

# Análise de Séries Temporais

Trabalho Individual

## T08

Nícolas Oliveira

# Sumário

<b>Sumário</b>	<b>1</b>
Contexto e Introdução	2
Base de Dados	2
<b>Estacionariedade da Série Temporal</b>	<b>4</b>
<b>Modelando a Série Temporal</b>	<b>5</b>
Modelo Holt-Winters	5
Modelo SARIMA	6
SARIMA #1 - ARIMA(1,1,1)(1,1,1)[12]	6
SARIMA #2 - ARIMA(1,0,1)(1,1,1)[12]	7
SARIMA autoarima - ARIMA(1,0,0)(2,1,0)[12]	8
<b>Heterocedasticidade</b>	<b>9</b>
<b>Os Resíduos</b>	<b>9</b>
<b>Previsão para os próximos 2 anos (2018 e 2019)</b>	<b>10</b>

## Contexto e Introdução

O agronegócio desempenha um importante papel na geração de riqueza no Brasil. Nas últimas décadas, o setor conviveu com inovações em produção. A Empresa Brasileira de Pesquisa Agropecuária – EMBRAPA, tem um papel relevante neste processo. O agronegócio efetua elevados investimentos que alimenta toda uma gama de cadeias econômicas e contribui a maior eficiência e eficácia na produção de alimentos. Neste quesito, o consumo de fertilizantes ocupa uma parcela significativa dos investimentos realizados.

Este documento tem por objetivo apresentar resultados da exploração de técnicas de séries temporais a fim de fornecer uma análise preditiva sobre o consumo de fertilizantes.

## Base de Dados

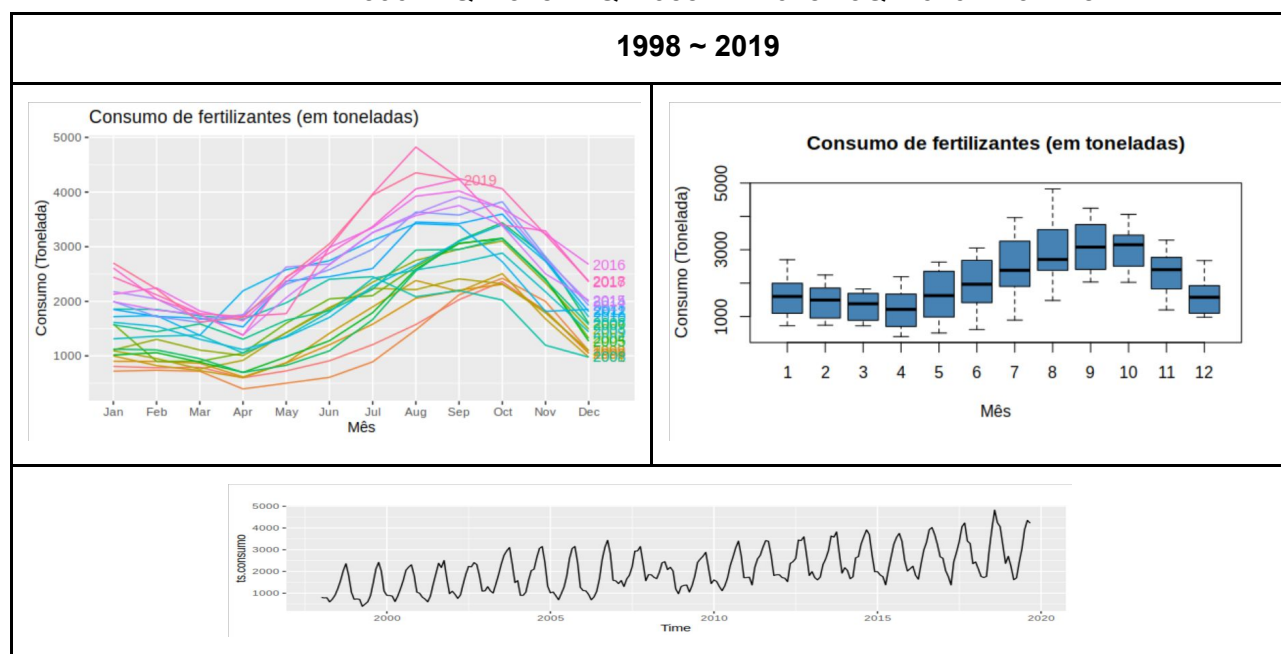
Para esta análise foram utilizados os dados do arquivo “Fertilizantes.xlsx”, que contempla a entrega de fertilizantes ao mercado em mil toneladas no período mensal de janeiro de 1998 a fevereiro de 2019. A fonte dos dados é o sítio da Associação Nacional para Difusão de Adubos – ANDA (<http://anda.org.br/estatisticas/>).

## Análise Exploratória dos Dados

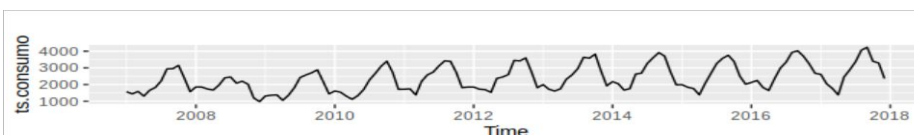
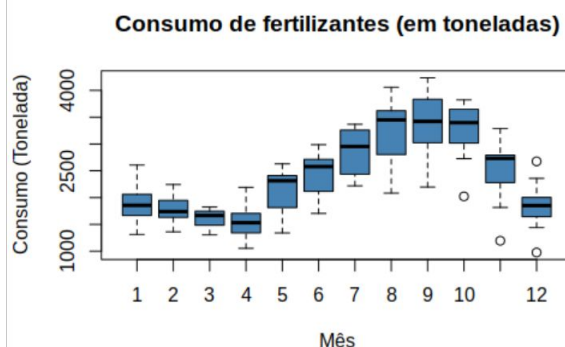
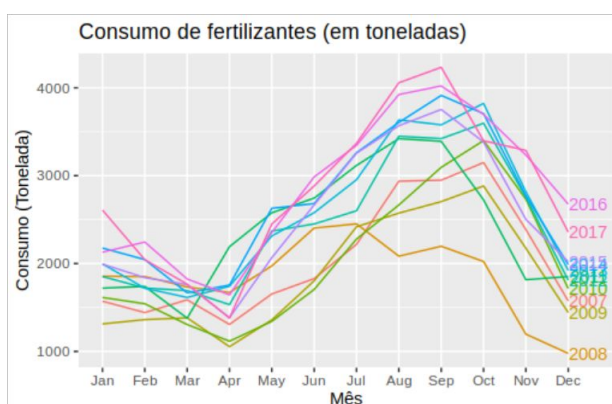
A análise exploratória de dados é uma abordagem à análise de conjuntos de dados de modo a resumir as características principais do conjunto analisado, frequentemente com métodos visuais.

Nesse contexto iniciamos identificando os tipos dos dados a serem trabalhados.

- **Date** (Qualitativa Ordinal) (Nenhum valor NA)
  - Min:1998-01 | Max: 2019-09
- **Consumo** (Variável Quantitativa Discreta) (Nenhum valor NA)
  - Min: 396 / 1Q: 1378 / 2Q: 1988 / M:2078 / 3Q: 2676 / Max: 4824



## 2007 ~ 2017

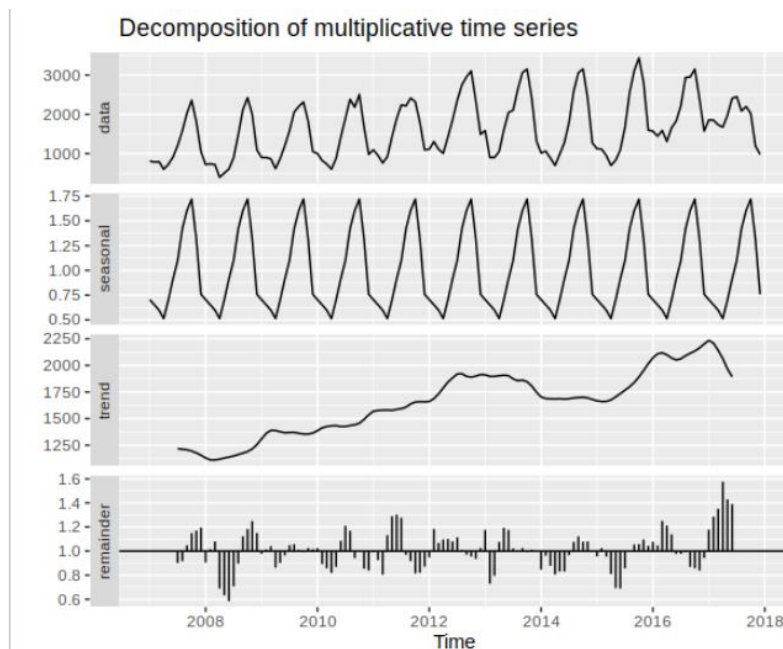


Pode-se observar um padrão sazonal de consumo ao longo do ano, onde é possível perceber um aumento do consumo em torno do 4º trimestre.

Além disso, ao analisar a janela 2007~2017 existem alguns outliers nos últimos trimestres, o que pode indicar algum evento pontual. Tal fato pode ter relação com os anos de 2008 e 2015, que foram atípicos quando analisados apenas visualmente.

O padrão dos dados observados não é tão óbvio quando tentamos observar a magnitude do padrão sazonal, o que trouxe dificuldade na escolha entre o Multiplicativo e Aditivo, quando olhamos apenas para a janela, contudo ao observar mais amostras concluímos se tratar de uma série visualmente multiplicativa.

Utilizando a janela de jan/2007 à dez/2017, podemos observar as principais componentes da série, assumindo se tratar de uma série multiplicativa.



A **Sazonalidade** explica, de certo modo, um padrão cíclico, que se repete com uma certa periodicidade. Descreve um comportamento que pode ser observado várias vezes dentro da série.

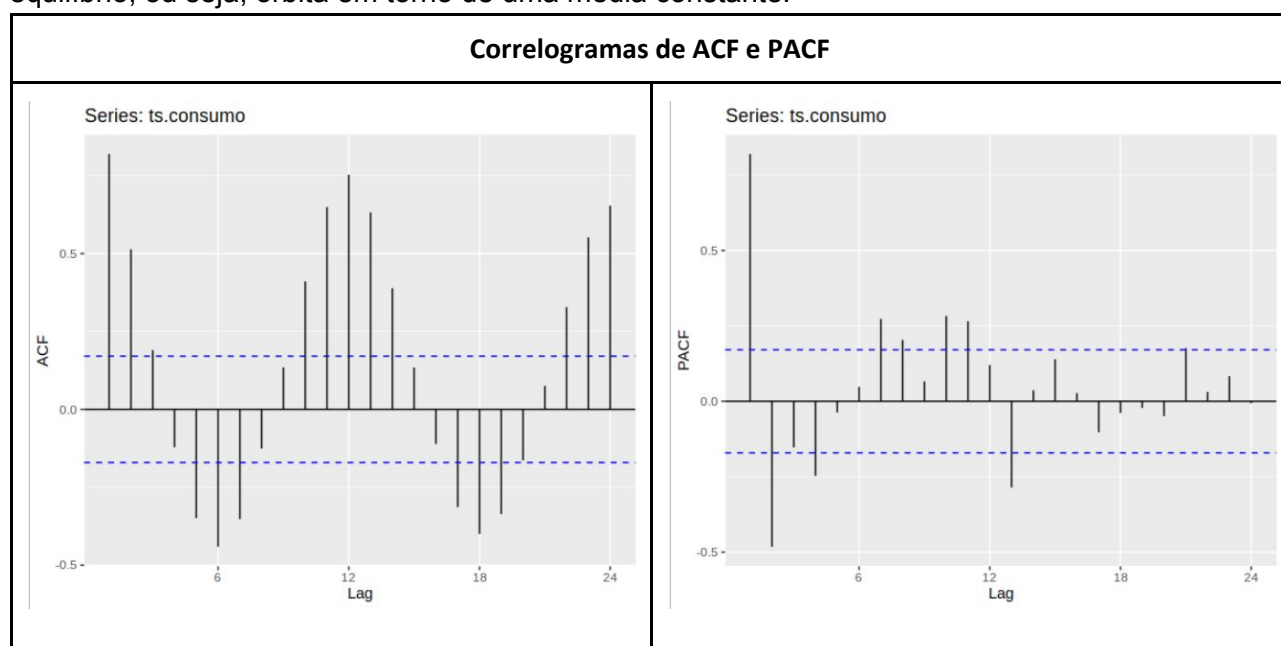
A componente **Tendência** observa se a série segue determinada direção, ou tendência, crescente ou decrescente, e não necessariamente linear.

A **Aleatoriedade** é um componente aleatório que corresponde ao Erro (ou ruído), é a variação aleatória resultante de erros de medição (parte não sistemática)

Após a decomposição é possível perceber claramente se tratar de uma série sazonal, com uma tendência crescente.

## Estacionariedade da Série Temporal

Visualmente podemos observar se tratar de uma série estacionária, pois apresenta um equilíbrio, ou seja, orbita em torno de uma média constante.



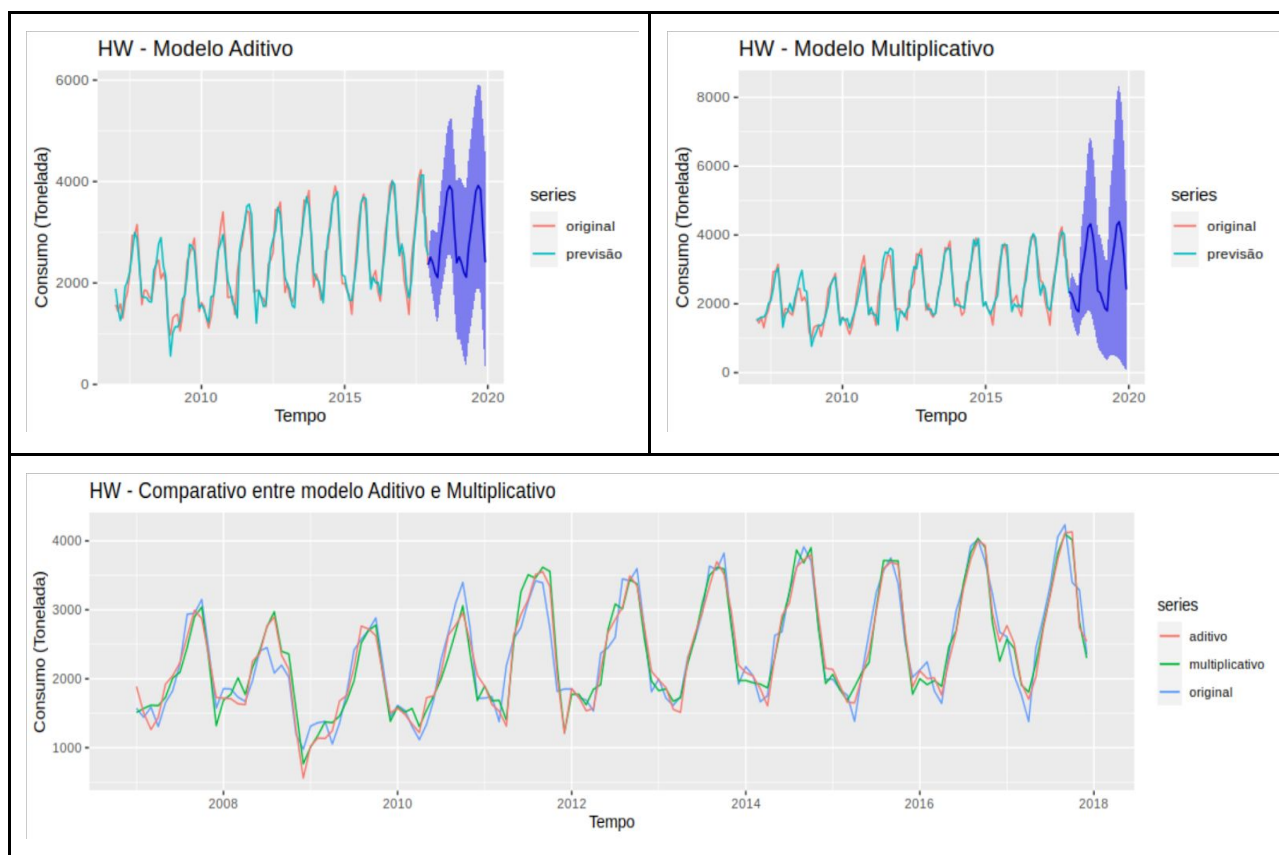
Os testes formais também mostraram evidências para considerar a série como estacionária.

Testes formais para verificar a estacionariedade da série		
<b>Augmented Dickey–Fuller Test</b>	p-value = 0.01 (Ho: Não Estacionária   Ha: Estacionária)	Estacionária
<b>Phillips-Perron Unit Root Test</b>	p-value = 0.01 (Ho: Não Estacionária   Ha: Estacionária)	Estacionária
<b>KPSS Test for Level Stationarity</b>	p-value = 0.01 (Ho: Estacionária   Ha: Não Estacionária)	Não Estacionária

# Modelando a Série Temporal

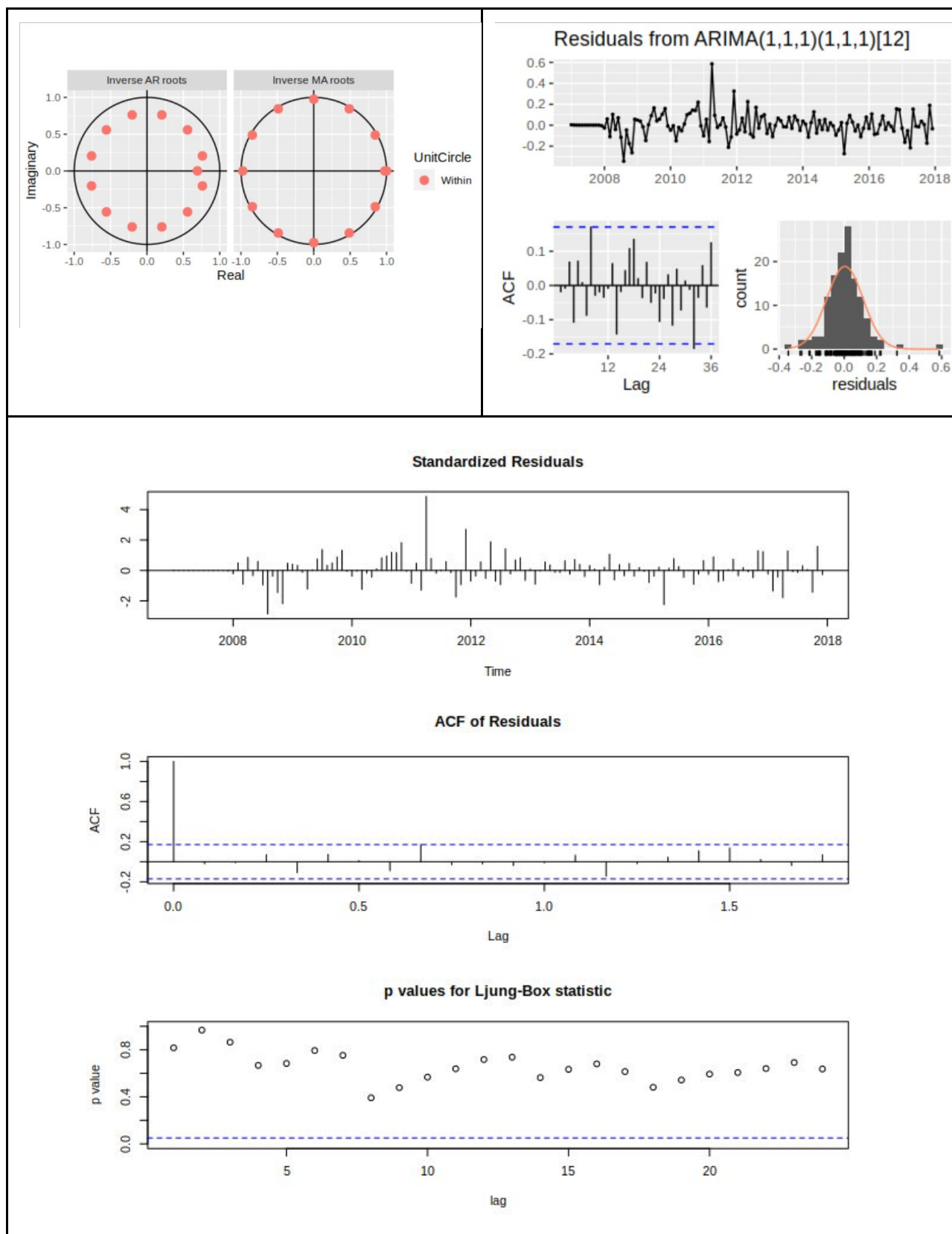
O intervalo de dados considerado foi de janeiro/2007 à dezembro/2017

## Modelo Holt-Winters

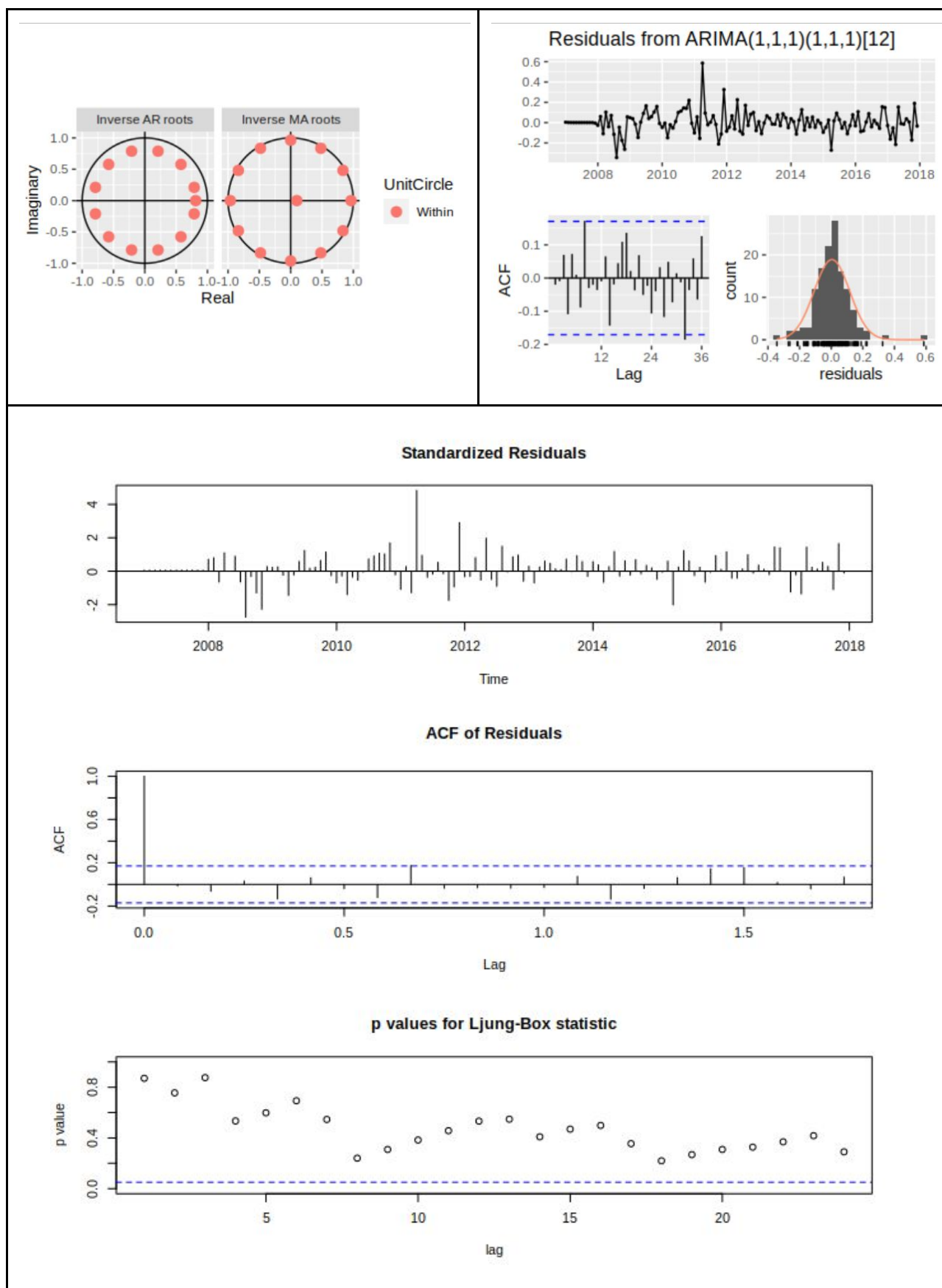


## Modelo SARIMA

SARIMA #1 - ARIMA(1,1,1)(1,1,1)[12]

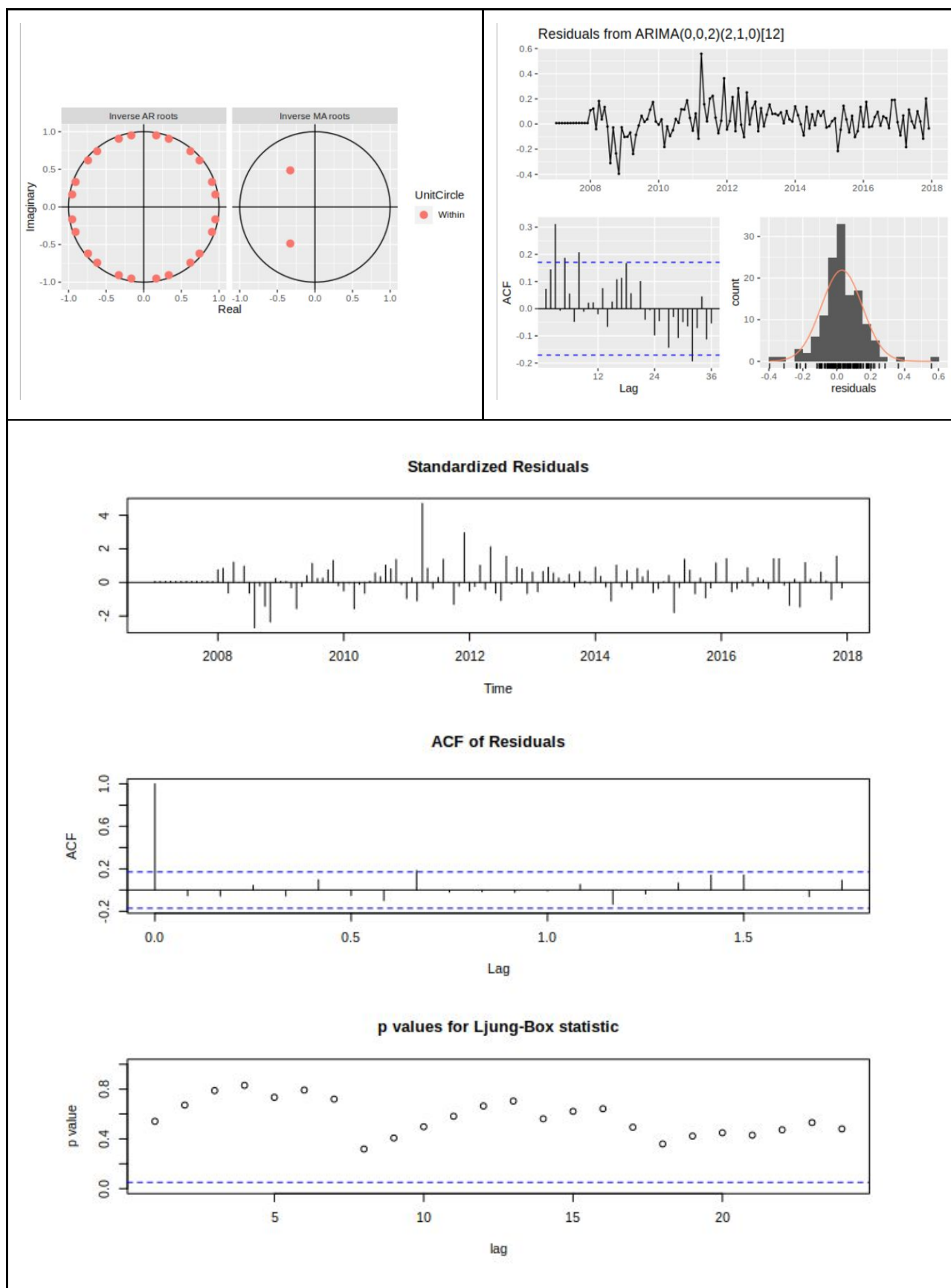


## SARIMA #2 - ARIMA(1,0,1)(1,1,1)[12]





# SARIMA *autoarima* - ARIMA(1,0,0)(2,1,0)[12]



## Heterocedasticidade

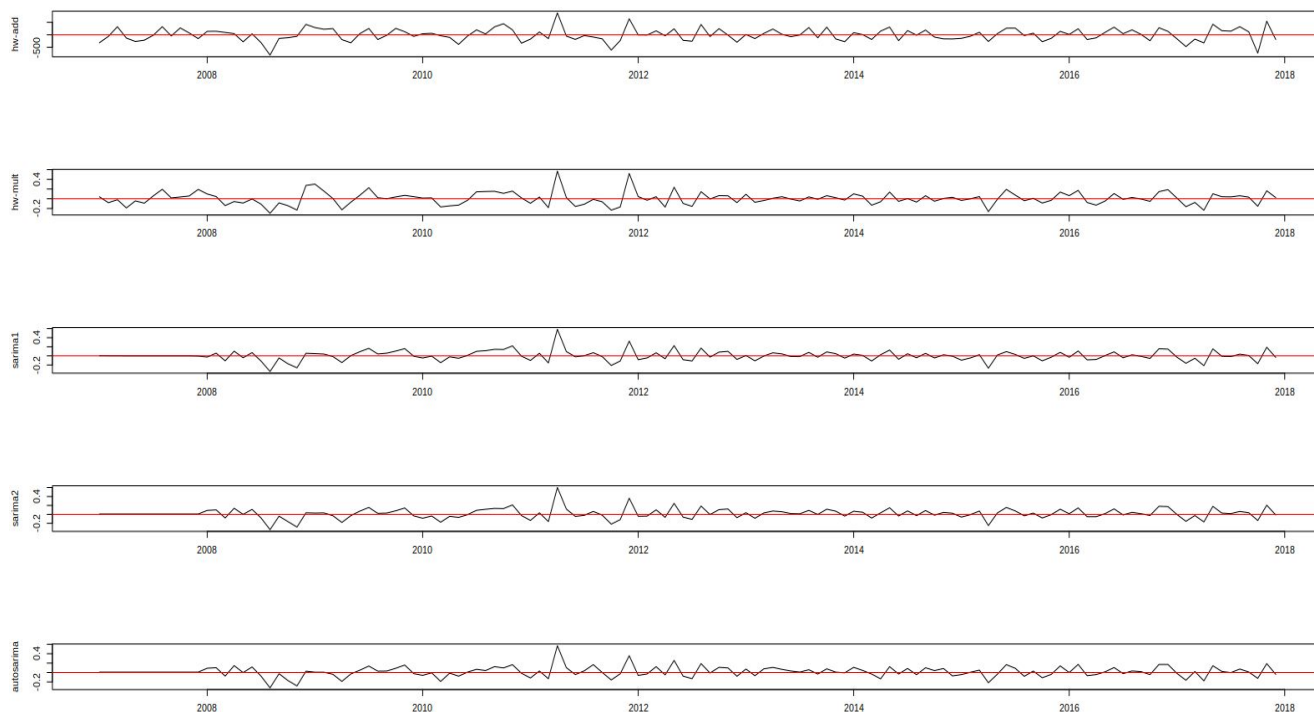
Heterocedasticidade é quando o modelo apresenta variâncias não iguais para todas as observações. Verificamos, por meio dos resíduos, se os dados regredidos encontram-se, ou não, homogeneamente concentrados em torno da regressão. Nenhum dos modelos apresentou evidências suficientes de efeitos de heterocedasticidade.

ARCH LM-test; Null hypothesis: no ARCH effects		
<b>HoltWinter Aditivo</b>	Chi-squared = 22.779, df = 24, p-value = 0.5329	Ho
<b>HoltWinter Multiplicativo</b>	Chi-squared = 33.822, df = 24, p-value = 0.08788	Ho
<b>SARIMA #1</b>	Chi-squared = 11.818, df = 24, p-value = 0.9819	Ho
<b>SARIMA #2</b>	Chi-squared = 14.83, df = 24, p-value = 0.9256	Ho
<b>SARIMA autoarima</b>	Chi-squared = 15.064, df = 24, p-value = 0.9189	Ho

## Os Resíduos

A ausência de heterocedasticidade pode ser percebida com base na análise dos resíduos nos modelos. Observa-se que há uma certa constância na variância dos resíduos, em nenhum dos modelos apresenta há aumento significativo na variância dos resíduos ao longo do tempo.

Logo, não há necessidade de correção da heterocedasticidade, que pode ser feito por meio de técnicas como a diferença.



## Avaliando os modelos propostos (AICc)

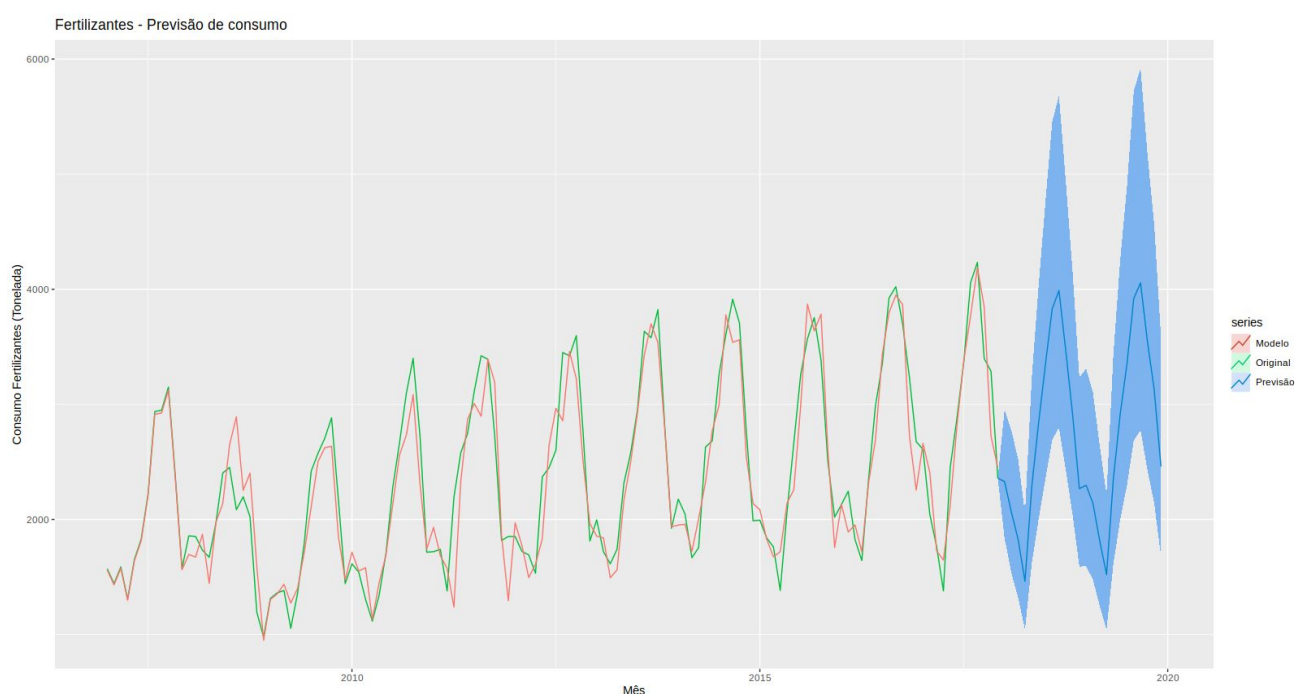
Com base no AICc o modelo que apresentou melhor taxa de AICc foi o Sarima, pois possui menor AICc.

Modelo	AICc	AIC	RMSE
HW - Aditivo	2141.080	2135.712	249.5421
HW - Multiplicativo	2186.532	2181.164	269.6934
SARIMA 1 $ARIMA(1,1,1)(1,1,1)[12]$	-144.14	-144.6717	247.7314
SARIMA 2 $ARIMA(1,0,1)(1,1,1)[12]$	-144.45	-144.9788	255.2942
SARIMA (autosarima) $ARIMA(1,0,0)(2,1,0)[12]$	-151.69	-152.0371	250.2959

## Previsão para os próximos 2 anos (2018 e 2019)

Com base no AICc adotou-se o modelo SARIMA (autoarima) para as previsões.

Temos no modelo o cenário otimista e pessimista, considerando 95% de confiança.



## Cenários

