

Universidade Federal de Alfenas - Unifal-MG  
Instituto de Ciências Sociais Aplicadas - ICSA

Caio Seidy de Oliveira Tanabe

**Prevendo o preço do milho:  
uma comparação entre quatro tipos de modelo**

Varginha-MG  
2019

Universidade Federal de Alfenas - Unifal-MG  
Instituto de Ciências Sociais Aplicadas - ICSA

Caio Seidy de Oliveira Tanabe

**Previendo o preço do milho:  
uma comparação entre quatro tipos de modelo**

Trabalho de conclusão de Piepex  
apresentado ao Instituto de  
Ciências Sociais Aplicadas da  
Universidade Federal de Alfenas  
como requisito parcial à obtenção  
do título de Bacharel em Ciências  
e Economia.

Orientador: Lincoln Frias

Varginha-MG  
2019

## Resumo

O objetivo deste artigo é elaborar um modelo para a previsão do preço do milho no Brasil. Antes disso, é feita uma revisão de literatura sobre a produção e o mercado nacional do milho, destacando a importância do grão como insumo básico na indústria agropecuária, a dificuldade da previsão de seu preço e as possibilidades de negociação dessa *commodity*. Os dados analisados são a série histórica preços do grão disponibilizada pelo Centro de Estudos Avançados de Economia Aplicada (CEPEA) da Esalq/USP, de 2005 a 2018 (sendo que o último ano foi usado como teste). Para identificar o modelo capaz de oferecer as melhores previsões, são comparados quatro tipos de modelo com o auxílio da linguagem de programação R: média histórica, *naive*, suavização exponencial (Holt-Winters) e o modelo autorregressivo integrado de médias móveis (ARIMA). Os resultados demonstram que as previsões mais próximas da realidade são aquelas do modelo ARIMA.

**Palavras-chave:** Milho; Previsão; Preço; ARIMA.

## Abstract

The purpose of this paper is to build a model to forecast the price of corn in Brazil. As a first, it reviews the literature on corn production and the national market for this crop, highlighting its relevance as a basic input in the agricultural industry, the difficulty of predicting its price and the possibilities of trading this commodity. The data analyzed are a historical price series from the Center for Advanced Studies in Applied Economics (Cepea) of Esalq/USP, from 2005 to 2018 (using the last year as the test set). To identify the model that offers the best forecasts, four model types are compared using the R programming language: historical mean, *naive*, exponential smoothing (Holt-Winters) and the integrated autoregressive mobile media model (ARIMA). The results show that the forecast from the ARIMA model were the closest to the observed values.

**Keywords:** Corn; Forecast; Price; ARIMA.

**Lista de Tabelas**

Tabela 1 - Medidas descritivas do preço do milho .....	20
Tabela 2 - Dez primeiros coeficientes de autocorrelação da série treino .....	24
Tabela 3 - Precisão dos modelos .....	25

## Lista de Gráficos

Gráfico 1 - Série histórica original do preço do milho .....	18
Gráfico 2 - Série histórica deflacionada do preço do milho .....	18
Gráfico 3 - Dados da série separados em treino e teste .....	19
Gráfico 4 - Série treino a ser estudada .....	19
Gráfico 5 - Decomposição, pelo modelo <i>STL</i> , da série treino .....	21
Gráfico 6 - Comparação sazonal entre os anos da série treino dos preços do milho .....	21
Gráfico 7 - Comparação entre meses ao longo da série com dados treino .....	22
Gráfico 8 - Correlação de defasagens .....	23
Gráfico 9 - Autocorrelação dos dados de treino .....	23
Gráfico 10 - Autocorrelação parcial dos dados de treino .....	24
Gráfico 11 - Comparação dos modelos de previsão aos dados de teste .....	25
Gráfico 12 - Comparação dos modelos de previsão aos dados de teste com recorte de observações .....	26
Gráfico 13 - Diagnóstico residual da série .....	27
Gráfico 14 - Previsão dos preços do milho pelo modelo ARIMA (2, 0, 1) .....	27
Gráfico 15 - Previsão dos preços do milho pelo modelo ARIMA (2, 0, 1) com recorte de observações .....	28

**Sumário**

<b>1. Introdução</b>	<b>7</b>
<b>2. Referencial bibliográfico</b>	<b>8</b>
2.1. Dificuldades na previsão de preço do milho	9
2.2. O funcionamento do mercado de milho	11
<b>3. Materiais e métodos</b>	<b>13</b>
3.1. A análise de séries temporais	13
3.2. Os dados analisados	17
<b>4. Resultados e discussão</b>	<b>20</b>
<b>5. Considerações finais</b>	<b>28</b>
<b>Referências</b>	<b>29</b>
<b>Anexo</b>	<b>34</b>

## 1. Introdução

O milho é um insumo de extrema importância tanto para alimentação humana quanto para a produção de ração animal (PAES, 2006). Além disso, pode-se considerar sua cadeia produtiva como uma das mais importantes para o agronegócio brasileiro (CALDARELLI; BACCHI, 2012).

De acordo com o Departamento de Agricultura dos Estados Unidos (USDA), os maiores produtores mundiais entre os anos-safra de 2007/08 e 2016/17 foram os Estados Unidos (35,5%), China (21,5%) e Brasil (8,1%), responsáveis por 57% da produção mundial do grão. Além disso, nesse mesmo período, o Brasil foi o terceiro colocado também entre os exportadores de milho (18%), atrás somente de Estados Unidos e Argentina. Somados ao quarto colocado, a Ucrânia, esses quatro países são responsáveis por 83% das exportações totais (NETO, 2018).

A posição privilegiada do Brasil dentre os maiores produtores e exportadores de milho do mundo é devida ao crescimento da produção do grão, que, segundo a Companhia Nacional de Abastecimento (CONAB), de 1976 a 2016 foi de aproximadamente 245%. Esse fato é explicado pelo aproveitamento de terras férteis de produtores que cultivam outras culturas, como é o caso da soja (CONAB, 2017).

Em agosto de 2019, segundo dados do Centro de Estudos Avançados em Economia Aplicada (CEPEA) da Esalq/USP, o preço à vista por saca de 60 kg de milho era de R\$36,33, conforme o indicador do milho Esalq/BM&FBOVESPA (CEPEA, 2019). O valor retratado tem periodicidade mensal e pode apresentar variações devido a elementos que possam alterar a oferta e a demanda do milho. Assim, pode-se dizer que o preço do milho é muito volátil devido a diversas variáveis que criam dificuldades aos agricultores na precificação de seu produto. Mais adiante, serão apresentadas algumas dessas variáveis que influenciam a formação de preços do grão.

Portanto, tendo em vista a importância do milho na economia nacional, este trabalho tem por objetivo utilizar modelos de séries temporais para modelagem e previsão do preço do milho no Brasil. A segunda seção será dividida em duas subseções que descrevem a dificuldade da previsão do preço do milho, além de retratar o funcionamento do mercado dessa *commodity*. Em seguida, a terceira seção apresenta os materiais e métodos utilizados.

Na quarta seção, serão expostos os resultados e discussão, seguidos das considerações finais na quinta e última seção.

## **2. Referencial bibliográfico**

Como dito anteriormente, o milho é caracterizado por destinar-se para consumo humano e animal, além de poder ser destinado para fabricação de etanol, o que nos últimos anos tem tido influência direta no aumento da produção do grão, principalmente no Brasil (MATTOSO et al., 2006; QUEIROZ, 2019).

No âmbito animal, auxilia como macroingrediente para a produção de rações. Estima-se que cerca de 70% do milho produzido no mundo é destinado ao consumo animal (PAES, 2006). Dentre seus destinos, destacam-se a avicultura, suinocultura e bovinocultura. Na bovinocultura, tem-se a produção de silagem para alimentação de vacas em produção de leite, além de ser utilizado para engorda do gado no período de inverno. Na criação de suínos e aves, o cereal é empregado na alimentação, em misturas com concentrados proteicos (MATTOSO et al., 2006).

Com as diversas formas de utilização, o grão passou a ser a cultura mais produzida do planeta, possuindo grande quantidade e disposição no mercado, e por isso tornou-se um ativo de alta liquidez para negócios (IMEA, 2015).

No que concerne à produção nacional do milho, tem-se que o plantio é realizado em duas etapas, a primeira e segunda safras. A primeira safra, ou plantio de verão, é realizada na época tradicional, durante o período chuvoso, com plantio ocorrendo de agosto a dezembro e a colheita prevista para novembro a março (EUZÉBIO, 2016). Já a segunda safra, também conhecida como “safrinha” ou “milho de sequeiro”, é plantada extemporaneamente, sem necessidade de irrigação (CRUZ et al., 2019). O plantio da “safrinha” ocorre no período de janeiro a março e a colheita entre abril e junho (EUZÉBIO, 2016). O aumento da produção está diretamente relacionado à implementação da segunda safra, que vem ganhando importância no agronegócio brasileiro nos últimos 40 anos devido a seu baixo custo.

Conforme dados disponibilizados pelo Instituto Brasileiro de Geografia e Estatística (IBGE), no período de 1990 a 2015 a região Centro-Oeste aparece como a principal região produtora do grão, tendo um crescimento de aproximadamente 520%. Esta expansão está



diretamente relacionada à implementação de novas tecnologias, aumento da produtividade e expansão das áreas cultivadas (CONAB, 2017). Além do destaque regional, o Mato Grosso aparece como o principal estado produtor no âmbito nacional desde a safra de 2011/12 e Sorriso-MT como o principal município produtor do cereal do Brasil (CONAB, 2017).

Vale lembrar que, da mesma maneira que a soja, o milho é insumo básico para a avicultura e a suinocultura, dois mercados muito competitivos internacionalmente e bastante relevantes para a economia brasileira.

## **2.1. Dificuldades na previsão de preço do milho**

O milho é classificado como uma *commodity*, ou seja, é um produto que possui qualidade e características uniformes, é produzido em grande escala, podendo ser estocado sem perda de qualidade e em que não há grande diferenciação do produto entre os diversos produtores. Além disso, seu preço é predominantemente determinado pela oferta e demanda internacional (WIKIPÉDIA, 2019). Essa classificação coloca-o como um cereal de determinantes de preço complexas, visto que esse valor varia conforme efeitos de sazonalidade, estoques e custos diversos (CORSINI; RIBEIRO, 2008).

A influência da sazonalidade no preço do milho é explicada por variações na temperatura e precipitação ou ainda pela ocorrência de cheias, secas, geadas entre outras causas naturais, que têm efeitos sobre temporadas de plantio e desenvolvimento da cultura e colheita. Mudanças climáticas desse porte podem ser prejudiciais para produtores, pois prejudicam lavouras inteiras, reduzindo a oferta do cereal. O clima, no entanto, pode ser favorável em determinados períodos, o que ajuda no cultivo da lavoura. Tal fato, explica a alta nos preços do grão em alguns meses do ano, quando a oferta é menor e baixa em outros, quando a produção é alta (PINO, 2014).

Outro fator que contribui para a precificação do milho são os estoques. Se porventura, da produção de períodos anteriores restarem estoques após as vendas, na próxima época de cultivo a quantidade inicial estocada irá aumentar (ABIMILHO, 2019). Com esse aumento, a oferta total eleva-se, o que resulta em uma queda no preço de venda. Em contrapartida, quanto menor o estoque inicial, menor a oferta total e maior o preço.

Além da oferta, a demanda apresenta-se como determinante direta da variação do preço do produto. O aumento da demanda é explicado pela crescente importação do cereal por

países como o Irã, que até a presente data, é o maior importador de milho brasileiro, segundo dados do Ministério da Indústria, Comércio Exterior e Serviços (MDIC, 2019). Com a demanda alta, o preço do grão tende a aumentar, o que é positivo para os produtores. Entretanto, a ocorrência de crises ou conflito de interesses entre países poderá resultar em uma redução da demanda da mercadoria e por consequência, do preço, por elevar os riscos do negócio.

Por último e não menos importante, há grande variabilidade nos custos de produção (fertilizantes, defensivos, combustível, frete etc.), os quais também possuem influência direta na determinação de preços por parte dos produtores. Tais gastos preocupam agricultores, que ao praticar gestão de custos de fatores de produção objetivam no mínimo cobri-los, o que evita prejuízos e resulta em lucros maiores (TIBULO; CARLI, 2014).

Como em sua maior parte é produzido em grande escala, o que o deixa vulnerável aos fatores apresentados anteriormente que influenciam diretamente a formação de seus preços, o milho está sujeito a variações em seu valor monetário (CORSINI; RIBEIRO, 2008). Com esta elevada variabilidade no preço, o mercado da *commodity* milho passou a dispor de uma nova modalidade de comercialização, os derivativos (contratos de vencimento futuro em que ganhos ou perdas dependem do comportamento de ativos subjacentes), que serão mais detalhados na próxima subseção (CALEGARI et al., 2012). Os contratos mais utilizados são os de mercado futuro, em que são feitos contratos de venda antecipada, com a fixação de um preço de venda por parte do produtor que recebe o valor do produto em data futura estimada, sem precisar ter em estoque a produção (TIBULO; CARLI, 2014).

Este tipo de comercialização é interessante para os agricultores que desejam assegurar preços, visto que esta fixação garante no mínimo a cobertura de seus custos de produção e despesas extras e ainda ter-se uma margem de lucro. O empecilho é que esse tipo de comércio nem sempre é um bom negócio para os agricultores. Para resolver tal problema, uma alternativa seria reduzir a incerteza no processo de tomada de decisões econômicas prevendo o comportamento de preços da *commodity* milho com auxílio de modelos de previsão de séries temporais (TIBULO; CARLI, 2014).

## 2.2. O funcionamento do mercado de milho

O milho pode ser negociado em quatro grandes mercados: no mercado físico, a termo, mercado futuro e mercado de opções. Esses são os grupos de operações praticadas em todo o mundo, indicadas especialmente em negociações de *commodities*, como a soja e o milho (FRANÇA JÚNIOR, 2015).

O mercado físico nada mais é do que a troca do produto por dinheiro, uma vez que se trata de uma troca imediata, ou seja, a entrega do produto e seu pagamento acontecem simultaneamente. Essa modalidade possui como característica apresentar alto nível de incerteza no que se refere ao comportamento dos preços, regularidade de suprimentos e qualidade de produtos (IMEA, 2015). Dessa forma, quando o produtor adota esse mecanismo, ele assume um risco que pode ocasionar insucesso nas suas operações, devido a incertezas relacionadas à oferta e demanda.

Uma forma de reduzir esse risco é adotar outro mecanismo de comercialização, como o mercado a termo, que possui diferenças do mercado físico nas transações, isto é, no mercado a termo as transações ocorrem em dois ou mais instantes no tempo. Algumas características desse contrato são que as partes negociam elementos da operação que irão ocorrer no futuro, ainda com algumas especificações contratuais, como a mercadoria negociada, a data de entrega, o local, o meio de transporte e pagamento, além de que o vendedor e comprador podem incorporar elementos no acordo (IMEA, 2015).

O funcionamento do mercado a termo se dá, por exemplo, na compra antecipada de milho por indústrias desse ramo. Antes da semeadura ou até mesmo durante o ciclo de plantio, o produtor pode optar por vender sua safra para a indústria processadora por um preço prefixado com pagamento em data futura. Dessa forma, o agricultor garante a venda da sua safra a um preço fixo. Por outro lado, a agroindústria tem a oportunidade de fazer planejamento de compras e recebimentos, com o intuito de aproveitar da melhor forma possível a capacidade de processamento da sua planta industrial (IMEA, 2015).

No entanto, apesar de ser uma maneira de vendedores e compradores eliminarem o risco de flutuação indesejada no preço (FRANÇA JUNIOR, 2015), pode haver empecilhos com relação a esse contrato, visto que há um risco de uma ou ambas as partes, por razões oportunistas, não o cumprirem, ainda que haja obrigações entre elas. Entretanto, o aspecto

positivo de um acordo como esse é que o agricultor não precisa retirar recursos do caixa para fechar o contrato, uma vez que a liquidação deste se dá somente no seu fechamento (IMEA, 2015).

O contrato a termo evoluiu e dele houve a formação do mercado futuro, um pouco mais complexo que seu progenitor. Algumas de suas diferenças estão na padronização acentuada dos contratos e negociação em bolsas organizadas que não permitem a inclusão de cláusulas por parte de vendedores e compradores, ou seja, tais contratos são registrados em bolsas de mercadorias e futuros. Além disso, nesse tipo de contrato, diferentemente do mercado a termo, há a especificação apenas de lugar, período de entrega, objeto transacionado (restritos à *commodities*) e lotes padrão (IMEA, 2015).

Em suma, o principal objetivo de um contrato futuro é a redução dos riscos observados em transações no mercado físico. Isto é, na compra ou venda de um contrato não há necessidade de avaliação de risco de seu descumprimento ou inspeção do produto comercializado, uma vez que, como já foi mencionado, esse é assegurado pela instituição responsável pela transação, que no caso brasileiro é a Bolsa de Mercadorias & Futuros (BM&F). A redução de risco é denominada *hedge* (proteção), que funciona através de mecanismos que asseguram a operação de fixação de preço para liquidação em uma data futura (IMEA, 2015).

Outro mercado onde o milho pode ser negociado é o de opções, que são contratos derivativos que têm por objetivo a transferência do risco de oscilação de preços do vendedor (produtor) ou comprador de uma *commodity* para outro agente do mercado, mediante o pagamento de um prêmio. Estes, funcionam como acordos que asseguram o direito de exercício de uma compra (*call*) ou de uma venda (*put*) de algum ativo (PEROBELLI, 2006). Essas negociações podem acontecer por um ativo físico (como mil toneladas de milho) ou um contrato futuro (IMEA, 2015).

Por exemplo, um produtor de milho que está sujeito ao risco de queda do preço de seu produto compra o direito de venda do grão a um preço de exercício que ao menos cubra seus custos. Dessa forma, caso ocorra uma queda dos preços, o agricultor pode exercer o direito de vender seu produto pelo preço prefixado, evitando prejuízos. Por outro lado, uma agroindústria que está sujeita ao risco de uma elevação do preço de seus insumos, pode comprar um direito de compra a um preço de exercício predeterminado. O valor do prêmio

pago por esses direitos possuem variações relacionadas ao tempo de vencimento do ativo e riscos do mercado (IMEA, 2015).

Apesar desses quatro grupos de mercado possíveis para o milho facilitarem as negociações, a falta de mecanismos para a previsão de preços dessa *commodity* pode aumentar o risco desses contratos. Por isso, é importante desenvolver modelos estatísticos que auxiliem nessa tarefa.

### 3. Materiais e métodos

Para a previsão do preço do milho foram comparados quatro modelos: média histórica, *naive*, suavização exponencial (*Holt-Winters*) e autoregressivo integrado de médias móveis (ARIMA). A análise foi feita por meio da linguagem de programação R (R CORE TEAM, 2019) e os códigos utilizados estão disponíveis no anexo.

#### 3.1. A análise de séries temporais

No modelo de média histórica, como o próprio nome diz, utiliza-se da média da série, ou seja, os valores a serem previstos são iguais à média dos valores da série estudada. Considera-se que os valores futuros serão iguais à média aritmética simples de todos os dados da série (HYNDMAN; ATHANASOPOULOS, 2018, seção 8.1):

$$\hat{y}_{t+h|t} = \frac{1}{T} \sum_{t=1}^t y_t ,$$

para  $h = 1, 2, \dots$ , em que  $h$  é o horizonte a ser previsto e onde  $t$  é o último período observado e  $T$  é a quantidade de observações da série temporal e  $\hat{y}$  é o resultado da previsão dos valores utilizando todos os valores observados na série  $(y_1, \dots, y_t)$ .

No método *naive* (ingênuo), os valores previstos são iguais ao último valor da série, isto é, para o presente estudo, a cotação do preço do milho prevista é igual ao último dado da série treino. De uma forma geral, observa-se que o método ingênuo assume que a observação

mais recente é a única importante e todas as observações anteriores não fornecem informações para o futuro. Isso pode ser pensado como uma média ponderada em que todo o peso é atribuído à última observação (HYNDMAN; ATHANASOPOULOS, 2018).

$$\hat{y}_{t+h|t} = y_t ,$$

para  $h = 1, 2, \dots$  e onde  $t$  é o último período observado.

Entre esses dois extremos apresentados, atribuir peso igual a todas as observações (método da média histórica) e todo o peso à última observação (método ingênuo), existe a suavização exponencial simples, adequada para a previsão de dados sem tendência clara ou padrão sazonal. Neste modelo, há a atribuição de pesos maiores, porém decrescentes, às observações mais recentes. Portanto, as previsões são calculadas utilizando médias ponderadas, em que os pesos diminuem exponencialmente, isto é, quanto mais no passado a observação estiver, menor o peso atribuído a ela (HYNDMAN; ATHANASOPOULOS, 2018, seção 8.1).

$$\hat{y}_{t+1|t} = \alpha y_t + \alpha(1 - \alpha)y_{t-1} + \alpha(1 - \alpha)^2 y_{t-2} + \dots ,$$

onde  $0 \leq \alpha \leq 1$  é o parâmetro de suavização, ou seja, a variável que controla a taxa na qual os pesos diminuem. A previsão para um passo à frente do tempo, isto é,  $t + 1$  é o cálculo da média ponderada de todas observações da série histórica,  $y_1, \dots, y_t$  (HYNDMAN; ATHANASOPOULOS, 2018, seção 8.1).

No método *Holt*, há uma extensão da suavização exponencial simples que permite a previsão de dados com tendência. Para isso, necessita-se de uma equação de previsão e duas outras de suavização, uma para o nível e a outra para tendência:

$$\text{Equação de previsão: } \hat{y}_{t+h|t} = l_t + hb_t$$

$$\text{Equação de nível: } l_t = \alpha y_t + (1 - \alpha)(l_{t-1} + b_{t-1})$$

$$\text{Equação de tendência: } b_t = \beta^*(l_t - l_{t-1}) + (1 - \beta^*)b_{t-1} ,$$

onde  $l_t$  indica uma estimativa do nível da série no tempo  $t$ ,  $b_t$  indica uma estimativa da tendência da série no tempo  $t$ , isto é, sua inclinação,  $\alpha$  é o parâmetro de suavização para o nível ( $0 \leq \alpha \leq 1$ ) e  $\beta^*$  é o parâmetro de suavização para a tendência ( $0 \leq \beta^* \leq 1$ ) (HYNDMAN; ATHANASOPOULOS, 2018, seção 8.2).

O método de Holt-Winters é a adaptação do método de suavização exponencial de Holt para lidar também com dados sazonais. Nesse modelo há quatro equações, sendo uma de previsão e três de suavização (uma para o nível, outra para a tendência e a última para o componente sazonal,  $s_t$ ). No modelo aditivo *Holt-Winters* tem-se:

$$\text{Equação de previsão: } \hat{y}_{t+h|t} = l_t + hb_t + s_{t+h-m(k+1)}$$

$$\text{Equação de nível: } l_t = \alpha(y_t - s_{t-m}) + (1 - \alpha)(l_{t-1} + b_{t-1})$$

$$\text{Equação de tendência: } b_t = \beta^*(l_t - l_{t-1}) + (1 - \beta^*)b_{t-1}$$

$$\text{Equação do componente sazonal: } s_t = \gamma(y_t - l_{t-1} - b_{t-1}) + (1 - \gamma)s_{t-m},$$

onde,  $k$  é a parte inteira de  $(h - 1)/m$ , o que assegura que a utilização das estimativas dos índices sazonais na previsão sejam provenientes do ano final da amostra. A equação de nível apresenta uma média ponderada entre a observação com ajuste de sazonalidade ( $y_t - s_{t-m}$ ) e a previsão não sazonal ( $l_{t-1} + b_{t-1}$ ) para o tempo  $t$ . A equação de tendência é idêntica ao modelo *Holt*. A equação do componente sazonal mostra uma média ponderada entre o índice de sazonalidade atual ( $y_t - l_{t-1} - b_{t-1}$ ) e o do mesmo subperíodo no ano passado ( $m$  períodos atrás) (HYNDMAN; ATHANASOPOULOS, 2018, seção 8.3).

O modelo autoregressivo integrado de médias móveis (ARIMA) tem sua estrutura dividida em três partes: autorregressiva, diferenciação e média móveis. Para entender a parte autorregressiva, faz-se uma analogia com o modelo de regressão múltipla, que de uma forma geral, é utilizado para prever a variável de interesse através do emprego de uma combinação linear de preditores. Já no modelo autoregressivo, a previsão da variável desejada se dá com a utilização de uma combinação linear de valores passados dessa mesma variável. A autorregressão tem esse nome pois é uma regressão variável nela mesma. A essa parte, a notação empregada é  $AR(p)$ , ou seja, um modelo autoregressivo de ordem  $p$  (número de valores defasados da variável) (HYNDMAN; ATHANASOPOULOS, 2018, seção 9.3).

O modelo de média móvel, em vez de usar valores passados da variável de previsão como em uma regressão, utiliza erros de previsão passados em uma abordagem semelhante à regressão, porém não usual, uma vez que não observa-se valores de ruído branco ( $\varepsilon_t$ ), ou seja, cada valor da série com média zero, variância constante e sem correlação serial. A notação dada à parte da média móvel do erro é  $MA(q)$ , isto é, um modelo de média móvel de ordem  $q$  (HYNDMAN; ATHANASOPOULOS, 2018, seção 9.4).

O modelo de diferenciação é utilizado para transformar em estacionária (série com média e variância constantes) uma série temporal não estacionária, para isso há um cálculo das diferenças entre observações consecutivas (HYNDMAN; ATHANASOPOULOS, 2018, seção 9.1), onde a integração (o “I” em ARIMA) aparece como o inverso da diferenciação.

A combinação dos modelos de diferenciação, autorregressão e de médias móveis resulta em um modelo denominado ARIMA ( $p, d, q$ ), onde:

- $p$  = ordem da parte autoregressiva;
- $d$  = grau da diferenciação envolvida;
- $q$  = ordem da parte da média móvel.

Para a escolha das melhores ordens de  $p$ ,  $d$  e  $q$ , primeiramente deve-se conhecer o *AIC*. Este é um acrônimo para critério de informação de Akaike, um indicador da melhor relação entre qualidade do ajuste e simplicidade do modelo, penalizando aumentos da complexidade do modelo (inclusão de novas variáveis explicativas) que não levem a melhorias substanciais de seu ajuste. Com a inclusão de correção para amostras pequenas tem-se o *AICc*. Dessa forma, os melhores parâmetros são escolhidos minimizando o *AICc*, ou seja, quanto menor o *AICc* melhor o modelo (WIKIPÉDIA, 2019; HYNDMAN; ATHANASOPOULOS, 2018, seção 6.5).

A escolha do melhor modelo de previsão deu-se ao reservar os 12 meses de 2018, apresentados na série histórica, para comparação dos valores previstos. Também é utilizado um critério de avaliação do melhor modelo de previsão: RMSE (*Root-mean-square error*) que é a raiz do quadrado do erro médio da diferença entre o valor previsto e o valor real:

$$RMSE = \sqrt{\frac{\sum_{t=1}^n (P_t - O_t)^2}{n}}$$



onde:

- $P_t$  é o valor previsto para o instante  $t$ ;
- $O_t$  é o valor da série no instante  $t$ ;
- $n$  é o número período de tempos.

A escolha do modelo ARIMA para previsão do preço do milho, justifica-se pela necessidade de um modelo que lidasse com as componentes tendência e (ou) sazonalidade tradicionalmente presentes em séries que envolvem produtos produzidos e consumidos mundialmente. Ajustar e escolher o melhor modelo tem por objetivo possibilitar a previsão de preços da *commodity* milho, com o intuito de auxiliar agricultores no comércio de seu produto.

### 3.2. Os dados analisados

Este trabalho utiliza dados referentes à série histórica do preço mensal do milho no Brasil no período de Agosto de 2004 a Agosto de 2019, no total de 181 observações. Tais dados foram disponibilizados pelo Centro de Estudos Avançados de Economia Aplicada (CEPEA) da Esalq/USP.

Um empecilho encontrado na comparação de valores no tempo é a taxa de crescimento do nível geral de preços, ou simplesmente inflação. Para isso, deve-se trabalhar sempre com a mesma base de preços eliminando os efeitos da inflação. Dessa forma, deflaciona-se os dados da série histórica do preço do milho, tendo como referência o mês julho de 2019, os códigos utilizados estão disponíveis no anexo. O Gráfico 1 apresenta a série original, já o Gráfico 2, a série deflacionada que será objeto de estudo para esse trabalho. Os dois gráficos possuem séries temporais de mesmo período, ou seja, de agosto de 2004 a agosto de 2019.

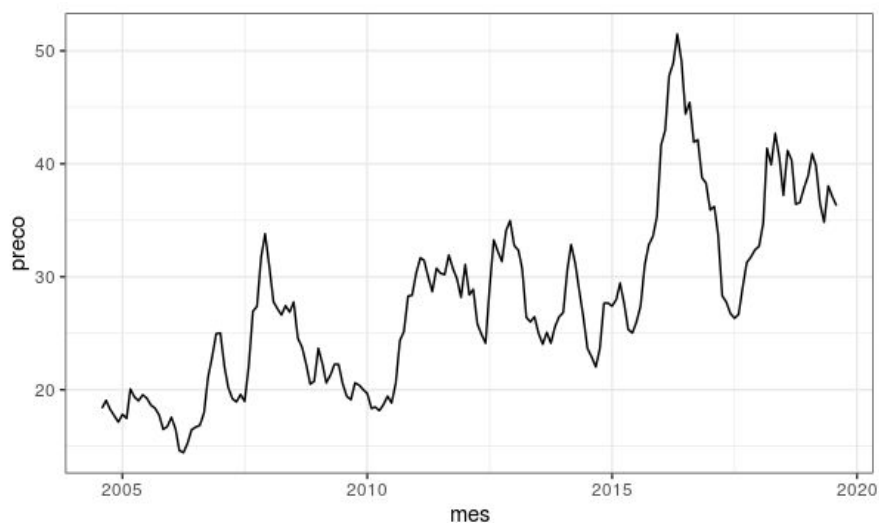


Gráfico 1 - Série histórica original do preço do milho.

Fonte: Elaboração própria.

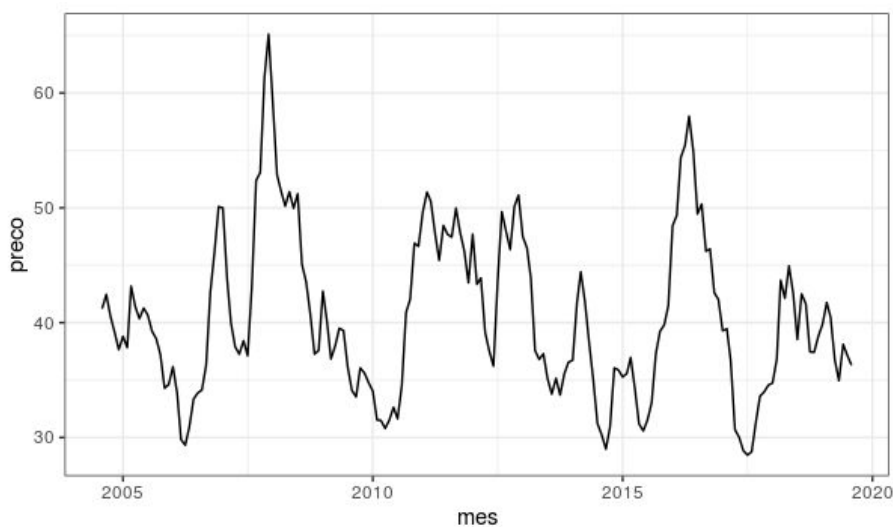


Gráfico 2 - Série histórica deflacionada do preço do milho.

Fonte: Elaboração própria.

Antes de construir o modelo e realizar testes com modelos estatísticos para previsão do preço do milho, houve a separação dos dados da série histórica entre treino e teste. Neste trabalho, são usados como dados de treino os valores de janeiro de 2005 a dezembro de 2017 e como teste, dados de janeiro a dezembro de 2018. No Gráfico 3 estão indicadas as duas séries:

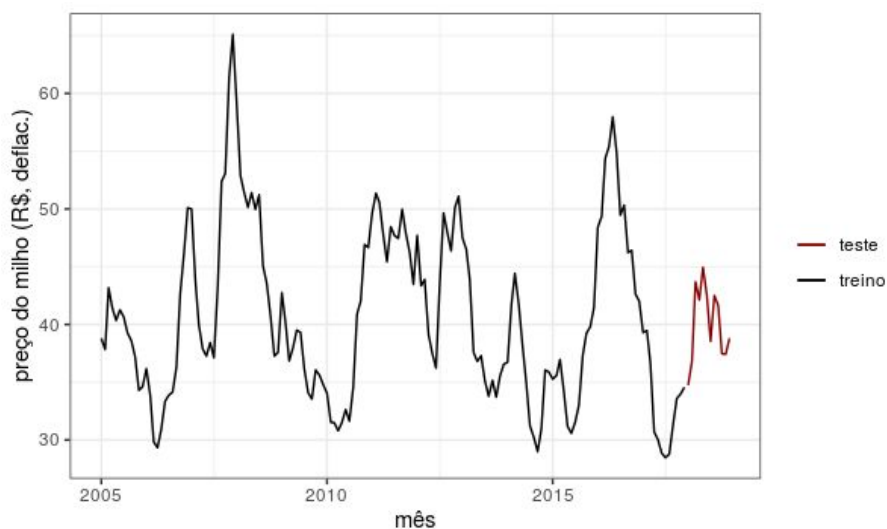


Gráfico 3 - Dados da série separados em treino e teste.

Fonte: Elaboração própria.

Como já mencionado, uma série temporal estacionária é aquela cujas propriedades estatísticas (média e variância) não dependem do momento em que ela série foi observada. O teste de raiz unitária de Kwiatkowski–Phillips–Schmidt–Shin (KPSS) indicou que a série de treino é estacionária, ou seja, não é necessário realizar nenhuma diferenciação.

A série treino estudada, portanto, é a série histórica original, de janeiro de 2005 a dezembro de 2017, deflacionada tendo como referência o mês julho de 2019, conforme mostrada no Gráfico 4:

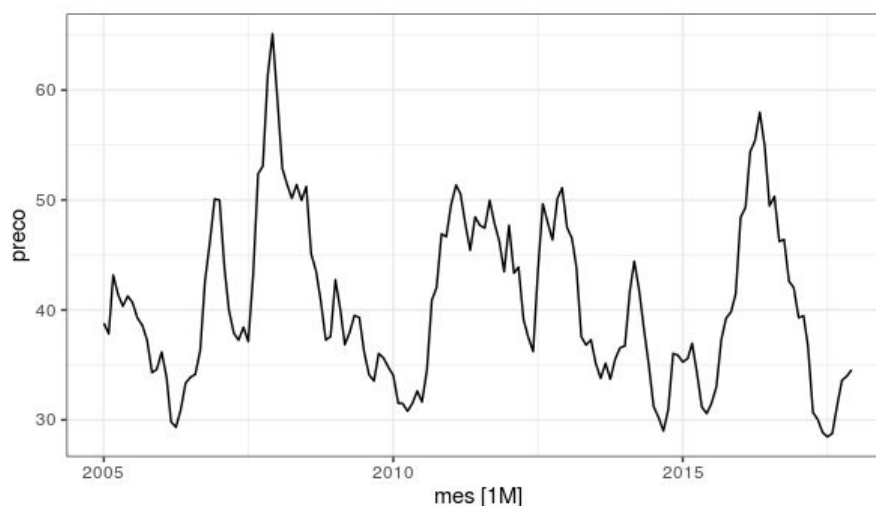


Gráfico 4 - Série treino a ser estudada.

Fonte: Elaboração própria.

Na próxima seção, serão mostrados os resultados obtidos a partir dos materiais e métodos apresentados.

#### 4. Resultados e discussão

Para melhor entendimento do preço do milho, são apresentadas na Tabela 1 algumas medidas descritivas da série estudada. A média de preços, R\$40,60, quando comparada ao valor mínimo que essa *commodity* já atingiu, R\$28,50, pode ser considerada um valor satisfatório, mas não o desejado, visto que produtores visam sempre maiores lucros, ou seja, almejam vender a preços sempre próximos do valor máximo, 65,1.

Tabela 1 - Medidas descritivas do preço do milho

Medida Descritiva	Resultado
Média	R\$40,60
Mediana	R\$39,20
Desvio Padrão	R\$7,10
Valor máximo	R\$65,10
Valor mínimo	R\$28,50

Fonte: Elaboração própria.

Ao estudar séries temporais, os dados podem apresentar uma diversidade de padrões. Devido a esse fato, geralmente é de grande utilidade dividir a uma série em vários componentes, cada um representando uma categoria de padrão subjacente.

As séries temporais normalmente são decompostas em três componentes: componente de tendência, componente sazonal e o resíduo. Um método versátil para decompor séries temporais é o *SLT* uma abreviação para “Decomposição sazonal e de tendências usando Loess”, sendo que Loess constitui-se em um método para estimar relações não lineares. A decomposição por *STL* foi desenvolvida em 1990 por Cleveland, Cleveland, McRae e Terpenning (HYNDMAN; ATHANASOPOULOS, 2018). A série após a decomposição pelo modelo de *STL* é apresentada no Gráfico 5:

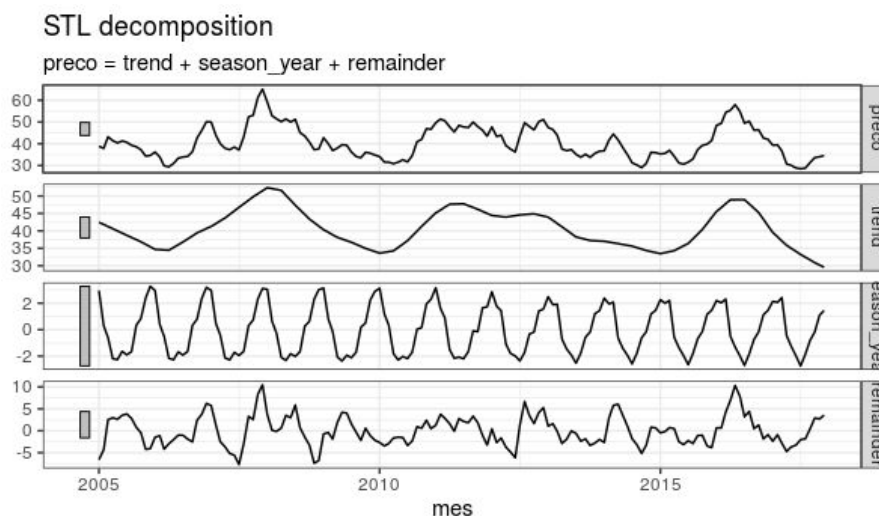


Gráfico 5 - Decomposição, pelo modelo *STL*, da série treino.

Fonte: Elaboração própria.

É possível notar no gráfico acima um padrão sazonal, no entanto, sua magnitude é pequena em relação àquela dos dados.

O Gráfico 6 é conhecido gráfico sazonal, pois subdivide a série em cada um dos subperíodos. Dessa maneira, é possível perceber um padrão próximo de uma parábola invertida, exceto nos anos 2008, 2011 e 2016. Essa queda recorrente do preço em determinadas épocas do ano pode ser explicada pelo período de safra do grão, quando a oferta aumenta.

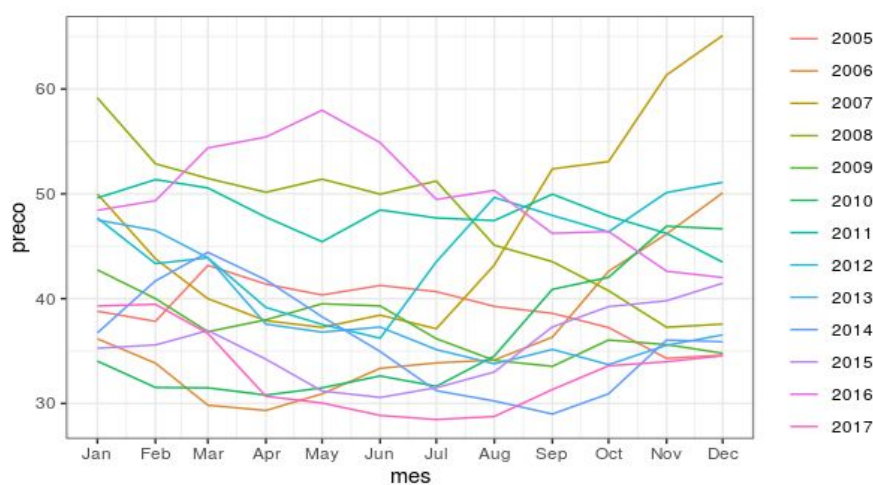


Gráfico 6 - Comparação sazonal entre os anos da série treino dos preços do milho.

Fonte: Elaboração própria.

Outra maneira de se observar padrões de sazonalidade em uma série temporal, é utilizar a plotagem de vários gráficos de séries temporais menores para cada mês separadamente de todos os anos observados. Nota-se no Gráfico 7 que a média confirma o que foi descrito anteriormente sobre o formato de parábola invertida do gráfico sazonal.

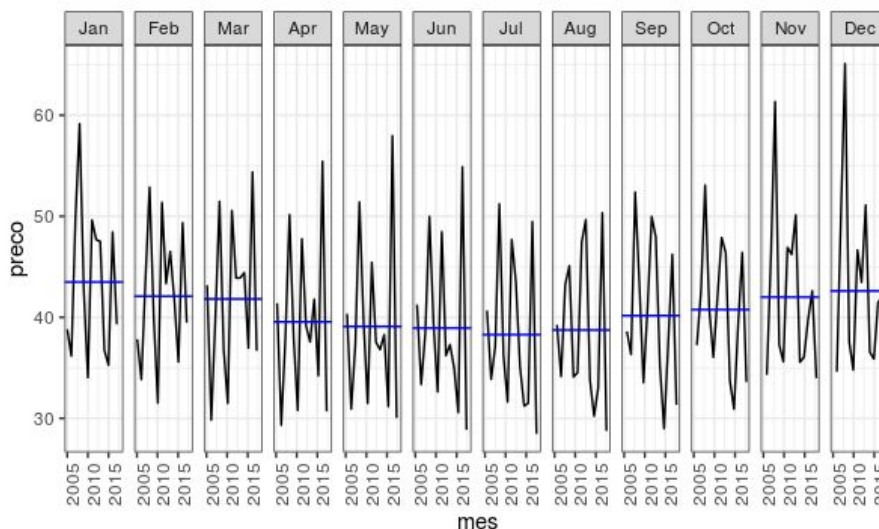


Gráfico 7 - Comparação entre meses ao longo da série com dados treino.

Fonte: Elaboração própria.

Os gráficos de defasagem podem ser utilizados para examinar a relação linear entre os dados de um período  $t$  e dados de  $t-1$ ,  $t-2$ ,  $t-3$ ...  $t-9$ . Por exemplo, com o Gráfico 8, é possível observar que quanto mais distantes no passado os valores estiverem, menor, em média, é essa relação. Em contrapartida, nas observações mais próximas no tempo ao período  $t$ , há uma relação linear mais forte entre os dados.

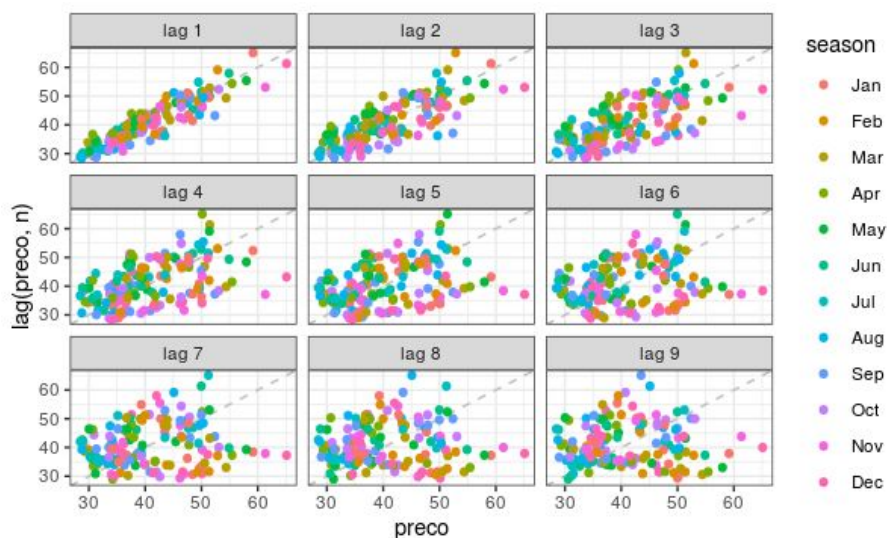


Gráfico 8 - Correlação de defasagens.

Fonte: Elaboração própria.

Assim como nos gráficos de defasagem, a autocorrelação mede a dependência linear entre os valores defasados de uma série temporal. Abaixo, no Gráfico 9, tem-se o gráfico da função de autocorrelação dos dados, em que é perceptível que a série estudada não é uma série de ruído branco, ou seja, mais de 5% dos picos da Função de Autocorrelação (ACF) estão fora de  $\pm 2/\sqrt{T}$ , onde  $T$  é o tamanho da série temporal. Esses limites são representados pelos traços azuis no gráfico.

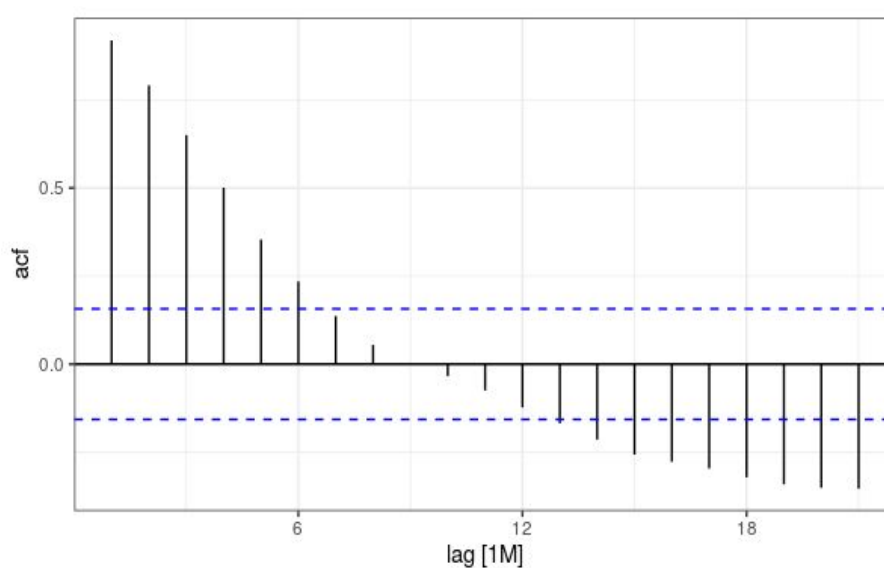


Gráfico 9 - Autocorrelação dos dados de treino.

Fonte: Elaboração própria.

Dentro da autocorrelação existem coeficientes que correspondem a cada painel no gráfico de defasagem. Por exemplo,  $1M$  mede a relação entre dados do período  $t$  e  $t-1$ ,  $2M$  mede a relação entre dados do período  $t$  e  $t-2$  e assim por diante. A Tabela 2 mostra alguns coeficientes de autocorrelação para os dados da série temporal dos preços do milho:

Tabela 2 - Dez primeiros coeficientes de autocorrelação da série treino.

1M	2M	3M	4M	5M	6M	7M	8M	9M	10M
0.92	0.79	0.65	0.50	0.35	0.24	0.14	0.05	0.00	-0.03

Fonte: Elaboração própria.

Na tabela acima, assim como foi observado no Gráfico 8, nota-se que quanto mais próximas no tempo ao período  $t$  as defasagens estiverem, maior o coeficiente de autocorrelação, isto é, uma dependência linear mais forte entre os dados.

No modelo ARIMA, geralmente, é impossível dizer, a partir de um gráfico de tempo quais os valores de  $p$  e  $q$  são mais apropriados para os dados. No entanto, por vezes é possível através de gráficos da Função de Autocorrelação (ACF) e da Função de Autocorrelação Parcial (PACF) determinar esses valores. No Gráfico 10, percebe-se que há uma correlação significativa nas defasagens 1 e 2, seguido por correlações que não são significativas. O padrão apresentado indica um termo autorregressivo de ordem 2.

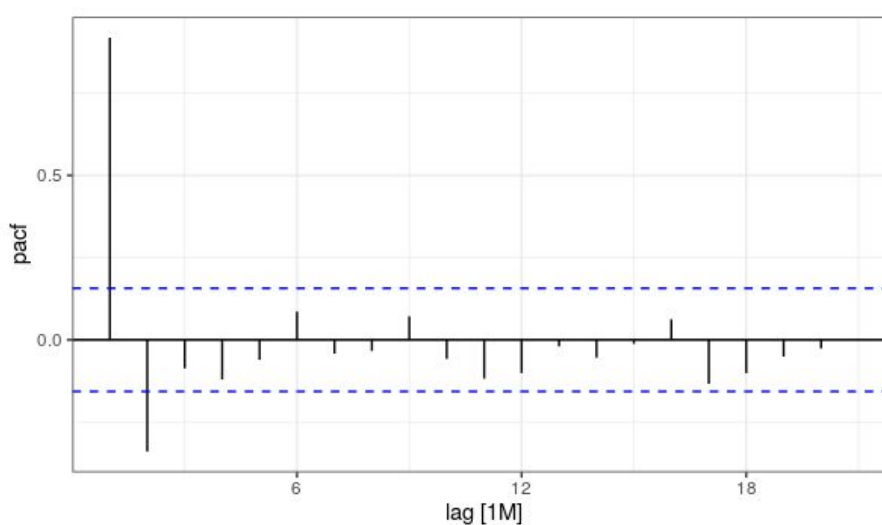


Gráfico 10 - Autocorrelação parcial dos dados de treino.

Fonte: Elaboração própria.



Uma estratégia para identificar o melhor modelo para prever o preço do milho usando a análise de séries temporais é a comparação entre modelos com o intuito de escolher aquele que mais se aproxima da série de teste (o preço mensal do milho de janeiro a dezembro de 2018). Os métodos escolhidos foram: média histórica, *naive*, suavização exponencial (*Holt-Winters*) e ARIMA.

Segundo a análise feita, nota-se que o modelo que mais aproximou das observações de teste foi o modelo ARIMA, que apresentou um *RMSE* (erro quadrático médio) menor na previsão nos dados de treino e o segundo menor *RMSE* na previsão nos dados de teste. A Tabela 3 apresenta a precisão das previsões para dados de treino e teste. Além disso, os Gráficos 11 e 12 apresentam a previsão dos modelos graficamente.

Tabela 3 - Precisão dos modelos

<b>Modelo</b>	<b>RMSE (treino)</b>	<b>RMSE (teste)</b>
Naive	3.00	6.33
Mean	7.54	3.10
Holt	2.91	12.2
ARIMA	2.75	3.41

Fonte: Elaboração própria.

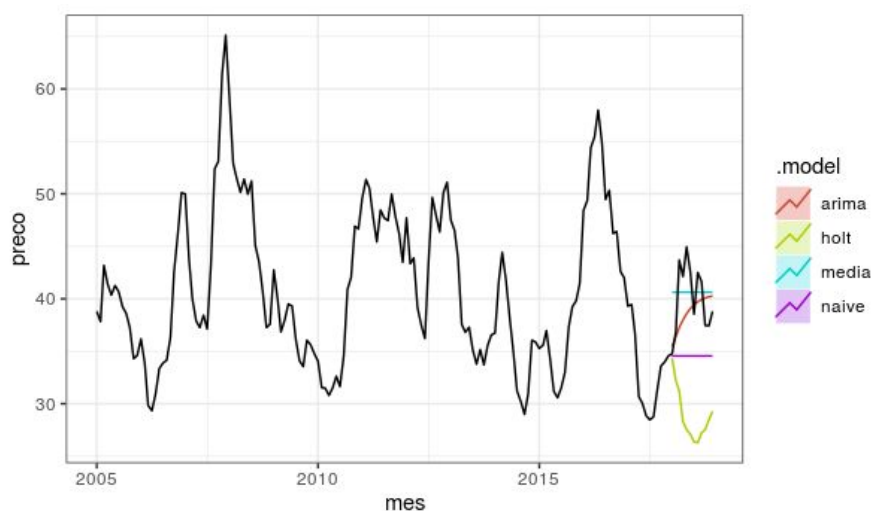


Gráfico 11 - Comparação dos modelos de previsão aos dados de teste.

Fonte: Elaboração própria.

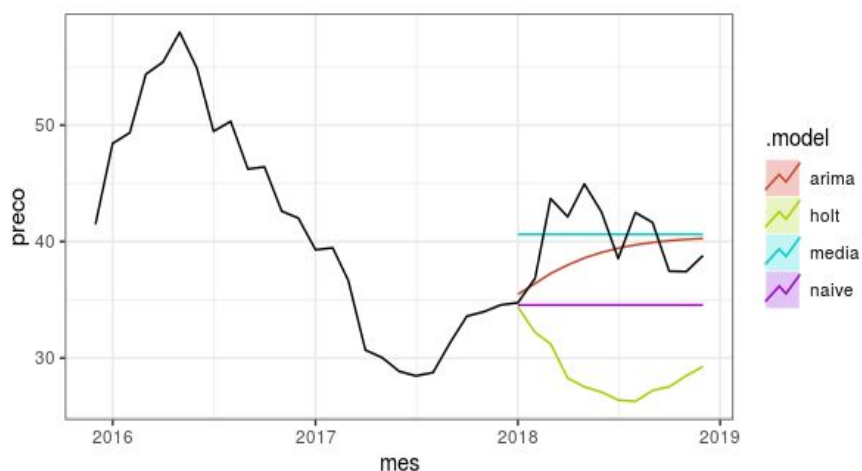


Gráfico 12 - Comparação dos modelos de previsão aos dados de teste com recorte de observações.

Fonte: Elaboração própria.

Para escolher os melhores parâmetros  $p$ ,  $d$  e  $q$  para o modelo ARIMA, deve-se minimizar o AICc (Critério de Informação de Akaike com correção para pequenas amostras). Isso foi feito de maneira automática utilizando a função ARIMA do pacote *fable* (FABLE, 2019), o qual seleciona o modelo com menor AICc. O modelo indicado pela função foi o ARIMA(2, 0, 1), confirmando o que havia sido sugerido pelo gráfico 13.

Ao analisar os resíduos da série histórica dos preços do milho, pode-se considerar que não há autocorrelação significativa, pois no correlograma abaixo há apenas uma defasagem que excede o limite de significância (o que é aceitável quando é utilizado o alfa de 0,05). Ou seja, apresenta ruído branco para os resíduos, o que sugere que as previsões são adequadas.

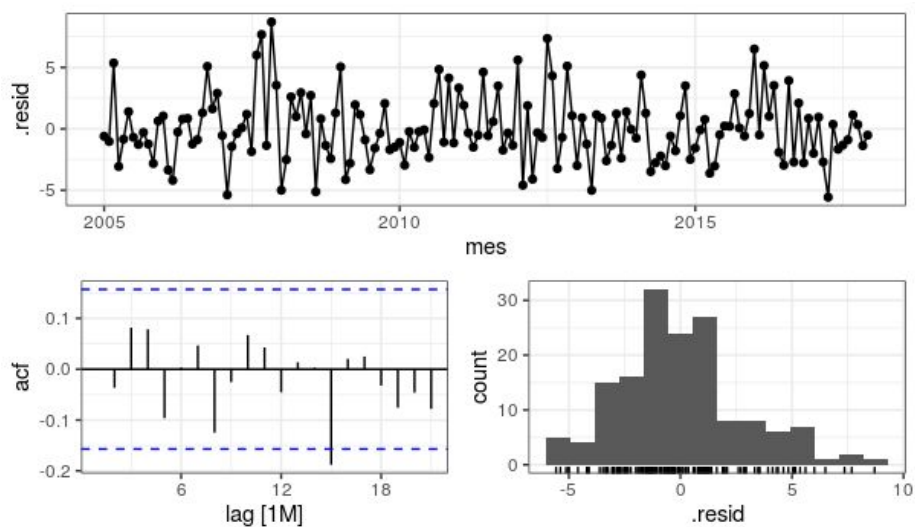


Gráfico 13 - Diagnóstico residual da série.

Fonte: Elaboração própria

Após escolher os parâmetros  $p$ ,  $d$  e  $q$ , tem-se as plotagens dos gráficos da previsão dos preços do milho pelo modelo ARIMA (2, 0, 1) (a linha azul) e a sua comparação com os dados de teste (a linha preta), com níveis de confiança de 80% e 95%. O Gráfico 14 apresenta toda a série temporal utilizada, enquanto o Gráfico 15 se restringe às observações mais recentes, para facilitar a visualização:

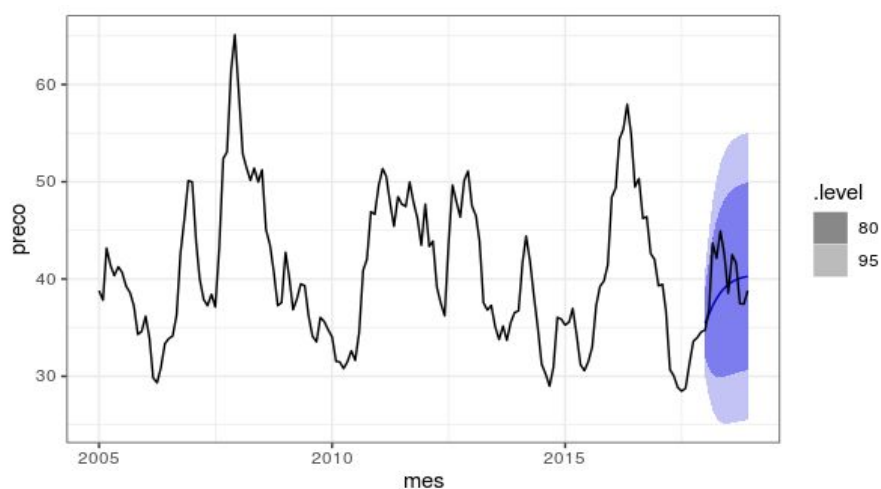


Gráfico 14 - Previsão dos preços do milho pelo modelo ARIMA (2, 0, 1).

Fonte: Elaboração própria.

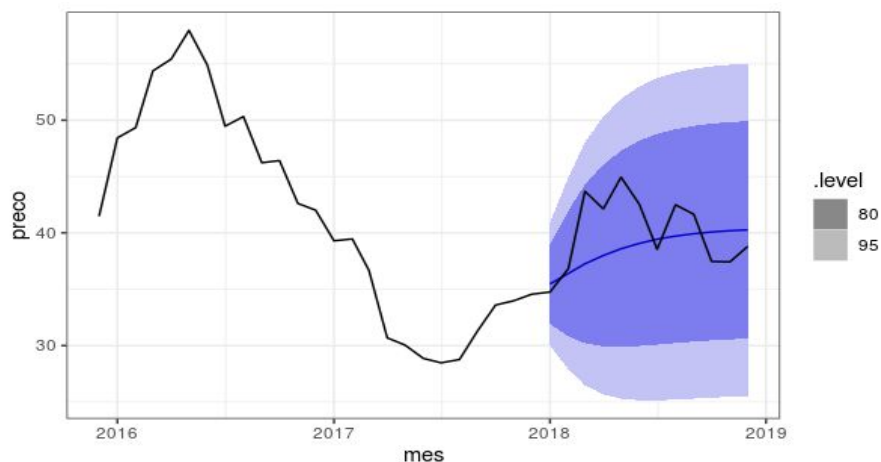


Gráfico 15 - Previsão dos preços do milho pelo modelo ARIMA (2, 0, 1) com recorte de observações.

Fonte: Elaboração própria.

Como se pode observar, ainda que as previsões não coincidam com os valores observados, elas foram capazes de captar os movimentos principais: uma elevação seguida de uma estabilização. Além disso, mesmo os movimentos mais sutis que escaparam ao modelo se mantiveram dentro do limite de confiança de 80%.

## 5. Considerações finais

Este trabalho apresentou um breve panorama sobre a produção e o mercado do milho, dando ênfase à variabilidade de seus preços. Além disso, foi mostrada a dificuldade de prever preços do grão, visto que existem variáveis que afetam diretamente seu valor de negociação, são elas: sazonalidade, estoques, oferta, demanda e custos diversos.

Devido à variabilidade de preços do milho e como forma de tentar proteger agricultores em contratos de fixação de preço futuro, foi feita uma comparação de modelos de previsão de preços para escolher aquele que mais se adequa à série histórica dos preços do milho utilizada como base deste estudo. Ao comparar quatro modelos (média histórica, naive, suavização exponencial e ARIMA), concluiu-se que o melhor modelo a ser utilizado é o modelo Autorregressivo Integrado de Médias Móveis (ARIMA), pois este apresentou o menor erro quadrático médio da diferença entre o valor previsto e o valor real. Vale lembrar que ao minimizar automaticamente o critério de informação de Akaike com correção para pequenas

amostras (AICc), o modelo escolhido foi o ARIMA (2, 0, 1). A partir de testes com os resíduos da série, como o gráfico ACF, notou-se ruído branco, o que indica que as previsões são consideradas adequadas.

Em suma, concluiu-se que o melhor modelo para previsão de preços da *commodity* milho, considerando como base a série histórica mensal dos preços do milho de janeiro de 2004 a dezembro de 2018, é o modelo ARIMA (2, 0, 1).

Entretanto, este trabalho possui diversas limitações. Por exemplo, é provável que as previsões fossem ainda melhores se houvesse uma investigação mais detalhada da sazonalidade da série e uma transformação logarítmica do preço, além da inclusão de novas variáveis aos modelos, tais como: temperatura, pluviosidade, produção em outros países, preço do dólar, preço da soja (por ser uma alternativa ao milho em diversas situações) etc.

## Referências

ASSOCIAÇÃO BRASILEIRA DAS INDÚSTRIAS DO MILHO (ABIMILHO). Estatísticas: Oferta e Demanda do Milho - Brasil. **Abimilho**, 2019. Disponível em: <<http://www.abimilho.com.br/estatisticas>>. Acesso em: 22 jun. 2019.

BRASIL. **Ministério da Indústria, Comércio Exterior e Serviços (MDIC)**, 2019.

Disponível em:

<<http://www.mdic.gov.br/comercio-exterior/estatisticas-de-comercio-exterior/comex-vis/frame-ppe?ppe=1240>>. Acesso em: 25 jul. 2019.

CALDARELLI, C. ; BACCHI, M. Fatores de influência no preço do milho no Brasil. **Nova Economia**, v. 22, n. 1, p. 141-164, 2012.

CALEGARI, I.; BAIGORRI, M.; FREIRE, F. Os derivativos agrícolas como uma ferramenta de gestão do risco de preço. **Custos e Agronegócio**, v. 8, p. 1-154, 2012. Número Especial.

CENTRO DE ESTUDOS AVANÇADOS EM ECONOMIA APLICADA (CEPEA).

Consultas ao banco de dados do site. **Cepea-Esalq/USP**, 2019. Disponível em:

<<https://www.cepea.esalq.usp.br/br/consultas-ao-banco-de-dados-do-site.aspx>> Acesso em: 28 agos. 2019.

COMPANHIA NACIONAL DE ABASTECIMENTO (CONAB). Séries históricas das safras.

**Conab.** Disponível em:

<https://www.conab.gov.br/info-agro/safras/serie-historica-das-safras?s tart=20> Acesso em: 29 out. 2018

CORSINI, F. ; RIBEIRO C. Dinâmica e previsão de preços de commodities agrícolas com o filtro de Kalman. In: XXVIII ENCONTRO NACIONAL DE ENGENHARIA DE PRODUÇÃO. **Anais ...** Rio de Janeiro, 2008. Disponível em:

<[http://www.abepro.org.br/biblioteca/enegep2008\\_tn\\_sto\\_074\\_526\\_10975.pdf](http://www.abepro.org.br/biblioteca/enegep2008_tn_sto_074_526_10975.pdf)>. Acesso em: 22 jun. 2019.

CRUZ, J.; PEREIRA FILHO, I.; DUARTE, A. Milho Safrinha. **Agência Embrapa de informação tecnológica**, 2019s.d.. Disponível em: <<http://www.agencia.cnptia.embrapa.br/gestor/milho/arvore/CONT000fya0krse02wx5ok0pvo4k3mp7ztkf.html>>.

Acesso em: 03 set. 2018.

EMPRESA BRASILEIRA DE PESQUISA AGROPECUÁRIA (EMBRAPA). Cultivo do Milho: Apresentação. **Sistema de Produção Embrapa**, 2012. Disponível em:

<[https://www.spo.cnptia.embrapa.br/conteudo?p\\_p\\_id=conteudoportlet\\_WAR\\_sistemasdeproducaolf6\\_1galceportlet&p\\_p\\_lifecycle=0&p\\_p\\_state=normal&p\\_p\\_mode=view&p\\_p\\_col\\_id=column-3&p\\_p\\_col\\_count=1&p\\_r\\_p\\_-76293187\\_sistemaProducaoId=3821&p\\_r\\_p\\_-996514994\\_topicoId=3715](https://www.spo.cnptia.embrapa.br/conteudo?p_p_id=conteudoportlet_WAR_sistemasdeproducaolf6_1galceportlet&p_p_lifecycle=0&p_p_state=normal&p_p_mode=view&p_p_col_id=column-3&p_p_col_count=1&p_r_p_-76293187_sistemaProducaoId=3821&p_r_p_-996514994_topicoId=3715)>. Acesso em: 29 out. 2018.

EUZÉBIO, P. Época de plantio do milho. **Folha Agrícola**, 2016. Disponível em:

<<http://folhaagricola.com.br/artigo/epoca-de-plantio-do-milho>>. Acesso em: 20 dez. 2018.

FABLE. **Fable**: tidy time series forecasting. Disponível em: [fable.tidyverts.org](http://fable.tidyverts.org) Acesso em: 02 dez. 2019.

FRANÇA JUNIOR, F. Quais as diferenças entre mercado físico, a termo, de futuro e de opções? Por Flávio França. **Notícias Agrícolas**, 2015. Disponível em: <<https://www.noticiasagricolas.com.br/artigos/artigos-principais/152271-quais-as-diferencas-entre-mercado-fisico-a-termo-de-futuro-e-de-opcoes-por-flavio-franca.html#.XUMhOehKjDd>>. Acesso em: 25 jul. 2019.

GUTH, T. Análise Mensal: Milho. **Conab**, 2018. Disponível em: <[https://www.conab.gov.br/info-agro/analises-do-mercado-agropecuario-e-extrativista/analise-s-do-mercado/historico-mensal-de-milho/item/download/15900\\_916dc0d0ebdd60644cdaef1edf68dde8](https://www.conab.gov.br/info-agro/analises-do-mercado-agropecuario-e-extrativista/analise-s-do-mercado/historico-mensal-de-milho/item/download/15900_916dc0d0ebdd60644cdaef1edf68dde8)>. Acesso em: 20 nov. 2018.

HYNDMAN, R.; ATHANASOPOULOS, G. **Forecasting: Principles and Practice**. Austrália, Otexts, 2018. Disponível em: <<https://otexts.com/fpp3/>>. Acesso em: 25 ago. 2019.

INSTITUTO MATO-GROSSENSE DE ECONOMIA AGROPECUÁRIA (IMEA). Entendendo o Mercado de Milho. **IMEA**, Brasil, 01 jul. 2015. Disponível em: <[http://www.imea.com.br/upload/pdf/arquivos/Paper\\_jornalistas\\_Milho\\_AO.pdf](http://www.imea.com.br/upload/pdf/arquivos/Paper_jornalistas_Milho_AO.pdf)>. Acesso em: 25 out. 2018.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE). Em maio, IBGE prevê safra 5,2% menor que a de 2017. **Agência de Notícias**, 2018. Disponível em: <<https://agenciadenoticias.ibge.gov.br/agencia-sala-de-imprensa/2013-agencia-de-noticias/releases/21412-em-maio-ibge-preve-safra-5-2-menor-que-2017.html>>. Acesso em: 03 set. 2018.

JHA, M. Brasil se beneficia de safra ruim de milho na Europa: Safras ruins de milho na Europa estão sugando para o continente montanhas de grão importado do Brasil. **Portal Uol-Economia**, 2017. Disponível em: <<https://economia.uol.com.br/noticias/bloomberg/2017/11/07/brasil-se-beneficia-de-safra-ruim-de-milho-na-europa.htm>>. Acesso em: 19 out. 2018.

MATIDA, K. A geada que mudou a economia agrícola do Paraná. **Folha de Londrina**. 19 jun. 2005. Disponível em: <<https://www.folhadelondrina.com.br/cadernos-especiais/a-geada-que-mudou-a-economia-agricola-do-parana-532777.html>>. Acesso em: 10 out. 2018.

MATTOSO, M. et al. Aspectos da produção e mercado de milho. **Informe Agropecuário**, Belo Horizonte, v. 27, n. 233, p.95-104, jul./ago. 2006. Disponível em: <<https://www.alice.cnptia.embrapa.br/alice/handle/doc/490281>>. Acesso em: 5 set. 2018.

NETO, A. A Cultura do Milho: análise dos custos de produção e da rentabilidade nos anos-safra 2007 a 2017. techreport Compêndio de Estudos Conab 14. **Diretoria de Política Agrícola e Informações, Superintendência de Informações do Agronegócio, Companhia Nacional de Abastecimento**, 2018. Disponível em: [https://www.conab.gov.br/institucional/publicacoes/compendio-de-estudos-da-conab/item/download/15977\\_22abe6563f37c285e7116eb716a02804](https://www.conab.gov.br/institucional/publicacoes/compendio-de-estudos-da-conab/item/download/15977_22abe6563f37c285e7116eb716a02804) Acesso em: 26 agos. 2019

PAES, M. Aspectos físicos, químicos e tecnológicos do grão de milho. **Circular Técnica**, Sete Lagoas, 2006.

PEROBELLI, F. Contratos de opções de milho. **AgroAnalysis**, v. 26, n. 6, p. 18-19, 2006.

PINO, F. Sazonalidade na Agricultura. **Revista de Economia Agrícola**, v. 61, n. 1, p.63-93, jan./jun. 2014.

QUEIROZ, A. Brasil já é o segundo maior exportador mundial de milho. **Portal Uol-Economia**, 2019. Disponível em: <<https://economia.uol.com.br/noticias/redacao/2019/05/10/brasil-ja-e-o-segundo-maior-exportador-mundial-de-milho.htm>>. Acesso em: 27 jul. 2019.

R CORE TEAM. **R**: A language and environment for statistical computing. 2019. Disponível em: <http://www.R-project.org/> Acesso em: 02 dez 2019.



SOUZA, A. et al. Estudo da produção do milho no Brasil. **South American Development Society Journal**, v. 4, n. 11, p.182-194, 2018. <http://dx.doi.org/10.24325/issn.2446-5763.v4i11p182-194>.

TIBULO, C.; CARLI, V. Previsão do preço do milho, através de séries temporais. **Scientia Plena**, v. 10, n. 10, 2014.

WIKIPÉDIA. Akaike Information Criterion. **Wikipédia, a enciclopédia livre**. 2019. Disponível em: <[https://en.wikipedia.org/wiki/Akaike\\_information\\_criterion#AICc](https://en.wikipedia.org/wiki/Akaike_information_criterion#AICc)>. Acesso em: 25 set. 2019.

WIKIPEDIA. Commodity. **Wikipédia, a enciclopédia livre**, 2019. Disponível em: <[pt.wikipedia.org/wiki/Commodity](http://pt.wikipedia.org/wiki/Commodity)>. Acesso em: 30 mai. 2019.

## Anexo

```

# remotes::install_github('robjhyndman/fpp3-package')
# remotes::install_github('meirelesff/deflateBR')
# install.packages(c('remotes', 'tidyverse', 'tsibble', 'feasts',
# 'RcppArmadillo', 'urca'))
# install.packages("fable", repos = "https://tidyverts.org/")
library(tidyverse)
library(readxl)
library(tsibble)
library(lubridate)
library(deflateBR)
library(fable)
library(feasts)
library(fabletools)
theme_set(theme_bw())

# preparação dos dados
df = read_excel('CEPEA_20190826101823.xls', skip=4, col_names=c('dia',
'reais', 'dolar'))
df$dia = dmy(df$dia)
df = df %>% select(dia, reais)

# transforma dados diários em mensais
df = df %>%
  group_by(mes=floor_date(dia, 'month')) %>%
  summarize(preco=mean(reais))

# deflaciona os dados
referencia = '07/2019'
df = df %>% mutate(precod = deflate(preco, mes, referencia))
df %>% ggplot(aes(mes, preco)) + geom_line()
df = df %>% select(mes, preco=precod)
df = as_tsibble(df, index=mes)
df$mes = yearmonth(df$mes)

# separação dos dados entre treino e teste
treino = df %>% filter(year(mes) > 2004 & year(mes) < 2018)
teste = df %>% filter(year(mes) == 2018)
df2 = df %>% filter(year(mes) > 2004 & year(mes) < 2019)
df3 = slice(df %>% filter(year(mes) < 2019), (n()-36):n())

ggplot() + geom_line(data=treino, aes(mes, preco, color='treino')) +
  geom_line(data=teste, aes(mes, preco, color='teste')) +
  scale_color_manual(values=c('darkred', 'black')) + labs(color='') +
  xlab('mês') + ylab('preço do milho (R$, deflac.)')

# resumo estatístico dos dados de treino e teste
treino$preco %>% summary()
treino$preco %>% sd()

```

```

teste$preco %>% summary()
teste$preco %>% sd()

df2 %>% arrange(-preco)
df2 %>% autoplot(preco)

# testes de estacionariedade
## h0: estacionário. Então, se valor-p < 0.05, é preciso diferenciar
treino %>% features(preco, unitroot_kpss)
## automatizando: quantas diferenciações são necessárias?
treino %>% features(preco, unitroot_ndiffs)

# a série a ser estudada
treino %>% autoplot()
# decomposição STL: tendência, sazonalidade e resíduos
treino %>% STL(preco, 'additive') %>% autoplot()
# comparação entre os anos
treino %>% gg_season(preco, 12)
treino$ano = treino$mes %>% year()
treino %>% index_by(ano) %>% summarise(media = mean(preco)) %>%
  arrange(-media)

# comparação entre os meses ao longo da série
treino %>% gg_subseries(preco, 12)
# correlação entre as defasagens
treino %>% gg_lag(preco, geom='point')
# autocorrelação
treino %>% ACF(preco)
treino %>% ACF(preco) %>% autoplot()
# autocorrelação parcial
treino %>% PACF(preco)
treino %>% PACF(preco) %>% autoplot()

# construção dos modelos
modelos = treino %>% model(
  naive = NAIVE(preco),
  media = MEAN(preco),
  holt = ETS(preco ~ error('A') + trend('A') + season('A')),
  arima = ARIMA(preco)
)

# previsões
previsoes = modelos %>% forecast(h=12)
previsoes %>% autoplot(df2, level=NULL)
previsoes %>% autoplot(df3, level=NULL)
augment(modelos)
tidy(modelos)
# precisão das previsões nos dados de treino
accuracy(modelos)
# precisão das previsões nos dados de teste
accuracy(previsoes, teste)

```