de unidades de desvio. Ou seja

$$\frac{Y_i}{S_Y} = \beta_0 + \beta_1 \frac{X_{1i}}{S_{X_1}} + \beta_2 \frac{X_{2i}}{S_{X_2}} + \dots + \beta_k \frac{X_{ki}}{S_{X_k}} + \varepsilon_i \quad \text{Equação 3.3.11}$$

Onde  $S_Y, S_{X_1}, S_{X_2}, \dots, S_{X_k}$  são os desvios das variáveis dependente e independentes, respectivamente.

Na prática, o coeficiente beta atua como medida da importância de cada variável explicativa, devido a padronização, eliminando o problema das unidades diferentes.

Sua interpretação se dá em módulo, ou seja, quanto maior o valor absoluto desse coeficiente, maior a importância da variável na regressão.

### 3.4 REDES NEURAIIS ARTIFICIAIS

Modelos de Redes Neurais Artificiais (RNA), muitas vezes chamados apenas de rede neural, é um modelo matemático inspirado pelas redes neurais biológicas, são capazes de lidar melhor com características de dados como não linearidade, melhor do que os modelos habituais (FLORES, 2009), considerado robusto. A partir dessa metodologia não paramétrica, apresenta facilidade para analisar sistemas dinâmicos como séries temporais não estacionárias (MULLER, 1995).

As redes neurais são usadas para modelar relações complexas entre entradas e saídas ou para encontrar padrões em dados. Uma rede neural consiste em um grupo de neurônios artificiais interligados e processa a informação através de uma abordagem conexionista à computação. Na maioria dos casos, uma rede neural é um sistema adaptável, que muda a sua estrutura durante uma fase de aprendizagem.

Nesse sentido, não é necessário que o tipo de relação entre variáveis dependentes e variáveis resposta seja, por exemplo, uma combinação linear, o que faz das redes neurais artificiais uma valiosa ferramenta quantitativa.

De todo modo, dados observados podem ser usados para treinar a rede neural, fazendo assim com que a rede neural "aprenda" uma aproximação da relação entre as variáveis independentes e dependentes de forma iterativa por meio da adaptação contínua dos seus parâmetros. Nesse trabalho, considerou-se um processo auto-regressivo de ordem 4 como input para treinar o modelo.

McCulloch e Pitts foram os pioneiros ao apresentar em 1943 um modelo inicial de RNA. Essa metodologia possui uma estrutura que pode ser descrita em 4 grupos: neurônios (nós), tipo de modelo de rede neural, as redes e o aprendizado. O neurônio é o principal elemento em uma RNA (MULLER, 1995), pois é responsável por executar todos os passos como entrada e saída de dados, e possui os seguintes elementos (Figura 11):

- Conexão: são os elos de ligação entre os neurônios, conhecidos na literatura como sinapses e vinculados a um peso  $w_{kj}$  (peso sináptico). Um sinal  $x_j$  é disparado na sinapse  $j$  que está ligada ao neurônio  $k$ .
- Junção aditiva: responsável por combinar os sinais  $x_j$  com seus respectivos pesos sinápticos  $w_{kj}$ .
- Função de ativação: seu papel é restringir ou limitar o intervalo permitido pela amplitude do sinal de saída.
- Viés: sua função é regular a entrada da função de ativação, aumentando ou diminuindo seu valor.

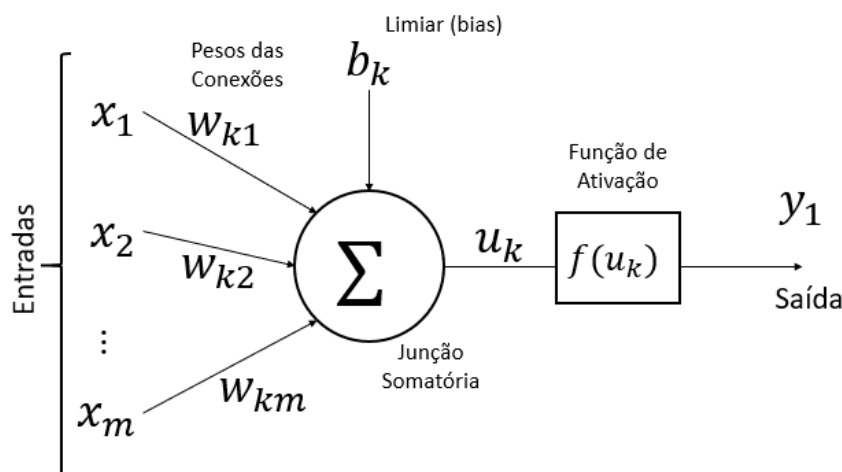


Figura 12 – Modelo Matemático de um Neurônio Artificial

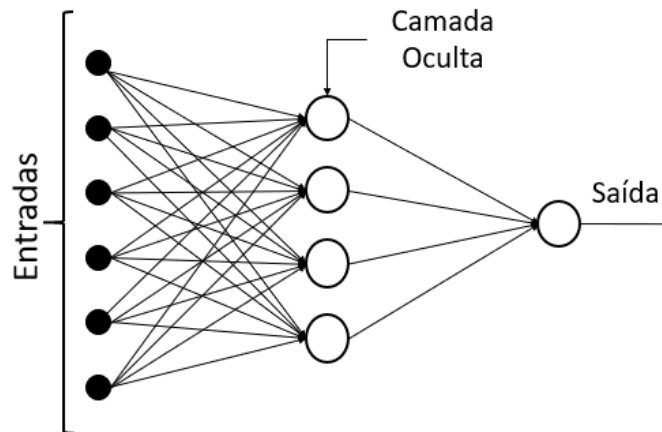
De forma geral, um neurônio pode ser definido a partir das equações 3.4.1 e 3.4.2.

$$y_k = f(u_k) = f\left(\sum_{j=1}^m w_{ij}x_j + b_k\right) \quad \text{Equação 3.4.1}$$

$$y_k = f(u_k) = f\left(\sum_{j=1}^m w_{ij}x_j\right) = f(w'X) \quad \text{Equação 3.4.2}$$

Nesse estudo, serão considerados os modelo RNA de múltiplas camadas, caracterizado por apresentar uma ou mais camadas ocultas de neurônios que não são visíveis como as camadas de entrada e saída (QI e ZHANG, 2001), e sua grande vantagem é sua capacidade de lidar com a complexidade dos sistemas, a partir da camada oculta e camada de saída.

Redes neurais multicamadas são arquiteturas em que neurônios são organizados em duas ou mais camadas (Figura 12), onde as entradas são conectadas aos elementos processadores básicos, que são interconectados com elementos de outras camadas e/ou saída de rede.



**Figura 13** - Redes Multicamadas

Uma fraqueza do método de redes neurais artificiais é a falta de procedimentos estabelecidos na literatura que permita realizar testes de significância estatística para os vários parâmetros que foram estimados.

### 3.5 TÉCNICAS DE COMBINAÇÃO DE PREVISÃO

Bates e Granger (1969) foram pioneiros na sistematização de um método de combinar os resultados dos modelos de previsão, com a finalidade de agregar as características mais relevantes de cada um dos modelos utilizados em uma determinada situação.

Segundo Costantini e Pappalardo (2010) a partir da combinação de previsão é possível melhorar a acurácia das previsões, pois consiste em um mecanismo com capacidade de captar diversos fatores que afetam cada técnica de previsão individual usada como base de obtenção da previsão combinada.

- Técnica de Variância Mínima

Este método de combinação propõe um esquema de combinação linear, cujos pesos são calculados a partir da matriz de covariância de erros de previsões individuais, dando-se assim maior peso à previsão que fornecer os menores erro quadráticos médios (Bates e Granger, 1969).

$$F_c = wF_1 + (1 - w)F_2 \quad \text{Equação 3.5.1}$$

Onde  $F_1$  e  $F_2$  são as previsões a serem combinadas resultantes dos modelos individuais,  $w$  o peso da previsão 1, e  $(1 - w)$  o peso da previsão 2.

O valor ótimo para  $w$ , conforme dito acima é determinado pela minimização da variância dos erros da combinação da previsão  $F_c$ , que é dada por

$$\sigma_c^2 = w^2\sigma_1^2 + (1 - w)^2\sigma_2^2 + 2\rho w\sigma_1(1 - w)\sigma_2 \quad \text{Equação 3.5.2}$$

Sendo  $\sigma_1$  e  $\sigma_2$  variâncias dos erros das previsões a serem combinadas, e  $\rho$  o coeficiente de correlação dos erros das previsões.

O valor mínimo de  $\sigma_c^2$  ocorre quando  $w$  assume

$$w = \frac{\sigma_2^2 - \rho\sigma_1\sigma_2}{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2} \quad \text{Equação 3.5.3}$$

Onde, observa-se que se  $\rho = 0$ ,

$$w = \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2} \quad \text{Equação 3.5.4}$$

Outra forma de realizar previsão a partir de combinação é técnica de combinação propõe combinar as previsões individuais através de média aritmética simples das previsões.

- Técnica de Média Aritmética Simples

Além de ser um dos métodos mais conhecidos e utilizados, é fácil de calcular, e seu sucesso pode estar associado à instabilidade dos pesos ao longo do tempo na matriz de covariância dos erros das previsões individuais.

Na equação 3.4.5 é apresentada a expressão para a combinação de  $n$  previsões individuais.

$$F_c = \frac{F_1 + F_2 + \dots + F_n}{n} \quad \text{Equação 3.5.5}$$

Makridakis e Winkler (1983) observaram que o uso de média das previsões traz consideráveis benefícios práticos em termos de melhoria da acurácia de previsão e diminuição de sua variabilidade.

- Técnica de Regressão por MQO

Método proposto por Granger e Ramanathan (1984), é uma combinação linear sem restrições nos pesos, porém com a adição de uma constante. Em detalhe, a combinação das previsões é calculada a partir de

$$F_c = w_1F_1 + w_2F_2 + \dots + w_nF_n + k \quad \text{Equação 3.5.6}$$

Onde  $F_c$  é a função combinada,  $F_i$  a previsão do modelo individual  $i = 1, \dots, n$ ,  $w$  são os pesos (parâmetros) dos modelos individuais, e  $k$  uma constante diferente de zero.

Os coeficientes  $w_1, w_2, \dots, w_n$  são estimados a partir do método de Mínimos Quadrados Ordinário (MQO), onde para realizar a regressão a variável dependente consiste dos valores observados dentro da amostra e as previsões individuais como variáveis explicativas.

### 3.6 MEDIDAS DE ACURÁCIA

Com o intuito de comparar os diferentes modelos de previsão e métodos de combinação de previsão serão considerados nesse estudo, o cálculo de seis medidas de acurácia, popularmente conhecidas como MAE, MSE, RMSE, MAPE, SMAPE, e U de Theil.

Essas medidas baseiam-se no cálculo do erro de previsão de um passo à frente no tempo  $t$ , dado pela diferença entre o valor de sequência de entrada no tempo ( $Y_i$ ) e a previsão no mesmo instante ( $\hat{Y}_i$ ). A seguir, é apresentada as formulas de cálculo de cada medida, onde  $n$  representa o número de previsões (Shcherbakov et all, 2013).

1. MAE - Erro Absoluto Médio (do inglês, Mean Absolute Error)

$$\frac{1}{n} \sum_{i=1}^n |Y_i - \hat{Y}_i| \quad \text{Equação 3.6.1}$$

2. MSE - Erro Quadrático Médio (do inglês, Mean Square Error)

$$\frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \quad \text{Equação 3.6.2}$$

3. RMSE - Raiz do Erro Quadrático Médio (do inglês, Root Mean Squared Error).

$$\sqrt{\frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2} \quad \text{Equação 3.6.3}$$

4. MAPE - Erro Percentual Absoluto Médio (do inglês, Mean Absolute Percentage Error)

$$\frac{1}{n} \sum_{i=1}^n \left| \frac{Y_i - \hat{Y}_i}{Y_i} \right| \times 100 \quad \text{Equação 3.6.4}$$



5. SMAPE - Erro Percentual Absoluto Médio Simétrico (do inglês, Symmetric Mean Absolute Percentage Error)

$$\frac{100\%}{n} \sum_{i=1}^n \frac{|Y_i - \hat{Y}_i|}{(|Y_i| + |\hat{Y}_i|)/2}$$

Equação 3.6.5

6. U de Theil

O Coeficiente U de Theil avalia a precisão levando em consideração o erro apresentado pelo modelo naïve e o modelo proposto, sendo determinado conforme especificação abaixo.

$$\frac{\sqrt{\frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2}}{\sqrt{\frac{1}{n} \sum_{i=1}^n (Y_i)^2 + \frac{1}{n} \sum_{i=1}^n (\hat{Y}_i)^2}}$$

Equação 3.6.6

Sua interpretação se dá da seguinte forma:

- se o coeficiente  $U = 1$ , o método naïve é tão eficiente quanto o método em avaliação.
- se o coeficiente  $U < 1$ , o método naïve é menos eficiente que o método em avaliação.
- se o coeficiente  $U > 1$ , o método naïve é mais eficiente que o método em avaliação.
- se o coeficiente  $U = 0$ , o método em avaliação é perfeito.

Em geral, espera-se que o modelo avaliado resulte em  $U < 1$ , ou seja, apresente um erro de precisão menor que um modelo ingênuo (naïve) (Makridakis, 1998).

A interpretação das outras medidas acima apresentadas se dá da seguinte forma: MAE representa o erro médio, MSE a variância do erro, RMSE a raiz da variância, ou seja, o desvio padrão do erro, MAPE representa o percentual do erro que o modelo incide na previsão.

## 4 RESULTADOS E DISCUSSÃO

Nessa seção apresenta-se os dados utilizados, suas características gerais e os resultados das estimativas e das previsões dos modelos individuais e das técnicas de combinação de previsões discutidas neste trabalho.

Ao aplicar os testes de estacionariedade, constou-se unanimemente pela não rejeição de  $H_0$ , isto é não rejeição da hipótese nula de que a série possui raiz unitária (não estacionariedade) ao nível de significância de 5% - Tabela 4, o que implica dizer, que se faz necessário utilizar da operação de diferença para que torná-las estacionárias.

**Tabela 4 - Testes de estacionariedade**

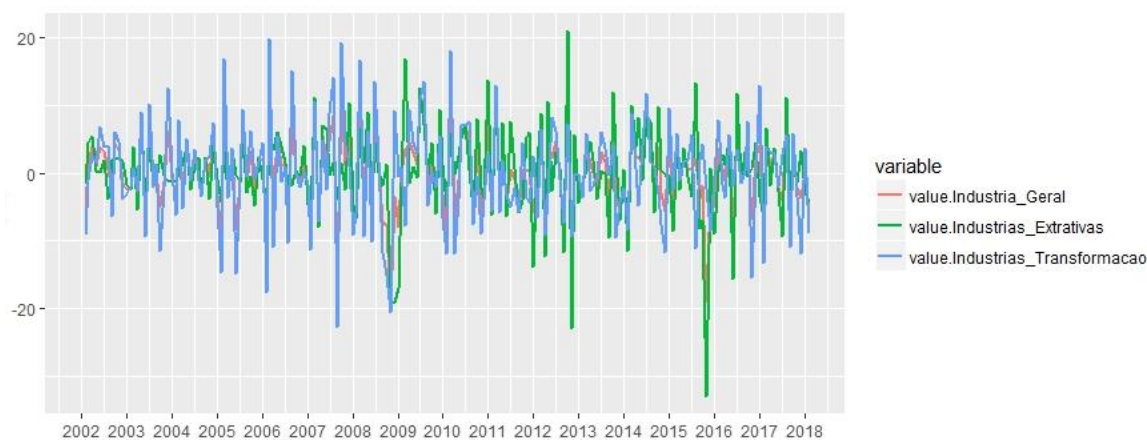
Série	Teste	Estatística de Teste	p-valor calculado
Indústria Geral	ADF	-0.3641	0,5390
	PP	-19,280	0,0133
Indústria Extrativa	ADF	-0,3143	0,5530
	PP	-7,892	0,2880
Indústria de Transformação	ADF	-0,6052	0,4620
	PP	-27,422	0,0100

Nesse sentido é necessário testar a série  $I(1)$  – possui ordem de integração 1 - pois concluímos que a série não é  $I(0)$ .

A tabela 5 apresenta o resultado dos testes ADF e PP para as séries com uma diferença, e dessa vez rejeita-se a hipótese de  $\pi = 1$ , isto é de possuir raiz unitária, ao nível de significância de 5%, e com isso, podemos concluir que as séries são  $I(1)$ .

**Tabela 5 - Testes de Estacionariedade na Série  $I(1)$**

Série	Teste	Estatística de Teste	p-valor calculado
Indústria Geral	ADF	-19,8662	0,0100
	PP	-275,0000	0,0100
Indústria Extrativa	ADF	-17,6054	0,0100
	PP	-251,0000	0,0100
Indústria de Transformação	ADF	-19,5923	0,0100
	PP	-239,0000	0,0100



**Figura 14** - Séries de Produção Industrial com uma diferença

Os tópicos a seguir apresentam os resultados obtidos em cada metodologia de previsão para cada uma das três séries em análise: Produção Geral, Produção Extrativa, e Produção de Transformação, respectivamente.

#### 4.1 INDÚSTRIA GERAL

Foram considerados os Índices Mensais de Produção Geral no período de janeiro de 2002 a dezembro de 2016, totalizando 168 observações para modelagem via Box-Jenkins, Holt-Winters e Redes Neurais Artificiais. Enquanto no Modelo Econométrico, devido a limitação na disponibilização dos dados das variáveis regressoras, foram considerados apenas os dados no período de janeiro de 2011 a dezembro de 2016, totalizando apenas 72 observações.

Na modelagem Box-Jenkins foram testados diversos modelos, e os oito melhores modelos obtidos são apresentados na Tabela 6, destacando-se o modelo SARIMA(2,1,0)(2,0,0)[12] que apresentou melhores medidas de acurácia, conforme apresentados na Tabela 7.

**Tabela 6 - Modelos Box-Jenkins para Produção Industrial Geral (jan-02 a dez-16)**

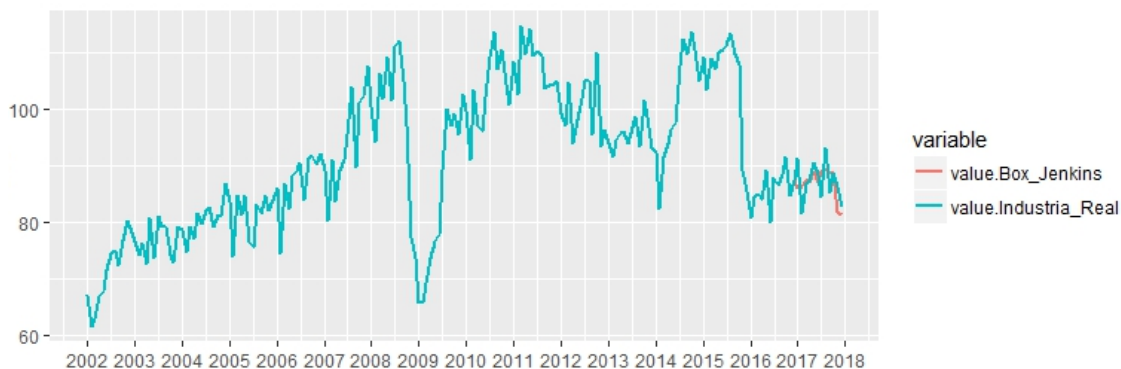
Modelo	Vetor	AR1	AR2	MA1	MA2	SAR1	SAR2
SARIMA(2,1,0)(2,0,0)[12]	Coeficiente	-0,2162	0,1444	[na]	[na]	0,2232	0,2954
	Desvio Padrão	0,0748	0,0739	[na]	[na]	0,0701	0,0726
SARIMA(2,1,0)(1,0,0)[12]	Coeficiente	-0,2272	0,1311	[na]	[na]	0,3268	[na]
	Desvio Padrão	0,0757	0,0740	[na]	[na]	0,0711	[na]
SARIMA(2,1,0)(0,0,0)[12]	Coeficiente	-0,2988	0,1084	[na]	[na]	[na]	[na]
	Desvio Padrão	0,0742	0,0743	[na]	[na]	[na]	[na]
SARIMA(1,1,0)(2,0,0)[12]	Coeficiente	-0,2541	[na]	[na]	[na]	0,2204	0,2901
	Desvio Padrão	0,0734	[na]	[na]	[na]	0,0705	0,0731
SARIMA(1,1,0)(1,0,0)[12]	Coeficiente	-0,2633	[na]	[na]	[na]	0,3203	[na]
	Desvio Padrão	0,0741	[na]	[na]	[na]	0,0717	[na]
SARIMA(2,1,1)(1,0,0)[12]	Coeficiente	0,0516	0,2080	-0,2833	[na]	0,3276	[na]
	Desvio Padrão	0,3419	0,1039	0,3436	[na]	0,0712	[na]
SARIMA(2,1,2)(1,0,0)[12]	Coeficiente	0,0418	0,1091	-0,2778	0,1020	0,3242	[na]
	Desvio Padrão	0,4310	0,2433	0,4310	0,2271	0,0716	[na]
SARIMA(1,1,1)(2,0,0)[12]	Coeficiente	-0,5482	[na]	0,3078	[na]	0,2224	0,2954
	Desvio Padrão	0,1561	[na]	0,1713	[na]	0,0703	0,0729

**Tabela 7 - Medidas de Acurácia para modelos Box-Jenkins – Indústria Geral (jan-17 a dez-17)**

Modelo Proposto	BIC	AIC	MAE	MSE	RMSE	MAPE	SMAPE	U-Theil
SARIMA(2,1,0)(2,0,0)[12]	1120,84	1104,902	265,6004	10,3117	3,2112	3,0661	0,0306	0,0184
SARIMA(2,1,0)(1,0,0)[12]	1131,02	1118,27	297,7819	13,4620	3,6691	3,4378	0,0342	0,0211
SARIMA(2,1,0)(0,0,0)[12]	1145,37	1135,80	287,7268	12,4678	3,5310	3,3354	0,0330	0,0203
SARIMA(1,1,0)(2,0,0)[12]	1119,43	1106,68	270,6777	10,2056	3,1946	3,1417	0,0312	0,0184
SARIMA(1,1,0)(1,0,0)[12]	1128,95	1119,38	298,4022	13,3460	3,6532	3,4583	0,0343	0,0210
SARIMA(2,1,1)(1,0,0)[12]	1135,72	1119,78	299,4303	13,4520	3,6677	3,4411	0,0344	0,0210
SARIMA(2,1,2)(1,0,0)[12]	1140,70	1121,57	297,3105	13,3303	3,6511	3,4178	0,0341	0,0209
SARIMA(1,1,1)(2,0,0)[12]	1122,23	1106,29	268,9243	10,3832	3,2223	3,1175	0,0310	0,0185

Onde os valores AIC e BIC são calculados considerando a performance dos modelos estimados no período de janeiro de 2002 a dezembro de 2016.

A partir da Figura 15 é possível visualizar o ajuste obtido para Indústria Geral com base do modelo melhor acurácia, SARIMA(2,1,0)(2,0,0)[12], e o intervalo de confiança pode ser visualizado a partir do APENDICE F.



**Figura 15** - Produção Industrial Geral: Realizado (2002-2017) x Previsto via Box-Jenkins (2017)

O resultado da modelagem via Holt-Winters, mesmo apresentando bons resultados, foi aquém do modelo SARIMA acima apresentado.

Neste método, os melhores resultados foram obtidos ao considerar o método multiplicativo, que apresentou menor erro da previsão 12 passos (Tabela 8).

**Tabela 8** - Medidas de Acurácia para modelos Holt-Winters – Indústria Geral (jan-17 a dez-17)

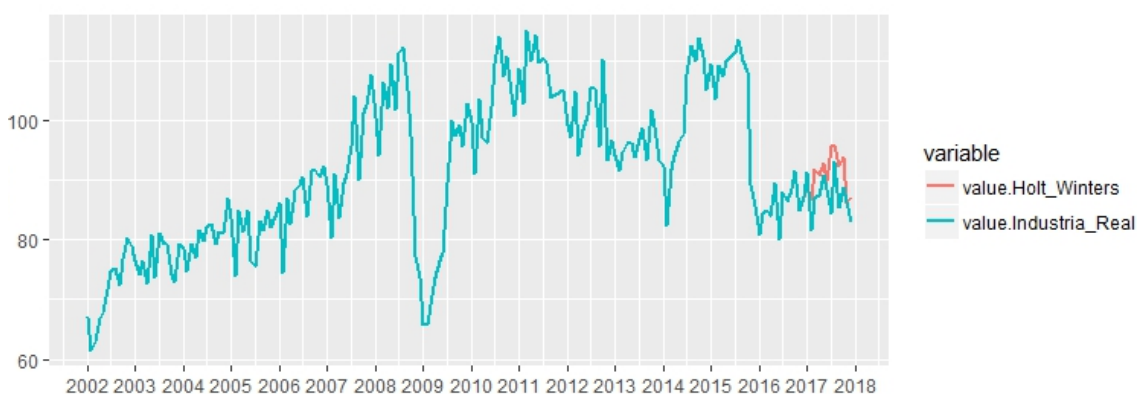
Modelo	MAE	MSE	RMSE	MAPE	SMAPE	U-Theil
Aditivo	440,0361	26,8950	5,1860	5,1017	0,0493	0,0290
Multiplicativo	418,4005	24,8859	4,9886	4,8557	0,0470	0,0280

As constantes de alisamento para a série  $Y_t$ , sendo  $Y_t$  a série dos índices de produção industrial no modelo selecionado por:  $\alpha = 0,7899$ ,  $\beta = 0$ , e  $\gamma = 0,8988$ . E os coeficientes estimados estão dispostos na Tabela 9.

**Tabela 9** – Coeficientes Modelo Holt-Winters Selecionado – Indústria Geral (jan-02 a dez-16)

Coeficientes	Valor Estimado
A	90,9252
B	[na]
C	[na]
S1	0,9740
S2	0,9529
S3	1,0101
S4	0,9976
S5	1,0214
S6	0,9876
S7	1,0504
S8	1,0553
S9	1,0152
S10	1,0311
S11	0,9467
S12	0,9598

A partir da Figura 15 é possível visualizar o ajuste obtido para Indústria Geral com base do modelo melhor acurácia via Holt-Winters. O intervalo de confiança 95% está disposto no APENDICE F.

**Figura 16** - Produção Industrial Geral: Realizado (2002-2017) x Previsto via Holt-Winters (2017)

No Modelo Econométrico foram avaliadas as seguintes variáveis regressoras: taxa de juros, taxa de inflação, taxa de cambio, índice de confiança do empresário industrial, utilização da capacidade instalada, percentual de desocupação, e paralisação da Samarco.

Considerando que em muitos estudos econômicos a variável resposta depende dos valores de períodos passados para explicar um comportamento futuro, considerou algumas defasagens de

modo a garantir essa interpretação. Tem-se ainda, segundo a teoria econômica que o modelo pode incluir também valores defasados na própria variável dependente (Gujarati, 2006).

Constatou-se a partir da Modelagem Econométrica, que o Índice de Produção Industrial Geral dependente do seu valor observado nos meses passados (1 e 2 defasagens), dependente do retorno da Samarco, da taxa de juros (1 e 2 defasagens), depende da taxa de câmbio (1 e 2 defasagens), o percentual de pessoas desocupadas também influencia, bem como o índice de confiança do empresário industrial (indústria extrativa e de transformação), e por fim, do percentual de utilização da capacidade instalada (Tabela 10).

Importante ressaltar que os modelos apresentados poderiam ser melhor especificados com a adição de novas variáveis. No entanto, como estamos pensando no aspecto preditivo, vamos considerar o modelo mais geral conforme apresentadas acima.

**Tabela 10** - Modelo Econométrico - Indústria Geral (jan-11 a dez-16)

Coeficientes	Estimativa	Erro Padronizado	t-valor	Pr(> t )		Coef. Beta
Intercepto	84,8870	33,1032	2,564	0,0127	*	8,8929
Ind Geral (1def)	0,1932	0,1055	1,832	0,0718	,	0,1910
Ind Geral (2def)	0,3083	0,1044	2,952	0,0044	**	0,3000
<b>Paralisacao_Samarco</b>	<b>-14,0001</b>	<b>3,9556</b>	<b>-3,539</b>	<b>0,0007</b>	***	<b>-0,5840</b>
Taxa de Juros (2def)	11,2271	5,0431	2,226	0,0296	*	0,2155
<b>Cambio (1def)</b>	<b>-6,4590</b>	<b>2,7452</b>	<b>-2,353</b>	<b>0,0218</b>	*	<b>-0,4880</b>
Perc Desocupacao	1,1426	0,7251	1,576	0,1201		0,2543
ICE Ind Transformacao	-0,4593	0,2004	-2,291	0,0253	*	-0,3640
UCI Ind Geral	0,5608	0,3275	1,712	0,0919	,	0,2243
Consumo Energia (2def)	-0,0037	0,0014	-2,608	0,0113	*	-0,2906

Signif, codes: 0 '\*\*\*' 0,001 '\*\*' 0,01 '\*' 0,05 '.' 0,1 ' ' 1

Foram apurados o coeficiente de determinação  $R^2$  e o coeficiente de determinação ajustado  $R^2_{ajustado}$ , resultado em **0,8166** e **0,7899** respectivamente.

Também foi considerado o resultado do teste F, que apresentou p-valor menor que 0,01%, o que significa dizer que a hipótese nula de os coeficientes serem todos iguais a zero é rejeitada.

Ao verificar a estatística de Durbin-Watson (**1,7769**), com p-valor = **0,04**, observou-se que os resíduos para os dados individualizados não são autocorrelacionados, ao nível de significância de 1% (não rejeita-se  $H_0$ ). Também nesse sentido, aferiu-se o teste de Kolmogorov-Sminorv nos resíduos obtendo-se uma estatística de teste igual a **0,0597** com p-valor de **0,94**, no qual não rejeita-se (ao nível de significância de 1%.) a hipótese nula de que os dados seguem uma distribuição normal.

O coeficiente beta (apresentado na Tabela 10) serve para medir a importância das variáveis independentes, pois eliminam o problema das unidades diferentes. Sabendo disso, conclui-se que a variável mais importante no modelo de regressão é a dummy paralização da samarco, e a segunda mais importante é a taxa de câmbio com 1 defasagem. A variável explicativa menos importante é o índice de confiança do empresário da indústria extrativa.

O coeficiente associado ao termo constante nos diz que, quando todas as variáveis explicativas do modelo são iguais a zero, a produção industrial geral tem o índice de 84,88. O período da paralização da Samarco influenciou negativamente o índice de produção industrial, ou seja, o índice de produção industrial teve uma queda nesse período de aproximadamente 14,00 unidades, se mantida constante as outras variáveis.

Para previsão 12 passos a frente, nota-se que não se poderia simplesmente utilizar os dados reais de 2017 das variáveis explicativas como input na regressão. Desse modo, considerou-se utilizar previsão para cada uma das 11 variáveis explicativas, conforme especificado a seguir:

- **Ind Geral (1def e 2 def):** considerou-se a previsão obtida via Box-Jenkins, modelo SARIMA(2,1,0)(2,0,0)[12], por apresentar menor erro se comparado com a modelagem Holt-Winters, a partir da tabela 11 é possível consultar os valores previstos para 1e 2 defasagens.

**Tabela 11** – Indústria Geral - Realizado x Previsto (2017)

Mês	Ind Geral (Real)	Ind Geral 1 def (Previsão)	Ind Geral 2 def (Previsão)
nov/16	84,9000		
dez/16	87,3000		
jan/17	91,4000	87,3000	84,9000
fev/17	81,6000	86,1593	87,3000
mar/17	87,0000	86,2397	86,1593
abr/17	87,5000	87,6358	86,2397
mai/17	90,7000	87,0759	87,6358
jun/17	88,4000	89,0033	87,0759
jul/17	84,4000	87,1372	89,0033
ago/17	93,1000	89,1324	87,1372
set/17	85,4000	89,4316	89,1324
out/17	88,7000	88,7855	89,4316
nov/17	86,4000	88,9717	88,7855
dez/17	82,8000	81,8915	88,9717

- **Paralisacao\_Samarco:** considerou-se todo ano de 2017 igual a 1, ou seja, sem retomada das atividades.



- **Taxa de Juros (2def):** realizou-se previsão 10 passos a frente utilizando as metodologias Box-Jenkins, Holt-Winters Aditivo e Multiplicativo (Resultados dispostos no Apêndice D), onde obteve-se melhor resultado a partir no Holt-Winters Aditivo.

Considerando-se defasagem igual a 2, considerou-se apenas os 10 passos a frente na previsão, pois deve-se considerar os meses de novembro e dezembro de 2016 para atualizar a previsão de janeiro e fevereiro de 2017, e a previsão de janeiro a outubro de 2017 (com base nos dados até dezembro de 2016 apenas), para atualizar a previsão de março a dezembro de 2017 (Tabela 12).

**Tabela 12** – Taxa de Juros (Selic) - Realizado x Previsto (2017)

Mês	Selic (Real)	Selic 2 def (Previsão)
nov/16	1,0383	
dez/16	1,1233	
jan/17	1,0861	1,0383
fev/17	0,8651	1,1233
mar/17	1,0521	1,0144
abr/17	0,7866	0,9067
mai/17	0,9271	0,9992
jun/17	0,8089	0,8927
jul/17	0,7979	0,8995
ago/17	0,8023	0,8870
set/17	0,6385	0,8889
out/17	0,6439	0,8855
nov/17	0,5682	0,8105
dez/17	0,5384	0,7907

Os testes de estacionariedade ADF e PP também foram aplicados a essa série de dados, onde também observou-se a necessidade e tomar uma diferença dos dados para tornar a série estacionaria (Tabela 13). Ou seja, é verificada a mesma ordem de integração I(1) da variável dependente em questão.

**Tabela 13** - Testes de Estacionariedade - Selic (jan-11 a dez-16)

Taxa de Juros	Teste	Estatística de Teste	p-valor calculado
Série Original	ADF	0,0310	0,650
	PP	0,1850	0,728
Série (1 Diferença)	ADF	-13,9192	<0,01
	PP	-101,0000	<0,01

- **Taxa de Cambio (1def):** realizou-se previsão 11 passos a frente utilizando as metodologias Box-Jenkins, Holt-Winters Aditivo e Multiplicativo (Resultados dispostos no Apêndice D), onde obteve-se melhor resultado a partir no Holt-Winters Multiplicativo.

Considerando-se defasagem igual a 1, considerou-se a previsão apenas de 11 passos a frente, pois deve-se considerar o realizado em dezembro de 2016 para atualizar a previsão de janeiro de 2017, e a previsão de janeiro a novembro de 2017 (com base nos dados até dezembro de 2016 apenas), para atualizar a previsão de fevereiro a dezembro de 2017 (Tabela 14).

**Tabela 14** – Taxa de Cambio - Realizado x Previsto (2017)

Mês	Cambio (Real)	Cambio 1 def (Previsão)
dez/16	3,3510	
jan/17	3,1970	3,3510
fev/17	3,1030	3,2899
mar/17	3,1270	3,1182
abr/17	3,1400	3,0660
mai/17	3,2090	3,0538
jun/17	3,2970	3,1973
jul/17	3,2050	3,3125
ago/17	3,1530	3,1796
set/17	3,1380	3,1259
out/17	3,1960	3,2269
nov/17	3,2570	3,2532
dez/17	3,2970	3,3255

Os testes de estacionariedade ADF e PP também foram aplicados a essa série de dados, onde também observou-se a necessidade e tomar uma diferença dos dados para tornar a série estacionaria (Tabela 15). Ou seja, é verificada a mesma ordem de integração I(1) da variável dependente em questão.

**Tabela 15** - Testes de Estacionariedade - Taxa de Cambio (jan-11 a dez-16)

Taxa de Cambio	Teste	Estatística de Teste	p-valor calculado
Série Original	ADF	1,6693	0,974
	PP	0,538	0,806
Série (1 Diferença)	ADF	-5,5037	<0,01
	PP	-38,800	<0,01

- **Perc Desocupacao:** os valores absolutos observados de janeiro a dezembro de 2016 foram replicados para 2017, ou seja, sem perspectiva de melhora na economia industrial.

Os testes de estacionariedade ADF e PP também foram aplicados a essa série de dados, onde também observou-se a necessidade e tomar uma diferença dos dados para tornar a série estacionaria (Tabela 16). Ou seja, é verificada a mesma ordem de integração I(1) da variável dependente em questão.

**Tabela 16** - Testes de Estacionariedade (Percentual de Desocupação) (jan-11 a dez-16)

Perc Desocupação	Teste	Estatística de Teste	p-valor calculado
Série Original	ADF	2,0500	0,99
	PP	0,9960	0,901
Série (1 Diferença)	ADF	-8,2900	<0,01
	PP	-75,3000	<0,01

- **ICE Ind Transformação:** os valores absolutos observados de janeiro a dezembro de 2016 foram replicados para 2017.

Os testes de estacionariedade ADF e PP também foram aplicados a essa série de dados, onde também observou-se a necessidade e tomar uma diferença dos dados para tornar a série estacionaria (Tabela 17). Ou seja, é verificada a mesma ordem de integração I(1) da variável dependente em questão.

**Tabela 17** – Testes de Estacionariedade – ICE Indústria de Transformação (jan-11 a dez-16)

ICE Ind Transformação	Teste	Estatística de Teste	p-valor calculado
Série Original	ADF	-1,022	0,311
	PP	-0,439	0,59
Série com 1 Diferença	ADF	-7,15	<0.01
	PP	-64,4	<0.01

- **UCI Ind Geral:** os valores absolutos observados de janeiro a dezembro de 2016 foram replicados para 2017.

Os testes de estacionariedade ADF e PP também foram aplicados a essa série de dados, onde também observou-se a necessidade e tomar uma diferença dos dados para tornar a série estacionaria (Tabela 18). Ou seja, é verificada a mesma ordem de integração I(1) da variável dependente em questão.

**Tabela 18 - Testes de Estacionariedade (UCI Indústria Geral) (jan-11 a dez-16)**

UCI Ind Geral	Teste	Estatística de Teste	p-valor calculado
Série Original	ADF	-0,669	0,437
	PP	-0,124	0,659
Série com 1 Diferença	ADF	-8,02	<0,01
	PP	-62,4	<0,01

- **Energia (2 def):** realizou-se previsão 10 passos a frente utilizando as metodologias Box-Jenkins, Holt-Winters Aditivo e Multiplicativo (Resultados dispostos no Apêndice D), onde obteve-se melhor resultado a partir no Holt-Winters Aditivo.

Considerando-se defasagem igual a 2, considerou-se apenas os 10 passos a frente na previsão, pois deve-se considerar os meses de novembro e dezembro de 2016 para atualizar a previsão de janeiro e fevereiro de 2017, e a previsão de janeiro a outubro de 2017 (com base nos dados até dezembro de 2016 apenas), para atualizar a previsão de março a dezembro de 2017.

**Tabela 19 – Consumo de Energia - Realizado x Previsto (2017)**

Mês	Energia (Real)	Energia 2 def (Previsão)
nov/16	13786,0000	
dez/16	13387,0000	
jan/17	13017,0000	13786,0000
fev/17	13323,0000	13387,0000
mar/17	13908,0000	12706,1200
abr/17	14007,0000	13347,7100
mai/17	13598,0000	13695,4600
jun/17	13919,0000	13813,3900
jul/17	14082,0000	13634,2200
ago/17	14197,0000	13444,8100
set/17	14146,0000	13548,2500
out/17	14173,0000	14008,8400
nov/17	14316,0000	13915,1300
dez/17	14073,0000	13928,8200

Os testes de estacionariedade ADF e PP também foram aplicados a essa série de dados, onde também observou-se a necessidade e tomar uma diferença dos dados para tornar a série estacionaria (Tabela 20). Ou seja, é verificada a mesma ordem de integração I(1) da variável dependente em questão.

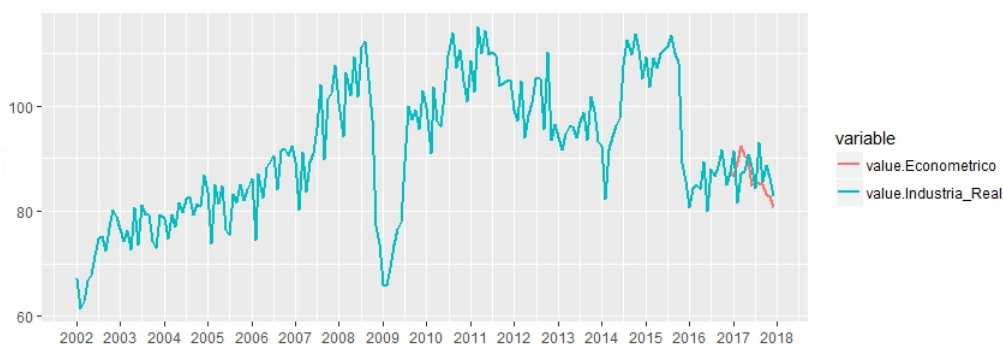
**Tabela 20** - Testes de Estacionariedade (Consumo de Energia) (jan-11 a dez-16)

UCI Ind Geral	Teste	Estatística de Teste	p-valor calculado
Série Original	ADF	-0,435	0,516
	PP	-0,083	0,669
Série com 1 Diferença	ADF	-7,802	<0,01
	PP	-54,000	<0,01

Logo, a partir da atualização do modelo econométrico considerado as previsões das variáveis regressoras para todo ano de 2017, a partir da Figura 17. A previsão pode ser visualizada a partir da Tabela 21.

**Tabela 21** - Medidas de Acurácia para Modelo Econométrico - Indústria Geral (jan-17 a dez-17)

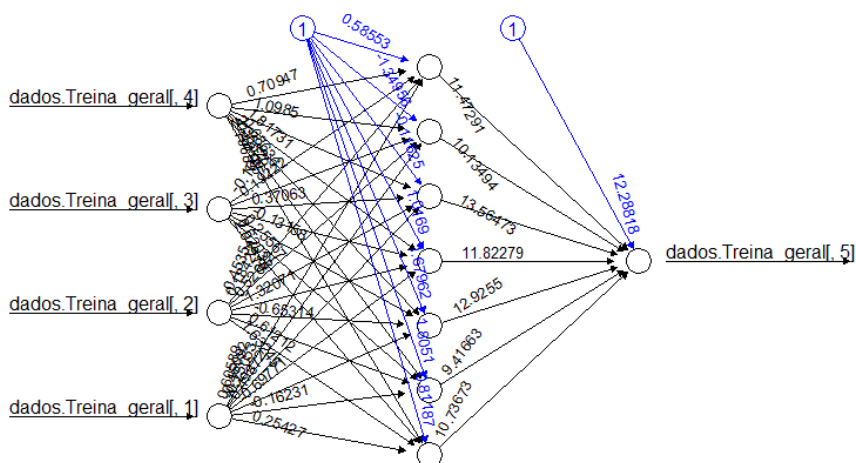
Modelo	MAE	MSE	RMSE	MAPE	SMAPE	U-Theil
Modelo Econométrico	380,8987	20,0028	4,4725	4,3531	0,0437	0,0257



**Figura 17** - Produção Industrial Geral: Realizado (2002-2017) x Previsto via Modelo Econométrico (2017)

Na modelagem via Redes Neurais Artificiais, foi considerado o modelo multicamadas, conforme pode ser visto na Figura 18.

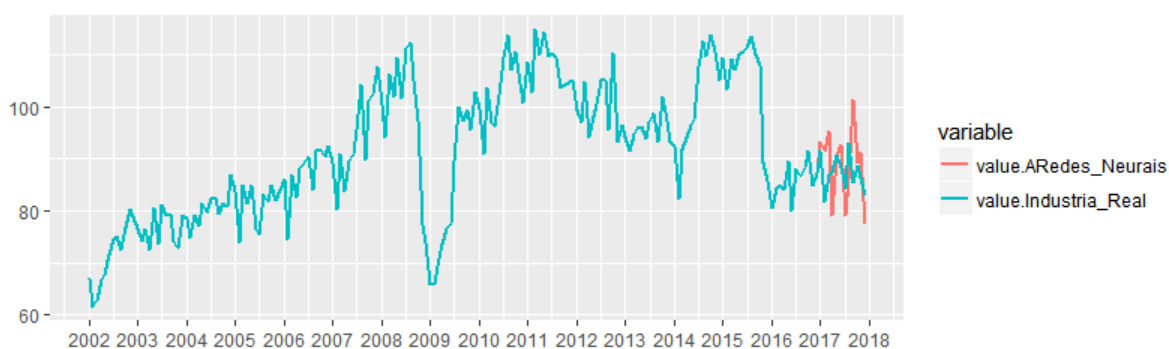
Considerou-se como entrada, 4 defasagens do próprio índice de produção industrial, imitando um processo autoregressivo, onde também se limitou os dados até 2016. As medidas de acurácia estão dispostas na Tabela 22.



**Figura 18** - Modelo Redes Neurais Artificiais - Indústria Geral (jan-02 a dez-16)

**Tabela 22** - Medidas de Acurácia via Redes Neurais - Indústria Geral (jan-17 a dez-17)

Modelo	MAE	MSE	RMSE	MAPE	SMAPE	U-Theil
Modelo RNA	598,4200	53,9624	7,3459	6,9494	0,0680	0,0416



**Figura 19** - Produção Industrial Geral: Realizado (2002-2017) x Previsto via Redes Neurais (2017)

A partir dos 4 principais modelos acima apresentados, a proposta do trabalho é aplicar métodos de combinação de previsão para melhoria de performance. Foram considerados 3

métodos: combinação por média, combinação por matriz de variância, e combinação a partir de técnica de regressão por MQO.

Os resultados de todas as previsões estão dispostos na Tabela 23, bem como os valores observados (reais) para o mesmo período.

**Tabela 23 - Previsões Indústria Geral (2017)**

Mês	Ind_Geral	BJ	HW	Econométrico	RNA	Comb_Media	Comb_Var	Comb_Reg
Jan	91,4000	86,1593	88,5625	86,5289	93,1369	87,0835	86,7740	86,3672
Feb	81,6000	86,2397	86,6471	89,3529	91,4897	87,4132	87,1588	85,7657
Mar	87,0000	87,6358	91,8474	92,3194	95,2831	90,6009	89,7912	87,6882
Apr	87,5000	87,0759	90,7107	90,2676	79,0796	89,3514	88,7089	87,2201
May	90,7000	89,0033	92,8703	89,6125	90,6484	90,4954	89,9963	88,3447
Jun	88,4000	87,1372	89,7988	84,8103	92,6725	87,2488	87,0871	86,9925
Jul	84,4000	89,1324	95,5035	85,7441	79,2320	90,1267	89,5952	89,2269
Aug	93,1000	89,4316	95,9503	85,3895	85,9207	90,2571	89,7515	89,4383
Sep	85,4000	88,7855	92,3032	85,0764	101,4572	88,7217	88,5500	88,1530
Oct	88,7000	88,9717	93,7503	83,3709	89,2114	88,6976	88,5016	88,6612
Nov	86,4000	81,8915	86,0786	83,0185	91,2230	83,6628	83,0915	84,6334
Dec	82,8000	81,3943	87,2674	80,5691	77,3828	83,0769	82,4349	84,9089

O método de combinação de previsão via regressão por MQO foi o que apresentou melhor resultado (Tabela 24), onde é observado melhor erro para todas as medidas de acurácia apresentadas. O segundo melhor modelo é o SARIMA, sendo observado valores similares de acurácia se comparado com os resultados obtidos via combinação por variância mínima e média simples (exceto RNA). Esses resultados vão de encontro com tantos outros estudos que apontam os ganhos de performance ao considerar métodos de combinação previsão para garantir uma estimativa pontual com melhor acurácia. A partir do APENDICE F é possível constatar o intervalo de confiança.

**Tabela 24 - Medidas de Acurácia - Todos Modelos- Indústria Geral (jan-17 a dez-17)**

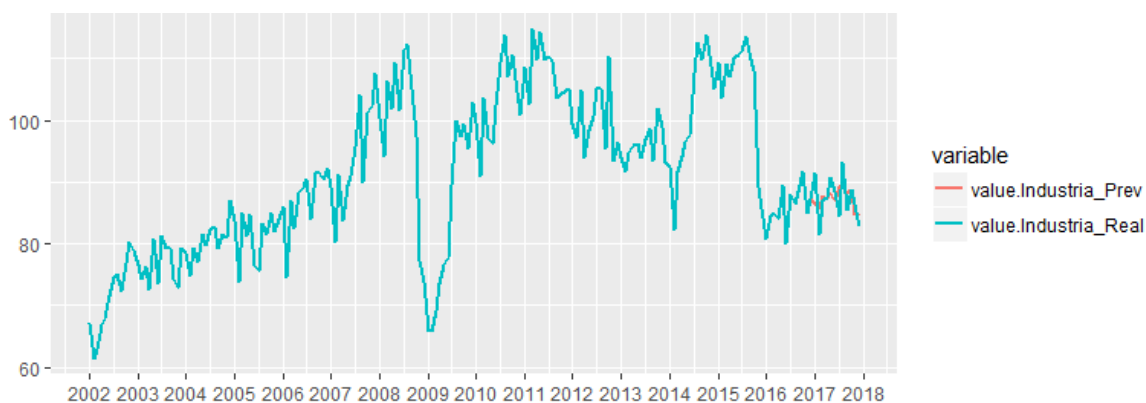
Modelos Propostos	MAE	MSE	RMSE	MAPE	SMAPE	U-Theil
BJ	265,6004	10,3117	3,2112	3,0661	0,0306	0,0184
HW	418,4005	24,8859	4,9886	4,8557	0,0470	0,0280
Econométrico	380,8987	20,0028	4,4725	4,3531	0,0437	0,0257
RNA	598,4200	53,9624	7,3459	6,9494	0,0680	0,0416
Comb_Media	258,7491	12,2249	3,4964	2,8987	0,0295	0,0199
Comb_Var	257,3528	10,1276	3,1824	2,9228	0,0295	0,0182
Comb_Reg	248,0572	8,1317	2,8516	2,8374	0,0284	0,0163
Comb_Media (exceto RNA)	267,2979	10,9102	3,3031	3,0262	0,0306	0,0188
Comb_Var (exceto RNA)	265,5781	10,3215	3,2127	3,0267	0,0304	0,0184

A Equação 4.4.1 apresenta os pesos estimados por MQO pra cada uma dos quatro métodos de previsão.

$$IG_k = 33,45 - 0,20J_k + 0,63HW_k + 0,06Eco_k + 0,12RNA_k \quad \text{Equação 4.1.1}$$

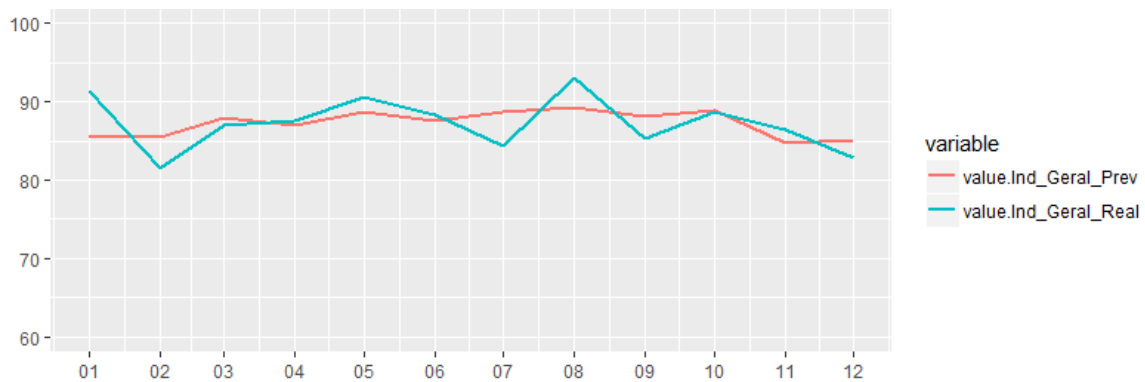
Onde  $IG_k$  representa o índice da indústria geral no  $k$  – *ésimo* mês de previsão, ou seja,  $k = 1, 2, \dots, 12$  representando os meses de janeiro a dezembro de 2017.  $BJ_k$  representa a previsão obtida a partir do modelo SARIMA para o mês  $k$ ,  $HW_k$  representa a previsão obtida a partir do método Holt-Winters para o mês  $k$ ,  $Eco_k$  representa a previsão obtida a partir do modelo econométrico para o mês  $k$ , e  $RNA_k$  representa a previsão obtida a partir do método de redes neurais artificiais para o mês  $k$ .

A partir das Figuras 20 e 21 é possível observar o ajuste obtido a partir do método de combinação de previsão via regressão por MQO.



**Figura 20** - Produção Industrial Geral: Realizado (2002-2017) x Previsto via Combinação Previsão (2017)





**Figura 21** – Previsão de Produção Industrial Geral (2017) - combinação de previsão via regressão por MQO

## 4.2 INDÚSTRIA EXTRATIVA

Foram considerados os Índices Mensais de Produção Extrativa no período de janeiro de 2002 a dezembro de 2016, totalizando 168 observações para modelagem via Box-Jenkins, Holt-Winters e Redes Neurais Artificiais. Enquanto no Modelo Econométrico, devido a limitação na disponibilização dos dados das variáveis regressoras, foram considerados apenas os dados no período de janeiro de 2011 a dezembro de 2016, totalizando apenas 60 observações.

Na modelagem Box-Jenkins foram testados diversos modelos, e os oito melhores modelos obtidos são apresentados na Tabela 25, destacando-se o modelo SARIMA(0,1,1)(0,0,2)[12] que apresentou melhores medidas de acurácia (Tabela 26).

**Tabela 25** - Modelos Box-Jenkins para Indústria Extrativa (jan-02 a dez-16)

Modelo	Vetor	AR1	AR2	MA1	MA2	SMA1	SMA2
SARIMA(0,1,2)(0,0,2)[12]	Coeficiente	[na]	[na]	-0,1461	0,1700	0,0606	0,1353
	Desvio Padrão	[na]	[na]	0,0749	0,0745	0,0786	0,0745
SARIMA(0,1,2)(0,0,1)[12]	Coeficiente	[na]	[na]	-0,1621	0,1663	0,0951	[na]
	Desvio Padrão	[na]	[na]	0,0743	0,0738	0,0685	[na]
SARIMA(1,1,2)(0,0,1)[12]	Coeficiente	-0,3195	[na]	0,1465	0,1360	0,0896	[na]
	Desvio Padrão	0,2714	[na]	0,2703	0,0874	0,0686	[na]
SARIMA(1,1,1)(0,0,2)[12]	Coeficiente	-0,5283	[na]	0,3353	[na]	0,0676	0,1334
	Desvio Padrão	0,1690	[na]	0,1800	[na]	0,0795	0,0732
SARIMA(0,1,1)(0,0,2)[12]	Coeficiente	[na]	[na]	-0,1486	[na]	0,0760	0,1298
	Desvio Padrão	[na]	[na]	0,0653	[na]	0,0809	0,0737
SARIMA(1,1,2)(0,0,2)[12]	Coeficiente	-0,3433	[na]	0,1849	0,1441	0,0530	0,1413
	Desvio Padrão	0,2540	[na]	0,2530	0,0865	0,0790	0,0743
SARIMA(0,1,1)(0,0,1)[12]	Coeficiente	[na]	[na]	-0,1565	[na]	0,1158	[na]
	Desvio Padrão	[na]	[na]	0,0654	[na]	0,0688	[na]
SARIMA(1,1,1)(0,0,0)[12]	Coeficiente	-0,5031	[na]	0,2838	[na]	[na]	[na]
	Desvio Padrão	0,1589	[na]	0,1669	[na]	[na]	[na]

**Tabela 26** - Medidas de Acurácia para modelos Box-Jenkins – Indústria Extrativa (jan-17 a dez-17)

Modelo Proposto	BIC	AIC	MAE	MSE	RMSE	MAPE	SMAPE	U-Theil
SARIMA(0,1,2)(0,0,2)[12]	1206,561	1190,628	466,0274	33,2409	5,7655	5,4840	0,0573	0,0350
SARIMA(0,1,2)(0,0,1)[12]	1204,605	1191,855	611,8495	49,3494	7,0249	7,0974	0,0745	0,0423
SARIMA(1,1,2)(0,0,1)[12]	1208,946	1192,946	590,7591	46,7458	6,8371	6,8703	0,0721	0,0412
SARIMA(1,1,1)(0,0,2)[12]	1207,724	1191,787	399,0413	26,4608	5,1440	4,7441	0,0494	0,0314
SARIMA(0,1,1)(0,0,2)[12]	1206,408	1193,658	369,1265	23,7901	4,8775	4,4086	0,0458	0,0298
SARIMA(1,1,2)(0,0,2)[12]	1210,546	1191,421	456,3069	32,1311	5,6684	5,3786	0,0561	0,0345
SARIMA(0,1,1)(0,0,1)[12]	1204,244	1194,682	473,2751	34,1322	5,8423	5,5785	0,0583	0,0355
SARIMA(1,1,1)(0,0,0)[12]	1202,898	1193,336	531,3787	39,4726	6,2827	6,2279	0,0651	0,0380



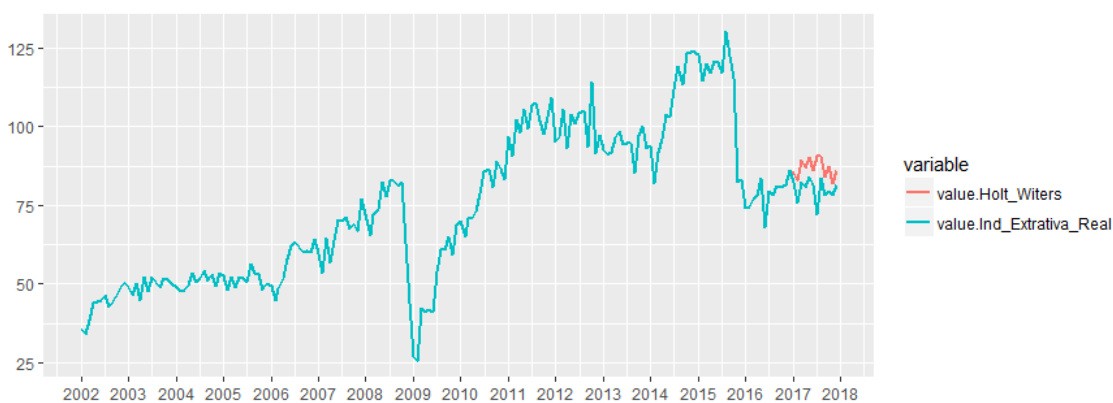
**Figura 22** – Indústria Extrativa: Realizado (2002-2017) x Previsto via Box-Jenkins (2017)

O resultado da modelagem via Holt-Winters, mesmo apresentando bons resultados, foi aquém do modelo SARIMA acima apresentado.

Neste método, conforme expectativa a priori, os melhores resultados foram obtidos ao considerar o método aditivo (Tabela 27).

**Tabela 27** - Acurácia para modelos Holt-Winters – Indústria Extrativa

Modelo	MAE	MSE	RMSE	MAPE	SMAPE	U-Theil
Aditivo (Série Original)	671,7721	59,8860	7,7386	8,5589	0,0808	0,0463
Multiplicativo (Série Original)	782,3883	83,0903	9,1153	9,9506	0,0930	0,0543



**Figura 23** – Indústria Extrativa: Realizado (2002-2017) x Previsto via Holt-Winters (2017)

As constantes de alisamento para a série  $\nabla^1 Y_t$ , sendo  $Y_t$  a série dos índices de produção industrial no modelo selecionado por:  $\alpha = 0,897$ ,  $\beta = 0$ , e  $\gamma = 1$ . E os coeficientes estimados estão dispostos na Tabela 28.

**Tabela 28** - Coeficientes Modelo Holt-Winters – Indústria Extrativa

Coeficientes	Valor Estimado
a	86,2974
b	[na]
c	[na]
S1	-1,2537
S2	-3,7018
S3	3,1505
S4	0,6496
S5	3,9200
S6	-0,3134
S7	4,3550
S8	4,4367
S9	-2,3591
S10	0,8719
S11	-4,5154
S12	-0,3975

No Modelo Econométrico também avaliadas como variáveis regressoras: taxa de juros, taxa de inflação, taxa de cambio, índice de confiança do empresário industrial, utilização da capacidade instalada, percentual de desocupação, e paralisação da Samarco.

Constatou-se a partir da Modelagem Econométrica, que o Índice de Produção Industrial Extrativa dependente do seu valor observado com uma e duas defasagens, dependente também do retorno da Samarco, da taxa de juros com uma defasagem e duas defasagens, depende do cambio com uma e duas defasagens, o percentual de pessoas desocupadas também influencia, bem como do índice de confiança do empresário industrial – apenas indústria extrativa (Tabela 29).

**Tabela 29** - Modelo Econométrico - Indústria Extrativa

Coeficientes	Estimativa	Erro Padronizado	t-valor	Pr(> t )		Coef. Beta
(Intercept)	95,5406	22,3603	4,2728	0,0001	***	6,6343
Ind Extrativas (1def)	0,2343	0,1050	2,2321	0,0291	*	0,2350
Ind Extrativas (2def)	0,1961	0,1004	1,9543	0,0550	,	0,1957
<b>Paralisacao Samarco</b>	<b>-23,6368</b>	<b>6,2258</b>	<b>-3,7966</b>	<b>0,0003</b>	<b>***</b>	<b>-0,6541</b>
Taxa de Juros (2def)	23,7365	6,3484	3,7390	0,0004	***	0,3020
<b>Cambio (1def)</b>	<b>-10,3623</b>	<b>4,0581</b>	<b>-2,5535</b>	<b>0,0131</b>	<b>*</b>	<b>-0,5197</b>
Perc Desocupacao	1,5356	0,9173	1,6740	0,0990	,	0,2265
<b>ICE Ind Extrativa</b>	<b>-0,9337</b>	<b>0,2646</b>	<b>-3,5295</b>	<b>0,0008</b>	<b>***</b>	<b>-0,5219</b>

Signif, codes: 0 '\*\*\*' 0,001 '\*\*' 0,01 '\*' 0,05 '.' 0,1 ' ' 1

Foram apurados o coeficiente de determinação  $R^2$  e o coeficiente de determinação ajustado  $R^2_{ajustado}$ , resultado em **0.8279** e **0.8091** respectivamente.

Também foi considerado o resultado do teste F, que apresentou p-valor menor que 0,01%, o que significa dizer que a hipótese nula de os coeficientes serem todos iguais a zero é rejeitada.

Ao verificar a estatística de Durbin-Watson (**1,8379**), com p-valor = **0,113**, observou-se que os resíduos para os dados individualizados não são autocorrelacionados ao nível de significância de 1%. Também nesse sentido, aferiu-se o teste de Kolmogorov-Sminorv nos resíduos obtendo-se uma estatística de teste igual a **0,0487** com p-valor de **0,9923**, no qual não rejeita-se a hipótese nula de que os dados seguem uma distribuição normal.

A partir da tabela 30, é possível verificar os coeficientes beta que apresentam os coeficientes das regressões padronizadas, e neste sentido, vale notar que a variável dummy que traduz a paralização da Samarco, e o índice de confiança relativo a indústria extrativa e a taxa de câmbio 1 mês são as variáveis mais importantes do modelo.

Para previsão 12 passos a frente, nota-se que não se poderia simplesmente utilizar os dados das variáveis explicativas como input na regressão. Desse modo, considerou-se utilizar previsão para todas 11 variáveis explicativas, conforme realizado na Indústria Geral. Neste sentido, acrescenta-se nessa seção, as previsões obtidas para Indústria Extrativa (1 e 2 defasagens), Taxa de Juros Selic (1 defasagem), e Taxa de Cambio (1 defasagem), pois todas demais regressoras selecionadas para o modelo econométrico da indústria extrativa estão dispostos na previsão da indústria geral apresentada na seção anterior.

- Indústria Extrativa (1 e 2 def): considerou-se a previsão obtida via Box-Jenkins, modelo SARIMA(0,1,1)(0,0,2)[12], por apresentar menor erro se comparado com a modelagem Holt-Winters, a partir da Tabela 30 é possível consultar os valores previstos para 1e 2 defasagens.

**Tabela 30** - Indústria Extrativa - Realizado x Previsto (2017)

Mês	Ind Ext (Real)	Ind Extrativa 1 def (Previsão)	Ind Extrativa 2 def (Previsão)
nov/16	81,500		
dez/16	85,900		
jan/17	82,400	85,9000	81,5000
fev/17	75,700	84,4778	85,9000
mar/17	82,400	83,6899	84,4778
abr/17	80,600	84,3882	83,6899
mai/17	84,200	83,9771	84,3882
jun/17	81,600	84,6881	83,9771
jul/17	72,200	83,5051	84,6881
ago/17	83,400	83,8281	83,5051
set/17	78,300	85,2192	83,8281
out/17	79,500	84,6402	85,2192
nov/17	78,100	83,4492	84,6402
dez/17	81,400	79,3064	83,4492

- **Taxa de Juros (1def):** realizou-se previsão 11 passos a frente utilizando as metodologias Box-Jenkins, Holt-Winters Aditivo e Multiplicativo (Resultados dispostos no Apêndice D), onde obteve-se melhor resultado a partir no Holt-Winters Aditivo.

Considerando-se defasagem igual a 2, considerou-se apenas os 10 passos a frente na previsão, pois deve-se considerar o mês de dezembro de 2016 para atualizar a previsão de janeiro de 2017, e a previsão de janeiro a novembro de 2017 (com base nos dados até dezembro de 2016 apenas), para atualizar a previsão de março a dezembro de 2017 (Tabela 31).

**Tabela 31** – Taxa de Juros (Selic) - Realizado x Previsto (2017)

Mês	Selic (Real)	Selic 1 def (Previsão)	Selic 2 def (Previsão)
nov/16	1,0383		
dez/16	1,1233	1,0383	
jan/17	1,0861	1,1233	1,0383
fev/17	0,8651	1,0144	1,1233
mar/17	1,0521	0,9067	1,0144
abr/17	0,7866	0,9992	0,9067
mai/17	0,9271	0,8927	0,9992
jun/17	0,8089	0,8995	0,8927
jul/17	0,7979	0,8870	0,8995
ago/17	0,8023	0,8889	0,8870
set/17	0,6385	0,8855	0,8889
out/17	0,6439	0,8105	0,8855
nov/17	0,5682	0,7907	0,8105
dez/17	0,5384	0,7421	0,7907

Os testes de estacionariedade ADF e PP também foram aplicados a essa série de dados (Tabela 13) onde foi verificada a mesma ordem de integração I(1) da variável dependente em questão.

- **Taxa de Cambio (2def):** realizou-se previsão 10 passos a frente utilizando as metodologias Box-Jenkins, Holt-Winters Aditivo e Multiplicativo (Resultados dispostos no Apêndice D), onde obteve-se melhor resultado a partir no Holt-Winters Multiplicativo (Tabela 32).

**Tabela 32** – Taxa de Cambio - Realizado x Previsto (2017)

Mês	Cambio (Real)	Cambio 1 def (Previsão)	Cambio 2 def (Previsão)
nov/16	3,3390		
dez/16	3,3510		
jan/17	3,1970	3,3510	3,3390
fev/17	3,1030	3,2899	3,3510
mar/17	3,1270	3,1182	3,2899
abr/17	3,1400	3,0660	3,1182
mai/17	3,2090	3,0538	3,0660
jun/17	3,2970	3,1973	3,0538
jul/17	3,2050	3,3125	3,1973
ago/17	3,1530	3,1796	3,3125
set/17	3,1380	3,1259	3,1796
out/17	3,1960	3,2269	3,1259
nov/17	3,2570	3,2532	3,2269
dez/17	3,2970	3,3255	3,2532

Os testes de estacionariedade ADF e PP também foram aplicados a essa série de dados (Tabela 15) onde foi verificada a mesma ordem de integração I(1) da variável dependente em questão.

- **ICE Ind Extrativa:** os valores absolutos observados de janeiro a dezembro de 2016 foram replicados para 2017.

Os testes de estacionariedade ADF e PP também foram aplicados a essa série de dados, onde também observou-se a necessidade e tomar uma diferença dos dados para tornar a série estacionaria (Tabela 33). Ou seja, é verificada a mesma ordem de integração I(1) da variável dependente em questão.

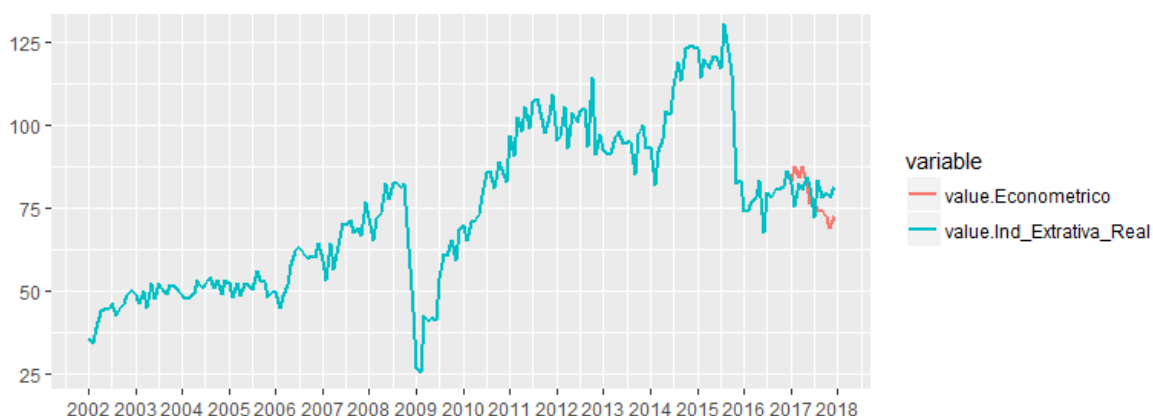
**Tabela 33** - Testes de Estacionariedade - ICE Indústria Extrativa (jan-11 a dez-16)

ICE Ind Extrativa	Teste	Estatística de Teste	p-valor calculado
Série Original	ADF	-1,086	0,288
	PP	-0,553	0,565
Série (1 Diferença)	ADF	-11,84	<0,01
	PP	-99,4	<0,01

Logo, apresenta-se a partir das previsões das variáveis regressoras para atualização do modelo econométrico para todo ano de 2017, as medidas de acurácia na Tabela 34. Em comparação com o resultado obtido via Holt-Winters, o Modelo Econométrico apresentou ótimos resultados.

**Tabela 34** - Medidas de Acurácia - Modelo Econométrico - Indústria Extrativa

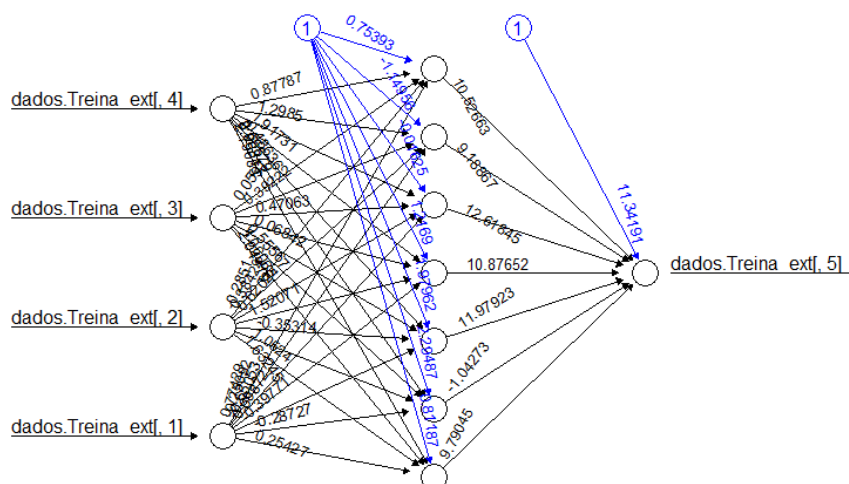
Modelo	MAE	MSE	RMSE	MAPE	SMAPE	U-Theil
Modelo Econométrico	611,3096	51,4193	7,1707	7,7408	0,0777	0,0449

**Figura 24** – Indústria Extrativa: Realizado (2002-2017) x Previsto via Modelagem Econométrica (2017)

Na modelagem via Redes Neurais Artificiais, foi considerado o modelo multicamadas, conforme pode ser visto na Figura 25.

Considerou-se como entrada, 4 defasagens do próprio índice de produção industrial, extrativa imitando um processo autoregressivo. E também se limitou os dados até 2016



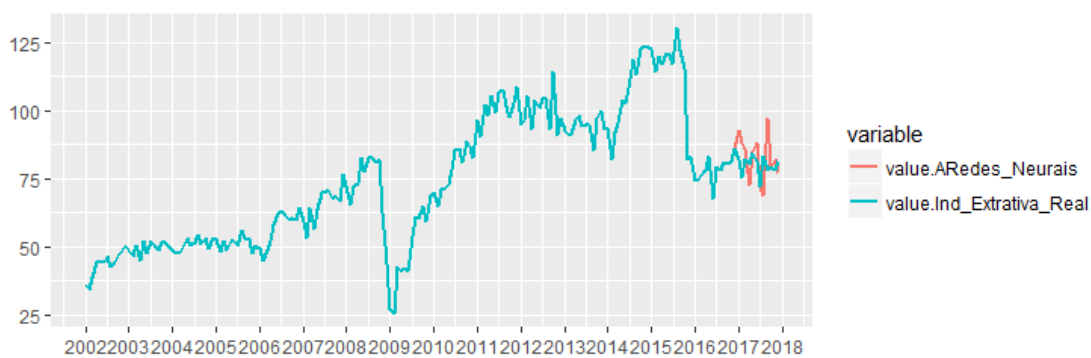


**Figura 25** - Modelo Redes Neurais Artificiais - Indústria Extrativa

Para tanto, a partir da Tabela 35 é possível constatar as medidas de acurácia.

**Tabela 35** - Medidas de Acurácia - Redes Neurais - Indústria Extrativa

Modelo	MAE	MSE	RMSE	MAPE	SMAPE	U-Theil
Modelo Redes Neurais	709,9747	83,3942	9,1320	8,8743	0,0864	0,0560



**Figura 26** – Indústria Extrativa: Realizado (2002-2017) x Previsto via Redes Neurais Artificiais (2017)

A partir dos 4 principais modelos acima apresentados, a proposta do trabalho é aplicar métodos de combinação de previsão para melhoria de performance. Foram considerados 3 métodos: combinação por média, combinação por matriz de variância, e combinação a partir de técnica de regressão por MQO.

Os resultados de todas previsões estão dispostos na Tabela 36, bem como os valores observados (reais) para o mesmo período.

**Tabela 36 - Previsões Indústria Extrativa**

Mês	Ind_Extr	BJ	HW	Econométrico	RNA	Comb_Media	Comb_Var	Comb_Reg
Jan	82,4000	84,4778	85,0438	84,4301	92,6228	86,6436	85,6555	80,2246
Feb	75,7000	83,6899	82,5957	88,6383	88,5483	85,8681	85,2006	78,4699
Mar	82,4000	84,3882	89,4480	84,3804	86,0608	86,0693	85,5459	81,8592
Apr	80,6000	83,9771	86,9471	89,3804	72,7930	83,2744	84,2052	78,6980
May	84,2000	84,6881	90,2175	83,8980	84,7538	85,8894	85,5504	82,0409
Jun	81,6000	83,5051	85,9841	78,7556	88,4635	84,1771	83,5992	80,3853
Jul	72,2000	83,8281	90,6525	78,5255	72,7316	81,4344	82,4726	80,7836
Aug	83,4000	85,2192	90,7342	73,0067	68,6568	79,4042	81,4040	79,8777
Sep	78,3000	84,6402	83,9384	74,3730	97,1443	85,0240	83,9586	80,1146
Oct	79,5000	83,4492	87,1694	72,9151	78,8838	80,6044	81,2591	79,6991
Nov	78,1000	79,3064	81,7820	68,8824	82,1548	78,0314	77,8954	78,0597
Dec	81,4000	79,8741	85,9000	73,3669	76,9494	79,0226	79,1976	79,5874

O método de combinação de previsão via regressão por MQO foi o que apresentou melhor resultado (Tabela 37), e o segundo melhor método seria a combinação por variância mínima, seguido pelo método de combinação por média. Esses resultados vão de encontro com tantos outros estudos que apontam os ganhos de performance ao considerar métodos de combinação previsão para garantir uma estimativa pontual com melhor acurácia.

**Tabela 37 - Medidas de Acurácia - Todos Modelos- Indústria Extrativa**

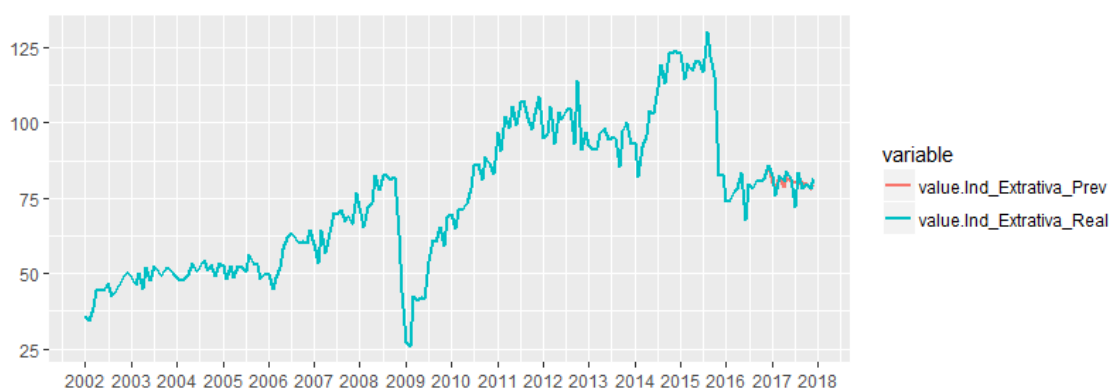
Modelos Propostos	MAE	MSE	RMSE	MAPE	SMAPE	U-Theil
BJ	369,1265	23,7901	4,8775	4,4086	0,0458	0,0298
HW	671,7721	59,8860	7,7386	8,5589	0,0808	0,0463
Econométrico	611,3096	51,4193	7,1707	7,7408	0,0777	0,0449
RNA	709,9747	83,3942	9,1320	8,8743	0,0864	0,0560
Comb_Media	404,3865	25,4032	5,0402	4,8324	0,0499	0,0309
Comb_Var	374,5832	23,2576	4,8226	4,4780	0,0464	0,0296
Comb_Reg	222,7862	9,5963	3,0978	2,7784	0,0282	0,0194

A Equação 4.2.1 apresenta os pesos estimados por MQO pra cada uma dos quatro métodos de previsão.

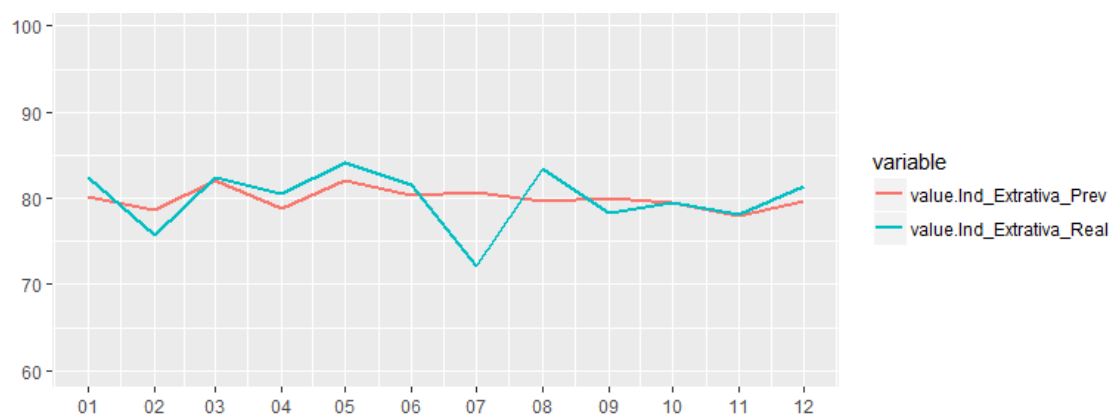
$$IE_k = 38,72 - 0,24BJ_k + 0,58HW_k + 0,03Eco_k + 0,14RNA_k \quad \text{Equação 4.2.1}$$

Onde  $IE_k$  representa o índice da indústria extrativa no  $k$  – *ésimo* mês de previsão, ou seja,  $k = 1, 2, \dots, 12$  representando os meses de janeiro a dezembro de 2017.  $BJ_k$  representa a previsão obtida a partir do modelo SARIMA para o mês  $k$ ,  $HW_k$  representa a previsão obtida a partir do método Holt-Winters para o mês  $k$ ,  $Eco_k$  representa a previsão obtida a partir do modelo econométrico para o mês  $k$ , e  $RNA_k$  representa a previsão obtida a partir do método de redes neurais artificiais para o mês  $k$ .

A partir da Figuras 27 e 28 é possível observar o ajuste obtido a partir do método de combinação de previsão via regressão por MQO.



**Figura 27-** Indústria Extrativa: Realizado (2002-2017) x Previsto via Combinação Previsão (2017)



**Figura 28** – Indústria Extrativa - combinação de previsão via regressão por MQO

### 4.3 INDÚSTRIA DE TRANSFORMAÇÃO

Foram considerados os Índices Mensais de Produção de Transformação no período de janeiro de 2002 a dezembro de 2016, totalizando 168 observações para modelagem via Box-Jenkins, Holt-Winters e Redes Neurais Artificiais. Enquanto no Modelo Econométrico, devido a limitação na disponibilização dos dados das variáveis regressoras, foram considerados apenas os dados no período de janeiro de 2011 a dezembro de 2016, totalizando apenas 60 observações.

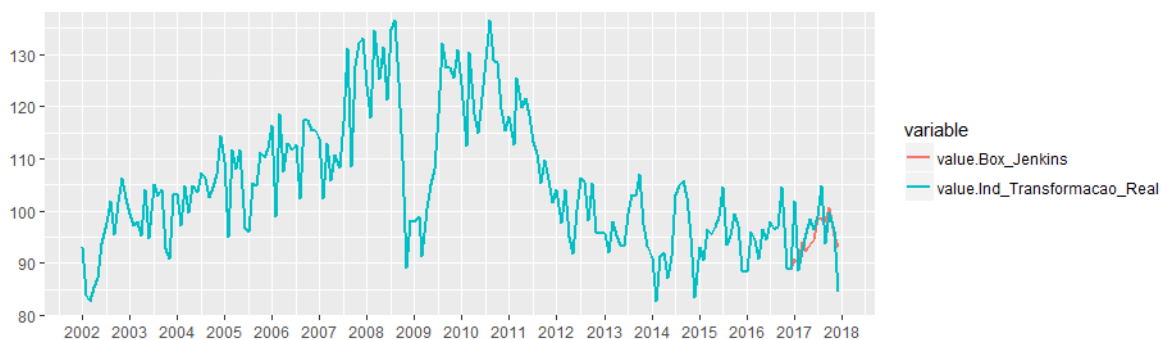
Na modelagem Box-Jenkins foram testados diversos modelos, e os oito melhores modelos obtidos são apresentados na Tabela 38, destacando-se o modelo SARIMA(2,1,2)(1,0,1)[12] que apresentou melhores medidas de acurácia (Tabela 39).

**Tabela 38 - Modelos Box-Jenkins para Produção Indústria De Transformação**

Modelo	AR1	AR2	MA1	MA2	SAR1	SAR2	SMA1
SARIMA(1,1,1)(0,0,0)[12]	0,5283	[na]	-0,8482	[na]	[na]	[na]	[na]
	0,1409	[na]	0,0971	[na]	[na]	[na]	[na]
SARIMA(0,1,1)(1,0,0)[12]	[na]	[na]	-0,3417	[na]	0,2239	[na]	[na]
	[na]	[na]	0,0754	[na]	0,0734	[na]	[na]
SARIMA(2,1,1)(1,0,1)[12]	0,6067	0,1413	-0,9355	[na]	0,9094	[na]	-0,7410
	0,0885	0,0804	0,0447	[na]	0,0624	[na]	0,1031
SARIMA(2,1,1)(1,0,0)[12]	0,5851	0,1707	-0,9465	[na]	0,2377	[na]	[na]
	0,0859	0,0789	0,0377	[na]	0,0761	[na]	[na]
SARIMA(2,1,1)(0,0,1)[12]	0,5583	0,1768	-0,9344	[na]	[na]	[na]	0,1536
	0,0869	0,0795	0,0411	[na]	[na]	[na]	0,0628
SARIMA(2,1,2)(1,0,1)[12]	1,0702	-0,1300	-1,4023	0,4023	0,9186	[na]	-0,7479
	0,3234	0,2835	0,3050	0,3048	0,0625	[na]	0,1081
SARIMA(2,1,2)(2,0,1)[12]	0,4006	0,2650	-0,7255	-0,1848	0,6934	0,1537	-0,6016
	1,6803	1,0517	1,7329	1,5503	0,1669	0,1008	0,1624
SARIMA(2,1,0)(1,0,1)[12]	-0,2875	-0,1077	[na]	[na]	0,9265	[na]	-0,7716
	0,0747	0,0753	[na]	[na]	0,0568	[na]	0,0983

**Tabela 39 - Medidas de Acurácia para modelos Box-Jenkins – Indústria De Transformação**

Modelo Proposto	BIC	AIC	mae	mse	rmse	mape	smape	u-theil
SARIMA(1,1,1)(0,0,0)[12]	1226,873	1217,3106	503,6399	38,1252	6,1746	5,4206	0,0530	0,0326
SARIMA(0,1,1)(1,0,0)[12]	1219,311	1209,6911	623,0968	50,4535	7,1031	6,8491	0,0660	0,0380
SARIMA(2,1,1)(1,0,1)[12]	1211,513	1192,3886	470,0347	34,9008	5,9077	5,1240	0,0494	0,0314
SARIMA(2,1,1)(1,0,0)[12]	1222,946	1207,0088	497,9156	37,4192	6,1171	5,3834	0,0525	0,0324
SARIMA(2,1,1)(0,0,1)[12]	1226,452	1210,5151	523,7975	40,3301	6,3506	5,6732	0,0552	0,0337
SARIMA(2,1,2)(1,0,1)[12]	1216,850	1194,5388	372,8539	24,9077	4,9908	3,9749	0,0393	0,0261
SARIMA(2,1,2)(2,0,1)[12]	1219,702	1194,2031	454,8345	33,7624	5,8105	4,9508	0,0477	0,0309
SARIMA(2,1,0)(1,0,1)[12]	1213,190	1197,2534	521,3838	39,3139	6,2701	5,7095	0,0550	0,0335



**Figura 29** – Indústria de Transformação: Realizado (2002-2017) x Previsto via Box-Jenkins (2017)

O resultado da modelagem via Holt-Winters, mesmo apresentando bons resultados, foi aquém do modelo SARIMA acima apresentado.

Neste método, conforme expectativa a priori, os melhores resultados foram obtidos ao considerar o método aditivo, que melhorou a performance da previsão 12 passos a frente em cerca de 50% em comparação com os demais modelos desenvolvidos segundo esta mesma metodologia (Tabela 40).

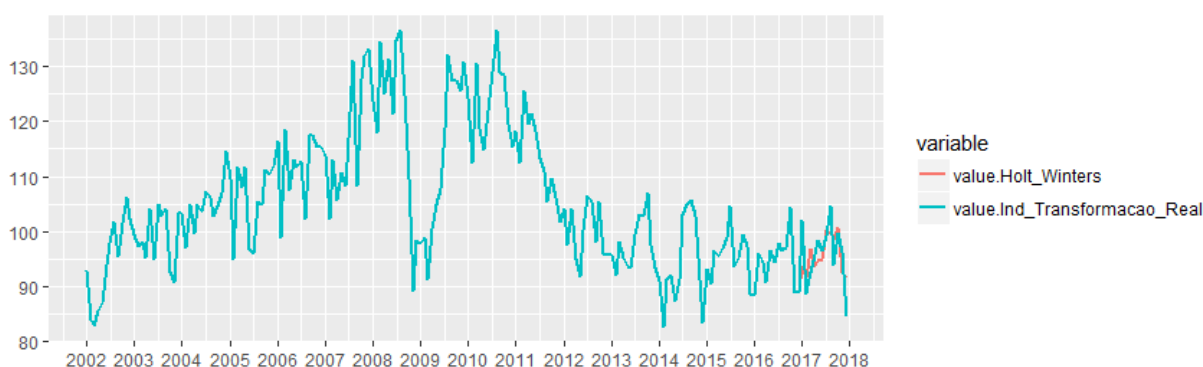
**Tabela 40** - Acurácia para modelos Holt-Winters – Indústria De Transformação

Modelo	MAE	MSE	RMSE	MAPE	SMAPE	U-Theil
Aditivo (Série Original)	378,3113	18,9993	4,3588	3,9774	0,0397	0,0227
Multiplicativo (Série Original)	373,7807	19,4818	4,4138	3,9209	0,0393	0,0231

As constantes de alisamento para a série  $\nabla^1 Y_t$ , sendo  $Y_t$  a série dos índices de produção industrial no modelo selecionado por:  $\alpha = 0,607$ ,  $\beta = 0,004$ , e  $\gamma = 0,466$ . E os coeficientes estimados estão dispostos na Tabela 41.

**Tabela 41** - Coeficientes Modelo Holt-Winters – Indústria De Transformação

Coeficientes	Valor Estimado
A	95,9233
B	0,2053
C	[na]
S1	-2,4500
S2	-4,7168
S3	0,1223
S4	-3,1206
S5	-2,3945
S6	-2,1665
S7	2,9285
S8	2,2149
S9	0,8021
S10	2,8249
S11	-5,6548
S12	-6,8771

**Figura 30** – Indústria de Transformação: Realizado (2002-2017) x Previsto via Holt-Hinters (2017)

No Modelo Econométrico também foram avaliadas como variáveis regressoras: taxa de juros, taxa de inflação, taxa de cambio, índice de confiança do empresário industrial, utilização da capacidade instalada, e percentual de desocupação.

Constatou-se a partir da Modelagem Econométrica, que o Índice de Produção Industrial de Transformação dependente do seu valor observado com uma e duas defasagens, da taxa de juros com uma defasagem e duas defasagens, do percentual de pessoas desocupadas também influencia, bem como o índice de confiança do empresário industrial – apenas indústria de Transformação (Tabela 42).

**Tabela 42** - Modelo Econométrico - Indústria de Transformação

Coeficientes	Estimativa	Erro Padronizado	t-valor	Pr(> t )		Coef. Beta
(Intercept)	76,480	29,322	2,608	0,011	*	8,7136
<b>Ind_Transformacao_1</b>	<b>0,680</b>	<b>0,097</b>	<b>7,020</b>	<b>0,000</b>	<b>***</b>	<b>0,6883</b>
Selic (1 def)	-10,737	8,522	-1,260	0,212		-0,2252
<b>Selic (2 def)</b>	<b>14,679</b>	<b>8,071</b>	<b>1,819</b>	<b>0,073</b>	.	<b>0,3065</b>
<b>Cambio (2 def)</b>	<b>-4,587</b>	<b>1,916</b>	<b>-2,394</b>	<b>0,020</b>	*	<b>-0,3764</b>
Consumo Energia (1 def)	-0,003	0,002	-1,539	0,129		-0,2120

Signif, codes: 0 '\*\*\*' 0,001 '\*\*' 0,01 '\*' 0,05 '.' 0,1 ' ' 1

Foram apurados o coeficiente de determinação  $R^2$  e o coeficiente de determinação ajustado  $R^2_{ajustado}$ , resultado em **0,6580** e **0,6321** respectivamente.

Também foi considerado o resultado do teste F, que apresentou p-valor menor que 0,01%, o que significa dizer que a hipótese nula de os coeficientes serem todos iguais a zero é rejeitada.

Ao verificar a estatística de Durbin-Watson (**1,9846**), com p-valor = **0,261**, observou-se que os resíduos para os dados individualizados não são autocorrelacionados ao nível de significância de 1%. Também nesse sentido, aferiu-se o teste de Kolmogorov-Sminorv nos resíduos obtendo-se uma estatística de teste igual a **0,0704** com p-valor de **0,842**, no qual não rejeita-se a hipótese nula de que os dados seguem uma distribuição normal.

A partir da tabela 42, é possível verificar os coeficientes beta que apresentam os coeficientes das regressões padronizadas, e neste sentido, as variáveis taxa de cambio defasada em 2 meses e Selic defasada em 2 meses são as variáveis mais importantes do modelo, perdendo apenas para primeira defasagem da própria série industrial de transformação.

Para previsão 12 passos a frente, nota-se que não se poderia simplesmente utilizar os dados das variáveis explicativas como input na regressão. Desse modo, considerou-se utilizar previsão para todas 11 variáveis explicativas, conforme realizado na Indústria Geral.

- Indústria de Transformação (1 e 2 def): considerou-se a previsão obtida via Box-Jenkins, modelo SARIMA(2,1,2)(1,0,1)[12] por apresentar menor erro se comparado com a modelagem Holt-Winters, a partir da tabela 43 é possível consultar os valores previstos para 1e 2 defasagens.

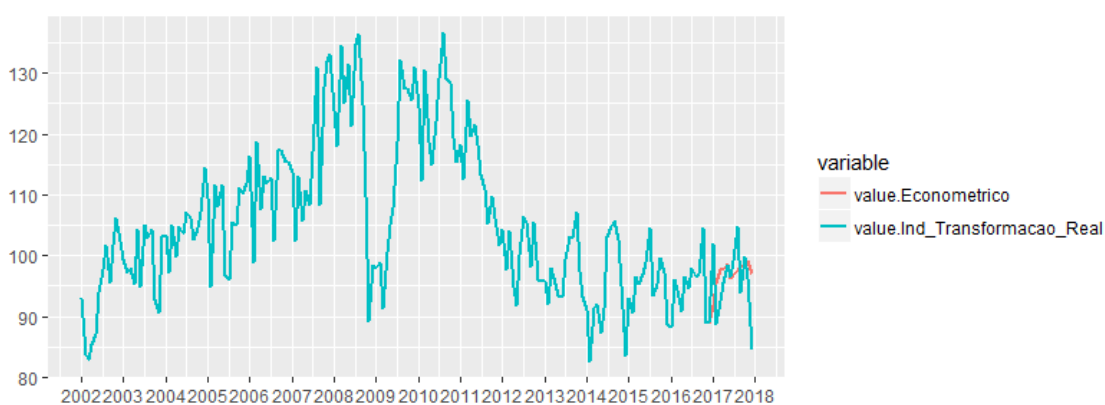
**Tabela 43** – Indústria de Transformação - Realizado x Previsto (2017)

Mês	Ind Transf (Real)	Ind de Trasnformação 1 def (Previsão)	Ind de Trasnformação 2 def (Previsão)
nov/16	89,0000		
dez/16	88,9000		
jan/17	101,9000	88,9000	89,0000
fev/17	88,7000	90,6887	88,9000
mar/17	92,5000	89,7273	90,6887
abr/17	95,6000	93,9993	89,7273
mai/17	98,4000	92,3130	93,9993
jun/17	96,4000	93,5579	92,3130
jul/17	99,0000	94,5047	93,5579
ago/17	104,7000	98,7805	94,5047
set/17	93,8000	98,4851	98,7805
out/17	99,7000	97,8427	98,4851
nov/17	96,2000	100,5730	97,8427
dez/17	84,4000	95,2519	100,5730

Logo, a partir da atualização do modelo econométrico foram consideradas as previsões das variáveis regressoras para todo ano de 2017, a partir da Tabela 44. Em comparação com os resultados obtidos via Box-Jenkins e Holt-Winters, o Modelo Econométrico apresentou ótimos resultados.

**Tabela 44** - Medidas de Acurácia - Modelo Econométrico – Indústria De Transformação

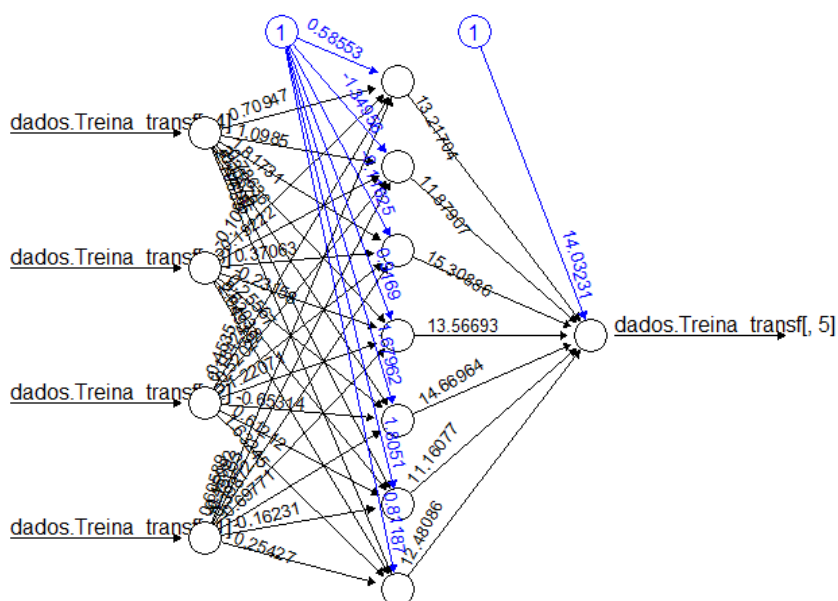
Modelo	MAE	MSE	RMSE	MAPE	SMAPE	U-Theil
Modelo Econométrico	457,2161	32,4954	5,7005	4,8244	0,0481	0,0297

**Figura 31** - Indústria de Transformação: Realizado (2002-2017) x Previsto via Modelagem Econométrica (2017)



Na modelagem via Redes Neurais Artificiais, foi considerado o modelo multicamadas, conforme pode ser visto na Figura 32.

Considerou-se como entrada, 4 defasagens do próprio índice de produção industrial, de Transformação imitando um processo autoregressivo. E também se limitou os dados até 2016.

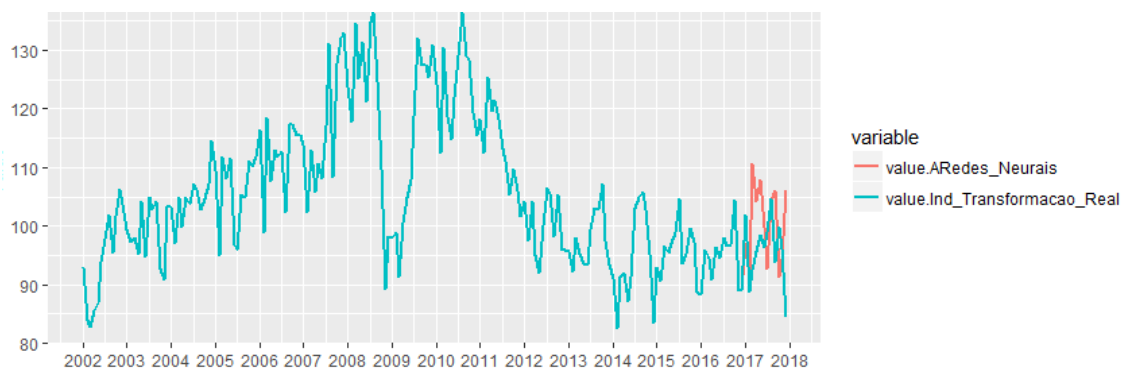


**Figura 32** - Modelo Redes Neurais Artificiais - Indústria de Transformação

Para tanto, a partir da Tabela 45 é possível constatar as medidas de acurácia.

**Tabela 45** - Medidas de Acurácia - Redes Neurais - Indústria de Transformação

Modelo	MAE	MSE	RMSE	MAPE	SMAPE	U-Theil
Modelo Redes Neurais	784,3013	95,5034	9,7726	8,2230	0,0807	0,0505



**Figura 33** - Indústria de Transformação: Realizado (2002-2017) x Previsto via Redes Neurais (2017)

A partir dos 4 principais modelos acima apresentados, a proposta do trabalho é aplicar métodos de combinação de previsão para melhoria de performance. Foram considerados 3 métodos: combinação por média, combinação por matriz de variância, e combinação a partir de técnica de regressão por MQO.

Os resultados de todas as previsões estão dispostos na Tabela 46, bem como os valores observados (reais) para o mesmo período.

**Tabela 46** - Previsões Indústria de Transformação

Mês	Ind_Trasnf	BJ	HW	Econométrico	RNA	Comb_Media	Comb_Var	Comb_Reg
Jan	101,9000	90,6887	93,6785	90,1787	96,1195	92,6664	92,1700	97,4867
Feb	88,7000	89,7273	91,6170	94,7575	100,2620	94,0909	92,4474	91,7190
Mar	92,5000	93,9993	96,6614	95,6486	102,8637	97,2933	96,1160	97,0350
Apr	95,6000	92,3130	93,6238	95,1576	85,8039	91,7246	92,9724	93,5068
May	98,4000	93,5579	94,5551	95,8793	97,9994	95,4979	94,8295	94,4815
Jun	96,4000	94,5047	94,9884	94,8498	99,0522	95,8488	95,1291	95,9441
Jul	99,0000	98,7805	100,2886	95,5177	84,9485	94,8838	97,5430	101,2031
Aug	104,7000	98,4851	99,7803	98,1648	96,1436	98,1435	98,7373	98,8740
Sep	93,8000	97,8427	98,5728	98,3806	114,7726	102,3922	99,5756	97,9043
Oct	99,7000	100,5730	100,8009	97,7901	98,9964	99,5401	99,9003	100,6867
Nov	96,2000	95,2519	92,5265	98,5292	103,5171	97,4562	95,5798	90,8096
Dec	84,4000	93,0819	91,5093	94,9880	86,3599	91,4848	92,3768	91,6492

O método de combinação de previsão via regressão por MQO foi o que apresentou melhor resultado (Tabela 47), e o segundo melhor método seria o combinação por variância mínima. Esses resultados vão de encontro com tantos outros estudos que apontam os ganhos de performance ao considerar métodos de combinação previsão para garantir uma estimativa pontual com melhor acurácia.

**Tabela 47** - Medidas de Acurácia - Todos Modelos- Indústria de Transformação

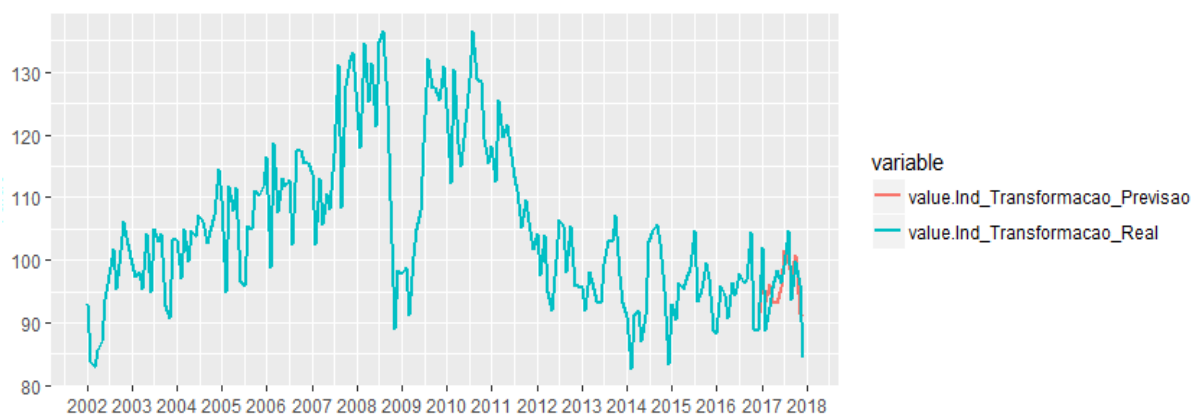
Modelos Propostos	MAE	MSE	RMSE	MAPE	SMAPE	U-Theil
BJ	372,8539	24,9077	4,9908	3,9749	0,0393	0,0261
HW	378,3113	18,9993	4,3588	3,9774	0,0397	0,0227
Econométrico	457,2161	32,4954	5,7005	4,8244	0,0481	0,0297
RNA	784,3013	95,5034	9,7726	8,2230	0,0807	0,0505
Comb_Media	454,2693	28,8830	5,3743	4,7553	0,0476	0,0280
Comb_Var	387,9583	23,1789	4,8145	4,0931	0,0408	0,0251
Comb_Reg	368,2881	17,2769	4,1566	3,8700	0,0388	0,0216

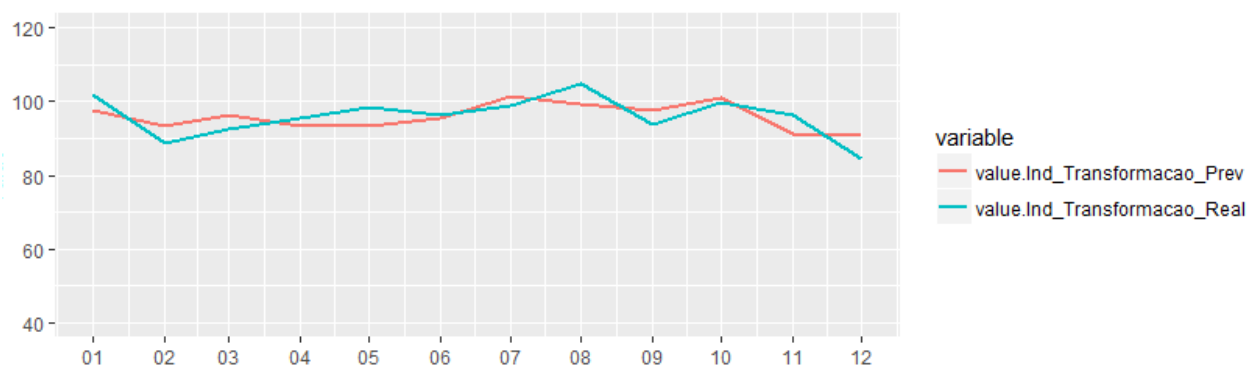
A Equação 4.3.1 apresenta os pesos estimados por MQO pra cada uma dos quatro métodos de previsão.

$$IT_k = 53,25 + 0,18BJ_k + 1,01HW_k - 0,78Eco_k + 0,03RNA_k \quad \text{Equação 4.3.1}$$

Onde  $IT_k$  representa o índice da indústria de transformação no  $k$  – *ésimo* mês de previsão, ou seja,  $k = 1, 2, \dots, 12$  representando os meses de janeiro a dezembro de 2017.  $BJ_k$  representa a previsão obtida a partir do modelo SARIMA para o mês  $k$ ,  $HW_k$  representa a previsão obtida a partir do método Holt-Winters para o mês  $k$ ,  $Eco_k$  representa a previsão obtida a partir do modelo econométrico para o mês  $k$ , e  $RNA_k$  representa a previsão obtida a partir do método de redes neurais artificiais para o mês  $k$ .

A partir da Figura 19 é possível observar o ajuste obtido a partir do método de combinação de previsão via regressão por MQO.

**Figura 34** - Indústria de Transformação: Realizado (2002-2017) x Combinação de Previsão (2017)



**Figura 35** - Previsão de Produção Industrial de Transformação - combinação de previsão via regressão por MQO

## 5 CONCLUSÃO

Este estudo propôs e avaliou diferentes métodos de previsão para o índice mensal de produção geral, extrativa e de transformação: Box-Jenkins, Holt-Winters, Modelo Econométrico e Redes Neurais Artificiais. A avaliação se deu por vários procedimentos de teste e de modo intensivo.

O poder preditivo dos diferentes modelos envolvendo essas variáveis foi avaliado a partir do uso de diferentes funções dos erros de previsão de fora da amostra preditiva, que são minimizados, conhecidas como medidas de acurácia.

Adicionalmente, consideramos também a combinação das previsões de diferentes modelos, seguindo a literatura de combinação de previsões iniciada por Bates e Granger (1969), considerando-se as combinações de previsão por média simples, variância e regressão por MQO. Complementarmente, também se avaliou os melhores resultados a partir das medidas de acurácia.

Diante da análise do modelo apresentado neste trabalho podemos perceber que todas as variáveis explicativas (apresentadas na modelagem econométrica), tais como taxa de juros (selic), taxa de câmbio, índice de confiança do empresário industrial, utilização da capacidade instalada, percentual de desocupação, consumo de energia no setor industrial, e dummy para informar período de paralisação da empresa Samarco, são significativas para mostrar o desempenho da produção industrial capixaba.

É interessante destacar que algumas variáveis foram utilizadas no modelo considerando sua defasagem e não no período corrente, mostrando o dinamismo da ciência econômica, pois todas as decisões tomadas pelos gestores de política econômica demandam um tempo até que surtam efeitos em todas as esferas da sociedade.

Observou-se que a combinação dos diferentes modelos de previsão, seja por média aritmética simples ou por média ponderada, gerou, de forma geral, os melhores modelos de previsão para fora da amostra.

Entretanto, houve casos em que modelos individuais, em especial, os obtidos via modelagem Box-Jenkins, tiveram um desempenho razoável e mesmo superior ao das combinações. Entretanto, diante dos resultados obtidos, nossa sugestão seria o uso dessas combinações de previsões como forma final de prever os índices de produção industrial capixaba.

Nossa lista de sugestões para aperfeiçoamentos futuros inclui avaliar a inclusão de novas variáveis para modelo econométrico, em especial, séries que reflitam estritamente a economia capixaba, para melhoria de performance dos modelos econométricos; inserir uma dummy que reflita quebra estrutural de novembro de 2015; e propor intervalo de confiança, além da previsão pontual.

## REFERÊNCIAS BIBLIOGRÁFICAS

- BATES, J.M.E.; GRANGER, C.W.J. (1969) The Combination of Forecasts. Operational Research Society. Vol. 20, No. 4.
- BONELLI, R.; PESSÔA, S.A. (2010) Desindustrialização no Brasil: um resumo da evidência.
- BOX, G.E.P; JENKINS, G.M. (1976) Time Series Analysis: Forecasting and Control. San Francisco: Holden-Day.
- BOX, G.E.P; JENKINS, G.M. (2008) Time Series Analysis: Forecasting and Control. 4 ed. San Francisco: Holden-Day.
- BROCKWELL, P.J, DAVIS, R.A. (2002) Introduction to Time series and Forecasting, 2a ed. New York: Springer.
- CHATFIELD, C. (1978) The Holt-Winters Forecasting Procedure. Journal of the Royal Statistical Society. Series C (Applied Statistics), Vol. 27, No. 3, pp. 264-279.
- CHATFIELD, C.; YAR, M. (1988) Holt-Winters Forecasting: Some Practical Issues. Journal of the Royal Statistical Society. Series D (The Statistician), Vol. 37, 129-140.
- CHATFIELD, C. (1989) The Analysis of Time Series: An Introduction. London, Chapman and Hall.
- COSTANTINI, M. PAPPALARDO, C. (2010) A hierarchical procedure for the combination forecasts. International Journal of Forecasting, V.26 n4, 725-743.
- DE LACERDA, A.C; BOCCHI, J.I.; REGO J.M.; BORGES M.A.; MARQUES, M.M. (2005) Economia Brasileira. Editora Saraiva. 4ª Edição.
- DICKEY, D. & FULLER, W. (1979), Distribution of the estimators for autoregressive time series with a unit root, Journal of the American Statistical Association 74(366), 427-431.
- FLORES, J.H.F. (2009) Comparação de Modelos MLP/RNA e modelos Box-Jenkins em séries temporais não lineares. Dissertação de Mestrado. Universidade Federal do Rio Grande do Sul.

Fuller, W. A. (1976), Introduction to the statistical time series, Wiley.

GELPER, S., FRIED, R.; CROUX, C. (2010). Robust forecasting with exponential and Holt-Winters smoothing, Journal of Forecasting, 29, 285-300.

GRANGER, C.W.; RAMANATHAN, R. (1984) Improved methods of combining forecasts. Journal of Forecasting, v.3, n.2, p.97–204.

GUJARATI, D. N (2006). Econometria básica. São Paulo. Editora Campos.

HAIR, J.R.; ANDERSON, R.E.; TATHAN, R.L.; BLACK, W.C. (2005) Análise Multivariada de Dados. 5 edição. Porto Alegre: Bookman.

HOLLAUER, G.; ISSLER, J.V.; NOTINI, H.H. (2008) Prevendo o crescimento da produção industrial usando um numero limitado de combinação de previsões. Economia Aplicada, São Paulo, V12, N2, P 117-198.

HOLT, C.C. (1957). Forecasting trends and seasonals by exponentially weighted averages, Carnegie Institute of Technology, Pittsburgh ONR memorandum no. 52.

INDICADORES CONJUNTURAIS DA INDÚSTRIA: produção / IBGE, Coordenação de Indústria. - 2. ed. - Rio de Janeiro: IBGE, 2015. (Relatórios metodológicos, ISSN 0101-2843 ; volume 31).

JUNIOR, M.A.L., MEDEIROS, R.L. (2014). O Desafio da Inovação Industrial no Espírito Santo. X Congresso Nacional de Excelência em Gestão.

FORTUNATO, D.O.B. (2011) Revista de História da Ufes – Dimensões. ISSN: 2179-8869; volume 27; pag 40-62.

LAWTON, J.H.; BIGNELL, D.E.; BOLTON, B.; BLOEMERS, G.F.; EGGLETON, P.; HAMMOND P.M.; HODDA, M.; HOLT, R.D.; LARSEN, T.B.; MAWDSLEY, N. A., STORK, N.E., SRIVASTAVA D. S.; WATT A.D. (1998). Biodiversity inventories, indicator taxa and effects of habitat modification in tropical forest. Nature volume 391, pages 72–76.



MARTINS, V.L.M. (2011) Comparação de Combinação de Previsões Correlacionadas e Não Correlacionadas com as Suas Previsões Individuais: Um Estudo com Series Industriais. Dissertação de Mestrado. Universidade Federal do Rio Grande do Sul.

MAKRIDAKIS, S.; WHEELWRIGHT, S. C.; HYNDMAN, R. J. (1998) Forecasting: methods and applications. John Wiley & Sons, 3 rd edition.

MATOS, R.S. (1998) Modelos Bayesianos de Previsão Aplicados à Análise do Comportamento da Produção Industrial de Minas Gerais. Universidade Federal de Juiz de Fora (UFJF). Minas Gerais.

MENEZES, L. M.; BUNN, D. W.; TAYLOR, J. W. (2000) Review of guidelines for the use of combined forecasts. *European Journal of Operational Research*, v.120, n.1, p.190–204.

MONTGOMERY, C.A.; SINGH, H. (1984) Diversification na Systematic Risk. *Strategic Management*. Volume 5.

MORETTIN, P.A., TOLOI, C.M.C.. (2004). *Análise de Series Temporais*. São Paulo: Edgard Blucher).

MORETTIN, P.A., TOLOI, C. (2006). *Análise de Series Temporais*. 2 ed. São Paulo: Edgard Blucher).

MOTA, F.C.M. (2002) *Integração e Dinâmica Regional: o caso capixaba (1960-2000)*. Tese de Doutorado – Universidade Estadual de Campinas, Instituto de Economia.

MULLER, B.; REINHARDT, J.; STRICKLAND, M.T. (1995) *Neural networks: an introduction*. 2ª edição. New York: Springer-Verlag.

NADER, Maria Beatriz. *Revista História Hoje*, Associação Nacional de História (ANPUH), 2007. (Volume 4, número 12)

PRATES, A.M.Q. (2010) O setor exportador do Espírito Santo nos anos recentes: Estrutura, Vantagens Comparativas e Comércio Intra-Industrial. Instituto de Ciências Econômicas Aplicadas. Universidade Federal de Alfenas (ICSA - UNIFAL – MG) Minas Gerais.

PEREIRA, Luiz Carlos Bresser. O empresário industrial e a revolução brasileira. *Revista de Administração de Empresas*, v. 3, n. 8, p. 11-27, 1963.

PHILLIPS, P. & PERRON, P. (1988), Testing for unit root in time series regression, *Biometrika* 75, 335-346.

PORTAL DA INDÚSTRIA, (2018) Disponível em: <<http://perfilestados.portaldaindustria.com.br/estado/es>>

VERSIANI, F.R. (1984). Industrialização: a década de 20 e a depressão. *Pesquisa e Planejamento Economico*. Vol 14. Número 1.

SHCHERBAKOV, M.V.; BREBELS, A.; SHCHERBAKOVA, N. L.; TYUKOV, A. P.; JANOVSKY, T.A.; KAMAIEV, V.A. (2013) A Survey of Forecast Error Measures. *World Applied Sciences Journal*, v.24, n.24 (Information Technologies in Modern Industry, Education & Society), p.171-176.

SILVA, L.H.A; CARMO, L.P.S; e FRANÇA, M.T.A. (2002) Modelo de Previsão da Produção Industrial Brasileira. Universidade Federal de Juiz de Fora – UFJF. Minas Gerais.

SUZIGAN, W; FURTADO, J. (2006). Política Industrial e desenvolvimento. (*Revista de Economia Política*. Volume 26. N2, pagina 163-185.

WINTERS, P.R. (1960). Forecasting Sales by Exponentially Weighted Moving Averages. *Management Science*, Vol 6, p. 324 – 342.

WEI, W.S. (1990) *Time Series Analysis: Univariate and Multivariate Methods*. AddisonWesley

## APENDICE A - Seção B da CNAE 2.0

Subclasse	Seção B - INDÚSTRIAS EXTRATIVAS (Denominação)
	<b>EXTRAÇÃO DE CARVÃO MINERAL</b>
0500-3/01	Extração de carvão mineral
0500-3/02	Beneficiamento de carvão mineral
	<b>EXTRAÇÃO DE PETRÓLEO E GÁS NATURAL</b>
0600-0/01	Extração de petróleo e gás natural
0600-0/02	Extração e beneficiamento de xisto
0600-0/03	Extração e beneficiamento de areias betuminosas
	<b>EXTRAÇÃO DE MINERAIS METÁLICOS</b>
	<b>Extração de minério de ferro</b>
0710-3/01	Extração de minério de ferro
0710-3/02	Pelotização, sinterização e outros beneficiamentos de minério de ferro
	<b>Extração de minerais metálicos não ferrosos</b>
	<b>Extração de minério de alumínio</b>
0721-9/01	Extração de minério de alumínio
0721-9/02	Beneficiamento de minério de alumínio
	<b>Extração de minério de estanho</b>
0722-7/01	Extração de minério de estanho
0722-7/02	Beneficiamento de minério de estanho
	<b>Extração de minério de manganês</b>
0723-5/01	Extração de minério de manganês
0723-5/02	Beneficiamento de minério de manganês
	<b>Extração de minério de metais preciosos</b>
0724-3/01	Extração de minério de metais preciosos
0724-3/02	Beneficiamento de minério de metais preciosos
	<b>Extração de minerais radioativos</b>
0725-1/00	Extração de minerais radioativos
	<b>Extração de minerais metálicos não ferrosos não especificados anteriormente</b>
0729-4/01	Extração de minérios de nióbio e titânio
0729-4/02	Extração de minério de tungstênio
0729-4/03	Extração de minério de níquel
0729-4/04	Extração de minérios de cobre, chumbo, zinco e outros minerais metálicos não ferrosos não especificados anteriormente
0729-4/05	Beneficiamento de minérios de cobre, chumbo, zinco e outros minerais metálicos não ferrosos não especificados anteriormente
	<b>EXTRAÇÃO DE MINERAIS NÃO METÁLICOS</b>
	<b>Extração de pedra, areia e argila</b>
0810-0/01	Extração de ardósia e beneficiamento associado
0810-0/02	Extração de granito e beneficiamento associado
0810-0/03	Extração de mármore e beneficiamento associado
0810-0/04	Extração de calcário e dolomita e beneficiamento associado
0810-0/05	Extração de gesso e caulim
0810-0/06	Extração de areia, cascalho ou pedregulho e beneficiamento associado
0810-0/07	Extração de argila e beneficiamento associado
0810-0/08	Extração de saibro e beneficiamento associado
0810-0/09	Extração de basalto e beneficiamento associado
0810-0/10	Beneficiamento de gesso e caulim associado à extração
0810-0/99	Extração e britamento de pedras e outros materiais para construção e beneficiamento associado

(continuação)

<b>Subclasse</b>	<b>Seção B - INDÚSTRIAS EXTRATIVAS (Denominação)</b>
	<b>Extração de outros minerais não metálicos</b>
	<b>Extração de minerais para fabricação de adubos, fertilizantes e outros produtos químicos</b>
0891-6/00	Extração de minerais para fabricação de adubos, fertilizantes e outros produtos químicos
	<b>Extração e refino de sal marinho e sal-gema</b>
0892-4/01	Extração de sal marinho
0892-4/02	Extração de sal-gema
0892-4/03	Refino e outros tratamentos do sal
	<b>Extração de gemas (pedras preciosas e semipreciosas)</b>
0893-2/00	Extração de gemas (pedras preciosas e semipreciosas)
	<b>Extração de minerais não metálicos não especificados anteriormente</b>
0899-1/01	Extração de grafita
0899-1/02	Extração de quartzo
0899-1/03	Extração de amianto
0899-1/99	Extração de outros minerais não metálicos não especificados anteriormente
	<b>ATIVIDADES DE APOIO À EXTRAÇÃO DE MINERAIS</b>
	<b>Atividades de apoio à extração de petróleo e gás natural</b>
0910-6/00	Atividades de apoio à extração de petróleo e gás natural
	<b>Atividades de apoio à extração de minerais, exceto petróleo e gás natural</b>
0990-4/01	Atividades de apoio à extração de minério de ferro
0990-4/02	Atividades de apoio à extração de minerais metálicos não ferrosos
0990-4/03	Atividades de apoio à extração de minerais não metálicos

Fonte: IBGE

## APENDICE B - Seção C da CNAE 2.0

Divisão	Seção C - INDÚSTRIAS DE TRANSFORMAÇÃO (Denominação)
10	Fabricação de produtos alimentícios
11	Fabricação de bebidas
12	Fabricação de produtos do fumo
13	Fabricação de produtos têxteis
14	Confecção de artigos do vestuário e acessórios
15	Preparação de couros e fabricação de artefatos de couro, artigos para viagem e calçados
16	Fabricação de produtos de madeira
17	Fabricação de celulose, papel e produtos de papel
18	Impressão e reprodução de gravações
19	Fabricação de coque, de produtos derivados do petróleo e de biocombustíveis
20 B	Fabricação de sabões, detergentes, produtos de limpeza, cosméticos, produtos de perfumaria e de higiene pessoal
20 C	Fabricação de outros produtos químicos
21	Fabricação de produtos farmoquímicos e farmacêuticos
22	Fabricação de produtos de borracha e de material plástico
23	Fabricação de produtos de minerais não metálicos
24	metalurgia
25	Fabricação de produtos de metal, exceto máquinas e equipamentos
26	Fabricação de equipamentos de informática, produtos eletrônicos e ópticos
27	Fabricação de máquinas, aparelhos e materiais elétricos
28	Fabricação de máquinas e equipamentos
29	Fabricação de veículos automotores, reboques e carrocerias
30	Fabricação de outros equipamentos de transporte, exceto veículos automotores
31	Fabricação de móveis
32	Fabricação de produtos diversos
33	Manutenção, reparação e instalação de máquinas e equipamentos

Fonte: IBGE

### APENDICE C – Lista de Atividades com Peso Normalizado

Código e Descrição das Atividades Descrição dos Produtos N= número de Produtos = 24	Peso dos Produtos			
	Peso de Seleção	Atividade	Peso Normalizado Seção	Indústria Geral
<b>12 Indústrias Extrativas</b>				
Gás natural	0,6592	0,0279	0,0279	0,0093
Minérios de ferro beneficiados (classificados, concentrados, pelotizados, sinterizados, etc)	20,9468	0,8867	0,8867	0,2965
Óleos brutos de petróleo	2,0172	0,0854	0,0854	0,0286
<b>15 Fabricação de produtos alimentícios e bebidas</b>				
Bombons contendo cacau	2,9872	0,4860	0,0635	0,0423
Bombons, caramelos, balas, pastilhas e outros confeitados, sem cacau	0,6062	0,0986	0,0129	0,0086
Café torrado e moído - inclusive aromatizado	0,1050	0,0171	0,0022	0,0015
Carnes de bovinos congeladas	0,3044	0,0495	0,0065	0,0043
Carnes de bovinos frescas ou refrigeradas	0,2306	0,0375	0,0049	0,0033
Chocolate em barras e em tabletes	0,4670	0,0760	0,0099	0,0066
Farinha de trigo	0,2426	0,0395	0,0052	0,0034
Leite esterilizado/ UHT/ Longa Vida	0,1709	0,0278	0,0036	0,0024
Massas alimentícias secas (macarrão, espaguete, etc)	0,1503	0,0245	0,0032	0,0021
Produtos de salamaria (salame, salsicha, lingüiça, mortadela, presunto, patês, etc)	0,3107	0,0506	0,0066	0,0044
Refrigerantes	0,5710	0,0929	0,0121	0,0081
<b>21 Fabricação de celulose, papel e produtos de papel</b>				
Pastas químicas de madeira, processo sulfato, branqueadas	14,6666	1,0000	0,3120	0,2076
<b>26 Fabricação de produtos de minerais não-metálicos</b>				
Cimentos Portland, exceto brancos	2,1586	0,4211	0,0459	0,0306
Granito talhado, serrado ou trabalhado de outro modo , inclusive chapas de granito para pias	1,0733	0,2094	0,0228	0,0152
Ladrilhos e placas (lajes) de cerâmica para pavimentação ou revestimento, esmaltados ( lado >= 7cm) (azulejos de cerâmica decorados)	0,8654	0,1688	0,0184	0,0123
Mármore e alabastro talhado, serrado ou trabalhado de outro modo, inclusive chapas de mármore para pias	0,2376	0,0464	0,0051	0,0034
Massa de concreto preparada para construção	0,3617	0,0706	0,0077	0,0051
Pias, banheiras, bidês e semelhantes para uso sanitário, de cerâmica, exceto porcelana	0,4293	0,0838	0,0091	0,0061

(continuação)

Código e Descrição das Atividades Descrição dos Produtos N= número de Produtos = 24	Peso de Seleção	Peso dos Produtos		
		Atividade	Seção	Indústria Geral
<b>27 Metalurgia básica</b>				
Ferro-gusa	0,5484	0,0260	0,0117	0,0078
Lingotes, blocos, tarugos ou placas de aços ao carbono	20,0933	0,9534	0,4274	0,2845
Perfis médios e pesados de aços ao carbono, laminados a quente	0,4327	0,0205	0,0092	0,0061

Fonte: IBGE

## APENDICE D – Previsão Variáveis Regressoras

- **SELIC (2 defasagens):**

Método Box-Jenkins				
Series: selic #2011a2016				
ARIMA(1,1,2)(0,1,1)[12]				
Coefficients:				
	ar1	ma1	ma2	sma1
	0.5175	-1.2112	0.7951	-0.5168
s.e.	0.1968	0.1461	0.1065	0.4620
sigma^2 estimated as 0.003821: log likelihood=80.03				
AIC=-150.05 AICc=-148.92 BIC=-139.66				

Holt-Winters Aditivo	
Holt-winters exponential smoothing with trend and additive seasonal component.	
Call: Holtwinters(x = selic #2011a2016, seasonal = c("additive"))	
Smoothing parameters: alpha: 0.2707111 beta : 1 gamma: 0.4643731	
Coefficients:	
	[,1]
a	1.035517679
b	-0.024259473
s1	0.003114185
s2	-0.080265438
s3	0.036435851
s4	-0.045772281
s5	-0.014744034
s6	-0.002998864
s7	0.023247063
s8	0.044066063
s9	-0.006674949
s10	-0.002262406
s11	-0.026506166
s12	0.069285704

Holt-Winters Multiplicativo	
Holt-winters exponential smoothing with trend and multiplicative seasonal component.	
Call: Holtwinters(x = selic #2011a2016, seasonal = c("multiplicative"))	
Smoothing parameters: alpha: 0.24802	



```
beta : 1
gamma: 0.4541692
```

```
Coefficients:
      [,1]
a    1.04769815
b   -0.01437759
s1   1.00377586
s2   0.91036544
s3   1.03323684
s4   0.96387168
s5   0.99985601
s6   1.00565054
s7   1.04375999
s8   1.05850507
s9   1.00403292
s10  1.00729574
s11  0.96684482
s12  1.05602232
```

Modelo Proposto	MAE	MSE	RMSE	MAPE	SMAPE	U-Theil
Box-Jenkins	25,8487	0,0903	0,3005	34,5475	0,2769	0,1546
<b>Holt-Winters Aditivo</b>	<b>8,7120</b>	<b>0,0094</b>	<b>0,0970</b>	<b>11,4094</b>	<b>0,1064</b>	<b>0,0554</b>
Holt-Winters Multiplicativo	14,2556	0,0284	0,1685	19,1925	0,1678	0,0924

- **CAMBIO (1 defasagem):**

#### Método Box-Jenkins

```
Series: Cambio #2011a2016
ARIMA(0,1,1)
```

```
Coefficients:
      ma1
      0.4937
s.e.  0.1014
```

```
sigma^2 estimated as 0.009363: log likelihood=65.44
AIC=-126.88   AICC=-126.71   BIC=-122.36
```

#### Método Holt-Winters Aditivo

```
Holt-winters exponential smoothing without trend and with additive seasonal component.
```

```
Call:
Holtwinters(x = Cambio #2011a2016, beta = FALSE, seasonal = c("additive"))
```

```
Smoothing parameters:
alpha: 1
beta : FALSE
gamma: 0.8756778
```

```
Coefficients:
      [,1]
```

```

a    3.27982986
s1  -0.01345486
s2  -0.12382986
s3  -0.07641319
s4  -0.03866319
s5   0.07017014
s6   0.11154514
s7  -0.10882986
s8  -0.08274653
s9   0.06250347
s10  0.06862847
s11  0.05992014
s12  0.07117014

```

#### Método Holt-Winters Multiplicativo

Holt-winters exponential smoothing without trend and with multiplicative seasonal component.

Call:

```
Holtwinters(x = Cambio#2011a2016, beta = FALSE, seasonal = c("multiplicative"))
```

Smoothing parameters:

```
alpha: 0.8458458
beta  : FALSE
gamma: 1
```

Coefficients:

```

      [,1]
a    3.1817598
s1   1.0339830
s2   0.9800110
s3   0.9636056
s4   0.9597809
s5   1.0048924
s6   1.0410857
s7   0.9993085
s8   0.9824348
s9   1.0141782
s10  1.0224414
s11  1.0451858
s12  1.0531908

```

Modelo Proposto	MAE	MSE	RMSE	MAPE	SMAPE	U-Theil
Box-Jenkins	12,3511	0,0184	0,1355	3,9110	0,0382	0,0209
Holt-Winters Aditivo	9,5723	0,0116	0,1076	3,0074	0,0295	0,0167
<b>Holt-Winters Multiplicativo</b>	<b>4,9962</b>	<b>0,0034</b>	<b>0,0585</b>	<b>1,5731</b>	<b>0,0157</b>	<b>0,0092</b>

- **CONSUMO DE ENERGIA (2 defasagens):**

Método Box-Jenkins
Series: Energia#2011a2016
ARIMA(0,1,0)(0,1,1)[12]
Coefficients:
smal
-0.5759
s.e.    0.1898
sigma^2 estimated as 35622: log likelihood=-394.76
AIC=793.52  AICC=793.73  BIC=797.67

Método Holt-Winters Aditivo
Holt-winters exponential smoothing without trend and with additive seasonal component.
Call:
Holtwinters(x = Energia#2011a2016, beta = FALSE, seasonal = c("additive"))
Smoothing parameters:
alpha: 0.8069873
beta : FALSE
gamma: 1
Coefficients:
[,1]
a    13639.983263
s1   -933.865939
s2   -292.273917
s3    55.475495
s4   173.404097
s5    -5.766984
s6   -195.175170
s7   -91.738130
s8   368.857182
s9   275.149201
s10  288.832574
s11  196.542629
s12  -252.983263

Método Holt-Winters Multiplicativo
Holt-winters exponential smoothing without trend and with multiplicative seasonal component.
Call:
Holtwinters(x = Energia_2011a2016, beta = FALSE, seasonal = c("multiplicative"))
Smoothing parameters:
alpha: 0.7962311

```
beta : FALSE
gamma: 1
```

```
Coefficients:
```

```
      [,1]
a  1.363243e+04
s1  9.369053e-01
s2  9.824225e-01
s3  1.004260e+00
s4  1.011315e+00
s5  9.994940e-01
s6  9.869434e-01
s7  9.932824e-01
s8  1.023826e+00
s9  1.017118e+00
s10 1.018356e+00
s11 1.012688e+00
s12 9.819967e-01
```

Modelo Proposto	MAE	MSE	RMSE	MAPE	SMAPE	U-Theil
Box-Jenkins	42598,1700	224577,8000	473,8964	3,0543	0,0311	0,0174
<b>Holt-Winters Aditivo</b>	<b>24491,2000</b>	<b>83931,1800</b>	<b>289,7088</b>	<b>1,7628</b>	<b>0,0178</b>	<b>0,0106</b>
Holt-Winters Multiplicativo	25958,7100	89031,4100	298,3813	1,8642	0,0189	0,0109

**APENDICE E – Rotina R Studio (Indústria Geral)**

```
## Pacotes
library("openxlsx")
library("ggplot2")
library("dplyr")
library("tidyr")
library("stringr")
library("forecast")
library("rugarch")
library("ForecastCombinations")
library("randtests")
library("aTSA")
library("nnet")
library("urca")
library("lmtest")
library("neuralnet")

#leitura dos dados - producao industrial
dados <- read.xlsx("C://Users/Taizi/Desktop/Modelagem/dados_r.xlsx")
attach(dados)
dados <- ts(dados,start=c(2002,1),frequency = 12)
Ind_G_x <- ts(Industria_Geral[1:180],freq = 12, start = 2002)
Ind_E_x <- ts(Industrias_Extrativas[1:180],freq = 12, start = 2002)
Ind_T_x <- ts(Industrias_Transformacao[1:180],freq = 12, start = 2002)

# Grafico das Séries Industriais
datas <- seq(as.Date(paste(c(start(dados),1), collapse="/")),
            by = "month", length.out = length(dados[,1]))
dados.df <- data.frame(date = datas, value = dados)
min <- as.Date("2002-1-1")
max <- NA

df <- dados.df %>%
```

```

select(date,value.Industria_Geral, value.Industrias_Extrativas,value.Industrias_Transformacao)
%>% gather(key = "variable", value = "value",-date)

#head(df)
p1 <- ggplot(df, aes(x = date, y = value)) + geom_line(aes(color = variable), size = 1)
p1 + scale_x_date(limits = c(min, max),date_labels = "%Y", date_breaks = "1 year")

# Grafico das Séries Industriais com 1 diferença
dados.df2 <- data.frame(date = datas[2:194], value = diff(dados))

df2 <- dados.df2 %>%
  select(date,value.Industria_Geral, value.Industrias_Extrativas,value.Industrias_Transformacao)
%>%
  gather(key = "variable", value = "value",-date)

p2 <- ggplot(df2, aes(x = date, y = value)) + geom_line(aes(color = variable), size = 1)
p2 + scale_x_date(limits = c(min, max),date_labels = "%Y", date_breaks = "1 year")

# TESTES DE ESTACIONARIEDADE
# H0: a série não é estacionária
# H1: a série é estacionária

summary(ur.df(Industria_Geral, type='none', lags=0)) # adf.test(Industria_Geral)
summary(ur.df(Industrias_Extrativas, type='none', lags=0)) # adf.test(Industrias_Extrativas)
summary(ur.df(Industrias_Transformacao, type='none', lags=0)) # adf.test(Industrias_Transformacao)
summary(ur.df(diff(Industria_Geral), type='none', lags=0)) # adf.test(diff(Industria_Geral))
summary(ur.df(diff(Industrias_Extrativas), type='none', lags=0)) # adf.test(Industrias_Extrativas)
summary(ur.df(diff(Industrias_Transformacao), type='none', lags=0)) #
adf.test(Industrias_Transformacao)
pp.test(Industria_Geral)
pp.test(Industrias_Extrativas)
pp.test(Industrias_Transformacao)

pp.test(diff(Industria_Geral))

```

```
pp.test(diff(Industrias_Extrativas))
pp.test(diff(Industrias_Transformacao))

#leitura dos dados - modelo econometrico
dados_2 <- read.xlsx("C://Users/Taizi/Desktop/Modelagem/dados_r_multivariada.xlsx")
dados_reg = dados_2[1:72,]
attach(dados_reg)

# verificando a ordem de integração das regressoras

#selic
summary(ur.df(Selic, type='none', lags=0))
pp.test(Selic)
summary(ur.df(diff(Selic), type='none', lags=0))
pp.test(diff(Selic))

#cambio
summary(ur.df(Cambio_4, type='none', lags=0))
summary(ur.df(diff(Cambio_4), type='none', lags=0))
pp.test(Cambio_4)
pp.test(diff(Cambio_4))

#Perc_Desocupacao
summary(ur.df(Perc_Desocupacao, type='none', lags=0))
summary(ur.df(diff(Perc_Desocupacao), type='none', lags=0))
pp.test(Perc_Desocupacao)
pp.test(diff(Perc_Desocupacao))

#ICE_Ind_Ext
summary(ur.df(ICE_Ind_Ext, type='none', lags=0))
summary(ur.df(diff(ICE_Ind_Ext), type='none', lags=0))
pp.test(ICE_Ind_Ext)
pp.test(diff(ICE_Ind_Ext))

#ICE_Ind_Transf
```

```

summary(ur.df(ICE_Ind_Transf, type='none', lags=0))
summary(ur.df(diff(ICE_Ind_Transf), type='none', lags=0))
pp.test(ICE_Ind_Transf)
pp.test(diff(ICE_Ind_Transf))

#UCI_Ind_Geral
summary(ur.df(UCI_Ind_Geral, type='none', lags=0))
summary(ur.df(diff(UCI_Ind_Geral), type='none', lags=0))
pp.test(UCI_Ind_Geral)
pp.test(diff(UCI_Ind_Geral))

## ESPECIFICAÇÃO DE MEDIDAS DE ACURACIA
mae <- function(real,prev){
  mae <- mean(abs(real - prev))*100
  return (mae)}
mse <- function(real,prev){
  mse <- mean( (real - prev) ^ 2)
  return (mse)}
rmse <- function(real,prev){
  rmse <- sqrt(mean( (real - prev) ^ 2))
  return (rmse)}
mape <- function(real,prev){
  mape <- mean(abs((real - prev)/real))*100
  return (mape)}
smape <- function(real,prev){
  smape <- mean(abs((real - prev)/((abs(real)+abs(prev))/2)))
  return (smape)}
u_ttheil <- function(real,prev){
  u_ttheil <- (sqrt(mean((real - prev)^2)))/(sqrt(mean(real^2))+sqrt(mean(prev^2)))
  return (u_ttheil)}

## COMBINAÇÃO DE PREVISAO
Forecast_comb <- function(obs, fhat, fhat_new= NULL,
                          Averaging_scheme=c("simple", "ols", "robust", "cls", "variance based", "best")) {

```



```

pckg = c("quantreg", "quadprog")
temp <- unlist(lapply(pckg, require, character.only=T))
if (!all(temp==1) ) {
  stop("This function relies on packages \"quadprog\" and \"quantreg\".
  Use ?install.packages if they are not yet installed. \n")
}
mat_err <- apply(fhat, 2, function(x) obs - x)
TT <- NROW(fhat) # numero de linhas
TT_new <- NROW(fhat_new)
p <- NCOL(fhat) # numero de columnas
pred <- NULL
weights <- matriz(ncol=p, nrow = TT) # matriz numero de columnas x numero de linhas
## Subroutine needed:
sq_er <- function(obs, pred) { mean( (obs - pred)^2 ) } #erro quadratico medio

if(length(Averaging_scheme) != 1) {
  stop("Pick only one of the following:
  c(\"simple\", \"ols\", \"robust\", \"variance based\", \"cls\", \"best\")")
}

## Different forecast averaging schemes
## simple
if(Averaging_scheme== "simple") {
  pred <- apply(fhat, 1, mean)
  weights <- matriz( 1/p, nrow = 1, ncol = p)
  if (!is.null(fhat_new)) { pred_new <- apply(fhat_new, 1, mean) }
  ## OLS weights
} else if (Averaging_scheme== "ols") {
  weights <- lm(obs ~ fhat)$coef
  pred <- t(weights %*% t(cbind(rep(1, TT), fhat)))
  if (!is.null(fhat_new)) { pred_new <- t(weights %*% t(cbind(rep(1, TT_new), fhat_new))) }
  ## Robust weights
} else if (Averaging_scheme== "robust") {
  weights <- rq(obs ~ fhat)$coef
  pred <- t(weights %*% t(cbind(rep(1,TT), fhat)))

```

```

if (!is.null(fhat_new)) { pred_new <- t(weights %*% t(cbind(rep(1, TT_new), fhat_new))) }
## Based on the variance of the errors. Inverse of MSE
} else if (Averaging_scheme=="variance based") {
temp = apply(mat_err^2, 2, mean)/sum(apply(mat_err^2, 2, mean))
weights <- (1/temp)/sum(1/temp)
pred <- t(weights %*% t(fhat))
if ( !is.null(fhat_new) ) { pred_new <- t(weights %*% t(fhat_new)) }
## Using constraint least squares
} else if (Averaging_scheme=="cls") {
## Subroutine needed:
cls1 = function(y, predictions){
Rinv <- solve(chol(t(predictions) %*% predictions))
C <- cbind(rep(1, NCOL(predictions)), diag(NCOL(predictions)))
b = c(1, rep(0, NCOL(predictions)))
d = t(y) %*% predictions
qp1 = solve.QP(Dmat= Rinv, factorized= TRUE, dvec= d, Amat= C, bvec = b, meq = 1)
weights = qp1$sol
yhatcls = t(weights %*% t(predictions))
list(yhat= yhatcls, weights= weights)
}
##
weights <- cls1(obs, fhat)$weights
pred <- t(weights %*% t(fhat))
if (!is.null(fhat_new)) { pred_new <- t(weights %*% t(fhat_new)) }
## Picking the best performing model up to date according to squared loss function
} else if (Averaging_scheme=="best") {
temp <- apply(fhat, 2, sq_er, obs= obs)
weights <- rep(0, p)
weights[which.min(temp)] <- 1
pred <- t(weights %*% t(fhat))
if (!is.null(fhat_new)) { pred_new <- t(weights %*% t(fhat_new)) }
}
if (is.null(fhat_new)) { pred_new <- NULL }
return( list(fitted= pred, pred = pred_new, weights = weights ) )
}

```

```
### MODELAGEM DA INDÚSTRIA GERAL
```

```
### BOX JENKINS
```

```
IndGeral_auto <- auto.arima(Ind_G_x)#, d = 1) # ARIMA(2,1,0)(2,0,0)[12]
```

```
IndGeral_m1 = arima(Ind_G_x,order=c(2,1,0),seasonal=list(order=c(2,0,0),period=12))
```

```
IndGeral_m2 = arima(Ind_G_x,order=c(2,1,0),seasonal=list(order=c(1,0,0),period=12))
```

```
IndGeral_m3 = arima(Ind_G_x,order=c(2,1,0),seasonal=list(order=c(0,0,0),period=12))
```

```
IndGeral_m4 = arima(Ind_G_x,order=c(1,1,0),seasonal=list(order=c(2,0,0),period=12))
```

```
IndGeral_m5 = arima(Ind_G_x,order=c(1,1,0),seasonal=list(order=c(1,0,0),period=12))
```

```
IndGeral_m6 = arima(Ind_G_x,order=c(2,1,1),seasonal=list(order=c(1,0,0),period=12))
```

```
IndGeral_m7 = arima(Ind_G_x,order=c(2,1,2),seasonal=list(order=c(1,0,0),period=12))
```

```
IndGeral_m8 = arima(Ind_G_x,order=c(1,1,1),seasonal=list(order=c(2,0,0),period=12))
```

```
IndGeral_r1 = predict(IndGeral_m1,n.ahead = 12)
```

```
IndGeral_r2 = predict(IndGeral_m2,n.ahead = 12)
```

```
IndGeral_r3 = predict(IndGeral_m3,n.ahead = 12)
```

```
IndGeral_r4 = predict(IndGeral_m4,n.ahead = 12)
```

```
IndGeral_r5 = predict(IndGeral_m5,n.ahead = 12)
```

```
IndGeral_r6 = predict(IndGeral_m6,n.ahead = 12)
```

```
IndGeral_r7 = predict(IndGeral_m7,n.ahead = 12)
```

```
IndGeral_r8 = predict(IndGeral_m8,n.ahead = 12)
```

```
a1 = c(
```

```
  IndGeral_m1$aic,
```

```
  IndGeral_m1$loglik,
```

```
  mae(IndGeral_r1$pred,Industria_Geral[181:192]),
```

```
  mse(IndGeral_r1$pred,Industria_Geral[181:192]),
```

```
  rmse(IndGeral_r1$pred,Industria_Geral[181:192]),
```

```
  mape(IndGeral_r1$pred,Industria_Geral[181:192]),
```

```
  smape(IndGeral_r1$pred,Industria_Geral[181:192]),
```

```
  u_ttheil(IndGeral_r1$pred,Industria_Geral[181:192]))
```

```
a2 = c(
```

```

IndGeral_m2$aic,
IndGeral_m2$loglik,
mae(IndGeral_r2$pred,Industria_Geral[181:192]),
mse(IndGeral_r2$pred,Industria_Geral[181:192]),
rmse(IndGeral_r2$pred,Industria_Geral[181:192]),
mape(IndGeral_r2$pred,Industria_Geral[181:192]),
smape(IndGeral_r2$pred,Industria_Geral[181:192]),
u_ttheil(IndGeral_r2$pred,Industria_Geral[181:192]))

```

a3 =c(

```

IndGeral_m3$aic,
IndGeral_m3$loglik,
mae(IndGeral_r3$pred,Industria_Geral[181:192]),
mse(IndGeral_r3$pred,Industria_Geral[181:192]),
rmse(IndGeral_r3$pred,Industria_Geral[181:192]),
mape(IndGeral_r3$pred,Industria_Geral[181:192]),
smape(IndGeral_r3$pred,Industria_Geral[181:192]),
u_ttheil(IndGeral_r3$pred,Industria_Geral[181:192]))

```

a4 =c(

```

IndGeral_m4$aic,
IndGeral_m4$loglik,
mae(IndGeral_r4$pred,Industria_Geral[181:192]),
mse(IndGeral_r4$pred,Industria_Geral[181:192]),
rmse(IndGeral_r4$pred,Industria_Geral[181:192]),
mape(IndGeral_r4$pred,Industria_Geral[181:192]),
smape(IndGeral_r4$pred,Industria_Geral[181:192]),
u_ttheil(IndGeral_r4$pred,Industria_Geral[181:192]))

```

a5 =c(

```

IndGeral_m5$aic,
IndGeral_m5$loglik,
mae(IndGeral_r5$pred,Industria_Geral[181:192]),
mse(IndGeral_r5$pred,Industria_Geral[181:192]),
rmse(IndGeral_r5$pred,Industria_Geral[181:192]),

```

```

mape(IndGeral_r5$pred,Industria_Geral[181:192]),
smape(IndGeral_r5$pred,Industria_Geral[181:192]),
u_ttheil(IndGeral_r5$pred,Industria_Geral[181:192]))

```

```
a6 =c(
```

```

  IndGeral_m6$aic,
  IndGeral_m6$loglik,
  mae(IndGeral_r6$pred,Industria_Geral[181:192]),
  mse(IndGeral_r6$pred,Industria_Geral[181:192]),
  rmse(IndGeral_r6$pred,Industria_Geral[181:192]),
  mape(IndGeral_r6$pred,Industria_Geral[181:192]),
  smape(IndGeral_r6$pred,Industria_Geral[181:192]),
  u_ttheil(IndGeral_r6$pred,Industria_Geral[181:192]))

```

```
a7 =c(
```

```

  IndGeral_m7$aic,
  IndGeral_m7$loglik,
  mae(IndGeral_r7$pred,Industria_Geral[181:192]),
  mse(IndGeral_r7$pred,Industria_Geral[181:192]),
  rmse(IndGeral_r7$pred,Industria_Geral[181:192]),
  mape(IndGeral_r7$pred,Industria_Geral[181:192]),
  smape(IndGeral_r7$pred,Industria_Geral[181:192]),
  u_ttheil(IndGeral_r7$pred,Industria_Geral[181:192]))

```

```
a8 =c(
```

```

  IndGeral_m8$aic,
  IndGeral_m8$loglik,
  mae(IndGeral_r8$pred,Industria_Geral[181:192]),
  mse(IndGeral_r8$pred,Industria_Geral[181:192]),
  rmse(IndGeral_r8$pred,Industria_Geral[181:192]),
  mape(IndGeral_r8$pred,Industria_Geral[181:192]),
  smape(IndGeral_r8$pred,Industria_Geral[181:192]),
  u_ttheil(IndGeral_r8$pred,Industria_Geral[181:192]))

```

## ### HOLT-WINTERS

```
hw1 <- HoltWinters(Ind_G_x,seasonal=c("additive"))
hw2 <- HoltWinters(Ind_G_x,beta = FALSE,seasonal=c("additive"))
hw2_diff <- HoltWinters(diff(Ind_G_x),beta = FALSE,seasonal=c("additive"))
hw3 <- HoltWinters(Ind_G_x,seasonal=c("multiplicative"))
hw4 <- HoltWinters(Ind_G_x,beta = FALSE,seasonal=c("multiplicative"))

hw1p <- predict(hw1, n.ahead = 12)
hw2p <- predict(hw2, n.ahead = 12)
hw2p_diff <- predict(hw2_diff, n.ahead = 12)
hw3p <- predict(hw3, n.ahead = 12)
hw4p <- predict(hw4, n.ahead = 12)

hw1acuracia = c(
  mae(Industria_Geral[181:192],hw1p),
  mse(Industria_Geral[181:192],hw1p),
  rmse(Industria_Geral[181:192],hw1p),
  mape(Industria_Geral[181:192],hw1p),
  smape(Industria_Geral[181:192],hw1p),
  u_ttheil(Industria_Geral[181:192],hw1p))

hw2acuracia = c(
  mae(Industria_Geral[181:192],hw2p),
  mse(Industria_Geral[181:192],hw2p),
  rmse(Industria_Geral[181:192],hw2p),
  mape(Industria_Geral[181:192],hw2p),
  smape(Industria_Geral[181:192],hw2p),
  u_ttheil(Industria_Geral[181:192],hw2p))

hw2_diff_acuracia = c(
  mae(Industria_Geral[181:192],Industria_Geral[181:192]+hw2p_diff),
  mse(Industria_Geral[181:192],Industria_Geral[181:192]+hw2p_diff),
  rmse(Industria_Geral[181:192],Industria_Geral[181:192]+hw2p_diff),
  mape(Industria_Geral[181:192],Industria_Geral[181:192]+hw2p_diff),
```

```
smape(Industria_Geral[181:192],Industria_Geral[181:192]+hw2p_diff),  
u_ttheil(Industria_Geral[181:192],Industria_Geral[181:192]+hw2p_diff))
```

```
hw3acuracia = c(  
  mae(Industria_Geral[181:192],hw3p),  
  mse(Industria_Geral[181:192],hw3p),  
  rmse(Industria_Geral[181:192],hw3p),  
  mape(Industria_Geral[181:192],hw3p),  
  smape(Industria_Geral[181:192],hw3p),  
  u_ttheil(Industria_Geral[181:192],hw3p))
```

```
hw4acuracia = c(  
  mae(Industria_Geral[181:192],hw4p),  
  mse(Industria_Geral[181:192],hw4p),  
  rmse(Industria_Geral[181:192],hw4p),  
  mape(Industria_Geral[181:192],hw4p),  
  smape(Industria_Geral[181:192],hw4p),  
  u_ttheil(Industria_Geral[181:192],hw4p))
```

```
### MODELO ECONOMETRICO
```

```
mod_geral = lm(Ind_Geral ~  
  Ind_Geral_1  
  + Ind_Geral_2  
  + Paralisacao_Samarco  
  + Selic_1  
  + Selic_2  
  + Cambio_4  
  + Perc_Desocupacao  
  + ICE_Ind_Ext  
  + ICE_Ind_Transf  
  + UCI_Ind_Geral)
```

```
summary(mod_geral)
```

```
dwtest(mod_geral)

#rstandard(mod_geral) # resíduos padronizados
ks.test(rstandard(mod_geral),pnorm,0,1)
#H0: os dados seguem uma distribuição normal
#H1: os dados não seguem uma distribuição normal

sd_Ind_Geral = Ind_Geral / sd(Ind_Geral)
sd_Ind_Geral_1 = Ind_Geral_1 / sd(Ind_Geral_1)
sd_Ind_Geral_2 = Ind_Geral_2 / sd(Ind_Geral_2)
sd_Ind_Geral_3 = Ind_Geral_3 / sd(Ind_Geral_3)
sd_Ind_Extrativas = Ind_Extrativas / sd(Ind_Extrativas)
sd_Ind_Extrativas_1 = Ind_Extrativas_1 / sd(Ind_Extrativas_1)
sd_Ind_Extrativas_2 = Ind_Extrativas_2 / sd(Ind_Extrativas_2)
sd_Ind_Extrativas_3 = Ind_Extrativas_3 / sd(Ind_Extrativas_3)
sd_Ind_Transformacao = Ind_Transformacao / sd(Ind_Transformacao)
sd_Ind_Transformacao_1 = Ind_Transformacao_1 / sd(Ind_Transformacao_1)
sd_Ind_Transformacao_2 = Ind_Transformacao_2 / sd(Ind_Transformacao_2)
sd_Ind_Transformacao_3 = Ind_Transformacao_3 / sd(Ind_Transformacao_3)
sd_Paralisacao_Samarco = Paralisacao_Samarco / sd(Paralisacao_Samarco)
sd_Selic = Selic / sd(Selic)
sd_Selic_1 = Selic_1 / sd(Selic_1)
sd_Selic_2 = Selic_2 / sd(Selic_2)
sd_Cambio_4 = Cambio_4 / sd(Cambio_4)
sd_Perc_Desocupacao = Perc_Desocupacao / sd(Perc_Desocupacao)
sd_ICE_Ind_Ext = ICE_Ind_Ext / sd(ICE_Ind_Ext)
sd_ICE_Ind_Transf = ICE_Ind_Transf / sd(ICE_Ind_Transf)
sd_UCI_Ind_transf = UCI_Ind_transf / sd(UCI_Ind_transf)
sd_UCI_Ind_Ext = UCI_Ind_Ext / sd(UCI_Ind_Ext)
sd_UCI_Ind_Geral = UCI_Ind_Geral / sd(UCI_Ind_Geral)

sd_mod_geral = lm(sd_Ind_Geral ~
                    sd_Ind_Geral_1
```



```

+ sd_Ind_Geral_2
+ sd_Paralisacao_Samarco
+ sd_Selic_1
+ sd_Selic_2
+ sd_Cambio_4
+ sd_Perc_Desocupacao
+ sd_ICE_Ind_Ext
+ sd_ICE_Ind_Transf
+ sd_UCI_Ind_Geral)

```

```
#coeficiente beta
```

```
summary(sd_mod_geral)$coefficients[,1]
```

```
# estimando Ind_Geral_1
```

```
Ind_Geral_1_prev = c(Ind_Geral[72],2,3,4,5,6,7,8,9,10,11,12)
```

```
Ind_Geral_1_prev[2:12] <- Industria_Geral[181:191]+hw2p_diff[1:11]
```

```
# estimando Ind_Geral_2
```

```
Ind_Geral_2_prev = c(Ind_Geral[71],Ind_Geral[72],3,4,5,6,7,8,9,10,11,12)
```

```
Ind_Geral_2_prev[3:12] <- Industria_Geral[181:190]+hw2p_diff[1:10]
```

```
# estimando Paralisacao_Samarco
```

```
Paralisacao_Samarco_prev = c(1,1,1,1,1,1,1,1,1,1,1,1)
```

```
Paralisacao_Samarco_prev = ts(Paralisacao_Samarco_prev[1:12],freq = 12, start = 2017)
```

```
# estimando Selic_1
```

```
Selic_1_aux = ts(Selic_1[1:72],freq = 12, start = 2011)
```

```
Selic_1_prev <- predict(HoltWinters(Selic_1_aux,seasonal=c("additive")), n.ahead = 12)
```

```
Selic_1_prev[1] = Selic[72]
```

```
# estimando Selic_2
```

```
Selic_2_prev = c(Selic[71],Selic[72],3,4,5,6,7,8,9,10,11,12)
```

```
Selic_2_prev[3:12] = Selic_1_prev[2:11]
```

```
# estimando Cambio_4
Cambio_4_aux = ts(Cambio_4[1:72],freq = 12, start = 2011)
Cambio_4_prev <- predict(HoltWinters(Cambio_4_aux,seasonal=c("additive")), n.ahead = 12)

# estimando Perc_Desocupacao_prev
Perc_Desocupacao_prev = Perc_Desocupacao[61:72]
Perc_Desocupacao_prev = ts(Perc_Desocupacao_prev ,freq = 12, start = 2017)

# estimando ICE_Ind_Ext
ICE_Ind_Ext_prev = ICE_Ind_Ext[61:72]
ICE_Ind_Ext_prev = ts(ICE_Ind_Ext_prev ,freq = 12, start = 2017)

# estimando ICE_Ind_Transf
ICE_Ind_Transf_prev = ICE_Ind_Transf[61:72]
ICE_Ind_Transf_prev = ts(ICE_Ind_Transf_prev ,freq = 12, start = 2017)

# estimando UCI_Ind_Geral
UCI_Ind_Geral_prev = UCI_Ind_Geral[61:72]
UCI_Ind_Geral_prev = ts(UCI_Ind_Geral_prev ,freq = 12, start = 2017)

mod_econometrico_geral_prev = ( summary(mod_geral)$coefficients[1]
  + summary(mod_geral)$coefficients[2]*Ind_Geral_1_prev
  + summary(mod_geral)$coefficients[3]*Ind_Geral_2_prev
  + summary(mod_geral)$coefficients[4]*Paralisacao_Samarco_prev
  + summary(mod_geral)$coefficients[5]*Selic_1_prev
  + summary(mod_geral)$coefficients[6]*Selic_2_prev
  + summary(mod_geral)$coefficients[7]*Cambio_4_prev
  + summary(mod_geral)$coefficients[8]*Perc_Desocupacao_prev
  + summary(mod_geral)$coefficients[9]*ICE_Ind_Ext_prev
  + summary(mod_geral)$coefficients[10]*ICE_Ind_Transf_prev
  + summary(mod_geral)$coefficients[11]*UCI_Ind_Geral_prev)
```

```

acuracia_eco_geral = c(
  mae(Industria_Geral[181:192],mod_econometrico_geral_prev),
  mse(Industria_Geral[181:192],mod_econometrico_geral_prev),
  rmse(Industria_Geral[181:192],mod_econometrico_geral_prev),
  mape(Industria_Geral[181:192],mod_econometrico_geral_prev),
  smape(Industria_Geral[181:192],mod_econometrico_geral_prev),
  u_ttheil(Industria_Geral[181:192],mod_econometrico_geral_prev))

#acuracia_eco_geral

#### REDES NEURAIIS
#http://pedrounb.blogspot.com.br/2013/01/redes-neurais-no-r-aplicacoes-em.html
#Dados para o neuralnet
t0<-as.numeric((Industria_Geral[1:180])) #Cinco meses antes
t0<-t0[-((length(t0)-4):length(t0))]
t1<-as.numeric(Industria_Geral[1:180]) [-1] #Quatro meses antes
t1<-t1[-((length(t1)-3):length(t1))]
t2<-as.numeric(Industria_Geral[1:180]) [-c(1,2)] #Três meses antes
t2<-t2[-((length(t2)-2):length(t2))]
t3<-as.numeric(Industria_Geral[1:180]) [-(1:3)] #Dois meses antes
t3<-t3[-((length(t3)-1):length(t3))]
t4<-as.numeric(Industria_Geral[1:180]) [-(1:4)] #Um meses antes
t4<-t4[-length(t0)]
t5<-as.numeric(Industria_Geral[1:180]) [-(1:5)] #Variável dependente

dados.Treina_geral <- cbind(t1,t2,t3,t4,t5)
set.seed(12345)
maxit<-as.integer(1000000)

nn <- neuralnet(dados.Treina_geral[,5]~dados.Treina_geral[,4]+
  +dados.Treina_geral[,3]+dados.Treina_geral[,2]+dados.Treina_geral[,1],
  data=dados.Treina_geral, hidden=7,threshold =1,stepmax= maxit)

#Apresenta os valores para os pesos
#nn$result.matriz

```

```

#Faz o gráfico do modelo
#plot(nn)

bd_valida_geral = c(Industria_Geral[176:180],Industria_Geral[181:192]+hw2p_diff)

#Dados para o neuralnet
t0_V<-as.numeric((bd_valida_geral)) #Cinco meses antes
t0_V<-t0_V[-((length(t0_V)-4):length(t0_V))]
t1_V<-as.numeric(bd_valida_geral) [-1] #Quatro meses antes
t1_V<-t1_V[-((length(t1_V)-3):length(t1_V))]
t2_V<-as.numeric(bd_valida_geral) [-c(1,2)] #Três meses antes
t2_V<-t2_V[-((length(t2_V)-2):length(t2_V))]
t3_V<-as.numeric(bd_valida_geral) [-c(1:3)] #Dois meses antes
t3_V<-t3_V[-((length(t3_V)-1):length(t3_V))]
t4_V<-as.numeric(bd_valida_geral) [-c(1:4)] #Um meses antes
t4_V<-t4_V[-length(t0_V)]
t5_V<-as.numeric(bd_valida_geral) [-c(1:5)] #Variável dependente
dados.Valida<-cbind(t1_V,t2_V,t3_V,t4_V,t5_V)

#Faz a previsão
#previsao<-compute(nn,dados.Valida[,1:4])
#Valores da previsao
#previsao.nn<-previsao$net.result

#Gera as estimativas para o Random Walk
epsilon<-(dados.Treina_geral[,5]-dados.Treina_geral[,4])
mu<-mean(epsilon)
sigma2<-var(epsilon)

#Faz a previsao usando o Random Walk
previsao.geral<-dados.Valida[,4]+rnorm(nrow(dados.Valida),+mu,sqrt(sigma2))

acuracia_rna_geral = c(
  mae(Industria_Geral[181:192],previsao.geral),

```

```

mse(Industria_Geral[181:192],previsao.geral),
rmse(Industria_Geral[181:192],previsao.geral),
mape(Industria_Geral[181:192],previsao.geral),
smape(Industria_Geral[181:192],previsao.geral),
u_ttheil(Industria_Geral[181:192],previsao.geral))

#acuracia_rna_geral

### COMBINAÇÃO DE PREVISAO

prevs_IndG =
cbind(IndGeral_r1$pred,Industria_Geral[181:192]+hw2p_diff,mod_econometrico_geral_prev,previsa
o.geral)

comb_simple <- Forecast_comb(Industria_Geral[181:192],fhat = prevs_IndG, fhat_new =
NULL,Averaging_scheme=c("simple"))
comb_variance <- Forecast_comb(Industria_Geral[181:192],fhat = prevs_IndG, fhat_new =
NULL,Averaging_scheme=c("variance based"))
comb_ols <- Forecast_comb(Industria_Geral[181:192],fhat = prevs_IndG, fhat_new =
NULL,Averaging_scheme=c("ols"))

tblgeral =
cbind(IndGeral_r1$pred,Industria_Geral[181:192]+hw2p_diff,mod_econometrico_geral_prev,previsa
o.geral,comb_simple$fitted, comb_variance$fitted,comb_ols$fitted)

erro1= c(
mae(comb_simple$fitted,Industria_Geral[181:192]),
mse(comb_simple$fitted,Industria_Geral[181:192]),
rmse(comb_simple$fitted,Industria_Geral[181:192]),
mape(comb_simple$fitted,Industria_Geral[181:192]),
smape(comb_simple$fitted,Industria_Geral[181:192]),
u_ttheil(comb_simple$fitted,Industria_Geral[181:192]))

erro2= c(
mae(comb_variance$fitted,Industria_Geral[181:192]),

```

```

mse(comb_variance$fitted,Industria_Geral[181:192]),
rmse(comb_variance$fitted,Industria_Geral[181:192]),
mape(comb_variance$fitted,Industria_Geral[181:192]),
smape(comb_variance$fitted,Industria_Geral[181:192]),
u_ttheil(comb_variance$fitted,Industria_Geral[181:192]))

erro3= c(
mae(comb_ols$fitted,Industria_Geral[181:192]),
mse(comb_ols$fitted,Industria_Geral[181:192]),
rmse(comb_ols$fitted,Industria_Geral[181:192]),
mape(comb_ols$fitted,Industria_Geral[181:192]),
smape(comb_ols$fitted,Industria_Geral[181:192]),
u_ttheil(comb_ols$fitted,Industria_Geral[181:192]))

# Grafico das Séries Real x Previsão

Ind_Geral_Real = Industria_Geral[181:192]
Ind_Geral_Prev = comb_ols$fitted[1:12]

teste = cbind(Ind_Geral_Real,Ind_Geral_Prev)

datas1 <- seq(as.Date(paste(c(start(hw2p_diff),1), collapse="/")),
             by = "month", length.out = length(hw2p_diff[,1]))

dados.dfgeral <- data.frame(date = datas1, value = teste)

min <- as.Date("2017-1-1")
max <- NA

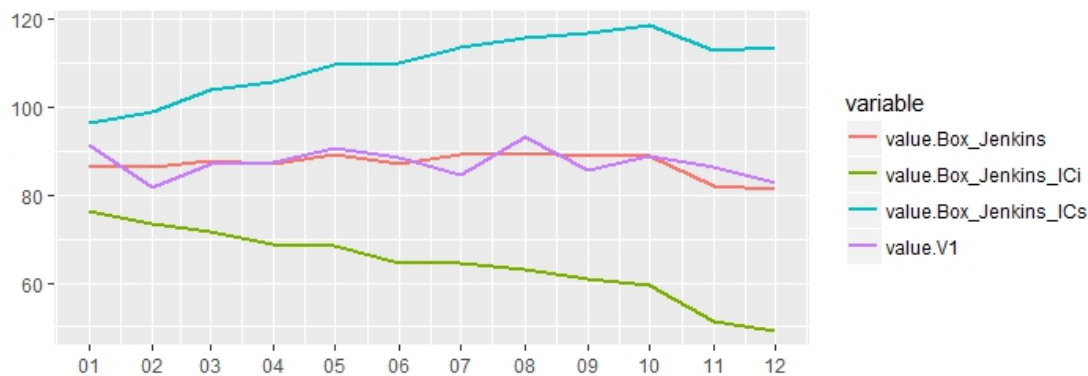
df <- dados.dfgeral %>%
  select(date,value.Ind_Geral_Real, value.Ind_Geral_Prev) %>%
  gather(key = "variable", value = "value",-date)

```

```
#head(df)
p1 <- ggplot(df, aes(x = date, y = value)) + geom_line(aes(color = variable), size = 1) +
scale_y_continuous(limits=c(40, 100))
p1 + scale_x_date(limits = c(min, max),date_labels = "%m", date_breaks = "1 month")
```

## APENDICE F – Intervalo de Confiança das Previsões

- **SELIC (2 defasagens):**



**Figura 36** - Produção Industrial Geral: Realizado x Previsto via Box-Jenkins com intervalo de confiança de 95% (jan-2017 x dez x 2017)

Para cálculo de um intervalo de confiança aproximado para a previsão combinada, considerou-se :

1- Dentre os métodos usados, o de redes neurais não tem pressupostos distribucionais. O método de Holt-Winters aditivo é linear e, apesar de sua estrutura adaptativa ser completamente deduzida sem suposições sobre distribuição probabilística, pode ser facilmente enquadrado em uma estrutura estatística. O método de Box & Jenkins, adota o pressuposto de normalidade, bem como a modelagem econométrica.

Exemplo: Moretin & Tolo (2006), Granger & Newbold (1986), Gardner (1985)

Moretin, P.A. e Tolo, Clélia M. de Castro (2004), *Análise de Séries Temporais*, Editora Edgard Blucher, São Paulo, 2004. [2a Edição de 2006].

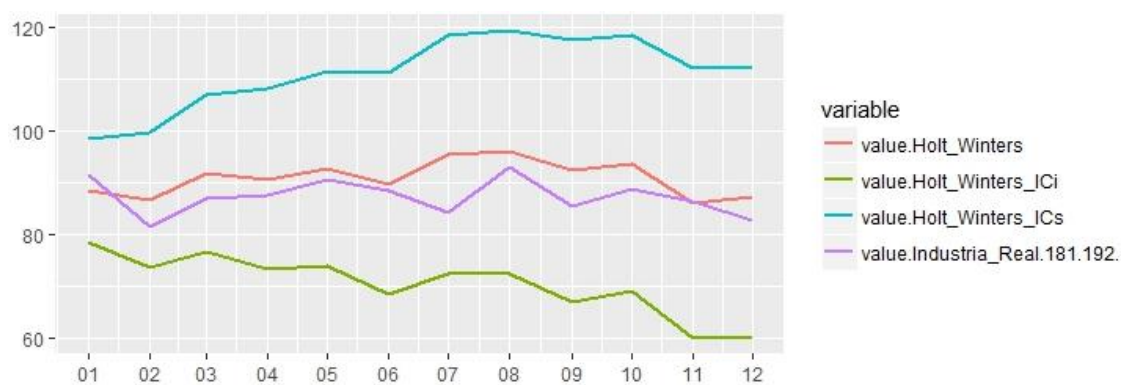
Gardner, E. S. Jr. [1985]. Exponential smoothing: the state of the art. 1. *Forecasting* 4, 1-28.

Granger, C. W. J. and Newbold, P. (1986). *Forecasting Economic Time Series*, 2nd ed. Academic Press, New York.



2- Supondo-se que as previsões obtidas pelas quatro metodologias adotadas são independentes, o que é razoável.

3- Utilizando-se resultados estatísticos bastante conhecidos conclui-se que a combinação das previsões, que é uma combinação linear das previsões individuais obtidas pelas metodologias, é também normal.



**Figura 37-** Produção Industrial Geral: Realizado x Previsto via Combinação de Previsão com intervalo de confiança de 95% (jan-2017 x dez x 2017)