



**Bruno Rial Farenzena**

**TRABALHO FINAL DE GRADUAÇÃO**  
**PREVISÃO DE PREÇO E QUANTIDADE PRODUZIDA DE ARROZ NO**  
**ESTADO DO RIO GRANDE DO SUL: PROPOSIÇÃO DE MODELOS ARIMA**

Santa Maria, RS

2021



**Bruno Rial Farenzena**

**PREVISÃO DE PREÇO E QUANTIDADE PRODUZIDA DE ARROZ NO  
ESTADO DO RIO GRANDE DO SUL:  
PROPOSIÇÃO DE  
MODELOS ARIMA**

Trabalho Final de Graduação apresentado ao curso de Ciências Econômicas, Área das Ciências Sociais, da Universidade Franciscana (UFN), como requisito parcial para a obtenção do grau de bacharel em Ciências Econômicas.

Orientadora: Prof.<sup>a</sup> Dr.<sup>a</sup> Denise Piper

Santa Maria, RS

2021

**Bruno Rial Farenzena**

**PREVISÃO DE PREÇO E QUANTIDADE PRODUZIDA DE ARROZ NO  
ESTADO DO RIO GRANDE DO SUL:  
PROPOSIÇÃO DE  
MODELOS ARIMA**

Trabalho Final de Graduação apresentado ao curso de Ciências Econômicas, Área das Ciências Sociais, da Universidade Franciscana (UFN), como requisito parcial para a obtenção do grau de bacharel em Ciências Econômicas.

---

Prof.<sup>a</sup> Dr.<sup>a</sup> Denise Piper – Orientadora (UFN)

---

Prof. Me. Fábio Nascimento – Avaliador (UFN)

---

Prof. Me. Mateus Sangoi Frozza – Avaliador (UFN)

Aprovado em ..... de ..... de 2021.



## RESUMO

O presente trabalho visa analisar o comportamento da produção e do preço do arroz em casca no estado do Rio Grande do Sul. Para tanto, procurou-se realizar as previsões a partir da metodologia ARIMA. O período analisado pela modelagem da série de produção de arroz compreende o período de 1971 a 2020 por meio do modelo ARIMA (0,1,1), que foi o mais eficiente dentre os modelos propostos e se mostrou útil ao apresentar boa projeção. O período analisado pela modelagem da série de preços de arroz em casca compreende o tempo de fevereiro de 2010 a julho de 2020 em dados mensais e o modelo ARIMA (1,0,1) estimado foi eficiente, apresentando projeção satisfatória. Os resultados obtidos fornecem uma ferramenta de análise para o mercado de arroz na medida em que demonstra a tendência de curto prazo servindo de auxílio à tomada de decisão de agentes que comercializam esse bem, formadores de políticas públicas e consumidores.

**Palavras-chave:** Séries temporais, modelagem ARIMA, arroz em casca, preço, produção.

## **ABSTRACT**

The present work aims to analyze the behavior of the production and the price of paddy rice in the state of Rio Grande do Sul. For this purpose, the forecasts were constructed based on the ARIMA methodology. The period analyzed by modeling the rice production series comprises the period of 1971 to 2020. The AR (1) model was the most efficient among the proposed models and proved to be useful by presenting a reasonable projection. The period analyzed by modeling the paddy rice series comprises the time from February 2010 to July 2020 in monthly data, the estimated ARIMA (1,1) model was efficient and presented a satisfactory projection. The results obtained provide an analysis tool for the rice market, insofar as it demonstrates the short-term trend, serving as an aid to the decision-making agents that sell this good, public policy makers and consumers.

**Keywords:** Time series, ARIMA modeling, paddy rice, price, production.

## **LISTA DE FIGURAS**

<b>FIGURA 1:</b> Diagrama da cadeia produtiva do arroz no Rio Grande do Sul.....	15
<b>FIGURA 2:</b> Diagrama Metodologia de Box-Jenkins.....	27



## LISTA DE GRÁFICOS

Gráfico 1: Série histórica de produção das safras de arroz de 1921/1922 a 2019/2020 no estado do Rio Grande do Sul.....	17
Gráfico 2: Indicador de preço mensal do arroz em casca ESALQ/SENAR-RS saca de 50kg nominal e real deflacionado pelo IGP-DI base 03/2021.....	18
Gráfico 3: Indicador do preço mensal do arroz em casca ESALQ/SENAR-RS saca de 50kg nominal preço corrente em dólar do dia.....	18
Gráfico 4: indicador do preço mensal do arroz em casca ESALQ/SENAR-RS saca de 50kg preços nominais e reais de 2020/01 a 2021/03.....	19
Gráfico 5: taxa de câmbio mensal referente ao preço de US\$ 1 no último dia do mês no período de 2010 a 2021.....	20
Gráfico 6: taxa de câmbio mensal referente ao preço de US\$ 1 no último dia do mês no período de 2020/01 a 2020/12.....	20
Gráfico 7: decomposição da série de tempo referente ao preço.....	31
Gráfico 8: FAC e FACP da série de produção (prod).....	32
Gráfico 9: FAC e FACP da série de produção (prod) em primeira diferença.....	32
Gráfico 10: FAC e FACP da série de preços (prer).....	33
Gráfico 11: FAC e FACP dos resíduos modelo de produção ARIMA (0,1,1) .....	6
Gráfico 12: distribuição dos resíduos modelo de produção ARIMA (0,1,1) .....	36
Gráfico 13: FAC e FACP dos resíduos do modelo de preços ARIMA (0,1,1) .....	37
Gráfico 14: distribuição dos resíduos do modelo de preços ARIMA (1,0,1) .....	37
Gráfico 15: previsto e observado da série de produção (prod) 1970-2025.....	39
Gráfico 16: previsto e observado da série de preço (prer) 2010-2021/04.....	39

## LISTA DE TABELAS

Tabela 1: Análise descritiva dos dados.....	30
Tabela 2: Resultados da transformação Box-Cox.....	30
Tabela 3: Resultado teste Dickey-Fuller aumentado (ADF).....	32
Tabela 4: Resultados das estimativas dos modelos ARIMA.....	34
Tabela 5: Análise dos resíduos.....	35
Tabela 6: Estatísticas de erro.....	36
Tabela 7: Resultados da previsão de produção (prod).....	38
Tabela 8: Resultados previsão de preços (prer).....	38

## LISTA DE SIGLAS

ARIMA – *Autoregressive Integrated Moving Average*  
IRGA – Instituto Riograndense do Arroz  
SINDARROZ – Sindicato das Indústrias de Arroz do Rio Grande do Sul  
USDA – *Departamento de Agricultura dos Estados Unidos*  
CEPEA – Centro de Estudos Avançados em Economia Aplicada  
Abiarroz – Associação Brasileira da Indústria do Arroz  
Emater – Empresa de Assistência Técnica e Extensão Rural  
ARCH - *Autoregressive Conditional Heteroskedasticity*  
GARCH – *Generalized Autoregressive Conditional Heteroskedasticity*  
SARIMA – *Seasonal Autoregressive Integrated Medium Average*  
IEA – Instituto de Economia Agrícola  
AIC – *Akaike Information Criteria*  
BIC – *Bayesian Information Criteria*  
MAPE – Percentual de Erro Médio Absoluto  
AR – Modelo Autorregressivo  
MA – Modelo de Médias Móveis  
FAC – Funções de autocorrelação  
FACP – Funções de autocorrelação parcial  
ADF – Aumentado Dickey Fuller  
AIC<sub>c</sub> – *Akaike Information Criteria corrected*  
IGP-DI – Índice Geral de Preços – Disponibilidade Interna  
FEALQ – Fundação de Estudos Agrários Luiz de Queiroz  
SENAR – Serviço Nacional de Aprendizagem Rural  
RMSE – Erro Padrão Residual Médio  
MAE – Erro Médio Absoluto  
MASE – Erro Médio Absoluto Escalado

## SUMÁRIO

<b>INTRODUÇÃO.....</b>	<b>12</b>
<b>1 O SETOR ORIZICOLA NO RIO GRANDE DO SUL.....</b>	<b>14</b>
1.1 As características do setor orizícola.....	14
1.2 As dinâmicas do preço e da produção do arroz no estado.....	16
1.3 Considerações finais.....	21
<b>2 PROPOSIÇÃO DOS MODELOS ARIMA PARA PREÇO E PRODUÇÃO.....</b>	<b>23</b>
2.1 O método de Box Jenkins aplicado ao mercado de <i>commodities</i> agrícolas.....	23
2.2 Aspectos Metodológicos.....	25
2.2.1 A metodologia Box-Jenkins.....	25
2.2.2 Descrição das Variáveis.....	28
<b>3. PROCESSO DE ESTIMAÇÃO E ANÁLISE DOS RESULTADOS.....</b>	<b>29</b>
3.1 Identificação.....	29
3.2 Estimativa.....	33
3.3 Verificação.....	34
3.4 Previsão.....	37
<b>4 Considerações Finais.....</b>	<b>40</b>
<b>5 CONCLUSÃO.....</b>	<b>41</b>
<b>REFERÊNCIAS.....</b>	<b>42</b>



## INTRODUÇÃO

O estado do Rio Grande do Sul é responsável por aproximadamente 70% da produção brasileira de arroz no ano de 2020, sendo um importante formador e referencial de preço. No âmbito internacional, o Brasil é o maior produtor e consumidor do cereal fora da Ásia, garantindo-lhe relevante papel no fornecimento do produto.

Nesse contexto, o presente trabalho visa avaliar a potencialidade da utilização de modelos de séries temporais na previsão de preço e produção de arroz em casca no estado do Rio Grande do Sul, apresentando um debate sobre a relevância do setor agroindustrial orizícola no referido estado, efetuando uma análise qualitativa de dados acerca do comportamento dos preços e da produção do arroz no supracitado estado entre os anos de 2005 a 2019, bem como estruturar um modelo autorregressivo de médias móveis capaz de realizar previsões referentes a preço e quantidade produzida de arroz no Rio Grande do Sul analisando os determinantes da dinâmica atípica do preço do arroz no ano de 2020. Essas providências poderão ser empregadas no planejamento e comercialização da *commodity*, gerando informações que darão suporte aos agentes produtores, consumidores e formadores de políticas públicas na tomada de decisão no momento de produzir, consumir, negociar ou promover investimentos direta ou indiretamente relacionados ao arroz.

Aproximadamente 10 milhões de toneladas de arroz por safra são produzidas no Brasil, o que define o país como o 11º maior produtor mundial e o maior produtor extra Ásia. O arroz gaúcho, em sua maioria irrigado, é de qualidade e produtividade superior à de tradicionais produtores do grão a nível mundial. A partir dessa magnitude, torna-se pertinente realizar os mais variados estudos referentes ao setor orizícola no estado do Rio Grande do Sul, seja tratando dos aspectos biológicos e agrônômicos ou mesmo sociais e econômicos relacionados ao arroz gaúcho.

Cabe destacar que a quantidade do cereal produzido e os níveis de preços internos e externos são de extrema significância para determinar como se dará a produção, o comércio e o consumo dele nas próximas safras. Assim sendo, o presente trabalho busca desenvolver um modelo preditivo, utilizando a metodologia de Box e Jenkins (1976) ou *Autorregressive Integrated Moving Average* (ARIMA). Dessa forma, o questionamento a ser respondido é: qual o modelo estatístico ARIMA que melhor se adapta para previsões de preço e de produção de arroz no Rio Grande do Sul?

Feitas essas considerações iniciais, pode-se afirmar que a previsão de preço e quantidade produzida de arroz no estado do Rio Grande do Sul por meio da proposição de modelos ARIMA é relevante porque poderão ser empregados no planejamento e comercialização da *commodity*, gerando informações que darão suporte os agentes produtores, consumidores e formadores de políticas públicas na tomada de decisão no momento de produzir, consumir, negociar ou promover investimentos relacionados direta ou indiretamente ao arroz.

## 2. O SETOR ORIZÍCOLA NO ESTADO DO RIO GRANDE DO SUL

### 2.1 AS CARACTERÍSTICAS DO SETOR ORIZÍCOLA

O arroz gaúcho representa cerca de 70% de sua produção e 14% de grãos no Brasil, ocupando posição de destaque do ponto de vista econômico e social, responsável por suprir a dieta básica da população com considerável aporte de calorias, proteínas e sais minerais (GOMES; MAGALHÃES JUNIOR, 2004), sendo considerado um dos alimentos mais importantes para a dieta humana, perdendo apenas para o trigo.

No Brasil estima-se que o consumo médio de arroz é de cerca de 25 kg por pessoa por ano (MAPA, 2015). O setor é destaque no mercado de trabalho do Rio Grande do Sul, representando uma importante fonte de emprego e renda pois, de acordo com dados de 2017 do Instituto Rio Grandense do Arroz (IRGA) e do Sindicato das Indústrias de Arroz do Rio Grande do Sul (SINDARROZ), as lavouras de arroz empregam 37 mil trabalhadores no estado e, na indústria de processamento, outros 15 mil trabalhadores, gerando mais de 50 mil empregos diretos.

Segundo o IRGA (2010), o registro da primeira produção de arroz no Rio Grande do Sul ocorreu no final do século XVIII. Tal fato foi realizado para garantir a subsistência dos moradores da capitania de São Pedro e, com a chegada dos imigrantes alemães em 1824, a produção passou a se manter regular e em maior escala a fim de manter os hábitos alimentares dos recém-chegados. Em relação ao beneficiamento do arroz, o IRGA (2010) relata o surgimento das primeiras industriais do seguinte modo:

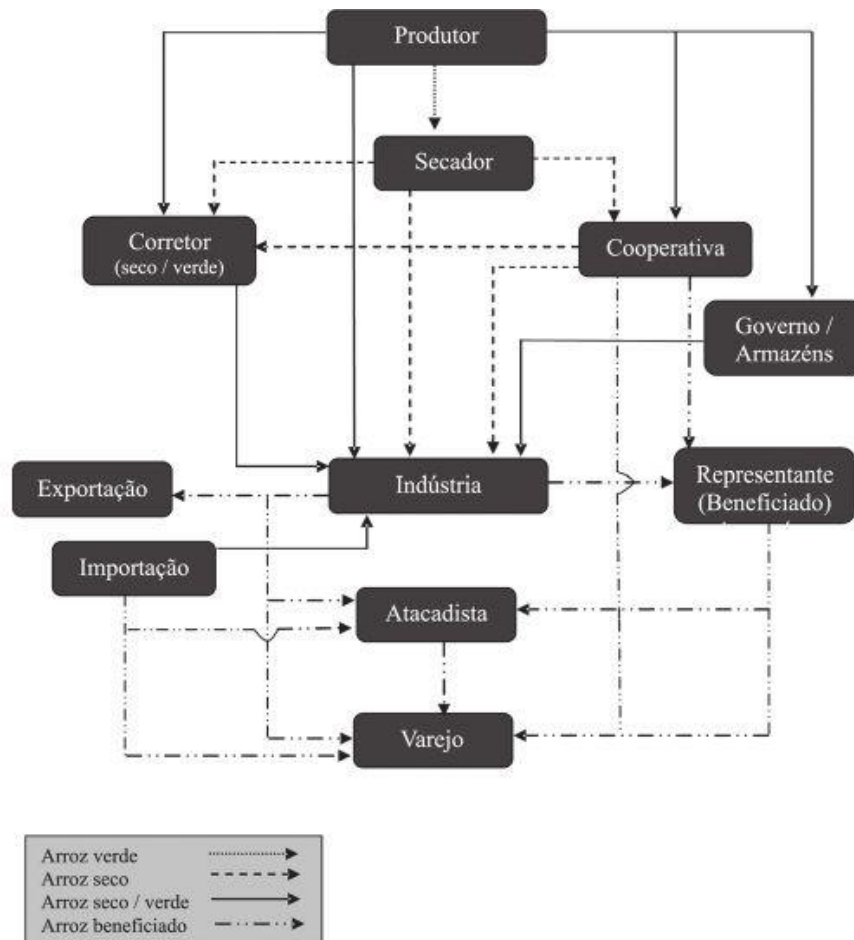
Em 1876, Cachoeira do Sul e Rio Pardo detinham a maior concentração de engenhos do grão no Rio Grande do Sul, como decorrência surgiu, em 1887, o primeiro engenho moderno, que era movimentado por uma máquina a vapor para a descasca do arroz. No ano seguinte foi desenvolvido o primeiro engenho movido a água.

Ainda de acordo com o IRGA (2010), a partir do surgimento dos engenhos e da irrigação mecanizada, o aumento da área e da produção do grão era crescente: na safra de 1951/52 a área plantada chegava a 229.542 mil hectares com uma produção de 591.987 toneladas de arroz, até então colhidos manualmente. Nas próximas décadas, é incrementada no estado a utilização de maquinários para o preparo do solo, plantio e processamento do grão, o que desencadeou uma maior produtividade média e estabilidade da lavoura de arroz gaúcha. A safra de 1969/70 demonstrou a consolidação do arroz gaúcho com 420.627 hectares plantados e produção de 1.456.980 ha. Na safra 2019/2020, o estado cultivou 934,538 mil hectares e teve produção igual a 7.839.113 toneladas.



Miranda *et al.* (2008) propõe descrever a cadeia agroindustrial do arroz do Rio Grande do Sul e desenvolver, a partir disso, uma estrutura para a cadeia desse cereal no estado. Através do contato com o IRGA (2010), a pesquisadora realiza um diagrama que representa os elos e fases da cadeia produtiva por meio da Figura 1:

**Figura 1:** Diagrama da cadeia produtiva do arroz no Rio Grande do Sul.



**Fonte:** Miranda *et al.* (2008).

A autora identifica alguns desafios que limitam o melhor funcionamento, produtividade e determinam os preços na cadeia de arroz, dentre eles a falta de coordenação do atacado e varejo e produtores que torna o setor industrial refém das oscilações de demanda; a influência desse setor na determinação dos preços do arroz em casca; a necessidade de abertura de novos canais de comercialização e aumento das exportações, sendo a taxa de câmbio uma variável essencial de competitividade. Somam-se a esses obstáculos o acréscimo da eficiência produtiva através da redução de custos de produção e comercialização e as necessidades de se construir mecanismos que reduzam o risco para os agentes através de contratos a termo, da criação de um mercado futuro para o arroz no Mercosul, da formalização das relações já existentes entre os elos ou do desenvolvimento de novas relações entre a cadeia agroindustrial.

De acordo com os dados do Departamento de Agricultura dos Estados Unidos *Foreign Agricultural Services/USDA* (2020), no cenário mundial, a produção e o consumo de arroz mantiveram-se ajustados nos últimos anos. Em 2018, a produção em milhões de toneladas e o consumo foram de, respectivamente, 481 e 495; em 2019 497 e 484; em 2020, a estimativa foi igual a 496 e 495 e, no ano seguinte, a projeção feita para a produção e o consumo foi equivalente a 502,6 e 499, respectivamente. No período de 2015 a 2020, conforme os dados da USDA (2020), o Brasil exportou 0,9 milhões de toneladas de arroz em 2015, 0,64 em 2016; 0,59 em 2017; 1,25 em 2018 e 0,95 em 2019 e 1,19 em 2020, ano que representou cerca de 2,74% das exportações mundiais.

Segundo os dados da Secretaria e Comercio Exterior (Secex) do Ministério da Economia, as importações no ano de 2020 atingiram o equivalente a 1,26 milhões de toneladas de arroz em casca, o maior volume em 17 anos. De acordo com o relatório anual do Centro de Estudos Avançados em Economia Aplicada (CEPEA-Esalq/USP), a flexibilização da Tarifa Externa Comum (TEC) para países de fora do Mercosul teve relevância no mercado, mas mostrou-se insuficiente para causar mudanças significativas de preços no curto prazo, devido a paridade de importação elevada por conta da variação cambial e da disponibilidade restrita do arroz em casca no Brasil.

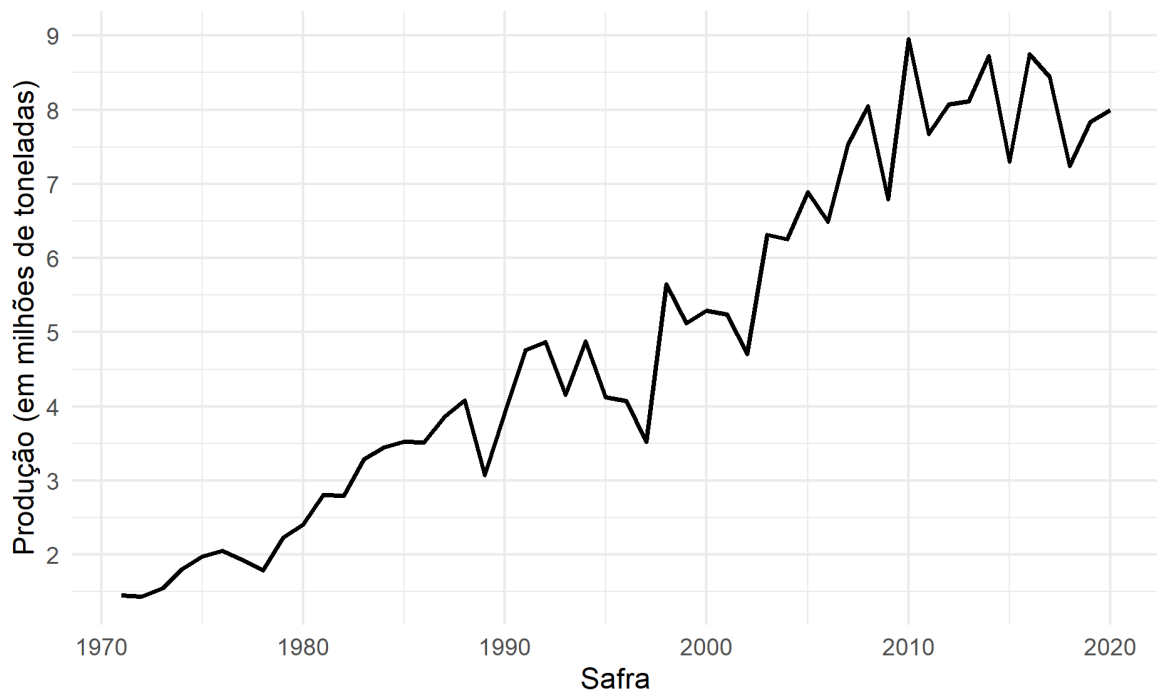
Posto isso, é notória a importância do supracitado cereal na formação histórica da agricultura no estado e na incorporação de tecnologias produtivas, auxiliando as lavouras a alcançarem o mesmo nível de produção dos maiores produtores mundiais do cereal (Sampaio e Furnestreu, 2003).

## 2.2 AS DINÂMICAS DO PREÇO E DA PRODUÇÃO DO ARROZ NO ESTADO DO RIO GRANDE DO SUL

O Gráfico 1 representa a série histórica de produção de arroz no Rio Grande do Sul no período de 1921 a 2020, sendo visualmente claro o aumento de produção ao longo do tempo, fato esse que pode ser explicado pelo desenvolvimento de novas tecnologias de cultivo adotadas pelos produtores gaúchos. Na última década da série, pode-se perceber uma variação bem grande na área plantada devido a introdução de novas culturas ou produtores abandonando o cultivo de arroz, entretanto os dados do IRGA (2010) mostram um aumento significativo de produtividade. Na safra de 2020/2021, a produtividade média do estado alcançou 9010 kg/ha, totalizando 8523429 toneladas, sendo a quarta maior produção da história.

No entanto, as três maiores produções foram obtidas sobre áreas acima de 1,1 milhão de hectares, enquanto a safra em questão foi colhida em 945.971 hectares. Notam-se também variações de produção em determinadas safras, as quais muitas vezes são explicadas por perda total de área plantada, principalmente por eventos climáticos, como nos anos de 1997, 2009 e 2015 em que se perdeu, respectivamente, 48.455; 35.273 e 31.324 hectares de acordo com o IRGA (2010).

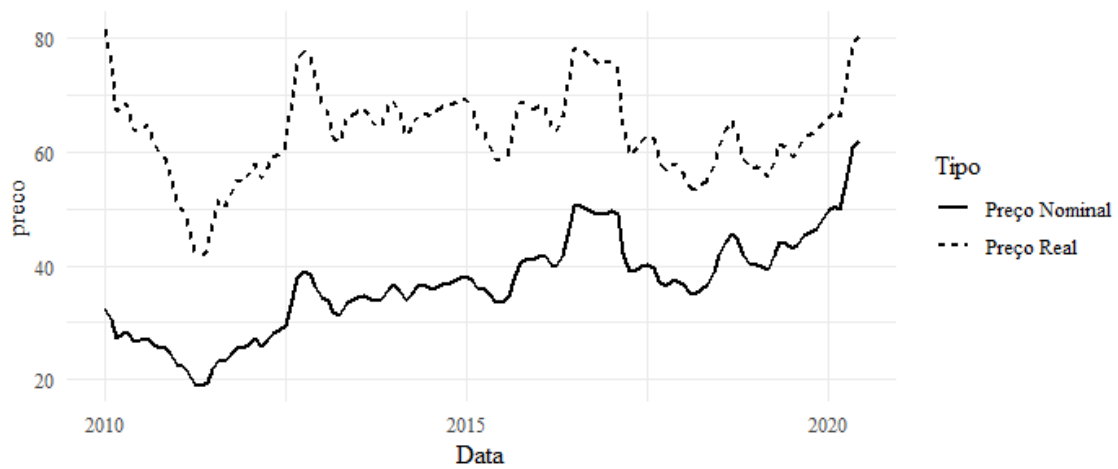
**Gráfico 1:** série histórica de produção das safras de arroz de 1921/1922 a 2019/2020 no estado do Rio Grande do Sul



**Fonte:** IRGA (2010)

O Gráfico 2 apresenta o índice do preço mensal do arroz em casca ESALQ/SENAR-RS saca de 50kg nominal e real deflacionado pelo IGP-dia com valores de 03/2021 para o período de 01/2010 a 07/2020. Nesse gráfico é perceptível que a oscilação do preço real acompanha a variação do preço nominal, cujo ponto máximo encontra-se no final do período analisado equivalendo a R\$ 64,67 e do preço real no início do espaço estudado em 2010, chegando em R\$82,21; o preço mais baixo registrado foi em maio de 2011 que, em valores nominais, apresentou média de R\$ 19,08 e R\$ 41,68 em valores reais.

**Gráfico 2:** indicador do preço mensal do arroz em casca ESALQ/SENAR-RS saca de 50kg nominal e real deflacionado pelo IGP-DI base 03/2021.

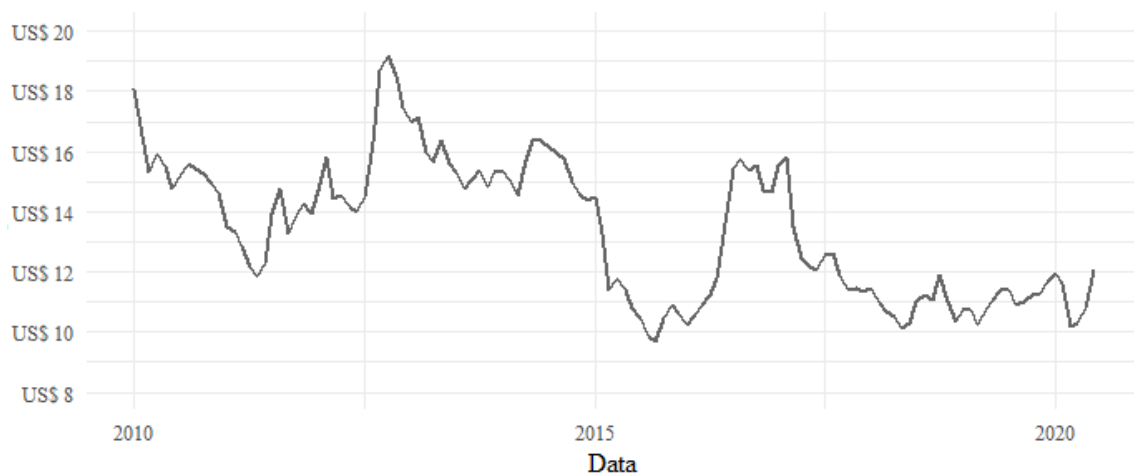


Fonte: USP/Esalq

**Fonte:** USP/Esalq.

O Gráfico 3 ilustra o indicador do preço de arroz em casca ESALQ/SENARRS a preços correntes em dólar para o período de 01/2010 a 06/2020. No referido gráfico, é perceptível a tendência de baixa ao longo da série devido à desvalorização cambial do real nesse período, sendo seu ponto mais alto no ano de 2013 e mais baixo no final de 2015 e começo de 2016, mantendo-se entre US\$ 10 e US\$ 12 nos anos seguintes.

**Gráfico 3:** indicador do preço mensal do arroz em casca ESALQ/SENAR-RS saca de 50kg nominal preço corrente em dólar do dia.

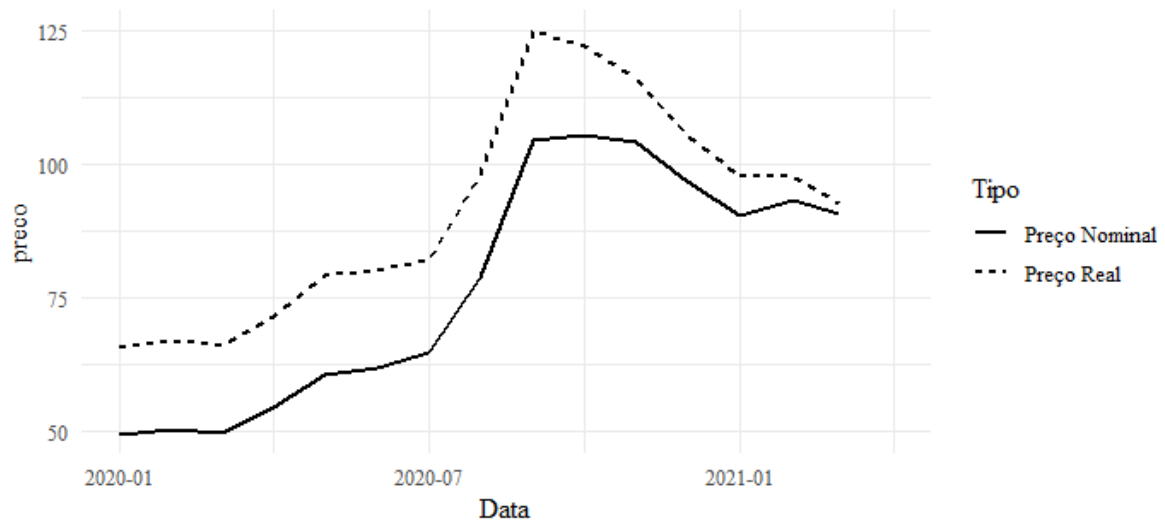


**Fonte:** USP/Esalq.

No Gráfico 4 é demonstrado o indicador do preço do arroz para o ano de 2020, ilustrando o aumento do preço do cereal no período. Embora não tenha sido o intuito do presente trabalho relatar todos os motivos que desencadearam o aumento, foi preciso restringir o período de análise da série temporal a julho de 2020 uma vez que, após esse tempo, ocorreu um aumento

significativo do indicador, ocasionando valores que fogem da normalidade e que pode causar anomalias nos resultados obtidos e viés nas previsões dos modelos, conhecidos na literatura como *outlier*. Silva (2011, p. 93) define esse termo como uma “observação aberrante, anormal, atípica, contaminante, dissimilar, estranha, extrema, discordante ou preocupante”. Além disso, Rosado (2006, p. 01) esclarece que “uma única observação (não detectada) pode destruir ou contrariar a conclusão de qualquer trabalho”.

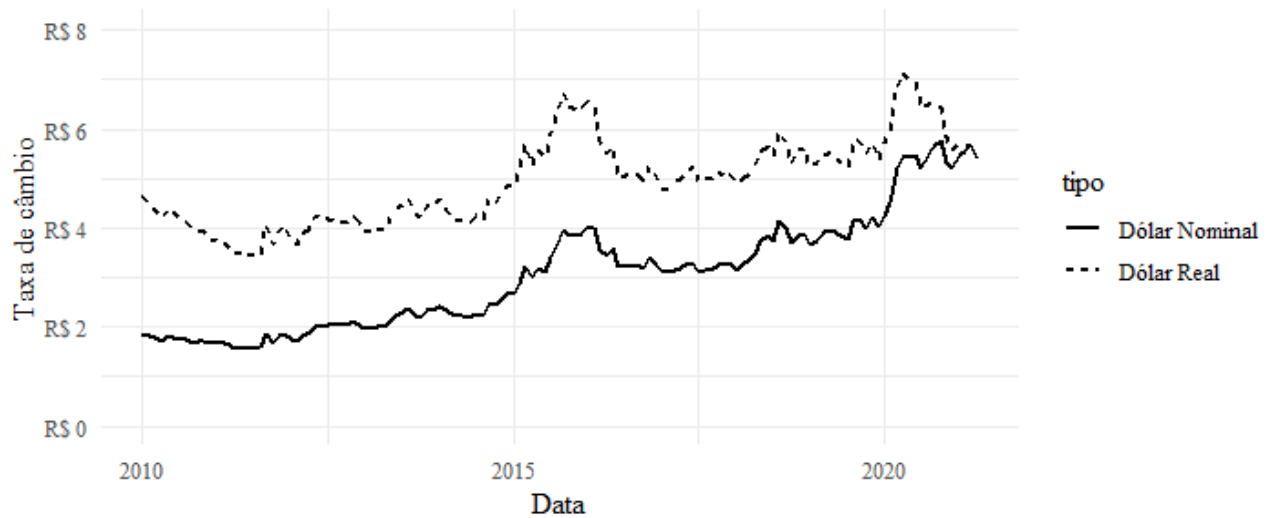
**Gráfico 4:** indicador do preço mensal do arroz em casca ESALQ/SENAR-RS saca de 50kg preços nominais e reais de 2020/01 a 2021/03



**Fonte:** Banco Central do Brasil, Boletim, Seção Balanço de Pagamentos (Bacen / Boletim / BP)

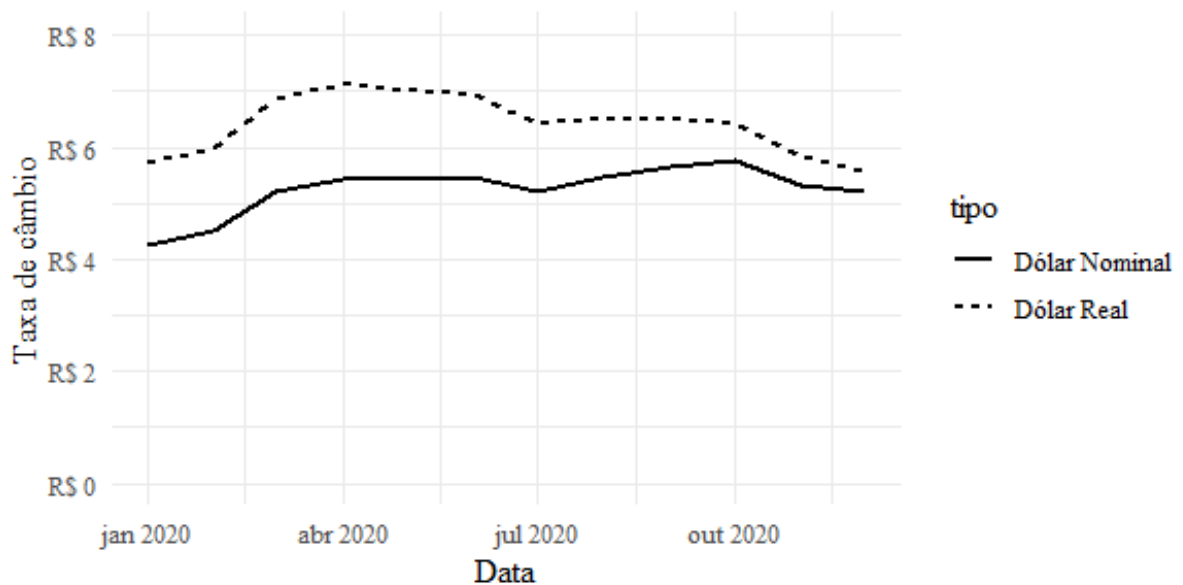
Os Gráficos 5 e 6 apontam a taxa de câmbio mensal referente ao preço de US\$ 1 no último dia do mês nos períodos de 2010 a 2021 e 2020, respectivamente.

**Gráfico 5:** Taxa de câmbio mensal referente ao preço de US\$ 1 no último dia do mês no período de 2010 a 2021.



**Fonte:** Banco Central do Brasil, Boletim, Seção Balanço de Pagamentos (Bacen / Boletim / BP).

**Gráfico 6:** Taxa de câmbio mensal referente ao preço de US\$ 1 do último dia de cada mês no período de 2020.



**Fonte:** Banco Central do Brasil, Boletim, Seção Balanço de Pagamentos (Bacen / Boletim / BP).

É relevante apresentar estudos que busquem explicar os causadores do aumento do preço do arroz no ano de 2020. Para isso, apresenta-se um trecho do relatório macroeconômico do CEPEA/Esalq-USP de junho de 2021 de Barros *et al.* (2021)

Em 2020, o mercado agropecuário no Brasil esteve firme. O volume agropecuário cresceu 2,5%, sendo que o volume de grãos avançou 4,3% e o da pecuária caiu 1%. O volume exportado de produtos agropecuários, por sua vez, aumentou 8,7% em 2020, deixando o mercado interno relativamente apertado.

Uma grande desvalorização cambial (perto de 30%) relacionada, principalmente aos desarranjos da política econômica interna e eventos externos de variada natureza incentivou a exportação, aproveitando a onda de aquecimento da demanda externa, que absorveu parte da produção nacional de *commodities*.

Os preços de *commodities* agropecuária aumentaram quase 9% em dólares em 2020, de acordo com o FMI. O ano foi também marcado por fortes movimentos nos preços do petróleo e seus derivados: uma queda profunda no primeiro semestre, quando a pandemia derrubou inesperadamente a demanda mundial, seguida de firme recuperação decorrente de retomada relativamente forte da economia global, graças a substanciais dispêndios e expansão monetária dos governos em geral.

Essa pesquisa explica parcialmente o que pode ter ocorrido no mercado orizicola do Rio Grande do Sul nesse período, entretanto deve-se observar as taxas de câmbio para o período analisado tal qual o volume exportado de arroz.

Com a desvalorização do real no ano de 2020, o mercado de arroz sofreu um choque de demanda, impulsionando o aumento do preço interno, de acordo com os números divulgados pela Abiarroz (Associação Brasileira da Indústria do Arroz) e Sindarroz-RS (Sindicato da Indústria do Arroz no Estado do Rio Grande do Sul) com base em informações do Ministério da Economia. Na série histórica do setor, as exportações de 2020 a março de 2021 só perdem para o ano comercial de 2011/12, quando as vendas externas somaram 2 milhões e 99 mil toneladas.

### 2.3 Considerações finais

Sabe-se a importância do cereal no desenvolvimento histórico do estado, na cadeia produtiva e na garantia nutricional da população, bem como o impacto que as variações de preço e produção podem vir a causar na relação dos agentes envolvidos na cadeia produtiva e consumidores. O avanço tecnológico e as melhores práticas produtivas possibilitaram um aumento substancial na produção gaúcha, impactando de maneira contínua o nível de preços.

Contudo, no curto prazo o impacto do comércio mundial, das políticas macroeconômicas e das variações cambiais induzem como nos demais setores produtivos e no mercado de *commodities* o comportamento dos preços, evidenciado no período de 2020 com a

desvalorização do real sendo o principal fator a explicar a alta do preço do arroz em casca no Rio Grande do Sul ao incentivar o aumento das exportações.

Desse modo, pode-se aferir a evidente importância dos supracitados grãos na nutrição populacional e na cadeia produtiva do estado. A inovação tecnológica e a adoção de melhores práticas de produção garantiram um aumento substancial na produção gaúcha, o que impactou continuamente os valores aplicados na comercialização dos produtos.

Entretanto, a curto prazo, o impacto do comércio mundial com suas políticas macroeconômicas e variações cambiais induzem, como nos demais setores produtivos e no mercado de *commodities*, uma alteração no comportamento dos preços. Tal modificação foi evidenciada principalmente no ano de 2020, sendo a desvalorização do real o principal fator a explicar a alta de preço do arroz em casca no Rio Grande do Sul incentivando o aumento das exportações.

No seguinte capítulo é feita a proposição dos modelos ARIMA, inicialmente apresentando trabalhos que aplicaram a metodologia ao mercado de *commodities*, em seguida é aplicada a metodologia de Box-Jenkins e, por fim, é realizada a estimação dos modelos e resultados.



### 3. PROPOSIÇÃO DOS MODELOS ARIMA PARA PREÇO E PRODUÇÃO

O presente capítulo toma, por objetivo central, explicitar a utilização da metodologia de Box Jenkins na previsão de series temporais, para posteriormente apresentar a aplicação do método nas series escolhidas para a consecução desse trabalho. Para tanto, a apreciação exposta na primeira subseção apresenta uma revisão bibliográfica com pesquisas que objetivaram realizar previsões aplicadas ao mercado de produtos agrícolas utilizando da metodologia ARIMA. A seção seguinte discorre-se acerca dos aspectos metodológicos, procedimentos e etapas necessárias para efetivar a metodologia e inferir sobre os dados de preço e produção.

Por fim, a terceira seção remete-se ao Processo de Estimação e Análise dos Resultados, sendo dividida em 4 subseções, delimitando as etapas de identificação; estimativa; verificação e previsão. O capítulo é encerrado com a apresentação das principais constatações legadas pela sua consecução.

#### 3.1 O MÉTODO DE BOX-JENKINS APLICADO AO MERCADO DE *COMMODITIES* AGRÍCOLAS

Conforme mencionado previamente, o mercado de arroz no Rio Grande do Sul, assim como os demais mercados de *commodities* agrícolas, é competitivo. Na literatura, são recorrentes estudos que investigam previsões de séries temporais de dados econômicos em mercados desse tipo a fim de comparar resultados de diferentes métodos. A presente revisão de literatura compila pesquisas que objetivaram realizar previsões aplicadas ao mercado de produtos agrícolas.

O estudo de Viana, Souza e Silveira (2009) analisa a confiabilidade da metodologia de Box e Jenkins (1976) em duas séries históricas de preços da bovinocultura de corte do Rio Grande do Sul de duas fontes diferentes. Na primeira série, apresentavam-se os preços médios pagos por quilograma de peso vivo do boi gordo e na segunda, valores médios pagos de janeiro de 1977 a dezembro de 2006 pelo consumidor, identificando uma tendência declinante de 1977 a 1994 e maior estabilidade de 1995 a 2006, concluindo que o modelo ARIMA pode vir a ser um bom mecanismo de apoio à tomada de decisão.

Marchezan e Souza (2010) empreendem um estudo empírico a fim de investigar as séries dos preços das três maiores culturas temporárias do estado do Rio Grande do Sul, sendo elas: soja, milho e arroz, bem como realizar previsões para o ano de 2007. Os dados-base são do período de 1995 a 2006 disponibilizados pela Empresa de Assistência Técnica e Extensão Rural - Emater/RS-ASCAR - e a escolha do melhor modelo para cada série se deu através do

uso dos critérios penalizadores AIC, SBC e Erro quadrado médio, sendo eles: ARIMA (0,1,2) (0,1,0) para soja; o modelo SARIMA (1,1,0) (1,0,1) para o arroz e o ARIMA (2,1,1) (0,1,0) para o milho.

Por sua vez, Arêdes e Pereira (2008) avaliaram a potencialidade do método ARIMA na previsão de preços do trigo no estado do Paraná. Os modelos de séries temporais estimados foram o ARIMA, o SARIMA, o ARCH, o GARCH e o TARARCH. Os resultados evidenciaram que todos esses são eficazes na previsão do preço do trigo, uma vez que os preços previstos são próximos aos observados, podendo ser utilizados como mecanismos de apoio à tomada de decisão na comercialização do trigo no estado com o intuito de diminuir seus riscos, elevando suas margens de retorno.

É relevante citar o estudo de Costa e Pavão (2009), em que foi proposto o modelo de Box e Jenkins (1976) para previsão do preço médio da carne de frango no varejo para o estado de São Paulo. Dessa forma, os dados que compreendem três períodos distintos entre 1995 e 2008 foram obtidos junto ao Instituto de Economia Agrícola (IEA) de São Paulo, uma vez que Costa e Pavão (2009) utilizam, para escolha do modelo mais adequado, o critério penalizador AIC e BIC. Após tornar a série estacionária, o mais compatível encontrado foi AR (1, 6, 30). Com base na análise de resultados, percebeu-se que os valores previstos ficaram bem próximos dos observados.

Em seu contexto, Darekar, Reddy e Amarander (2017) fazem uso da metodologia ARIMA para a previsão dos preços de arroz na Índia para a safra de 2018 utilizando dados mensais coletados pela AGmark do período de janeiro de 2006 a dezembro de 2016. Foram empregados, para uma seleção do modelo mais adequado para a previsão, o critério de MAPE, AIC e BIC, encontrando o ARIMA (0, 1, 0) (0,0, 2) como o melhor modelo, concluindo que há uma tendência de aumento de preço para ano futuro e uma demanda maior por áreas de produção. Dessa forma, enfatizou-se a limitação do modelo de representar a realidade dos preços de mercado, já que o modelo desconsidera outras variáveis exógenas que afetam o preço diário.

Sangsefidi *et al.* (2015) também utiliza a metodologia ARIMA para realizar previsão de preço de alguns produtos agrícolas no Irã, incluindo batata, cebola, tomate e bovinos, usando dados semanais desde o início de 2007 a 2015. Os resultados encontrados foram comparados aos modelos de Heterocedasticidade Condicional Autorregressiva (ARCH), sendo que a estimativa do método ARIMA adapta-se melhor do que modelos ARCH ao que se propõe.

Pérez e Pire (2015) aplicam o modelo ARIMA em seu estudo a fim de realizar uma previsão do preço internacional de arroz, fazendo uso de dados mensais do período

compreendido entre junho de 2002 e novembro de 2012. Os referidos pesquisadores encontram, como modelo mais adequado, a previsão para os próximos 14 períodos o ARIMA (1,0,1), concluindo que reflete o comportamento do modelo estocástico gerado pela série de dados, podendo ser bastante útil para previsões que auxiliem no desenho de políticas públicas e estratégias confiáveis.

### 3.1.1 Aspectos Metodológicos

No presente trabalho estima-se modelos uni variados de previsão para preço e produção de arroz no estado do Rio Grande do Sul utilizando da metodologia de Box-Jenkins (1976).

#### 3.1.2 A Metodologia Box-Jenkins

Conforme Moretin e Tolo (1981), uma série temporal é qualquer conjunto de observações ordenadas no tempo, sendo compostas por quatro elementos:

1. Tendência: verifica o sentido de deslocamento da série ao longo de vários anos.
2. Ciclo: movimento ondulatório que, ao longo de vários anos, tende a ser periódico.
3. Sazonalidade: movimento ondulatório de curta duração, geralmente inferior a um ano, sendo associada, na maioria dos casos, a mudanças climáticas.
4. Ruído aleatório ou erro: compreende a variabilidade intrínseca aos dados e não pode ser modelado.

De modo geral, há quatro diferentes modelos para realização de previsões utilizando séries temporais econômicas: modelo de regressão única, de regressão com equações simultâneas, modelo autorregressivo integrado de média móvel e modelo autorregressivo vetorial (GUJARATI, 2000).

A metodologia ARIMA de Box e Jenkins (1976) é dividida em quatro etapas:

Etapa 1 - Identificação: São identificados os filtros  $p$ ,  $d$  e  $q$ , em que  $p$  é o termo autorregressivo (AR),  $q$  o termo média móvel (MA) e  $d$  representa o número de diferenças para tornar a série estacionária.

Para a identificação do modelo, deve-se levar em consideração as funções de autocorrelação (FAC) e autocorrelação parcial (FACP). Para entender detalhes da funcionabilidade da FAC e FACP, ver Moretin e Tolo (2006).

Com o objetivo de verificar a estacionariedade das séries visando não tornar os modelos explosivos aplica-se, nesse estudo, o teste Dickey Fuller Aumentado (ADF), o qual se pode perceber que um processo estocástico será estacionário se possuir média e variância independentes do tempo, contudo terá covariâncias dependentes das diferenças entre os

instantes de tempo. Para mais informações acerca do teste e da raiz unitária, ver Dickey e Fuller (1979).

Etapa 2 - Estimativa: os parâmetros dos modelos são estimados. Nesse caso, pode-se utilizar o método de mínimos quadrados lineares ou não-lineares nos parâmetros. Uma vez estimados, a escolha dos melhores modelos será baseada em critérios de informação estatística de comparação. Os critérios utilizados serão:

- a) Critério de Informação de Akaike (AIC);
- b) Critério de Informação de Akaike (AICc);

Suas fórmulas são:

$$AIC = T \ln + 2n$$

$$AIC_c = AIC + \frac{2k(k+1)}{n-k-1} = \frac{2kn}{n-k-1} - 2 \ln(L)$$

Em que  $n$  é o número de parâmetros estimados,  $T$  o número de observações utilizadas e  $2n$  é a soma do quadrado dos resíduos e  $k$  consiste nos graus de liberdade.

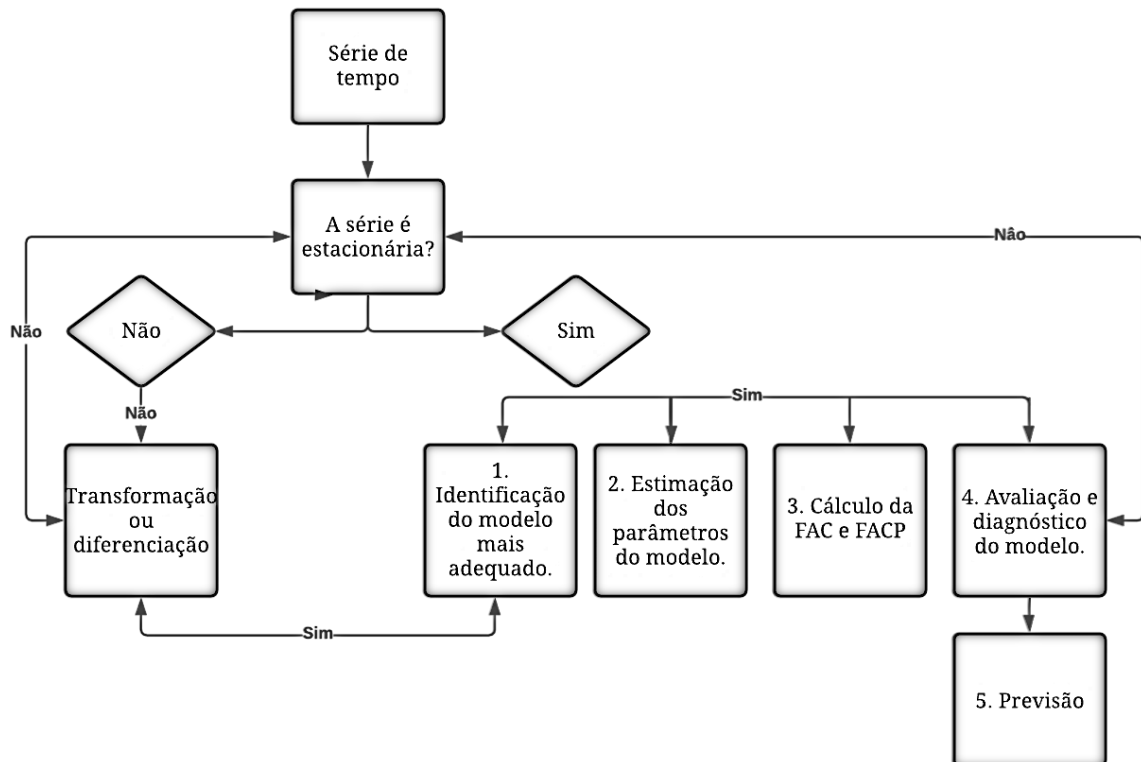
A regra de escolha é de que o melhor modelo é aquele que apresenta o menor AIC e AICc entre os candidatos. De acordo com Gujarati e Porter (2008), o critério de Akaike tem como função ponderar a qualidade da regressão e sua complexidade.

Etapa 3 - Verificação: os modelos estimados são avaliados de acordo com o seu ajuste. Averigua-se, nessa etapa, se os resíduos são ruídos brancos. Caso negativo, um novo modelo deve ser estimado. O teste que se aplica à análise dos resíduos é o teste Ljung-Box.

Etapa 4 - Previsão: por fim, deve-se verificar a potencialidade de previsão do modelo.

Para Fischer (1982), as previsões advindas da metodologia ARIMA possuem um limite do horizonte preditivo, de modo que os modelos possuem a característica geral de reverter a média quando o horizonte de previsão aumenta. Por isso, a previsão desses modelos deve ser feita no curto prazo, Granger e Newbold (1986), asseveram que a metodologia Box-Jenkins é usual em séries com mais de cinquenta observações. Em geral, a melhor forma de explicar a metodologia de Box e Jenkins (1976) é através do diagrama apresentado pela Figura 2:

**Figura 2:** Diagrama Metodologia de Box-Jenkins



**Fonte:** autor.

Todas as etapas do procedimento geral de ajuste dos modelos ARIMA aos dados são realizadas no *software R* (R CORE TEAM, 2019). O pacote *forecast*, desenvolvido por Hyndman e Khandakar (2008) e que automatiza todo o processo de previsão é usado no presente trabalho. Pacotes consistem em coleções de funções, dados e códigos compilados que ampliam as capacidades originais do *software*. Como os pacotes e o próprio *software* são de código aberto, eles estão disponíveis à verificação da comunidade científica com relação a sua correção, eficácia etc.

Para deflacionar os dados, utiliza-se do pacote *deflateBR* (Meireles, 2018), que automatiza três processos para correlacionar uma série. Primeiro, são extraídos dados de um índice - no caso, o do IGP-DI - desenvolvidos pela FGV. Segundo, o autor compatibiliza as datas mensais desses índices com as datas dos valores nominais da série. Por fim, analisa e retorna os valores corrigidos, possibilitando a comparação de valores de diferentes datas.

Nesse estudo, se faz necessário, nas duas séries temporais analisadas, realizar a transformação dos dados através do método de transformação de *Box-Cox*, de acordo com

Osborne (2010), representa a potencial melhor prática quando se deseja a normalização dos dados ou a equalização da variância, o que é visto como uma panaceia para corrigir simultaneamente a normalidade, a linearidade e reduzir a homocedasticidade.

A transformação de *Box-Cox* foi feita pelo *software* através do pacote *forecast* de Hyndman e Khandakar (2008). Outros estudos também fizeram uso da transformação logarítmica com objetivo de estabilizar a variância de séries temporais (SILVA, 2019; DOM et al., 2013).

A fórmula da transformação Box-Cox é:

$$y_t^* = \left\{ \frac{y_t^\lambda}{\lambda} \right\}, \text{ para } \lambda \neq 0$$

$$y_t^* = \log_e(y_t), \text{ para } \lambda = 0$$

Nessa fórmula,  $y$  é o dado real no momento  $t$ ,  $y_t^*$  é os dados transformados no tempo  $t$ , e  $\lambda$  é o valor mínimo do erro quadrático médio residual. A transformação na Eq. (3) é válida apenas para séries positivas,  $y_t > 0$ .

Os parâmetros  $p$ ,  $d$ ,  $q$  e  $\lambda$  são todos determinados automaticamente pelo pacote *forecast* versão 8.14 com o *software R* versão 4.1.11. Uma vez ajustados ambos os modelos aos dados de teste, o *software* faz previsões para poucos períodos a frente. Dadas as previsões e os dados de teste, o pacote *forecast* (Hyndman e Khandakar, 2008) calcula automaticamente medidas de erros de previsão e define o melhor modelo para previsão utilizando dos Critério de Informação de Akaike (AIC), Critério de Informação de Akaike corrigido (AICc).

### 3.1.3 Descrição das Variáveis

Para a consecução das estimativas de preço e produção, faz-se uso de duas bases de dados de periodicidade mensal e anual respectivamente, a base de dados da série histórica de preços do arroz, é disponibilizada na página do CEPEA/ESALQ-USP<sup>1</sup> referente ao indicador diário do arroz em casca ESALQ/SENARRS saca de 50kg, convênio entre a Fundação de Estudos Agrários Luiz de Queiroz (FEALQ) e o Serviço Nacional de Aprendizagem Rural (SENAR), com responsabilidade de execução pelo Centro de Estudos Avançados em Economia

---

<sup>1</sup> Acesso em: <https://www.cepea.esalq.usp.br/br/indicador/arroz.aspx>.

Aplicada (CEPEA) do departamento de Economia, Administração e Sociologia da escola superior de agricultura Luiz de Queiroz, da Universidade de São Paulo.

A série histórica da produção do arroz é obtida na página do Instituto Riograndense do arroz (IRGA)<sup>2</sup>, instituição que disponibiliza informações da área e produção das safras de arroz de 1921/1922 a 2020/2021 no estado do Rio Grande do Sul.

As variáveis envolvidas no processo de estimativas uni variadas da produção e de preços foram arquitetadas do seguinte modo:

a)  $prod =$  Índice de Produção (em toneladas) de arroz cultivado no estado do Rio Grande do Sul referente ao período de 1970 a 2019, considerando o ano de plantio como referência para o ano da safra. Periodicidade: Anual. Fonte: IRGA.

b)  $pre =$  Média mensal do Índice do Preço diário do Arroz em Casca (em reais por saca de 50 kg, tipo 1, 58/10) posto à vista na indústria do Rio Grande do Sul (Prazo de Pagamento descontado pela taxa CDI/CETIP), referente ao período de janeiro de 2010 a julho de 2020<sup>3</sup>. Periodicidade: Mensal. Fonte: ESALQ/SENAR

c)  $prer =$  variável pre deflacionada pelo Índice Geral de Preços - Disponibilidade Interna (IGP-DI) para valores de março de 2021. Fonte: Fundação Getúlio Vargas (FGV).

## 3.2 PROCESSO DE ESTIMAÇÃO E ANÁLISE DOS RESULTADOS

### 3.2.1 Identificação

Na presente etapa, é testada a necessidade de uma transformação da série para que essa se torne estacionária, condição essencial na construção de modelos ARIMA. Quando a série temporal observada apresenta tendência e comportamento não-sazonal, aplica-se a transformação e diferenciação com o objetivo de estabilizar a variância e remover a tendência. No modelo de Box-Jenkins, são usadas as funções de autocorrelação (FAC) e autocorrelação parcial (FACP) para identificar a ordem do modelo. Dessa forma, o modelo escolhido é aquele que o p-valor dos parâmetros é estatisticamente significativo e seus resíduos são pequenos, sendo aleatoriamente distribuídos, não contendo nenhuma informação útil.

---

<sup>2</sup>Acesso em: <https://irga-admin.rs.gov.br/upload/arquivos/201909/19141756-producao-rs-x-br.pdf>.

<sup>3</sup> A metodologia completa do índice é disponibilizada na página do CEPEA/ESALQ-USP, em: <<https://www.cepea.esalq.usp.br/br/metodologia/metodologia-do-arroz-em-casca-esalq-senar-rs.aspx>>.

A Tabela 1 apresenta uma análise descritiva dos dados utilizados no processo de estimação, a série de preços real (deflacionada pelo IGP-DI, a série de Produção e a de preços nominal) projetada as medidas de centralidade e dispersão dos dados, auxiliando no entendimento e visualização geral das séries.

**Tabela 1:** Análise descritiva dos dados

	prer	pre	Prod
Mínimo	41,68	19,08	1.347.493
Máximo	82,21	64,67	8.953.598
Média	63,53	36,93	4.801.348
Mediana	63,92	36,74	4.431.156
Desvio-Padrão	8,01	8,82	2.349.906
Variância	64,20	77,87	5.522.057

**Fonte:** autor.

A Tabela 2 exibe o resultado da transformação de Box-Cox e, de acordo com a função autoarima do pacote *forecast* de Hyndman e Khandakar (2008), possibilita a estimação do modelo mais adequado segundo o critério de escolha AICc para o objetivo de estabilizar a variância e distribuição das séries temporais.

**Tabela 2:** Resultados da transformação de Box-Cox

transformação de Box-Cox	Lambda ( $\lambda$ )
prod	-0.3124922
prer	0.9681823

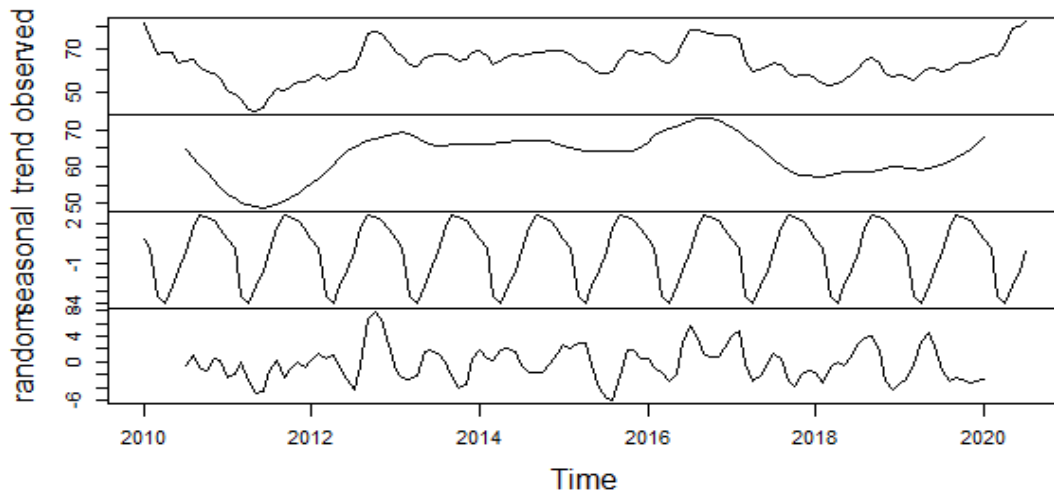
**Fonte:** autor.

O Gráfico 7 apresenta a série de preços (prer) e sua decomposição sazonal e parte aleatória, sendo analisados através da função *decompose* do pacote *forecast*. Nota-se que tais preços oscilam em torno de um valor médio e tem um padrão periódico bem definido segundo a sazonalidade aditiva e não tendencial.

Entretanto, o modelo em questão não considera a existência significativa de sazonalidade, não sendo necessário seu uso na série em questão. Vale ressaltar que, no caso de uma série temporal de periodicidade anual, não haverá componente sazonal, como no caso da série de produção (prod).



**Gráfico 7:** decomposição da série de tempo referente ao preço.



**Fonte:** autor

Para realização do teste de raiz unitária, utiliza-se a função `ur.df` do pacote `urca` (2016) implementado por Bernhard Pfaff, sendo considerado um modelo de passeio aleatório com deriva e tendência para a realização do teste. Para interpretar a saída do teste de Dickey Fuller Aumentado (ADF) apresentado na Tabela 3, entende-se que a hipótese nula é rejeitada para  $\tau \leq$  que o valor crítico  $\tau_c$ , caso seja rejeitada temos que a série é estacionária, ou seja, não tem raiz unitária. Em caso positivo, entende-se que a série é estacionária, ou seja, não tem raiz unitária; se o  $\tau \geq$  que o valor crítico  $\tau_c$ , então pode-se concluir que ela é não estacionária.

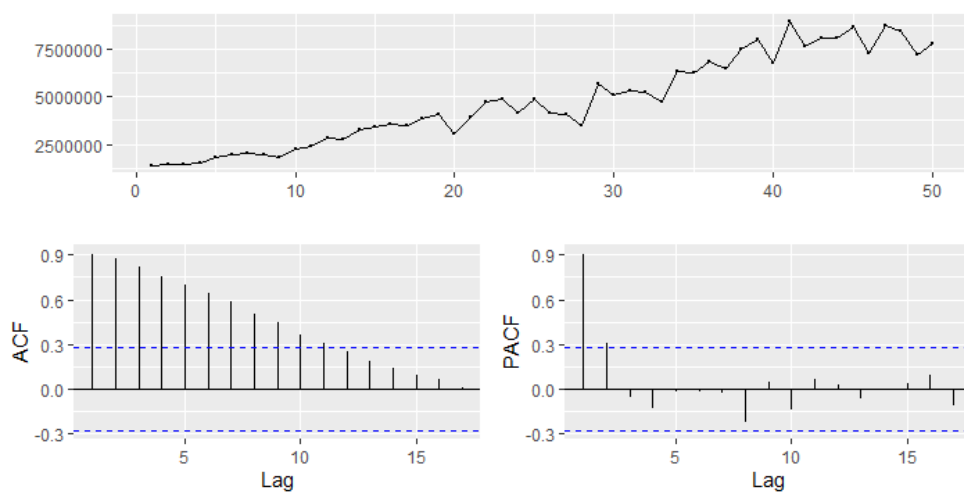
Os valores contidos na tabela mostram que a série de produção (`prod`) não rejeita a hipótese nula do teste, inferindo que o processo estocástico possui raiz unitária, sendo suficiente para a garantia da hipótese de estacionariedade a diferenciação em primeira diferença.

No caso da série de preços (`prer`), utilizou-se da função `auto.arima` do pacote `forecast` para a realização do teste de estacionariedade para a variável, nesse caso é rejeitada a hipótese nula de que a série possui raiz unitária regular com significância de 5%, logo a série é estacionária em nível, não sendo necessária a diferenciação.

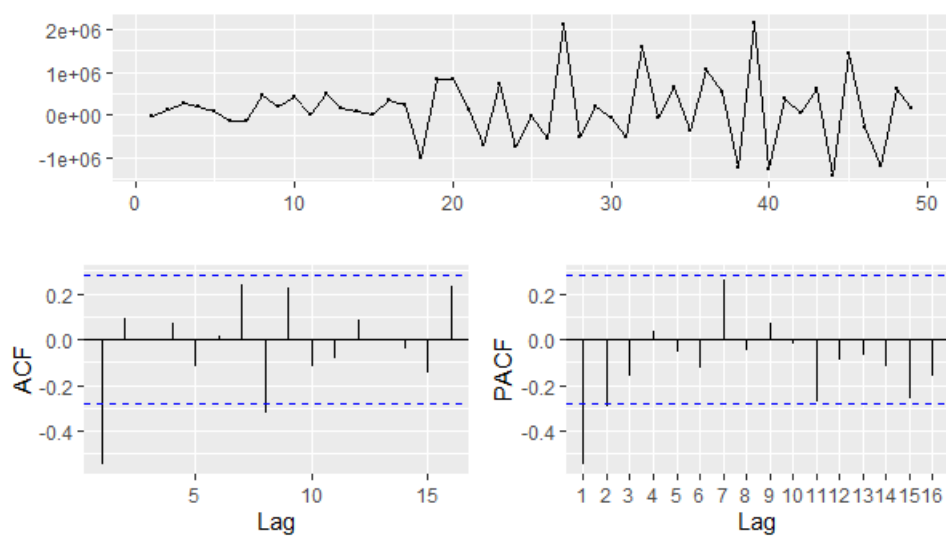
**Tabela 1:** Resultado teste Dickey-Fuller aumentado (ADF)

Variável	ADF	Valor Crítico (5%)	Interpretação
<i>prod</i>	-0.5493	-3.50	H0
<i>prer</i>	_____	-1.95	Rejeita H0

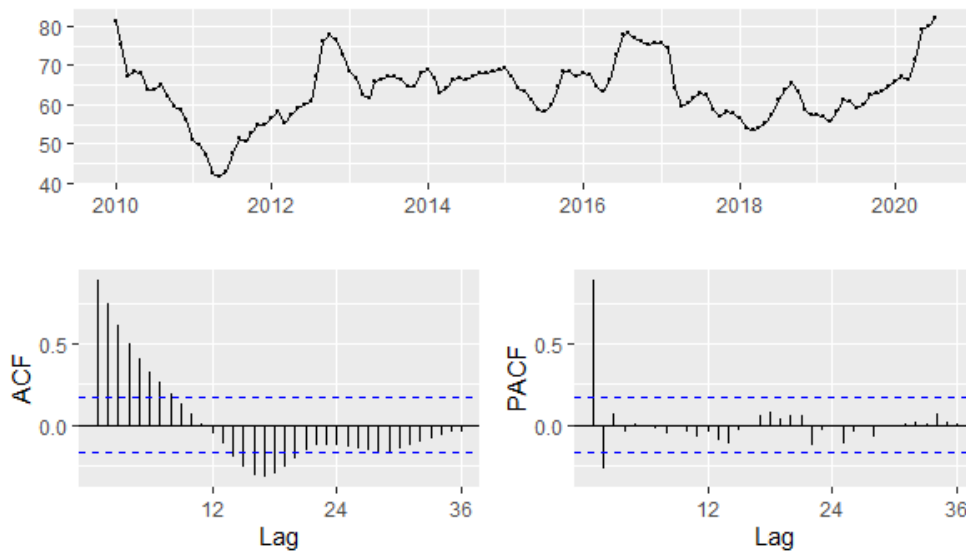
Fonte: autor

**Gráfico 8:** FAC e FACP da série de produção (prod)

Fonte: autor

**Gráfico 9:** FAC e FACP da série de produção (prod) em primeira diferença

Fonte: autor

**Gráfico 10:** FAC e FACP da série de preços (prer)

**Fonte:** autor

Por meio dos resultados da identificação pode-se aferir que, através da função “auto.arima” do pacote *forecast* de Hyndman e Khandakar (2008), estimou-se os filtros  $p$  e  $q$  e a partir do teste de estacionariedade Dickey Fuller Aumentado (ADF) o filtro  $d$ , onde  $p$  é o atraso autorregressivo,  $d$  ordem de diferenciação e  $q$  é o parâmetro de média móvel, representando respectivamente AR, I e MA.

Constata-se, portanto, que o melhor modelo ajustado para série de produção é o ARIMA (0,1,1) e para a série de preços é o ARIMA (1,0,1), representados nas equações genéricas abaixo:

$$Dprod_t = \mu + \beta_0 \varepsilon_t + \beta_1 \varepsilon_{t-1}$$

$$prer_t = \theta + d_1 prer_{t-1} + \beta_0 \varepsilon_t + \beta_1 \varepsilon_{t-1}$$

Em que  $\varepsilon$  é um termo de erro do tipo ruído branco.

### 3.2.2 Estimativa

Na presente etapa ocorre a estimativa dos parâmetros. Para tal, fez-se uso da função “auto.arima” do pacote *forecast* de Hyndman e Khandakar (2008), função essa que tem a capacidade de estimar todas as possibilidades de modelos significativos de acordo com as restrições preestabelecidas e escolher o melhor modelo de acordo com o critério escolhido pelo autor. No presente trabalho foram escolhidos os critérios de informação de akaike (AIC) e de informação de akaike corrigido (AICc).

**Tabela 4:** Resultados das estimativas dos modelos ARIMA.

VARIÁVEIS	MODELOS	PARÂMETROS			AIC	AICC
PROD	ARIMA (0,1,1)	Ma1	Intercepto		-	-
		-0.4257 (0.1321) **	0,00032 (0,00017) .		223.55	223.03
PRER	ARIMA (1,0,1)	Ar1	Ma1	Intercepto	543.21	543.53
		0.9083 (0.0420) ***	0.7778 (0.0590) ***	58.8200 (3.2769) ***		

**Fonte:** autor

( ) erro padrão, códigos de Significância: 0 ‘\*\*\*\*’; 0.001 ‘\*\*\*’ 0.01; ‘\*\*’ 0.05; ‘.’ 0.1

$$Dprod_t^{\wedge} = 0,00032 - 0.4257_{\varepsilon_{t-1}}$$

$$Prer_t^{\wedge} = 58.82 + 0.9083prer_{t-1} + 0.7778_{\varepsilon_{t-1}}$$

### 3.2.3 Verificação

Na presente etapa da análise de uma série temporal, pode-se perceber o ajuste do modelo aos dados com um nível de significância pré-estabelecido efetuando-se um estudo dos correlogramas dos resíduos. Assim, se o modelo for adequado aos dados, esses resíduos são distribuídos em torno de zero com variância de forma bem ajustada, normal, constante e aleatória, sendo preferencialmente independentes.

Na Tabela 5 encontra-se a análise dos resíduos dos modelos com os testes utilizados e suas interpretações. Para teste de ruído branco, utiliza-se o Teste Ljung-Box, cuja hipótese nula é de que os resíduos são independentes ( $H_0$ : p-valor < 0,05).

Pode-se verificar que, no caso da série de produção do modelo ARIMA (0,1,1) com média diferente de 0, o p-valor é maior que o nível de significância predeterminado de 5%, indicando que os resíduos são independentes. Ao avaliar o resultado do teste para o caso da série de preços modelo ARIMA (1,0,1) com média diferente de 0, o p-valor igual a 0.9387 rejeita a hipótese nula de que há dependência dos resíduos, comprovando o que já havia sido observado nas funções de autocorrelação nos gráficos 9 e 10 de que o modelo se adapta bem aos dados.

Considerando que o principal interesse no ajuste de séries temporais são valores futuros,

é essencial a realização dos testes de heteroscedasticidade e análise de normalidade nos resíduos dos modelos. A suposição de normalidade dos resíduos padronizados é feita pelo teste de Jarque-Bera no modelo produção, entretanto o modelo de preços não apresenta resíduos normalmente distribuídos. O teste ARCH-LM nos resíduos do modelo indica que a heteroscedasticidade condicional não está presente em ambas as séries.

**Tabela 5:** Análise dos resíduos

<b>Modelo</b>	<b>Teste</b>	<b>Estatística do teste</b>	<b>p-valor</b>		<b>Interpretação</b>
prod	Box-Ljung	0.011002	0.9165	$h_0$	Resíduos independentes
ARIMA (0,1,1)	arch-LM	Qui-quadrado 6.5438	0.8862	$h_0$	Não existe estrutura de variância condicional heteroscedástica – ARCH
Com deslocamento	Jarque-Bera teste	0.19562	0.901	$h_0$	Resíduos normalmente distribuídos
prer	Box-Ljung	0.0059	0.2258	rejeita $h_0$	Resíduos independentes
ARIMA (1,0,1)	arch-LM	Qui-quadrado 6.3894	0.8952	$h_0$	Não existe estrutura de variância condicional heteroscedástica – ARCH
Média diferente de 0	Jarque-Bera teste	25,51		rejeita $h_0$	Resíduos não tem distribuição normal
		JB	0.0055		

**Fonte:** autor

Na Tabela 6 são apresentadas as principais estatísticas de erro para análise da qualidade das previsões dos modelos estimados. Essas medidas figuram como critérios de comparação do ajuste de determinado modelo aos dados de uma série histórica (SALLES, 2005), podendo ser de natureza absoluta ou relativa.

O Erro Padrão Residual Médio (RMSE), que é o desvio padrão da amostra da diferença entre o previsto e o realizado, independe da escala, possibilitando a comparação entre os modelos; o Erro Médio Absoluto (MAE), média da diferença absoluta entre o previsto e realizado, depende de escala, não sendo possível comparar com modelos de diferentes bases de dados; o Percentual de Erro Médio Absoluto (MAPE), consistindo na diferença absoluta percentual de erro entre o previsto e o realizado, sendo que o método preferido é o que possui o MAPE mais baixo; o Erro Médio Absoluto Escalado (MASE), que é uma medida da precisão das previsões, é o erro absoluto médio dos valores de previsão dividido pelo erro absoluto médio da previsão ingênua em uma etapa da amostra. Ao comparar métodos de previsão, o método com o MASE mais baixo é o método preferido, trazendo a Função de autocorrelação dos erros

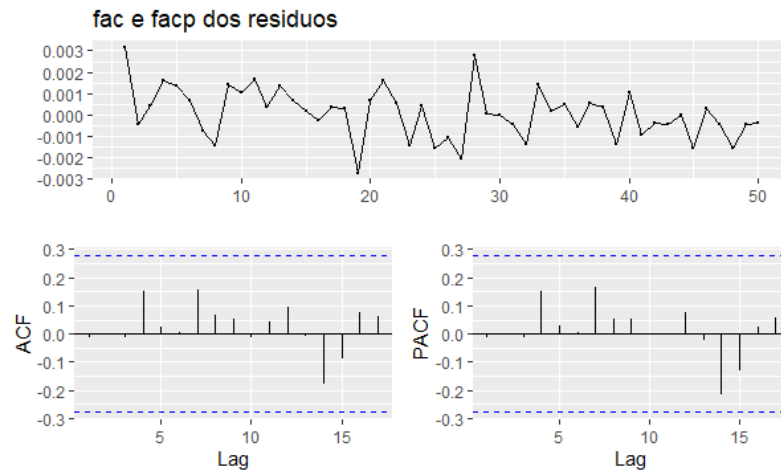
no lag 1 (FAC1).

**Tabela 6:** Estatísticas de erro.

VARIÁVEL	RMSE	MAE	MAPE	MASE	FAC1
prod	697543.4	517665.3	10.87%	0.9273806	-0.1394625
prer	2.238836	1.653141	2.63%	0.189769	-0.00651073

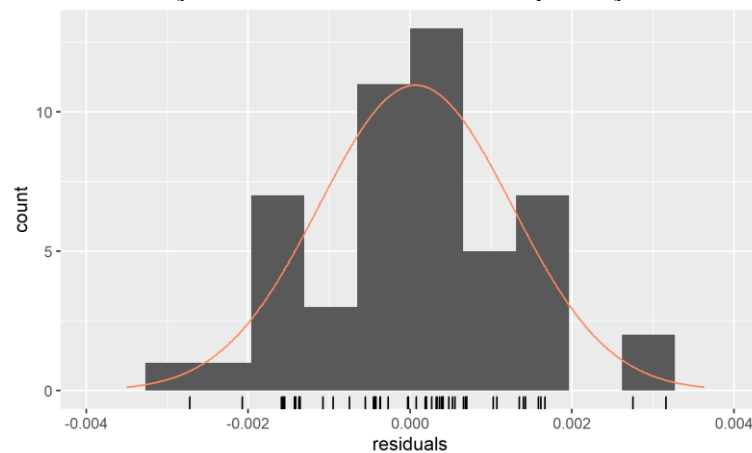
Fonte: autor.

**Gráfico 11:** FAC e FACP dos resíduos modelo de produção ARIMA (0,1,1) com deslocamento



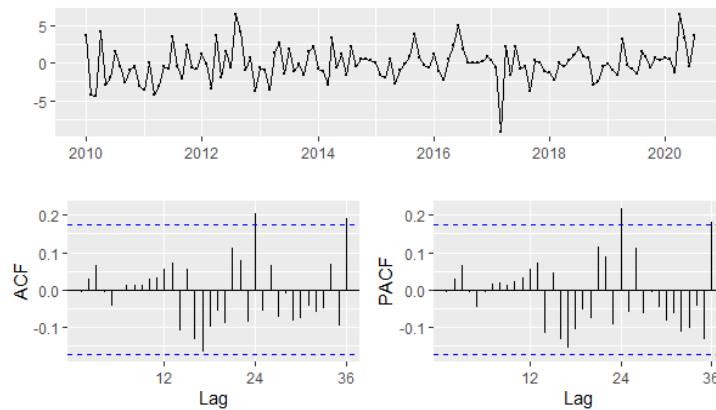
Fonte: autor.

**Gráfico 12:** distribuição dos resíduos modelo de produção ARIMA (0,1,1).



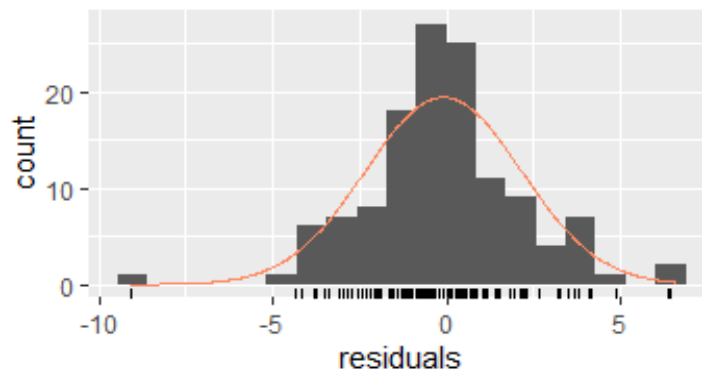
Fonte: autor

**Gráfico 13:** FAC e FACP dos resíduos do modelo de preços ARIMA (1,0,1).



**Fonte:** autor.

**Gráfico 14:** distribuição dos resíduos do modelo de preços ARIMA (1,0,1).



**Fonte:** autor

Os resultados expressos nessa subseção indicam que o comportamento dos resíduos está dentro do esperado e as medidas de erro corroboram que se pode realizar previsões a partir dos modelos estimados.

#### 3.2.4 Previsão

O procedimento de previsões realizado neste trabalho evidencia o modelo que apresenta a melhor capacidade preditiva, uma vez que realiza sobreavisos para cada modelo possível, com reajuste a cada passo, verificando o modelo em que essas se aproximam mais dos valores reais da série temporal.

Os resultados expostos na Tabela 7 indicam que o modelo para a série de produção ARIMA (0,1,1) obteve um resultado de previsão satisfatório apresentando ponto de previsão de 8,5 milhões de toneladas, cerca de 53 mil a mais que a produção observada pelo IRGA (2010) para a safra 2020/2021, valor que está inserido no intervalo de confiança proposto pelo modelo.

**Tabela 7:** Resultados da previsão de produção (prod).

Mês	Ponto de previsão	Observado	Inferior	Superior
			95%	95%
2020	8576722	8523429*	6152984	12424480
2021	8996400	-	6118807	13945434

**Fonte:** autor

\*estimado pelo IRGA.

Na Tabela 8 encontram-se os resultados das previsões de preço feitas pelo modelo ARIMA (1,0,1) para o período de agosto de 2020 a junho de 2021 para um intervalo de confiança de 95%, apresentando um bom desempenho comparado aos valores observados no período, com exceção dos meses de setembro a dezembro de 2020, quando os valores observados foram os mais altos da série de tempo analisada, historicamente extrapolando o intervalo de confiança predeterminado.

**Tabela 8:** Resultados previsão de preços (prer)

Mês	Ano	Ponto de previsão	Observado *	Inferior 95%	Superior 95%
Ago	2020	83.60392	78,94	79.16131	88.04653
Set	2020	82.01208	104,39	73.30387	90.72030
Out	2020	80.56622	105,38	69.51577	91.61667
Nov	2020	79.25295	103,98	66.59226	91.91364
Dez	2020	78.06011	96,85	64.21120	91.90902
Jan	2021	76.97666	90,47	62.21932	91.73400
Fev	2021	75.99256	88,08	60.52588	91.45925
Mar	2021	75.09871	86,05	59.07043	91.12699
Abr	2021	74.28683	87,03	57.80964	90.76402
Mai	2021	73.54940	83,25	56.71086	90.38794
Jun	2021	72.87960	73,15	55.74869	90.01052

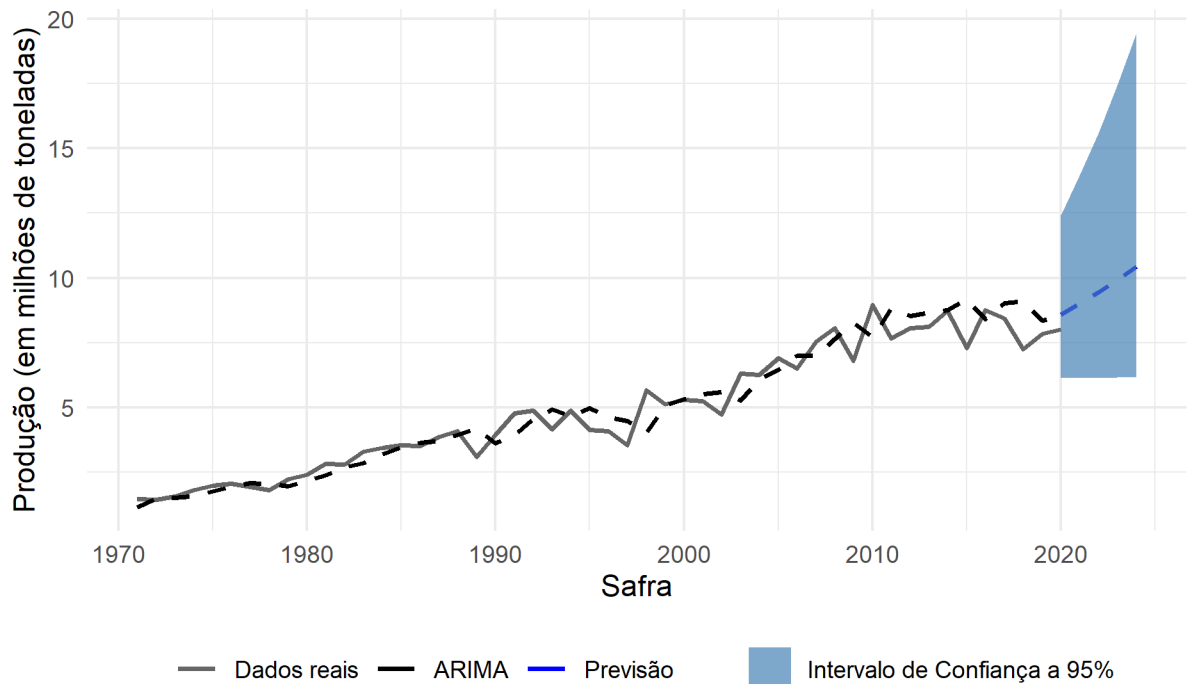
**Fonte:** autor

\* Média mensal do Índice do Preço diário do Arroz em Casca (em reais por saca de 50 kg, tipo 1, 58/10. Fonte: ESALQ/SENAR.

A comparação entre os valores obtidos na estimação e os realizados estão nos Gráficos 14 e 15. Observa-se no gráfico 15 que as previsões obtidas pelo ARIMA (0,1,1) da série de produção são próximas dos valores observados, assim como as obtidas pelo processo ARIMA (1,0,1) se aproximam dos valores percebidos da série dos preços do arroz em casca.

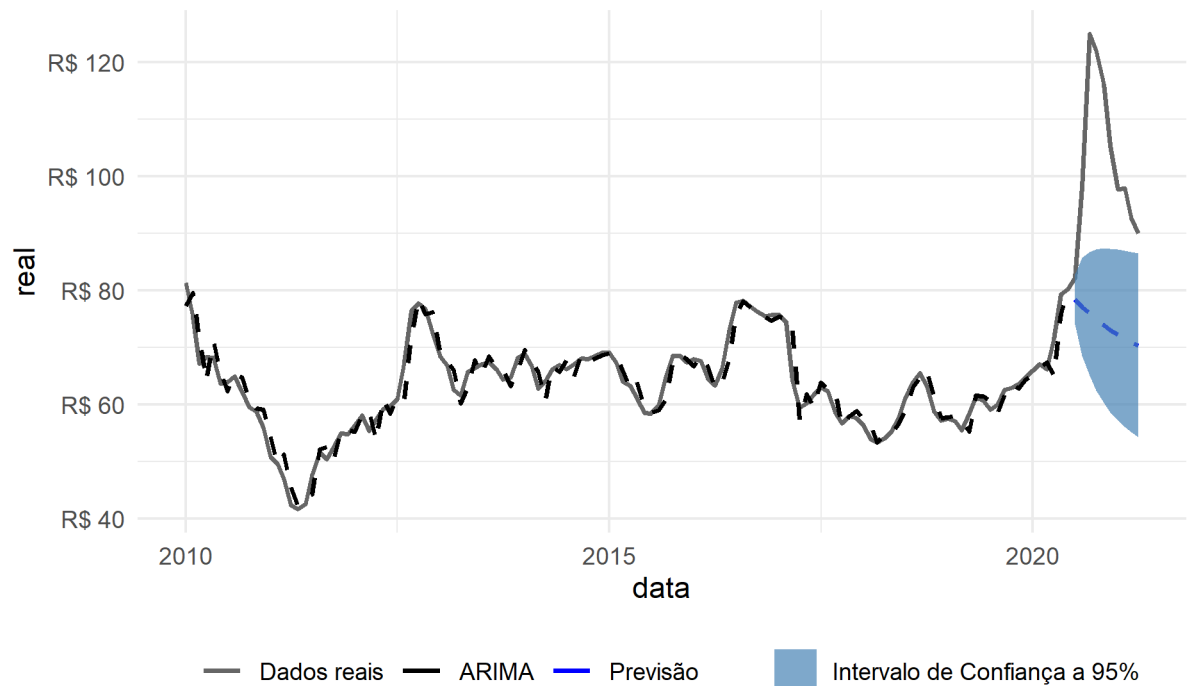


**Gráfico 15:** previsto e observado da série de produção (prod) 1970-2025.



Fonte: autor.

**Gráfico 16:** previsto e observado da série de preço (prer) 2010-2021/04.



Fonte: autor.

Portanto, pode-se aferir que os gráficos expostos no presente trabalho representam uma boa ferramenta para a visualização dos resultados dessa pesquisa evidenciando o bom

desempenho de previsão dos modelos propostos, considerando suas restrições de informações e potencialidade de suposição.

#### **4 Considerações Finais**

Ao analisar os resultados da previsão para a série de produção através do modelo ARIMA (0,1,1), pode-se aferir que apresentam parâmetros significativos, sendo possível asseverar que o critério de estacionariedade são atendidas após a diferenciação e as hipóteses de normalidade e homoscedasticidade, aceitando-se também a hipótese nula do teste de autocorrelação dos resíduos, confirmando que esses são independentes. Ao analisar as medidas de erro de previsão, reporta-se uma diferença de 10,87% entre os valores previstos e realizados (MAPE) e um erro padrão residual médio (RMSE) igual a 697.543,4.

Pode-se apontar que a série de preços através do modelo ARIMA (1,0,1) para previsão do indicador de preços do arroz foi eficiente e apresentou projeção próxima aos dados observados. Por se tratar de uma série estacionária em nível, não é necessário realizar a diferenciação, pois apresenta alta significância dos parâmetros, atende ao pressuposto de dependência dos resíduos e da homoscedasticidade, já o teste de Jarque-Bera rejeitou a hipótese nula de normalidade dos resíduos. O modelo tem medidas de erro igual a 2.63% para diferença entre o previsto e realizado (MAPE) e erro padrão residual médio (RMSE) de 2.238836.

Contudo, os resultados obtidos fornecem uma ferramenta de análise para o mercado na medida que demonstram uma tendência de preços para um horizonte de curto prazo, servindo de auxílio aos agentes do setor. Portanto, pode-se afirmar que os modelos estimados cumpriram com os objetivos do trabalho, sendo eficientes e apresentando projeções satisfatórias no curto prazo.

## 5. CONCLUSÃO

A metodologia de previsões de Box-Jenkins é muito ampla, flexível e altamente subjetiva, também sendo considerada uma importante ferramenta econométrica que pode auxiliar no processo de tomada de decisão e planejamento futuro por parte dos agentes atuantes em diversos mercados inclusive no mercado de arroz, já que uma antecipação do preço e da produção do cereal em questão gera uma série de informações benéficas aos interessados nesse mercado. Entretanto, apesar da modelagem ARIMA apresentar um adequado poder de previsão de curto prazo, por ser uma metodologia aplicada a séries uni variadas, é preciso ter cautela ao realizar previsões baseadas em valores passados.

Os referidos valores podem apresentar quantias diferentes das reais observadas, pois choques podem afetar a trajetória dessa variável em séries econômicas. Previsões não consideram fatos não esperados que possam atingir seu preço ou produção no futuro, porém isso deve ser levado em consideração, sobretudo quando se analisa o mercado de *commodity*, que está sujeito a diversas oscilações.

Desse modo, uma perspectiva futura para continuidade do estudo desse tema seria a exploração de outros métodos econométricos, modelos mais robustos com melhores performances.

## REFERÊNCIAS

ARÊDES, A. F. de; PEREIRA, M. W. G. Potencialidade da utilização de modelos de Séries temporais na previsão do preço do trigo no estado do Paraná. **Revista de Economia Agrícola**, São Paulo, v. 55, n.1, p.63-76, jan./jun. 2008. Disponível em: <http://www.iea.agricultura.sp.gov.br/ftpiea/publicar/REA2-nota.pdf>. Acesso em: 29 nov 2020.

BARROS, G.S.A.C. *et. al.* **O que provocou o repique da inflação em 2020? Centro de Estudos Avançados em Economia Aplicada**, v. 1, n. 1, abril, 2021. Disponível em: [Especial repique inflação\(1\).pdf \(usp.br\)](#) Acesso em: 30 jun 2021.

BOX, G. E. P.; JENKINS, G. M. *Time series analysis: forecasting and control*. San Francisco: Holden Day, 1976.

COSTA, J.; PAVÃO, A. Modelo de previsão de Box-Jenkins para o preço médio da carne de frango no varejo para o estado de São Paulo. **Revista Teoria e Evidência Econômica**, v. 15, n. 33, 1 set. 2009. Acesso em: 19 nov 2020.

DAREKAR, A.; REDDY, A.; AMARANDER, A. *Forecasting of Common Paddy Prices in India*. **Journal of Rice Research**, v. 10, n. 1, p.71-75, jun. 2017. Disponível em: <https://ssrn.com/abstract=3064080> ou <http://dx.doi.org/10.2139/ssrn.3064080>. Acesso em 21 nov 2020.

DOM, N. C. *et al.* *Generating temporal model using climate variables for the prediction of dengue cases in Subang Jaya, Malaysia*. **Asian Pacific Journal of Tropical Disease**, v. 3, n. 5, p. 352–361, 2013. Disponível em: <https://www.researchgate.net/publication/259173506> *Generating temporal model using climate variables for the prediction of dengue cases in Subang Jaya Malaysia*. Acesso em 21 nov 2020.

DICKEY, D. A.; FULLER, W. A. *Likelihood ratio statistics for autoregressive time series with a unit root*. **Econometrica**, 49, 1057-1072, 1981. Disponível em: [\(PDF\) The Likelihood Ratio Statistics For Autoregressive Time Series With a Unit Root \(researchgate.net\)](#). Acesso em 21

nov 2020.

FISCHER, S. **Séries univariantes de tempo metodologia de Box e Jenkins**. Porto Alegre: Fundação de Economia e Estatística, 1982.

GIL, A. C. **Métodos e Técnicas de Pesquisa Social**. Sétima Edição, Editora Atlas, 2019.

GOMES, A. S.; MAGALHÃES JÚNIOR, A. M. **Arroz Irrigado no sul do Brasil**. Brasília: Embrapa Informação Tecnológica, 2004. 899 p.

GRANGER, C. W. J.; NEWBOLD, P. *Forecasting economic time series*. 2. ed. Academic Press, 1986.

GUJARATI, D. N. **Econometria Básica**. São Paulo: Makron Books, 2000.

HALL, R. E.; LIEBERMAN, M. **Microeconomia**: princípios e aplicações. São Paulo: Pioneira Thomson Learning Edições, 2003.

HYNDMAN R. J.; KHANDAKAR, Y. *Automatic Time Series Forecasting: the forecast package for R*. **Journal of Statistical Software**, 27(3), 2008. Disponível em: [\(PDF\) Automatic Time Series Forecasting: The forecast Package for R \(researchgate.net\)](#). Acesso em 21 nov 2020.

**Indicador do arroz em casca ESALQ/SENAR-RS**. Disponível em: <https://www.cepea.esalq.usp.br/br/indicador/arroz.aspx>. Acesso em: 1 jan 2021.

INSTITUTO RIO-GRANDENSE DO ARROZ - IRGA. **Arroz irrigado no RS**: área, produção e rendimento. Disponível em: <https://irga-admin.rs.gov.br/upload/arquivos/201909/19141756-producao-rs-x-br.pdf>. Acesso em: 20 nov 2020.

MANKIW, N. G. **Introdução à economia**. São Paulo: Thomson Learning Edições, 2006.

MARCHEZAN, A.; SOUZA, A. M. Previsão do preço dos principais grãos produzidos no Rio Grande do Sul. **Ciência Rural**, Santa Maria, v. 40, n. 11, p. 2368-2374, nov./dez. 2010.

Disponível em: <http://dx.doi.org/10.1590/S0103-84782010001100019>. Acesso em: 20 nov 2020.

**Metodologia de preços do arroz ESALQ/SENAR.** Disponível em: <https://www.cepea.esalq.usp.br/br/metodologia/metodologia-do-arroz-em-casca-esalq-senars.aspx>. Acesso em: 1 jan 2021.

MARCONI, M. A.; LAKATOS, E. M. **Metodologia científica**. 6 ed. São Paulo: Atlas, 2011

MARTINEZ, E. Z.; SILVA, E. A. S. da. *Predicting the number of cases of dengue infection in Ribeirão Preto, São Paulo State, Brazil, using a SARIMA model*. **Caderno de Saúde Pública**, v. 27, n. 9, p. 1809–1818, 2011<sup>a</sup>. Acesso em 21 nov 2020.

MEIRELES, F. *DeflateBR: Deflate Nominal Brazilian Reais*. **R package version 1.1.2**, 2018. Disponível em: <https://CRAN.R-project.org/package=deflateBR>, Acesso em 15 fev. 2021.

MIRANDA, S. H. G. *et al.* A cadeia agroindustrial orizícola do Rio Grande do Sul. Disponível em: [5113-43093-1-PB.pdf](#). Acesso em 21 nov 2020.

MORETTIN, P.A.; TOLOI, C. M. **Análise de Séries Temporais**. São Paulo, Blucher, 2006.

PEREZ, S.G.C.; PIRE, R. *Estimación del precio internacional del arroz bajo en modelo ARIMA*. *Revista Mexicana de Ciências Agrícolas* v. 11, p. 2083-2089 mai. 2015. Disponível em [SSRN: http://cienciasagricolas.inifap.gob.mx/editorial/index.php/agricolas/article/view/776](https://www.ssrn.com/sol3/papers.cfm?abstract_id=2648888). Acesso em 21 nov 2020.

SANGSEFIDI, S. J. *ET. AL.* *Forecasting The Prices Of Agricultural Products In Iran With Arima And Arch Models*. **International Journal Of Advanced And Applied Sciences** v.2 n.11, p. 54-57, 2015. Disponível em: <http://www.science-gate.com/IJAAS/Articles/2015-2-11/08%202015-2-11-pp.54-57.pdf>. Acesso em: 29 nov 2020.

SAMPAIO, M. H. A.; FÜRSTENAU, V. **Agropecuária do Rio Grande do Sul – 1990-00**. Porto Alegre, Documentos FEE n.53, 2003. 76p.

SILVA, A. C. R. de. **Metodologia da pesquisa aplicada à contabilidade**: orientações de estudos, projetos, relatórios, monografias, dissertações, teses. São Paulo: Atlas, 2003.

SILVA, R.G.L. **Avaliação da Precisão dos Modelos ARIMA com e sem Transformação Estabilizadora da Variância na Previsão de Séries Temporais Anuais**. 2019. 10 f. Trabalho de Conclusão de Curso (Graduação em Gestão da Informação) – Universidade Federal de Uberlândia, Uberlândia, 2019.

PINDYCK, R. S.; RUBINFELD, D. L. **Microeconomia**. 5 ed. São Paulo: P. Hall, 2005.

ROSADO, F. **Outliers em dados estatísticos**. Lisboa: Sociedade Portuguesa de Estatística, 2006.

RSTUDIO. **RStudio: integrated development environment for R**. Boston: RStudio. Disponível em: <http://www.rstudio.com>. Acesso em 18 dez. 2020.

VIANA, J. G. A.; SOUZA, R. S.; SILVEIRA, V. C. P. Evolução dos preços históricos da bovinocultura de corte do Rio Grande do Sul: tendência e comportamento dos preços em nível de produtor e consumidor. **Ciência e Agrotecnologia**, Lavras, v. 33, n. 4, p. 1109-1117, jul./ago. 2009. Disponível em: Evolução dos preços históricos da bovinocultura de corte do Rio Grande do Sul: tendência e comportamento dos preços em nível de produtor e consumidor (scielo.br). Acesso em 21 nov 2020.

WAQUIL, P.D.; MIELE, M. SCHULTZ, G. **Mercados e Comercialização de Produtos Agrícolas**. Universidade Aberta do Brasil – UAB/SEAD/UFRGS. Porto Alegre: Editora da UFRGS, 2010. Disponível em: <https://lume.ufrgs.br/handle/10183/56447>. Acesso em: 19 nov 2020.

